Й. Й. Білинський, К. В. Огородник, М. Й. Юкиш

ЕЛЕКТРОННІ СИСТЕМИ

Міністерство освіти і науки, молоді та спорту України
Вінницький національний технічний університет

Й. Й. Білинський, К. В. Огородник, М. Й. Юкиш

ЕЛЕКТРОННІ СИСТЕМИ

Навчальний посібник

Вінниця
ВНТУ
2011

УДК 621.38.061
ББК 32.85я73
Б 61

Рекомендовано до друку Вченою радою Вінницького національного технічного університету Міністерства освіти і науки, молоді та спорту України (протокол № 7 від 24 лютого 2011 р.)

Рецензенти:

В. П. Манойлов, доктор технічних наук, професор
О. М. Шинкарук, доктор технічних наук, професор
М. А. Філинюк, доктор технічних наук, профессор

Білинський, Й. Й.

Б61 Електронні системи: навчальний посібник / Й. Й. Білинський, К. В. Огороднік, М. Й. Юкиш. – Вінниця : ВНТУ, 2011. – 208 с.

У навчальному посібнику розглядаються основи теорії електронних систем. Зокрема розглядаються питання: інформаційних оцінок електронних систем, моделювання сигналів та завад, передачі, прийому та реєстрації інформації у електронних системах.

Метою навчального посібника є познайомити читача з сучасними електронних системами та принципами їх побудови. Значна увага приділяється обробці та поданню інформаційних сигналів у електронних системах.

Посібник розроблено відповідно до навчальної програми з дисципліни «Електронні системи» та розраховано для студентів спеціальностей 7.05080101 – «Мікро- та наноелектроніка», 7.05080201 – «Електронні прилади та пристрої».

УДК 621.38.061
ББК 32.85я73

© Й. Білинський, К. Огороднік, М. Юкиш, 2011

ЗМІСТ

ВСТУП 5

1 ЗАГАЛЬНІ ВІДОМОСТІ ПРО ЕЛЕКТРОННІ СИСТЕМИ 6
1.1 Поняття про електронну систему та її характеристики 6
1.2 Класифікація електронних систем 8
1.3 Оптико-електронні системи, основні тенденції розвитку 14
1.4 Приклади електронних систем 17

2 ІНФОРМАЦІЙНІ ОЦІНКИ ЕЛЕКТРОННИХ СИСТЕМ 21
2.1 Інформація та її характеристики 23
2.2 Ентропія джерела інформації 27
2.3 Передача інформації без завад 30
2.4 Передача інформації із завадами 32

3 МОДЕЛЮВАННЯ СИГНАЛІВ ТА ЗАВАД В ЕЛЕКТРОННИХ СИСТЕМАХ 36
3.1 Перетворення сигналів та їх спектральні характеристики 36
3.2 Види модуляції сигналів 48
3.3 Випадкові сигнали та їх характеристики 59
3.4 Білий шум 68
3.5 Проходження випадкових сигналів через ЕС 70

4 ДИСКРЕТИЗАЦІЯ, КВАНТУВАННЯ ТА КОДУВАННЯ СИГНАЛІВ В ЕЛЕКТРОННИХ СИСТЕМАХ 72
4.1 Цифрове подання сигналів 72
4.2 Дискретизація сигналів 79
4.3 Квантування сигналів 82
4.4 Аналого-цифрові та цифро-аналогові перетворювачі 87

5 ПЕРЕДАЧА ІНФОРМАЦІЇ В ЕЛЕКТРОННИХ СИСТЕМАХ 106
5.1 Типова система передачі даних 106
5.2 Канали зв'язку. Аналогові і цифрові канали 108
5.3 Кодування інформації 110
5.3.1 Завадостійке кодування 111
5.3.2 Коди із заданою виправляючою здатністю 115
5.3.3 Систематичні коди 118

5.4 Коди з виявленням помилок 123
5.4.1 Код із перевіркою на парність 123
5.4.2 Код з прямим повторенням 128
5.4.3 Коди з інверсним повторенням (інверсні коди) 130
5.4.4 Кореляційний код 133
5.5 Коди з виявленням і виправленням помилок 136
5.5.1 Код Хеммінга 137
5.5.2 Циклічні коди 142
5.5.3 Рекурентні коди 150
5.6 Кодування із стисненням інформації 152
5.6.1 Код Шеннона-Фано 152
5.6.2 Код Хаффмена 156
5.6.3 Арифметичне кодування 158
5.6.4 Динамічне кодування методом Хаффмена 164
5.6.5 Динамічне кодування методом FGK 168
5.6.6 Динамічне кодування методом Віттера 170
5.7 Методи стиснення зображень 172

6 ПРИЙМАННЯ ІНФОРМАЦІЇ В ЕЛЕКТРОННИХ СИСТЕМАХ 180
6.1. Оптимальне приймання та завадостійкість 180
6.2. Виявлення та приймання інформації 197

СЛОВНИК ТЕРМІНІВ 206
СПИСОК СКОРОЧЕНЬ 209
СПИСОК ЛІТЕРАТУРИ 210

ВСТУП

Стрімкий розвиток науки та техніки, зростання об’ємів інформації, що виробляється людьми в процесі їх життєдіяльності, масове під’єднання користувачів персональних комп’ютерів до глобальних інформаційних мереж, всеохоплююча інформатизація та автоматизація усіх сфер людського життя призводить до бурхливого розвитку електроніки і, як наслідок, до розвитку, створення та постійного вдосконалення електронних систем різного призначення. У сучасному світі жодна з галузей людської діяльності не обходиться без використання певних електронних систем.

У основі роботи будь-якої електронної системи лежить обробка та перетворення отриманого інформаційного або керуючого сигналу та виведення результатів обробки. Тому метою навчального посібника є ознайомити читача з сучасними електронними системами та принципами їх побудови. Значна увага приділяється саме обробці та поданню інформаційних сигналів у електронних системах.

В основу посібника покладено курс лекцій з дисципліни «Електронні системи».

Посібник складається з шести розділів, що містять загальні відомості про електронні системи, інформаційні оцінки електронних систем, моделювання сигналів та завад в електронних системах, принципи дискретизації, квантування та кодування сигналів, відомості про системи передачі, прийому, реєстрації та відображення інформації.

Посібник адресований для широкого кола читачів, що займаються розробкою електронних систем та розрахований для студентів спеціальностей 7.05080101 – «Мікро- та наноелектроніка», 7.05080202 – «Електронні прилади та пристрої» й може бути корисним студентам інших спеціальностей.

1 ЗАГАЛЬНІ ВІДОМОСТІ ПРО ЕЛЕКТРОННІ СИСТЕМИ

1.1 Поняття про електронну систему та її характеристики

Слово *система* (англ. *system*) походить від грецького „складений” і пояснюється як група різних предметів, які об’єднані таким чином, що утворюють єдине ціле та функціонують узгоджено та підпорядковані єдиній формі управління.

Це означення може бути застосоване й до електронної системи (ЕС) в її найзагальнішому вигляді: сукупність електронних компонентів, що пов’язані між собою й діють як одне ціле завдяки спеціальним сигналам управління та виконують задану функцію.

Іншими словами: *електронна система* (англ. *electronic system*) – це будь-який електронний вузол, блок, прилад або комплекс, що виконує обробку інформації.

Таким чином, будь-який пристрій від однокаскадного підсилювача до найскладнішої мікропроцесорної системи може розглядатися як електронна система. Але між мікропроцесорною системою та однокаскадним підсилювачем є суттєва різниця в плані опису деталей кожної із систем. Мікропроцесорну систему у вигляді компонент опорів, конденсаторів, транзисторів, якими описується підсилювач, описати практично неможливо. Тому для опису такої системи необхідно виконати групування ряду компонент у функціональні блоки. Такий функціональний блок носить назву чорний ящик або підсистема.

Електронна підсистема (англ. *electronic subsystem*) – це частина системи, до складу якої входять більше ніж один елемент (проміжний елемент поділу системи) і яка має певне функціональне призначення нижчого рівня, ніж система.

Чорний ящик характеризується його функцією, а не внутрішньою структурою або принципом дії. Крім цього і саму систему можна розглядати як чорний ящик, що виконує певну функцію або набір функцій. Узагальнена електронна система у вигляді чорного ящика, призначеного для обробки вхідних і видачі вихідних сигналів, наведена на рис. 1.1.

Рисунок 1.1 – Узагальнена схема електронної системи як чорний ящик

Елемент електронної системи – це неподільна частина системи.

Часто виникає проблема одночасно мати інформацію про сотні або тисячі одно- і (або) різнорідних фізичних величин, яку потрібно отримати без участі людини, використовуючи технічні засоби, які є досить складними автоматичними системами.

Під автоматизацією розуміють використання пристрою, який за заданою програмою без участі людини виконує всі операції в отриманні, перетворенні, передачі й розподіленні енергії, матеріалів та інформації.

Автоматизована електронна система (англ. *automated electronic system*) – це сукупність керованого об’єкта та автоматичних пристроїв, у якій частину функцій керування виконує людина (оператор).

Автоматична електронна система (англ. *automatic electronic system*) – це сукупність керованого об’єкта і автоматичних пристроїв, що виконують певну задачу згідно із заданою програмою без участі людини.

За призначенням автоматичні системи поділяються на такі класи:

- інформаційно-вимірювальні;
- контролю;
- ідентифікації або розпізнавання образів;
- керування.

Будь-яка система характеризується вирішуваною нею задачею, швидкодією, гнучкістю та надмірністю.

Задача (англ. *task*) – це набір функцій, виконання яких потрібно від електронної системи.

Швидкодія (англ. *speed*) – це показник швидкості виконання електронною системою її функцій.

Гнучкість (англ. *flexibility*) – це здатність системи налаштовуватися під різні задачі.

Надмірність (англ. *excessiveness*) – це показник ступеня відповідності можливостей системи до розв'язуваної даною системою задачі.

Задачі, що вирішує ЕС, відображають мету її функціонування. Досягається ця мета за допомогою вирішення конкретних функціональних задач. Функціональні задачі зручно розділити на чотири рівні.

Перший рівень: задачі збору та попередньої обробки *інформації* (англ. *information*). До них відносять:

– квантування аналогових сигналів за часом та рівнем;

– попередню цифрову чи аналогову фільтрацію сигналів;

– перетворення та обчислення спектрів;

– нормалізацію, підсилення чи послаблення сигналів;

– зміну рівнів сигналів;

– перетворення струму в напругу.

Другий рівень:

– геометричні перетворення систем координат;

– визначення змінних параметрів вимірюваних процесів;

– сумісне розв’язування лінійних та нелінійних систем рівнянь;

– рішення систем диференціальних рівнянь;

– інтерполяція та екстраполяція функцій, що обчислюються;

– цифро-аналогові перетворення.

Третій рівень: розв’язування задач обробки інформації. До таких задач відносять задачі оптимізації, тобто пошуку екстремумів деякої функції чи декількох змінних, на які накладено певні обмеження.

Четвертий рівень: розв’язування задач подання інформації в зручному для сприйняття оператором вигляді.

1.2 Класифікація електронних систем

Вхідними та вихідними *сигналами* (англ. *signal*) електронної системи можуть бути *аналогові сигнали* (англ. *analog signal*), одиничні *цифрові сигнали* (англ. *digital signal*), цифрові коди, послідовності цифрових кодів. Відповідно системи можуть бути аналоговими, цифровими або комбінованими, тобто аналого-цифровими.

Якщо система аналого-цифрова, то вхідні аналогові сигнали перетворюються в послідовності кодів вибірок за допомогою аналого-цифрового перетворювача (АЦП), а вихідні аналогові сигнали формуються з послідо-

вності кодів вибірок за допомогою цифро-аналогового перетворювача (ЦАП). Обробка й зберігання інформації виконуються в цифровому вигляді.

За своєю будовою електронні системи поділяють на системи на “жорсткій логіці” та мікропроцесорні системи.

Характерною особливістю традиційної цифрової системи на відміну від мікропроцесорної є те, що алгоритми обробки й зберігання інформації в ній жорстко зв'язані зі схемотехнікою системи. Тобто, зміна цих алгоритмів можлива тільки шляхом зміни структури системи, заміни електронних вузлів, що входять у систему, і/або зв'язків між ними. Саме тому традиційна цифрова система часто називається системою на "жорсткій логіці".

Будь-яка система на "жорсткій логіці" обов'язково є спеціалізованою системою, налаштованою винятково на одну задачу або (рідше) на декілька близьких, заздалегідь відомих задач. Це має свої безперечні переваги.

По-перше, спеціалізована система (на відміну від універсальної) ніколи не має апаратурної надмірності, тобто кожен її елемент обов'язково працює на повну потужність (звичайно, якщо ця система грамотно спроектована).

По-друге, саме спеціалізована система може забезпечити максимально високу швидкодію, тому що швидкість виконання алгоритмів обробки інформації визначається в ній тільки швидкодією окремих логічних елементів й обраною схемою шляхів проходження інформації. А саме, логічні елементи завжди мають максимальну на даний момент швидкодією.

Але в той же час великим недоліком цифрової системи на "жорсткій логіці" є те, що для кожної нової задачі її необхідно проектувати й виготовляти заново. Це процес тривалий, дорогий та потребує високої кваліфікації виконавців. Шлях подолання цього недоліку досить очевидний: необхідно побудувати таку систему, що могла б легко адаптуватися під будь-яку задачу, перебудовуватися з одного алгоритму роботи на інший без зміни апаратури. І задавати той або інший алгоритм шляхом уведення в систему додаткової керуючої інформації, програми роботи системи (рис. 1.2). Тоді система стане універсальною або програмованою, не жорсткою, а гнучкою. Саме це й забезпечує мікропроцесорна система.

Мікропроцесор (англ. *microprocessor*) – програмно-керований пристрій, призначений для обробки цифрової інформації й керування проце-

сом цієї обробки, виконаний у вигляді однієї (або декількох) інтегральної схеми з високим ступенем інтеграції електронних елементів.

Зменшення вартості, споживаної потужності й габаритних розмірів, підвищення надійності й продуктивності мікропроцесорів сприяли значному розширенню сфери їхнього використання. Поряд із традиційними обчислювальними системами вони все частіше стали використовуватися в задачах керування та обробки. При цьому перед мікропроцесором ставилася задача програмного управління різними периферійними об'єктами в реальному масштабі часу.

Спрощена структурна схема мікропроцесорної системи управління має вигляд (рис.1.2).

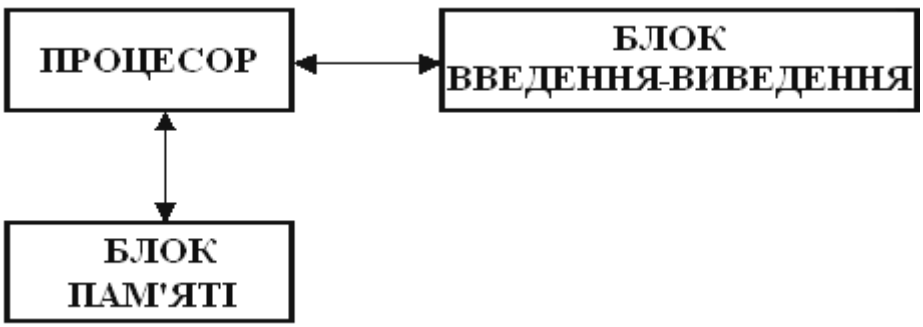


Рисунок 1.2 – Структурна схема мікропроцесорної системи

На процесор покладається задача виконання всіх програмних дій, необхідних відповідно до алгоритму роботи пристрою. У блоці пам'яті зберігаються команди програми функціонування процесора, а також значення констант і змінних величин, що беруть участь в обчисленнях. Блок введення-виведення виконує функцію з’єднання мікропроцесорної системи з об'єктом керування.

Широке використання мікропроцесорної техніки саме для завдань керування привело до появи на ринку спеціалізованих мікропроцесорних пристроїв, орієнтованих на подібного роду застосування. Особливістю цих мікросхем є те, що крім власне процесора, на цьому ж кристалі розташована й система введення-виведення, що дозволяє знизити функціональну складність і габаритні розміри мікропроцесорної системи керування. Подібні пристрої називаються мікроконтролерами.

Мікроконтролер (англ. *microcontroller*) – обчислювально-керуючий пристрій, призначений для виконання функцій логічного контролю й керування периферійним устаткуванням, виконаний у вигляді однієї ВІС, що сполучає в собі мікропроцесорне ядро і набір вбудованих пристроїв введення-виведення.

У сучасній перетворювальній техніці мікроконтролери виконують не тільки роль безпосереднього керування напівпровідниковим перетворювачем за рахунок вбудованих спеціалізованих периферійних пристроїв, але й роль цифрового регулятора, системи захисту й діагностики, а також системи зв'язку з технологічною мережею вищого рівня.

Останнім часом з'явився ряд мікроконтролерів, спеціалізованих для завдань керування напівпровідниковими перетворювачами. Їхнє обчислювальне ядро, побудоване, як правило, на базі так званих "процесорів цифрової обробки сигналів", адаптовано на виконання рекурентних поліноміальних алгоритмів цифрового регулювання. Вбудовані периферійні пристрої містять у собі багатоканальні генератори ШІМ-сигналів, аналого-цифрові перетворювачі, блоки векторних перетворень координат, таймери-лічильники тощо. Прикладами таких пристроїв можуть служити мікроконтролери ADMC330 фірми Analog Devices, TMS320C240 фірми Texas Instruments, 56800 фірми Motorola, векторний співпроцесор ADMC200 фірми Analog Devices.

Але будь-яка універсальність обов'язково призводить до надмірності. Адже вирішення максимально важкої задачі потребує набагато більше засобів, ніж розв’язання максимально простої задачі. Тому складність універсальної системи повинна бути такою, щоб забезпечити розв’язання найважчої задачі, а при розв’язанні простої задачі система буде працювати далеко не на повну силу, буде використовувати не всі свої ресурси. І чим простіша розв'язувана задача, тим більша надмірність і тем менш виправданою стає універсальність. Надмірність веде до збільшення вартості системи, зниження її надійності, збільшення споживаної потужності й т. д.

Крім того, універсальність, як правило, призводить до істотного зниження швидкодії. Оптимізувати універсальну систему необхідно так, щоб кожна нова задача розв’язувалася максимально швидко. Загальне правило таке: чим більша універсальність, гнучкість, тим менша швидкодія. Більше того, для універсальних систем не існує таких задач, які б вони розв’язували з максимально можливою швидкодією.

Таким чином можна зробити такий висновок. Системи на "жорсткій логіці" ефективні там, де розв'язувана задача не змінюється тривалий час, де потрібна найвища швидкодія, де алгоритми обробки інформації граничні прості. А універсальні, програмовані системи ефективні там, де часто змінюються розв'язувані задачі, де висока швидкодія не є занадто важливою, де алгоритми обробки інформації досить складні.

Однак за останні десятиліття швидкодія універсальних (мікропроцесорних) систем дуже виросла (на кілька порядків). До того ж великий обсяг випуску мікросхем для цих систем привів до різкого зниження їхньої вартості. У результаті область застосування систем на "жорсткій логіці" різко звузилася. Більш того, високими темпами розвиваються зараз програмовані системи, призначені для розв’язання одної задачі або декількох близьких задач. Вони вдало об’єднують у собі як переваги систем на "жорсткій логіці", так і програмованих систем, забезпечуючи досить високу швидкодію й необхідну гнучкість.

В науково-технічній літературі системи, що використовують у своєму складі ЕОМ в залежності від області їх застосування, а також від класу задач, які розв’язуються ними, називаються комп’ютерними або комп’ютеризованими системами.

Таким чином, *комп’ютеризовані системи* (англ. *computer aided system*) – це системи, під якими розуміють сукупність технічних засобів, яка має в своєму складі керуючу ЕОМ, засоби збору, перетворення, передачі, відображення, а також спеціальне математичне та програмне забезпечення, що виконує весь комплекс обробки інформації.

Вважається, що ЕОМ треба вводити до складу автоматичної системи тоді, коли кількість отриманої інформації стає дуже великою, а обробка інформації ускладнюється або стає занадто тривалою у часі й не відповідає вимогам або швидкодії системи. В таких системах передбачені підсистеми мультиплексування інформації, вимірювальні канали та канали зв’язку.

Узагальнена схема комп’ютеризованої системи наведена на рис. 1.3.

Вимірювальний канал (англ. *measuring channel*) – це засіб вимірювання, який є складовою частиною вимірювальних систем у вигляді сукупності засобів вимірювальних операцій, засобів (каналів або ліній) зв’язку для послідовних перетворень і передачі вимірювальної інформації.

Канал зв’язку (англ. *channel of communication*) – це сукупність передавача, приймача та лінії зв’язку, призначених для передачі інформації від відправника (джерела) до адресата.

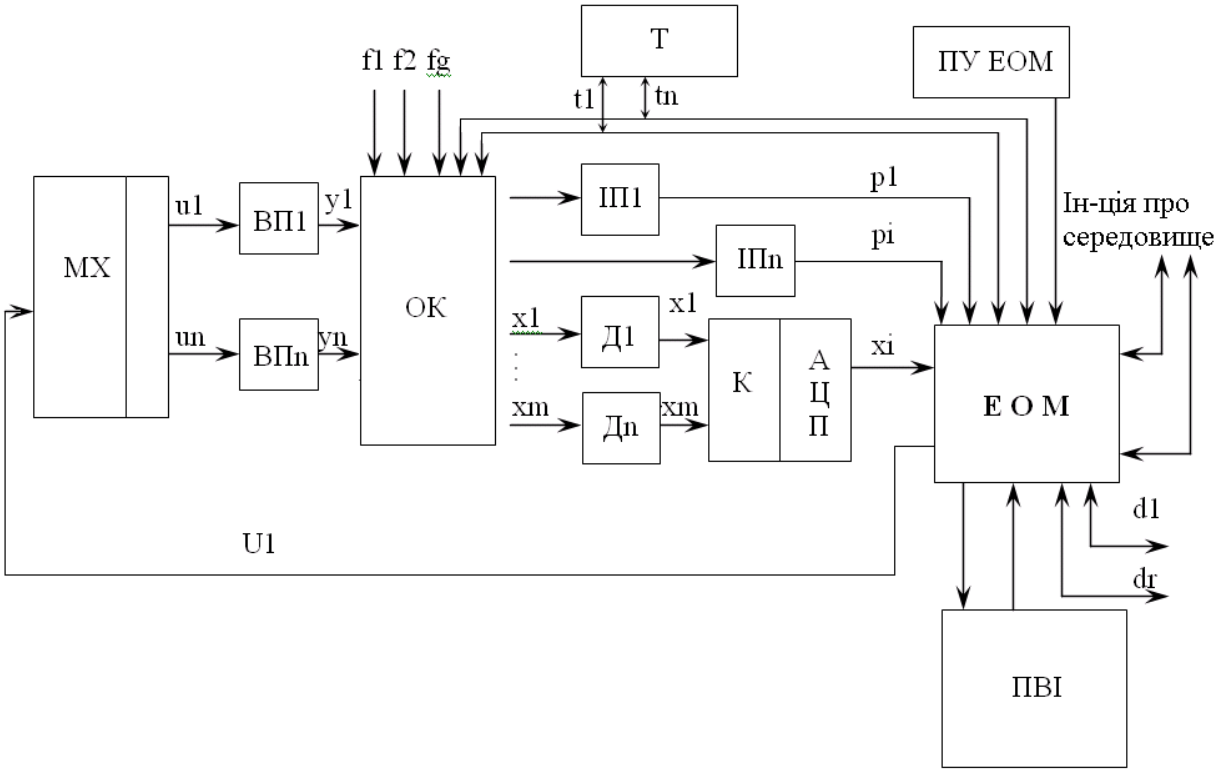


Рисунок 1.3 – Узагальнена схема комп’ютеризованої системи:
МХ – мультиплексор; ОК – об’єкт керування (дослідження);
Д1, Дn – датчики; Хі,Хп – інформація; ВП – виконавчий пристрій;
ІП – індикатори повідомлень; u1, un – сигнал керування; Т – таймер;
ПУ ЕОМ – пульт управління ЕОМ; ПВІ – пристрій
відображення інформації.

Лінія зв'язку (англ. *communication line*) – це фізичне середовище, по якому передаються сигнали.

Програмований канал зв’язку або інтерфейс (англ. *interface*) – це узгодженість обміну інформацією та правил обміну інформацією, під якою розуміють електричну, логічну й конструктивну сумісність пристроїв, що беруть участь в обміні.

1.3 Оптико-електронні системи, основні тенденції розвитку

Комп’ютеризовані оптико-електронні системи (КОЕС) дозволяють скоротити обсяг оброблюваної інформації при збереженні якості зображення, автоматизувати процес вимірювання, що приводить до підвищення точності та усунення суб’єктивних помилок спостерігачів.

До переваг КОЕС у порівнянні з традиційними приладами можна віднести:

– високу чутливість;

– стійкість до електромагнітних завад і радіаційних впливів;

– відсутність електричних кіл, наявність гальванічної розв'язки між чутливим й виконавчими елементами;

– пожежо- та іскробезпеку;

– стійкість до агресивних впливів, термічну стабільність й корозійну стійкість;

– різноманітність геометричних форм і розмірів;

– можливість створення інтегральних оптичних первинних вимірювальних перетворювачів;

– дистанційність вимірювань, можливість сумісної роботи з волоконно-оптичними лініями зв'язку.

До основних тенденцій розвитку КОЕС відносяться:

– подальше вдосконалення параметрів і характеристик окремих елементів і пристроїв (приймачів і джерел випромінювання, оптичних вузлів, електронних елементів);

– підвищення експлуатаційної надійності окремих елементів і пристроїв в цілому;

– мікромініатюризація, зменшення маси, габаритів, енергії, яка споживається окремими ланками інформаційного каналу;

– удосконалення методів попередньої обробки даних.

Дія вимірювального каналу (ВК) КОЕС базується на прийомі і перетворенні електромагнітного випромінювання в різних діапазонах оптичної області спектра, тобто в ультрафіолетовій, видимій та інфрачервоній частинах.

Одна з можливих узагальнених схем оптико-електронного вимірювального каналу КОЕС наведена на рис. 1.4, основними елементами якої є оптична система та приймач сигналу.

Розробка ВК передбачає обізнаність з особливостями об'єкта, умовами і режимами роботи. Важливою умовою успішного створення і ефективного використання таких систем є повнота уявлення про фізико-метрологічні властивості об'єкта.

Різноманітність різних за своєю природою об'єктів і різномасштабність полів фізичних величин зумовили створення і використання різноманітних КОЕС, що дало змогу удосконалити класифікацію за ознаками, наведеними на рис. 1.5.

Для розв'язання задач виявлення оптичних сигналів і образів, на-

приклад зображення досить складного за формою, спектром й іншими ознакам об’єкта, що знаходиться на складному фоні, у сучасних КОЕС майже повсюдно використовується такий опис об’єктів (сукупність оброблюваних сигналів), що містить лише обмежене число відмінних ознак.

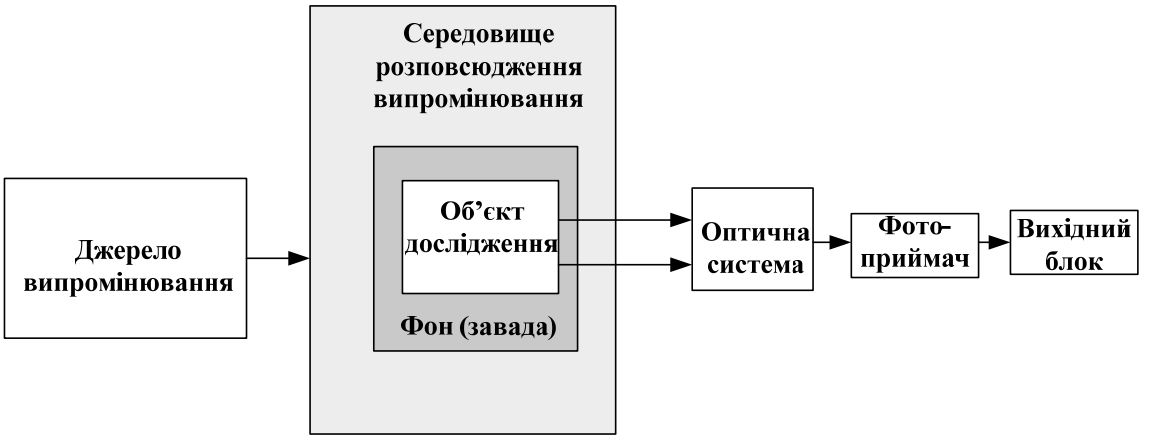


Рисунок 1.4 – Узагальнена структурна схема ВК оптико-електронної системи

Вибір ознак, які найбільш істотно відрізняють даний клас об’єктів (зображень, сигналів), є найважливішою задачею при розробці КОЕС та удосконаленні існуючих. Тому дуже важливо відібрати таку мінімальну кількість ознак, яка б забезпечувала задані показники якості роботи КОЕС, але не ускладнювала їхню конструкцію і тим самим не знижувала надійність роботи систем і не здорожчувала їхнє виробництво та експлуатацію.

Найбільш часто використовуються такі групи ознак:

– геометричні, виділення й обробка яких залежить насамперед від просторової роздільної здатності вимірювального каналу КОЕС; до цих ознак відносяться розміри і форма зображення; гістограми розподілів кутів, хорд, довжин сторін; площа, центр мас, периметр, компактність, орієнтація головної осі, геометричні моменти; просторово-частотні спектри Фур’є і Мелліна; функції Уолша тощо;

– спектральні, виділення й обробка яких залежить від спектральної роздільної здатності каналу КОЕС; до них відносяться поглинальна, випромінювальна і відбивна здатність, колір, колір/яскравість тощо;

– енергетичні, що характеризуються відношенням сигнал/шум;

– динамічні, що використовують інформацію про зміну координат об’єкта, про швидкість його переміщення тощо.

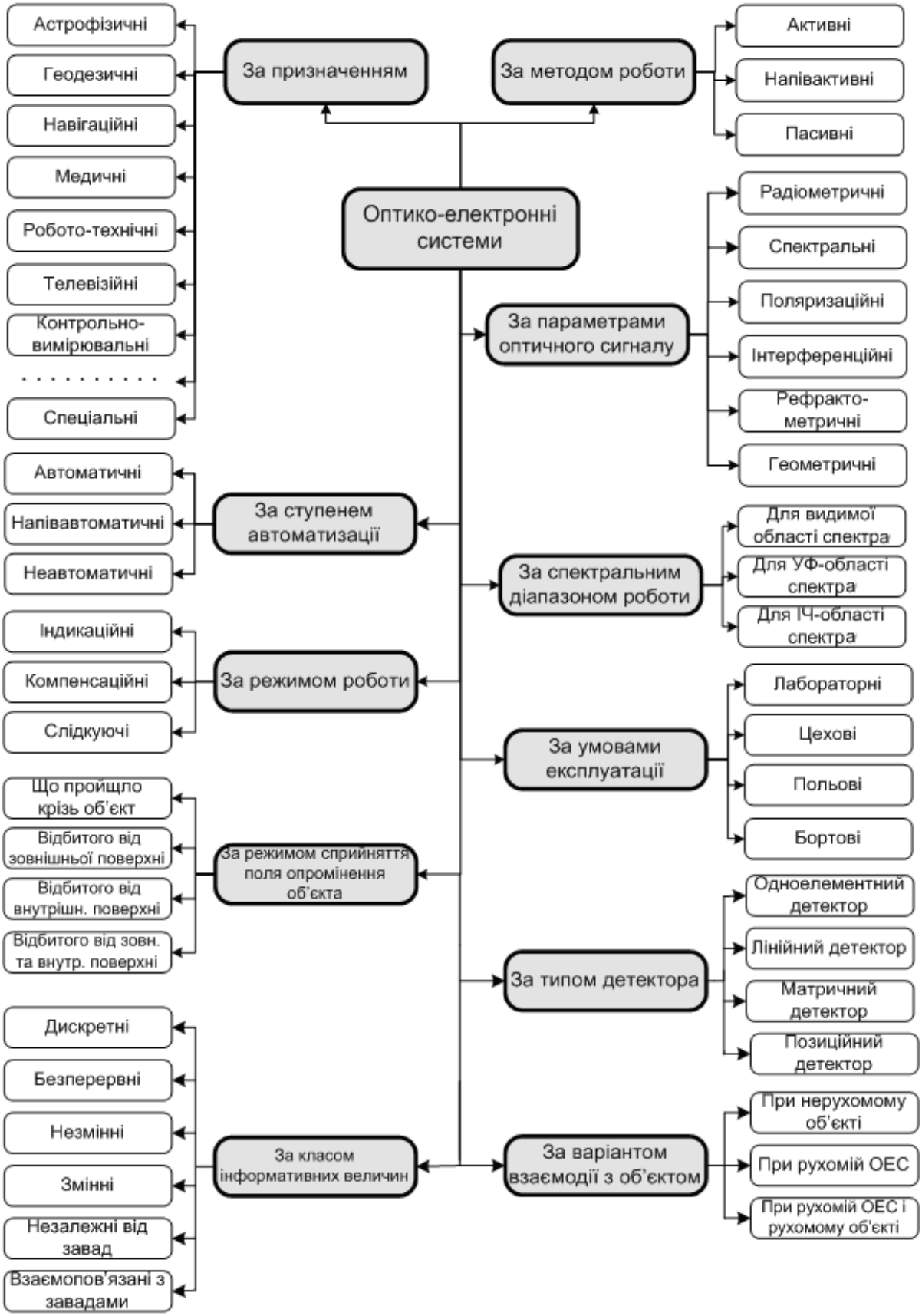


Рисунок 1.5 – Класифікація комп'ютеризованих оптико-електронних систем

КОЕС вимірювання геометричних параметрів об’єктів, які використовують для своїх задач розподіл інтенсивності отриманого зображення об’єкта, як правило, виділяють певним чином ранжовані елементи зображення. До таких елементів відноситься зображення світлової плями, яка може бути самостійним елементом зображення або складовою частиною елементів більш високого рангу. Інформаційним елементом є також світлова пляма, яка також може бути самостійним елементом або частиною більш високого рангу, що потребує вимірювання її параметрів. Специфічним елементом зображення є лінія, яка може розглядатися або як область, довжина якої значно переважає ширину, або як подвійна межа.

Наступним за рангом і за величиною є перепад інтенсивності, який являє собою елемент більш високого порядка, межа якого розділяє дві сусідні області інтенсивності. Якщо перепад одновимірний, то межа має певну протяжність і форму. В результаті виділення даного перепаду формується контур об’єкта, що має відмінну від фону інтенсивність.

1.4. Приклади електронних систем

Більшість сучасних електронних систем є вбудованими. Робот, такий як марсохід, є вбудованою системою. Стільниковий телефон, PDA, портативний мультимедіа програвач є вбудованими пристроями. Навіть електрична зубна щітка являє собою вбудованну систему. Невеликий мікроконтролер у зубній щітці забезпечує програмне керування швидкістю й індикацію стану заряду батареї. Сучасні автомобілі можуть містити понад сто вбудованих мікроконтролерів. У цілому вбудовані системи складають більшу частину світового виробництва мікропроцесорів.

Як видно з таблиці 1.1, вбудовані системи можна знайти в різноманітних продуктах, включаючи літаки і військові системи, біомедицинські системи, автомобілі, комунікації, комп'ютерні пристрої, електронні інструменти, домашня електроніка, промислове устаткування, офісні машини, персональні пристрої, роботи та інтелектуальні іграшки. Вбудовані електронні системи можна знайти всюди.

Конструктори вбудованих систем часто зіштовхуються зі складними проектними задачами, оскільки вони повинні бути надійними. Багато з них не можуть підлягати ремонту і не можуть бути перезавантажені. Програмне забезпечення неможливо обновити в багатьох вбудованих пристроях. Багато систем мають жорсткі конструкційні обмеження за продуктивністю і споживанням енергії. Деяким системам необхідно працювати від ба-

тареї тривалий період часу. У багатьох застосуваннях існують обмеження реального часу і багато систем мають обмежену пам'ять і обчислювальну потужність.

Таблиця 1.1 – Приклади вбудованих систем

Авіаційні та військові системи	Автопілоти літаків, авіоніка і навігаційні системи, системи автоматичної посадки, системи наведення, керування двигуном
Біомедицинські системи	Системи комп'ютерної томографії й ультразвукового дослідження, моніторинг пацієнтів, кардіостимулятори
Автомобілі	Керування двигуном, антиблокувальні гальмові системи, керування подушками безпеки, керування системою обігріву і кондиціонування повітря навігація GPS, супутникове радіо, системна діагностика
Комунікації	Комунікаційні супутники, мережні маршрутизатори, комутатори, концентратори
Споживча електроніка	телевізори, духовки, програвачі DVD, стереосистеми, системи безпеки, керування поливом газонів, термостати, фотокамери, автовідповідачі, TV-декодери
Пристрої в/для комп'ютера	Клавіатури, миші, принтери, сканери, дисплеї, модеми, пристрої жорстких дисків, пристрої DVD, графічні плати, пристрої USB
Електронні інструменти	Системи збору даних, осцилографи, вольтметри, генератори сигналів, логічні аналізатори
Промислове устаткування	Керування ліфтами, системи спостереження, роботи, верстати з ЧПК, програмувальні логічні контролери, промислові системи автоматизації і керування
Офісні машини	факси-апарати, копіри, телефони, калькулятори, касові апарати
Персональні пристрої	стільникові телефони, переносні програвачі MP3, персональні цифрові помічники (PDA), електронні наручні годинники, портативні відеоігри, цифрові камери, системи GPS
Роботи	Промислові роботи, автономні транспортні засоби, космічні дослідницькі роботи
Іграшки	системи відеоігор, іграшки-роботи типу "Aibo", "Furby", "Elmo"

Електронні системи реального часу

Системи реального часу (англ. *real time system*) повинні реагувати на зовнішні параметри введення і створювати нові результати виведення за обмежений час, як показано на рис. 1.6. Час відповіді повинен бути обмежений. Дуже тривалий час відповіді може призвести до відмови систем реального часу.

Ілюстративним прикладом системи реального часу є контролер автомобільної повітряної подушки безпеки. Коли сенсори руху повітряної подушки (акселерометри) розпізнають зіткнення, системі необхідно зреагувати, розкриваючи повітряну подушку протягом 10 мс або ж система не спрацює потрібним чином. На високій швидкості із затримкою більш 10 мс водій уже зіштовхнеться з кермовим колесом до того, як розкриється подушка.

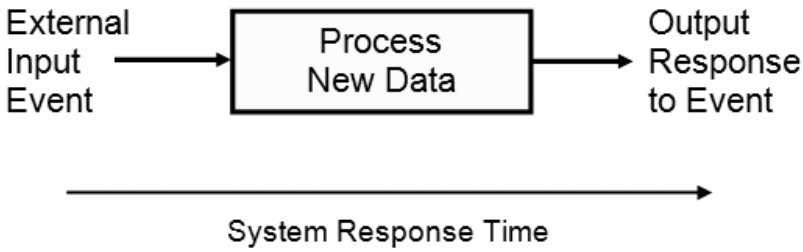


Рисунок 1.6 – Система реального часу

У “м'якій” системі реального часу пріоритети мають критично важливі задачі. М'яка система реального часу зазвичай задовольняє обмеження відгуку реального часу. Прикладом типової м'якої системи реального часу є програвач мультимедіа. Програвач може іноді пропустити відеокадр або аудіозвук і користувач може це навіть не помітити.

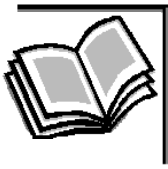
У жорсткій системі реального часу новий результат виведення завжди повинен бути обчислений у зазначених границях часу або система не спрацює. Як приклад жорсткої системи реального часу розглянемо систему дистанційного керування кермом (тобто, керовану комп'ютером). У системі керування польотом літака, коли льотчик переміщає штурвал керування, рулі керування польотом повинні у відповідь перемістити дуже швидко або ж літак втратить стійкість і впаде. Щоб забезпечити безпеку FAA постійно перевіряє і сертифікує реакцію в реальному часі керованих комп'ютером симуляторів польоту і літаки.

Багато вбудованих систем є системами реального часу з декількома входами і виходами. Кілька подій відбуваються незалежно одна від одної. Програмування таких ситем спрощується при поділі задач, але це потребує

від центрального процесора постійного переключення між різними задачами. Таке переключення та поділ часу процесора між декількома задачами зазвичай забезпечують багатозадачні операційні системи. Вони забезпечують також елементи синхронізації, необхідні для координації дій між різними задачами, що виконуються паралельно.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Які основні характеристики електронних систем?
2. На які рівні поділяють функціональні задачі в електронних системах?
3. У чому відмінність мікропроцесорної системи від системи на “жорсткій логіці”?
4. Які недоліки електронних систем на “жорсткій логіці”?
5. Узагальнена структурна схема ВК оптико-електронної системи.
6. Наведіть основні переваги КОЕС.
7. У чому відмінність каналу зв’язку від лінії зв’язку?
8. Які групи ознак найбільш часто використовують при класифікації ЕС?
9. Наведіть узагальнену структурну схему ВК КОЕС.
10. Що означає поняття “вбудована система”?
11. У чому специфіка електронних систем реального часу?



2 ІНФОРМАЦІЙНІ ОЦІНКИ ЕЛЕКТРОННИХ СИСТЕМ

Поняття інформації відноситься до основних понять науки і тісно пов'язане з такими поняттями, як «інформаційні ресурси», «інформаційні технології», «інформаційне забезпечення», «інформаційні системи», «інформаційні структури» тощо. Для того, щоб дати найбільш повне визначення інформації, необхідно звернути увагу на те, що в її основі лежить взаємозалежність пари об'єктів – джерела і споживача інформації. Джерелами інформації, насамперед, є природні об'єкти: люди, тварини, рослини, планети та ін. Разом із цим в міру розвитку науки і техніки джерелами інформації стають наукові експерименти, машини, апарати, технологічні процеси. Значний також перелік об'єктів, що є споживачами інформації: люди, тварини, рослини, різноманітні технічні пристрої тощо.

Виходячи зі сказаного, дають таке означення інформації: **інформація** – «сукупність відомостей про всілякі об'єкти, явища, процеси». Більш змістовним є твердження: інформація є вибір одного варіанта (або декількох) із багатьох можливих і рівноправних. Слово «рівноправних» означає, що усі варіанти, із яких робиться вибір, мають щось спільне, тобто належать одній множині. У наукових працях розглянуто три основних варіанти вказаного вибору: якщо вибір підказаний (або вказаний), то мова йде про одержання (рецепцію) інформації; якщо вибір зроблений самостійно і випадково, то говорять про виникнення інформації; і, нарешті, якщо вибір однозначно визначений ситуацією або попередніми подіями (тобто вибору, по суті, немає), то про інформацію взагалі говорити не доводиться.

Основними *формами подання інформації є символьна, текстова і графічна*.

Символьна форма ґрунтується на використанні символів – літер, цифр, знаків і т. д., є найбільш простою і практично застосовується тільки для передачі сигналів про різні події.

Більш складною є текстова форма подання інформації. Тут, як і в попередній формі, використовуються символи – літери, цифри, математичні знаки. Однак інформація закладена не тільки в цих символах, але й у їх сполученні, порядку проходження. Завдяки взаємозв'язку символів і відображенню мови людини текстова інформація надзвичайно зручна і широко

використовується в повсякденному житті.

Найбільш місткою, але і найбільш складною, є графічна форма подання інформації. Образи природи, фотографії, креслення, схеми, малюнки мають велике значення в нашому житті і містять величезну кількість інформації. І хоча інформація не має ні маси, ні геометричних розмірів, жодних фізичних або хімічних властивостей, проте для її існування обов'язкова наявність якогось матеріального об'єкта, що передає або зберігає інформацію. Таких об'єктів досить багато, і їхня кількість увесь час зростає.

Можливі різні класифікації за різними ознаками. Як приклад можна навести одну з класифікацій соціальної інформації, згідно з якою виділяються два основних класи – масова (загальна) і спеціальна (системна) інформація.

Масова інформація – соціальна інформація, адресована всім членам суспільства незалежно від їх становища і роду занять.

Спеціальна інформація адресована не всім членам суспільства, а певним соціальним групам (наприклад, ученим даного фаху, економістам, робітникам якоїсь професії і т. д.). Для сприйняття цієї інформації необхідний початковий запас спеціальних знань і володіння професійною мовою. Назвемо деякі найбільш важливі різновиди спеціальної соціальної інформації.

1. Наукова інформація – утворюється в результаті науково-дослідної діяльності. Наукову інформацію можна визначити як передане в інформаційному процесі наукове знання.

2. Технічна інформація – створюється в сфері техніки і призначена для вирішення технічних задач (розробки нових технічних пристроїв, машин, матеріалів і т. п.). Структура і властивості наукової і технічної інформації дуже близькі, тому ці два види часто об'єднують терміном «науково-технічна інформація».

3. Технологічна інформація – безпосередньо використовується в сфері матеріального виробництва для створення матеріальних благ (продуктів харчування, одягу, машин тощо).

4. Планово-економічна інформація – наприклад, про стан і перспективи розвитку промислових підприємств – використовується для планування і керування виробництвом.

Кожне повідомлення містить як умовну, так і безумовну інформацію. Визначити, яка інформація є умовною і яка безумовною, не так просто, як це здається на перший погляд.

Тут відіграють роль такі обставини:

- по-перше, умовна інформація має тенденцію до уніфікації, що природно, оскільки цінність її при цьому зростає;

- по-друге, уніфікована умовна інформація часто сприймається як безумовна;

- по-третє, найбільш цікавим і гострим залишається питання про умовність (або безумовність) інформації в природничих науках; прийнято вважати, що, вивчаючи природу, ми одержуємо (рецепіюємо) об'єктивну (тобто безумовну) інформацію.

Відомі дві основних форми існування інформації – **статична** (у вигляді записів на папері, стрічці, диску, фотопапері тощо) та **динамічна** – під час її передавання. Потрібно зауважити, що процес фізичного перевезення чи пересування носія інформації (листа, магнітної стрічки, диска, касети тощо) не відноситься до динамічної форми існування інформації. Якщо дані передаються каналом зв’язку, то у кожній точці каналу під час передавання процес змінюється в часі і так само змінюється вплив зовнішніх факторів на сигнали, що несуть в собі інформацію. При фізичному перевезенні цього не відбувається, хоча дані, що зафіксовані на носію, теж підпадають під вплив зовнішніх факторів і можуть руйнуватися з часом. Таким чином, статичною цю форму можна назвати відносно. Більш точне визначення – квазістатична.

Інформація, що зберігається на носію, може зчитуватись, передаватись, знов записуватись, тобто вона може багаторазово переходити з однієї форми існування до іншої.

Основна проблема – передавання інформації з найменшими втратами. При цьому необхідно оцінювати інформацію кількісно.

2.1 Інформація та її характеристики

Перша спроба ввести науково обґрунтовану міру інформації була зроблена в 1927 році Р. Хартлі (Англія). Він запропонував та обґрунтував кількісну міру, яка дозволяє порівнювати спроможність різних систем передавати інформацію. Ця міра підходить і для систем зберігання інформації, тому вона є відправною точкою для створення теорії інформації.

Природною вимогою, що висувається до інформаційної міри, є вимога адитивності, тобто, кількість інформації, яка може бути збережена у двох однакових комірках, повинна бути удвічі більшою за ту, що зберіга-

ється в одній з них.

Якщо одна комірка для зберігання інформації має m можливих станів, то дві таких комірки будуть мати m^2 можливих станів, а n однакових комірок – m^n можливих станів. Це саме стосується і кількості можливих повідомлень. Якщо символ може прийняти значення «0» або «1», то з одного символу можуть бути одержані 2 повідомлення, з двох символів – 4, з трьох – 8 тощо. Таким чином кількість можливих повідомлень визначається кількістю символів, що входять до слова n та кількістю можливих станів символу m: m^n. Тому Р. Хартлі ввів логарифмічну міру інформаційної ємності:

$$C = \log m .$$

Така міра задовольняє вимогу адитивності. Ємність засобу, що складається з n комірок і має m^n станів, дорівнює ємності однієї комірки, помноженої на їх кількість:

$$C_n = \log m^n = n \log m .$$

За одиницю вимірювання інформаційної ємності вибрана двійкова одиниця – *біт* (binary digit – двійковий знак), що дорівнює ємності однієї комірки з двома можливими станами. Інформаційна ємність C у двійкових одиницях в загальному випадку визначається як:

$$Ca_2 = k_a \log_a m ,$$

де k_a – коефіцієнт, що залежить від основи логарифму a.

При використанні для зберігання інформації десяткових комірок більш зручно користуватись десятковими логарифмами. В цьому випадку:

$$K_{10} = log_{\ 2}\ 10 \approx 3{,}32,$$

тобто, одна десяткова комірка за інформаційною ємністю дорівнює 3,32 двійковим. Одиниця вимірювання кількості інформації в цьому випадку – **діт**.

Якщо від джерела інформації каналом зв'язку передається повідомлення про подію, апріорна імовірність якої на передавальному боці дорівнювала p_1, то після приймання повідомлення апостеріорна імовірність цієї події для приймача інформації дорівнює p_2. Збільшення кількості інформації з урахуванням логарифмічної міри складає:

$$\Delta I = \log\left(\frac{p_2}{p_1}\right) = \log p_2 - \log p_1 .$$

Для ідеального каналу зв'язку (без завад та спотворень) приймання інформації є вірогідною подією, тобто імовірність p_2 дорівнює одиниці:

$$\Delta I = -\log p_1 . \qquad (2.1)$$

Чим меншою буде імовірність p_1, тим більшою буде невизначеність результату, тобто тим більша кількість інформації вміщується у прийнятому повідомленні.

Значення p_1 знаходиться у межах $0 < p_1 < 1$, тобто ΔI завжди додатна величина. Якщо припустити, що може передаватися n_a символів S_a, що відповідають події A, n_b символів S_b, що відповідають події B тощо, а всього m різних символів. Символи S_a, S_b тощо являють собою алфавіт з різних m символів. Сума усіх символів q складає:

$$q = n_a + n_b + ...$$

Згідно з (2.1) приймання символу S_a дає кількість інформації:

$$\Delta I = -\log p_a ,$$

де p_a – імовірність події A.

Тоді у n_a символах міститься кількість інформації $n_a(- \log p_a)$. Загальна кількість інформації складає:

$$I = (-n_a \log p_a - -n_b \log p_b - ...) = \sum_{i=1}^{m} n_i \log p_i . \qquad (2.2)$$

Вираз для визначення середньої кількості інформації, що припадає на один символ, можна отримати, розділивши (2.2) на q:

$$I_i = \sum_{i=1}^{m} \frac{n_i}{q} \log p_i . \qquad (2.3)$$

У (2.3) відношення $\frac{n_i}{q}$ є апріорною імовірністю появи символу S_a для великих значень n_i та q, $\frac{n_i}{q}$ – імовірність символу S_b тощо.

Тоді:

$$\lim_{q \to \infty} \left(\frac{n_s}{q} \right) = p_i .$$

При цьому сума імовірностей: $p_a + p_b + ... = 1$, оскільки одна з усіх m подій A, B, ... відбувається обов'язково (повна імовірність подій).

Таким чином, можна отримати вираз для середньої кількості інформації на один символ:

$$I_i = \sum_{i=1}^{m} \frac{n_i}{q} \log p_i ,$$

де p_i – імовірність і-того символу.

ПРИКЛАД

Необхідно підняти вантаж на певний поверх 16-ти поверхового будинку (нумерація поверхів 0-15, N = 16). Скільки біт інформації повністю визначають завдання?

$$I = Log_2 16 = 4 .$$

Отже, 4 біти інформації необхідні й достатні для повного зняття невизначеності вибору. У цьому можна переконатися застосуванням логіки вирахування з послідовним розподілом навпіл інтервалів станів. Наприклад, для 9-го поверху:

1) вище 7-го поверху? Так = 1;

2) вище 11-го поверху? Ні = 0;

3) вище 9-го поверху? Ні = 0;

4) вище 8-го поверху? Так = 1.

Підсумок: поверх номер 9 або 1001 у двійковому численні, чотири двійкових розряди.

Якщо в наведеному прикладі на поверхах є по 4 квартири з нумерацією на кожному поверсі 0-3 (M=4), то при адресації вантажу у квартиру буде потрібно ще 2 біти інформації. Такий же результат одержимо, якщо замість незалежної нумерації поверхів і квартир на поверхах (два джерела невизначеності) ми будемо мати тільки наскрізну нумерацію квартир (одне узагальнене джерело):

$$I = Log_2 N + Log_2 M = Log_2 16 + Log_2 4 = Log_2 (NM) = Log_2 64 = 6 ,$$

тобто, кількість інформації відповідає вимозі адитивності: невизначеність об'єднаного джерела дорівнює сумі невизначеностей вихідних джерел, що відповідає інтуїтивній вимозі до інформації: вона повинна бути однозначною, а її кількість повинна бути тією самою незалежно від способу задання.

2.2 Ентропія джерела інформації

Ентропія являє собою логарифмічну міру безладдя стану джерела повідомлень і характеризує середній ступінь невизначеності стану цього джерела.

Ентропія H визначається з формулою згідно з теоремою К. Шеннона, на основі якої, середня кількість інформації, що припадає на один символ, визначається

$$H = \sum_{i=1}^{m} \frac{n_i}{q} \log p_i \, .$$

В інформаційних системах невизначеність знижується за рахунок прийнятої інформації, тому чисельно ентропія H дорівнює кількості інформації I, тобто є кількісною мірою інформації.

У загальному випадку, відповідно до теорії ймовірностей, джерело інформації однозначно й повно характеризується ансамблем станів $U = \{(u_1),(u_2),...(u_n)\}$ з імовірностями станів відповідно $\{p(u_1), p(u_2),... p(u_n)\}$ за умови, що сума ймовірностей всіх станів дорівнює 1.

Міра кількості інформації, як невизначеності вибору дискретним джерелом стану з ансамблю U, запропонована К. Шенноном в 1946 році й одержала назву ентропії дискретного джерела інформації або ентропії кінцевого ансамблю:

$$H(U) = - \sum_{n=1}^{N} p_n \log_2 p_n.$$

Ступінь невизначеності стану об'єкта (або так званого джерела інформації) залежить не тільки від числа його можливих станів, але й від імовірності цих станів. При нерівноймовірних станах варіанти вибору для джерела обмежуються.

Отже, основні властивості ентропії.

1. Ентропія є величиною дійсною й від'ємною, тому що значення ймовірностей p_n перебувають в інтервалі 0-1, значення $\log p_n$ завжди від'ємні, а значення $-p_n \log p_n$ відповідно додатні.

2. Ентропія – величина обмежена, тому що при $p_n \Rightarrow 0$ значення $-p_n \log p_n$ також прямує до нуля, а при $0 < p_n < 1$ обмеженість суми всіх доданків очевидна.

3. Ентропія дорівнює 0, якщо ймовірність одного зі станів джерела інформації дорівнює 1, і тим самим стан джерела повністю визначено (імовірності інших станів джерела дорівнюють нулю, тому що сума ймовірнос-

тей повинна бути 1).

4. Ентропія джерела із двома станами u_1 й u_2 при зміні співвідношення їхніх імовірностей $p(u_1) = p$ й $p(u_2) = 1 - p$ визначається виразом

$$H(U) = -\left[p \log p + (1-p) \log p\right]$$

і досягає максимуму при рівності ймовірностей.

Так, якщо із двох можливих станів імовірність одного з них дорівнює 0,999 то ймовірність іншого стану відповідно дорівнює 1 – 0,999 = 0,001.

5. Ентропія об'єднаних статистично незалежних джерел інформації дорівнює сумі їх ентропій.

6. Ентропія максимальна при рівній імовірності всіх станів джерела інформації: $H(U)_{\max} = -(\frac{1}{N}) \log \frac{1}{N} = \log N$.

В цьому окремому випадку кількісна міра Шеннона збігається з мірою Хартлі. Якщо повідомлення нерівноймовірні, то середня кількість інформації, що вміщується в одному повідомленні, буде меншою.

При використанні двійкової системи з рівними імовірностями виникнення «0» та «1», згідно із формулою Шеннона:

$$H = -0{,}5 \cdot \log_2 0{,}5 - 0{,}5 \cdot \log_2 0{,}5 = 1.$$

Ентропія, а разом з нею і кількість інформації, дорівнюють нулю у випадках, коли $p_1 = 0$, або $p_1 = 1$.

Графік ентропії наведено на рисунку 2.1.

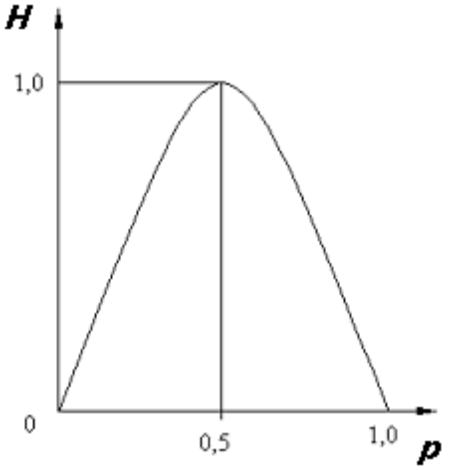


Рисунок 2.1 – Ентропія H для двох можливих станів з імовірностями p_1 та $(1 - p_1)$

Для каналів передачі дискретних сигналів (дискретні каналу зв'язку) використають поняття технічної й інформаційної швидкості передачі

даних.

Під *технічною швидкістю передачі* (англ. *technical speed of transfer*) розуміють число елементарних сигналів (символів), переданих по каналу за одиницю часу. Найпростіший елементарний символ – однополярний електричний імпульс тривалістю *T*. У дискретних каналах використовують, як правило, двополярні імпульси, позитивні на першій половині інтервалу *T* і негативні на другій половині. Це дозволяє підтримувати нульовий потенціал кабелю й виконувати тактову синхронізацію прийому-передачі сигналів. Одиницею міри технічної швидкості $V_t=1/T$ служить БОД – один символ у секунду. Смуга пропускання каналу зв'язку обмежується граничною частотою $F_{гр}$ за рівнем загасання сигналу до рівня статистичних перешкод, при цьому значення технічної швидкості передачі даних не може бути вище $F_{гр}$ без спеціальних пристроїв виділення інформаційних сигналів.

При відомій технічній швидкості Vt швидкість передачі інформації вимірюється в бітах у секунду й задається співвідношенням:

$$V = V_t H(s)$$

де *H(s)* – ентропія символу.

Для двійкових дискретних символів [0, 1] при постійній амплітуді імпульсів значення *H(s)* дорівнює 1. При числі *L* можливих рівноймовірних рівнів амплітуди імпульсів (рівень завад менший різниці рівнів амплітуд імпульсів) значення *H(s)* дорівнює *log L*.

Інформаційна ємність сигналу або повна кількість інформації в сигналі *S* (повідомленні, кодовій послідовності/слові) визначається повною кількістю $N = t/T$ ентропії символів у бітах на інтервалі задання сигналу *t*:

$$I(s) = NLogL = (t/T)LogL .$$

Збільшення числа рівнів *L* збільшує пропускну здатність каналів зв'язку, але ускладнює апаратуру кодування даних і знижує завадостійкість.

Для неперервних сигналів передача по каналах зв'язку можлива тільки за умови, що максимальна інформаційна частота в сигналі F_{max} не перевищує граничної частоти $F_{гр}$ передачі сигналів каналом зв'язку. Для оцінки інформаційної ємності безперервного сигналу виконаємо його дискретизацію з інтервалом $t = 1/2F_{max}$. Як установлено Котельниковим і Шенноном, за миттєвими відліками неперервного сигналу з таким інтервалом дискретизації аналоговий сигнал може бути відновлений без втрати

інформації.

При повній тривалості сигналу T_s число відліків:

$$N = {T_s}/{t} = 2F_{\max}T_s .$$

Визначимо максимально можливе число вибірок у кожному відліку при наявності шуму в каналі із середньою потужністю $P_ш = 2$. При середній потужності сигналу $P_s = s^2$:

$$L = \sqrt{\frac{s^2 + \delta^2}{\delta^2}} = \sqrt{1 + \frac{s^2}{\delta^2}} .$$

Інформаційна ємність сигналу:

$$I(s) = 2F_{\max}T_s LogL .$$

Інформаційні можливості сигналу зростають із розширенням його спектра й перевищенням його рівня над рівнем завад.

2.3 Передача інформації без завад

Ємність каналу (англ. *channel capacity*) – гранична швидкість передавання інформації цим каналом:

$$C = \lim_{T\to\infty}\left(\frac{\log q}{T}\right),$$

де q – кількість елементарних інформативних повідомлень, що передається за час T.

Якщо сигнали передаються зі швидкістю S імпульсів за секунду, тобто:

$$S = \frac{1}{\tau},$$

де τ – час передавання одного імпульсу;

то за час T можна передати n імпульсів:

$$n = \frac{T}{\tau} = ST .$$

Для двійкового каналу, що пропускає лише елементарні сигнали «0» та «1», максимальна кількість комбінацій елементарних сигналів, яка може бути передана за час T, складає:

$$q = 2^n = 2^{ST} .$$

Тоді ємність бінарного каналу зв'язку визначається:

$$C = \lim_{T\to\infty}\left(\frac{Log_2 q}{T}\right) = \frac{Log_2 2^{ST}}{T} = S ,$$

тобто, чим меншою буде тривалість імпульсу $\tau = {1}/{S}$, тим більшою буде

ємність каналу *C*. Для недвійкового каналу:

$$q = m^{ST},$$

де m – кількість символів у алфавіті;

і ємність каналу:

$$C = \lim_{T\to\infty}\left(\frac{Logq}{T}\right) = \frac{Log(m^{ST})}{T} = SLogm.$$

Ємність каналу зв'язку *C* може бути виражена у бітах на символ. Якщо до входу каналу підключене джерело повідомлень з ентропією на символ, що дорівнює ємності каналу зв'язку, то джерело інформаційно узгоджене з каналом. Якщо ентропія джерела менша ніж ємність каналу, то ємність каналу використовується не повністю (канал інформаційно недовантажений).

Узгодження джерела з каналом є досить складною справою і реалізовується за допомогою статистичного кодування. К. Шеннон показав, що інформаційне узгодження, яке досягається статистичним кодуванням, аналогічне енергетичному узгодженню внутрішнього опору електричного генератора з навантаженням за допомогою трансформатора для передавання від генератора максимальної потужності. Тут маємо на увазі узгодження джерела з каналом зв'язку за допомогою пристрою кодування з метою максимального використання ємності каналу.

2.4 Передача інформації із завадами

Завади (англ. *interference*) або *шуми* (англ. *noise*) у каналі зв'язку суттєво ускладнюють передавання інформації. На приймальному боці немає впевненості, що той чи інший елемент повідомлення прийняті у тому вигляді, в якому вони були передані. Тому під час передавання каналом із завадами виникають дві проблеми:

– підвищення ефективності передавання;

– підвищення вірогідності (завадозахищеності) передавання.

Ці проблеми до певної міри протилежні.

Якщо за рахунок впливу шуму був прийнятий елемент повідомлення *j*, в той час, як був переданий елемент *i*, то збільшення інформації можна визначити:

$$\Delta I_{ij} = \log\frac{1}{p_i} - \log\frac{1}{p_j(i)} = \log\frac{p_j(i)}{p_i},$$

де p_i – апріорна імовірність передавання елемента *i*;

$p_j(i)$ – умовна імовірність приймання елемента j в той час, як був переданий елемент i.

Якщо шуми досить великі $(p_j(i) = p_i,\ \Delta I = \log 1 = 0)$, повідомлення, що приймається, не вміщує інформації і приймання його не змінює початкових знань. За умови відсутності шуму: $p_j(i) = 1$, якщо $i = j$, або $p_j(i) = 0$, якщо $j \neq i$. В цьому випадку:

$$\Delta I = \log \frac{1}{p_i} = -\log p_i .$$

Пропускна здатність каналу з шумами (у двійкових одиницях на символ) дорівнює середньому за всіма і та j значенню приросту інформації:

$$R_c = \sum_{j,i} p_i p_j(i) \Delta I_{ij} = H_i - H_j(i) = H_j - H_i(j),$$

де $H_i = \sum_i p_i \log p_i$ – ентропія джерела;

$H_j = \sum_j p_j \log p_j$ – ентропія повідомлень на приймальному боці;

$H_i(j) = \sum_{i,j} p_i(j) \log p_i(j)$, $H_j(i) = \sum_{i,j} p_j(i) \log p_j(i)$ – умовні ентропії.

Для каналу з шумами швидкість передавання інформації (у бітах за секунду):

$$v = SR_c ,$$

де S – кількість символів, що передаються за секунду.

$$R_c = H_i - H_j(i) .$$

Тоді:

$$v = S(H_i - H_j(i)) . \qquad (2.4)$$

ПРИКЛАД

Якщо швидкість передавання інформації каналу складає 1000 біт/с, а дія завад викликає помилку у 1% символів, то за умови, що імовірності передавання «0» та «1» однакові, ентропія є максимальною $H = 1$ (біт/символ). Імовірність того, що під час передавання «0» приймається «1» складає $p_1(0) = 0{,}01$. Відповідно до цього інші умовні імовірності будуть $p_0(0) = 0{,}99$; $p_0(1) = 0{,}01$; $p_1(1) = 0{,}99$. Розраховані ентропії складають: $H_i = 1$, $H_i(j) = H_j(i) = 0{,}081$. Згідно з (2.4) швидкість передавання інформації каналом із шумами:

$v = 1000 \cdot (1 - 0{,}081) = 919$ (біт/с).

Таким чином швидкість передавання інформації під впливом шуму зменшується більш різко, ніж кількість правильно переданих символів.

К. Шеннон доказав, що якщо ентропія джерела інформації не перевищує пропускної здатності каналу, тобто $H \le C$, то існує код, який забезпечує передавання інформації каналом із шумами з якою завгодно малою частотою помилок або з якою завгодно малою невірогідністю. При $H > C$ такого коду не існує, тобто передавання без помилок неможливе. К. Шеннон визначив також максимальну швидкість передавання інформації:

$$v_k = f_m \log(1 + \frac{P_c}{P_ø}) \ , \qquad (2.5)$$

де f_m – смуга частот каналу;

P_c – середня потужність сигналу;

$P_ш$ – середня потужність білого шуму.

Таким чином можна передавати інформацію, якщо швидкість передавання інформації не перевищує максимальної швидкості каналу. Для випадку, коли середня потужність білого шуму набагато більша за середню потужність сигналу, тобто $P_c >> P_ш$, у формулі (2.5) одиницею можна знехтувати:

$$v_k = f_m \log(\frac{P_c}{P_ø}) .$$

Максимальна кількість інформації, яка може бути передана за певний час T:

$$V_{\max} = f_m T \log(\frac{P_c}{P_ø}) . \qquad (2.6)$$

Оскільки ця величина може бути зображена у вигляді паралелепіпеда, то вона отримала назву *об'єму сигналу*. Таким чином можна змінювати окремі параметри сигналу, не змінюючи його об'єм. Якщо до виразу (2.6) підставити потенційні можливості каналу передавання (час, на який канал надається користувачу, виділена йому смуга частот і максимальна потужність сигналу, що може передаватися каналом), то параметр матиме назву *ємність каналу*. Для передавання сигналу каналом зв'язку необхідно щоб об'єм сигналу був не менший, ніж ємність каналу, тобто необхідно виконати умову:

$$V_c \le V_k . \qquad (2.7)$$

Якщо ж дана умова не виконується, то сигнал передати цим каналом зв'язку неможливо. Може бути ситуація, при якій дана умова виконується, але смуга частот, на яку розрахований канал, менша за смугу частот сигналу або час, який виділено на передавання інформації менший, ніж необхідно, тобто умова (2.7) розпадається на систему:

$$\begin{cases} f_c \le f_k \\ T_c \le T_k \\ h_c \le h_k \end{cases}.$$

Одним з найбільш поширених способів перетворення сигналу є варіювання величинами f_c та T_c за їх незмінним добутком. Потужність сигналу, як правило, не збільшується.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Що являють собою статична та динамічна форми існування інформації?
2. Що являють собою міри інформації за Хартлі та Шенноном?
3. Які одиниці вимірювання інформації?
4. Що таке ентропія?
5. Що таке ємність каналу передавання?
6. Що таке інформаційне узгодження джерела повідомлення з каналом зв'язку?
7. В чому полягає статистичне кодування? Які його переваги та недоліки?
8. Що розуміють під поняттям "шум"?
9. Чим визначається пропускна здатність каналу зі шумами?
10. Чим визначається швидкість передавання інформації для каналу з шумами?
11. Яка максимальна швидкість передавання інформації?
12. Що таке об'єм сигналу? Як об'єм сигналу пов'язаний з ємністю каналу?

3 МОДЕЛЮВАННЯ СИГНАЛІВ ТА ЗАВАД В ЕЛЕКТРОННИХ СИСТЕМАХ

3.1 Перетворення сигналів та їх спектральні характеристики

Інформація сприймається споживачем через сигнали, які генеруються або безпосередньо джерелом, або за допомогою генераторів, що не зв'язані з джерелом (локація). Сигнали мають енергетичну та матеріальну основу – світлові потоки, звукові хвилі, теплове випромінювання, зображення точок, неоднорідності електричного і магнітного полів. Сигналами можуть бути документи, книжки, графіки, таблиці і т. ін. Отже, сигнали різняться за фізичною природою. В електронних системах сприйняття сигналів різної фізичної природи здійснюється первинними перетворювачами, які перетворюють інформаційний параметр в сигнал, що є зручним для подальшої обробки.

Важливий елемент інформаційного процесу – передача інформації. Вона відбувається за допомогою повідомлень – послідовності сигналів в часі або просторі, відповідно повідомлення будуть часові або просторові. Повідомлення складаються із матеріальних елементів, якими вони представлені. Це є абетка повідомлення (знаки, символи, колір, частота, інтенсивність і т. ін.), що містить певну кількість елементів. Передача інформації здійснюється за допомогою переносників інформації. Матеріальну основу переносника інформації складає певний фізичний об'єкт або процес, який характеризується набором параметрів. Один із цих параметрів може стати інформаційним, якщо він змінюється під впливом джерела інформації (модуляція параметра). Власне, після модуляції переносник перетворюється в носія інформації.

Як переносники інформації використовуються коливання різної природи, зокрема – гармонічні коливання, включно з нульовою частотою. В електронних системах знайшли поширення переносники у вигляді коливань електричного струму або напруги – електричні сигнали. Якщо частота коливань дорівнює нулю, є лише один інформаційний параметр – рівень струму або напруги. Якщо частота коливань відрізняється від нуля, інформаційними параметрами можуть бути амплітуда, частота або фаза коливань.

Отже, сигнал – змінна фізична величина, що забезпечує передачу інформації лінією зв'язку.

Загалом всі сигнали розподіляються на детерміновані (невипадкові) і випадкові.

Детермінований сигнал характеризується визначеністю його значень в будь-які моменти часу (задаються певною визначеною функцією часу). Випадковий сигнал характеризується тим, що його значення в будь-який момент часу є випадковими величинами. Випадковими можуть бути як корисні (інформаційні) сигнали, так і шкідливі сигнали – завади, які перешкоджають сприйняттю інформаційних сигналів.

Як приклад детермінованого сигналу можна навести синусоїдальну хвилю, яка графічно описується синусоїдою. Синусоїда – це функція часу t, яка записується так:

$$f(t) = A\sin(\omega t + \theta),$$

де величину сигналу визначає коефіцієнт A, названий амплітудою;

ω – кутова частота;

θ – початкова фаза.

Синусоїда разом з трикутним, пилкоподібним, прямокутним та іншими сигналами належить до періодичних сигналів, тобто до таких сигналів, які повторюють свою форму через деякий сталий проміжок часу.

Будь-який складний періодичний сигнал може бути поданий за допомогою ряду Фур'є як сума простих гармонічних коливань. Сукупність простих гармонічних коливань, на які може бути розкладений складний періодичний сигнал, називається його спектром.

Розподіл амплітуд гармонік за частотою називають амплітудно-частотним спектром або скорочено амплітудним спектром, а розподіл їхніх початкових фаз за частотою – фазочастотним спектром або фазовим спектром.

Лінії дискретного спектра мають розмірність амплітуди сигналу. Безперервний спектр указує на розподіл амплітуд по всьому спектрі й має розмірність щільності амплітуд сигналу.

Якщо спектр сигналу є необмеженим, то при визначенні ширини нехтують гармоніками, амплітуди яких невеликі й не перевищують певного (заданого) рівня. Найбільш часто користуються рівнем 0,707 за амплітудою або 0,5 за потужністю від максимального значення.

Перетворення Фур'є

Перетворення Фур'є – інтегральне перетворення однієї комплекснозначної функції дійсної змінної на іншу, тісно пов'язане з перетворенням Лапласа та аналогічне розкладу у ряд Фур'є для неперіодичних функцій. Це перетворення розкладає дану функцію на осциляторні функції. Використовується для того, щоби розрахувати спектр частот для сигналів змінних у часі (таких як мова або електрична напруга). Перетворення названо на честь французького математика Жана Батиста Жозефа Фур'є, який ввів це поняття в 1822 році.

Перетворення Фур'є функції $f(t)$ математично визначається як комплексна функція $F(\omega)$, яка задається інтегралом

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt.$$

Обернене перетворення Фур'є задається виразом

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

Серед властивостей перетворення Фур'є можна відмітити такі.

Якщо задані інтегровані функції $f(x)$, $g(x)$ та $h(x)$ і їх відповідні перетворення Фур'є $\hat{f}(\xi)$, $\hat{g}(\xi)$ та $\hat{h}(\xi)$, тоді перетворення має такі властивості:

а) лінійність;

для довільних комплексних чисел a та b, якщо $h(x) = af(x) + bg(x)$, тоді

$$\hat{h}(\xi) = a\hat{f}(\xi) + b\hat{g}(\xi).$$

б) трансляція;

Для довільного дійсного числа x_0, якщо $h(x) = f(x - x_0)$, тоді

$$\hat{h}(\xi) = e^{-2\pi i x_0 \xi} \hat{f}(\xi).$$

в) модуляція;

Для довільного дійсного числа ξ_0, якщо $h(x) = e^{2\pi i x \xi_0} f(x)$, тоді

$$\hat{h}(\xi) = \hat{f}(\xi - \xi_0).$$

г) масштабування;

Для нерівного нулю дійсного числа a, якщо $h(x) = f(ax)$, тоді

$$\hat{h}(\xi)=\frac{1}{|a|}\hat{f}\left(\frac{\xi}{a}\right).$$

Випадок $a=-1$ приводить до властивості обернення часу, згідно з якою: якщо $h(x)=f(-x)$, тоді $\hat{h}(\xi)=\hat{f}(-\xi)$.

д) спряження;

Якщо $h(x)=\overline{f(-x)}$, тоді $\hat{h}(\xi)=\overline{\hat{f}(-\xi)}$.

Зокрема, якщо f дійсне, тоді має місце умова дійсності $\hat{f}(-\xi)=\overline{\hat{f}(\xi)}$.

е) згортка.

Якщо $h(x)=(fg)(x)$, тоді $\hat{h}(\xi)=\hat{f}(\xi)\cdot\hat{g}(\xi)$.

Перетворення Фур'є застосовуються для отримання частотного спектру неперіодичної функції, наприклад, електричного сигналу, тобто для подання сигналу у вигляді суми гармонічних коливань. При цьому використовується властивість згортки.

Нехай відгук системи на подразнення у вигляді сигналу $f(t)$ буде

$$g(t)=\int_0^\infty \alpha(\tau)f(t-\tau)d\tau,$$

де $\alpha(\tau)$ – певна функція.

Такий запис означає, що відгук системи залежить не тільки від моментального значення збурення, а також від того подразнення, яке було певний час тому, і яке змінило стан системи.

Застосовуючи перетворення Фур'є до обох частин рівняння, отримуємо

$$G(\omega)=\int_{-\infty}^{\infty}e^{-i\omega t}\int_0^\infty \alpha(\tau)f(t-\tau)d\tau dt=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-i\omega t}\int_0^\infty \alpha(\tau)\int_{-\infty}^{\infty}F(\omega')e^{i\omega'(t-\tau)}d\omega' d\tau dt.$$

Оскільки

$$\int_{-\infty}^{\infty}e^{i(\omega'-\omega)t}dt=2\pi\delta(\omega'-\omega),$$

де $\delta(x)$ – дельта-функція Дірака, інтегрування дає $G(\omega)=A(\omega)F(\omega)$,

де $A(\omega)=\int_0^\infty \alpha(\tau)e^{i\omega\tau}$.

Важливим висновком з цього перетворення є те, що вихідний спектр отримується з вхідного простим множенням на функцію відклику системи $A(\omega)$.

Спектри періодичних і неперіодичних сигналів

Відомо, що будь-яка періодична функція, яка задовольняє умови Діріхле, може бути подана у вигляді нескінченної у загальному випадку суми гармонічних складових – рядом Фур'є. Умова Діріхле полягає у тому, що: функція $x(t)$ повинна бути обмеженою, кусочно-неперервною та мати протягом періоду скінченне число екстремумів.

Відомо дві форми розкладання в ряд Фур'є: тригонометрична й комплексна. Тригонометрична форма розкладання виражається у вигляді

$$x(t) = \frac{1}{2} A_0 + \sum_{k=1}^{\infty} A_k \cos(K\omega_0 t - \varphi_k),$$

де $\frac{1}{2} A_0$ – постійна складова функції $x(t)$;

$A_k \cos(K\omega_0 t - \varphi_k)$ – k-та гармонічна складова;

$A_k, K\omega_0, \varphi_k$ – амплітуда, частота та початкова фаза k-тої гармонічної складової;

$\omega_0 = \frac{2\pi}{T}$ – частота основної (першої) гармоніки;

Т – період зміни функції $x(t)$.

В математичному відношенні зручніше оперувати комплексною формою ряду Фур'є, поданою у вигляді

$$x(t) = \frac{1}{2} \sum_{k=-\infty}^{\infty} A_k \exp\{jK\omega_0 t\},$$

де $A_k = A_k \exp\{-j\varphi_k\}$ – комплексна амплітуда гармонічної складової частоти $\omega_k = K\omega_0$.

Комплексна амплітуда визначається через тимчасову функцію $x(t)$ за допомогою формули

$$A_k = \frac{2}{T} \int_{t_1}^{t_1+T} x(t) \exp\{-jK\omega_0 t\} dt.$$

Сукупність амплітуд і відповідних частот гармонік прийнято називати спектром амплітуд.

Сукупність початкових фаз і відповідних частот гармонік назива-

ють спектром фаз.

На рис. 3.1. подані графічні зображення спектра амплітуд і спектра фаз періодичного сигналу.

Окремі спектральні складові у графічному зображенні спектра амплітуд називають спектральними лініями.

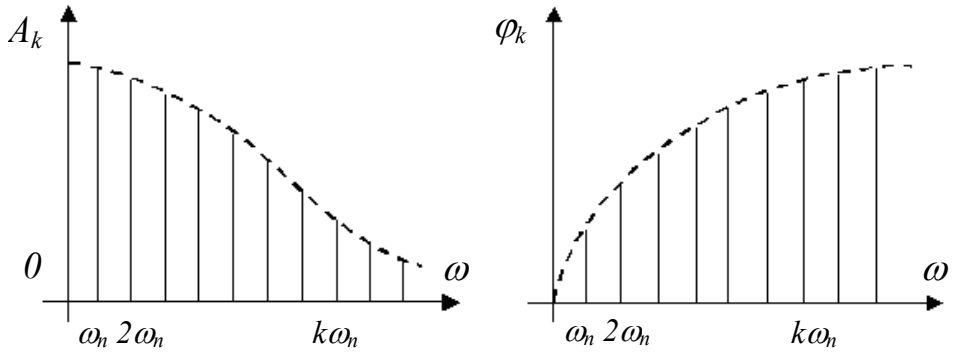


Рисунок 3.1 – Графічні зображення спектра амплітуд і спектра фаз періодичного сигналу

Будь-який неперіодичний сигнал можна розглядати як періодичний, період зміни якого дорівнює нескінченності. У зв'язку з цим розглянутий раніше спектральний аналіз періодичних процесів може бути узагальнений і на неперіодичний сигнал.

Розглянемо як буде змінюватись спектр неперіодичного сигналу при необмеженому збільшенні періоду зміни сигналу. При збільшенні періоду T інтервали між суміжними частотами в спектрі сигналу і амплітуди спектральних складових зменшується і в границі при $T \to \infty$ стають нескінченно малими величинами. При цьому спектральний розклад неперіодичного сигналу відображається рядом Фур'є.

Комплексна форма неперіодичного сигналу має вигляд

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega)^* \exp(j\omega t) d\omega,$$

де $S(j\omega) = S(\omega)\exp(j\phi(\omega))$ – спектральна щільність сигналу;

$S(\omega) = |S(j\omega)|$ – амплітудно-частотна характеристика сигналу;

$\phi(\omega)$ – фазочастотна характеристика сигналу.

Попередній вираз називається формулою оберненого перетворення Фур'є.

Подання неперіодичної функції інтегралом Фур'є можливе при ви-

конанні таких умов:

1) функція $x(t)$ задовольняє умову Діріхле;

2) функція $x(t)$ абсолютно інтегрована

$$\int_{-\infty}^{\infty} |x(t)|dt < \infty .$$

Таким чином, спектр неперіодичного сигналу, на відміну від спектра періодичного сигналу, є суцільним і являє собою суму нескінченної кількості гармонічних складових із нескінченно малими складовими.

Амплітуди гармонічних складових можуть бути подані у такому вигляді

$$dA = \frac{1}{2\pi} S(j\omega)d\omega ,$$

звідки спектральна щільність визначається виразом

$$S(j\omega) = 2\pi \frac{dA}{d\omega} .$$

Спектральна щільність пов'язана з функцією часу через пряме перетворення Фур'є

$$S(j\omega) = \int_{-\infty}^{\infty} x(t)^* \exp(-j\omega t)dt .$$

Спектральна щільність однозначно відображає неперіодичний сигнал і задовольняє умову: $\lim_{\omega\to\infty} S(\omega) = 0$.

Модуль спектральної щільності є парною, а аргумент непарною функцією частоти

$$S(\omega) = S(-\omega); \quad \varphi(\omega) = -\varphi(-\omega) .$$

Спектри одиничних та періодичних імпульсних послідовностей

Аналіз перехідних процесів в колі при дії на нього складної (негармонічної) ЕРС класичним методом виявляється досить складним. Більш зручними в таких випадках є методи, основані на спектральному поданні зовнішньої ЕРС і принципі суперпозиції. Тому, перш ніж переходити до аналізу таких перехідних процесів, розглянемо спектри деяких найбільш важливих для радіотехніки періодичних та неперіодичних ЕРС.

Будь-яку функцію $f(t)$, задану в інтервалі $t_1 < t < t_2$, і яка періодично повторюється з частотою $\Omega = \frac{2\pi}{T}$, де T – період повторення, та таку,

що задовольняє умову Діріхле, можна подати рядом Фур'є, який можна записати або в тригонометричній формі

$$f(t) = \frac{C_0}{2} + \sum_{K=1}^{\infty} C(K\Omega)\cos\left[K\Omega t - \varphi(K\Omega)\right],$$

де $K = 1,2,3\ldots$,

або в комплексній формі

$$f(t) = \frac{1}{2}\sum_{K=-\infty}^{+\infty} \overline{C}(K\Omega)e^{jK\Omega t} ,$$

де $\overline{C}(K\Omega) = C(K\Omega)e^{-j\varphi(K\Omega)}$ – комплексна амплітуда K-ї гармонічної складової;

$C(K\Omega)$ – модуль цієї величини, чи просто амплітуда;

$\varphi(K\Omega)$ – початкова фаза K-ї гармонічної складової.

Величина $\frac{C_0}{2}$ – середнє за період значення функції $f(t)$ чи постійна складова складної ЕРС

$$\frac{C_0}{2} = \frac{1}{T}\int_0^T f(t)dt .$$

Розглянемо основні величини, що характеризують спектр складної періодичної ЕРС. Комплексна амплітуда, яка входить в ряд Фур'є, визначається як

$$\overline{C}(K\Omega) = \frac{2}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} f(t)e^{-jK\Omega t}\cdot dt .$$

Залежність модуля комплексної амплітуди від частоти зображають у вигляді графіка, який називається амплітудно-частотним спектром (рис. 3.2, а). Тут кожній частоті $K\Omega$ відповідає лінія, величина якої $C(K\Omega)$ вказує амплітуду гармонічної складової.

Залежність початкової фази від частоти $\varphi(K\Omega)$ зображають у вигляді фазочастотного спектра (рис. 3.2, б). Як видно з рис. 3.2, для складної періодичної ЕРС, спектр є лінійчатим чи дискретним. Тут лінії розміщенні по шкалі частот так, що відділені на відстань Ω, яка дорівнює частоті повторення ЕРС. Зберігаючи незмінним амплітудно-частотний спектр, але змінюючи вигляд фазочастотного спектра, ми тим самим змінюємо форму складної періодичної ЕРС, яка зображена цими спектрами. При розгляді спектра складної ЕРС часто обмежуються одним графіком, на якому

зображено амплітудно-частотний спектр з вказуванням фаз гармонічних складових.

Рисунок 3.2 – Амплітудно-частотний та фазочастотний спектри

Теорема про суму спектрів і теорема запізнення дозволяють обчислити спектр групи однакових рівновідстаючих імпульсів. Нехай є два однакових імпульси $f_1(t)$ і $f_2(t) = f_1(t-\tau)$, розділених інтервалом часу τ. Можна записати спектральну функцію другого імпульсу через спектральну функцію першого:

$$\overline{S}_2(\omega) = \overline{S}_1(\omega) e^{-j\omega\tau}.$$

Тоді на основі виразу для спектральної функції суми двох імпульсів отримаємо:

$$\begin{aligned}
\overline{S}(\omega) &= \overline{S}_2(\omega) + \overline{S}_1(\omega) = \overline{S}_1(\omega)(1 + e^{-j\omega\tau}) = \\
&= \overline{S}_1(\omega)\left[1 + \cos\omega\tau - j\sin\omega\tau\right] = \\
&= \overline{S}_1(\omega)\sqrt{(1+\cos\omega\tau)^2 + \sin^2\omega\tau}\ \ e^{-j\psi(\omega)} = \\
&= \overline{S}_1(\omega) 2\cos\frac{\omega\tau}{2} e^{-j\psi(\omega)},
\end{aligned}$$

де $\psi(\omega) = \operatorname{arctg}\dfrac{\sin\omega\tau}{1+\cos\omega\tau}$.

Зі збільшенням кількості імпульсів спектр групи імпульсів наближається за структурою до лінійчатого спектра періодичної послідовності імпульсів.

Практично всі канали зв’язку мають обмежену смугу пропускання. Отже, при передачі сигналу через реальний канал зв’язку може бути передана лише частина його частотного спектра.

За практичну ширину спектра сигналу приймають діапазон частот, в межах якого знаходиться найбільш вагома частина спектра сигналу. Вибір практичної ширини спектра сигналу визначається двома критеріями:

енергетичним критерієм та критерієм допустимих спотворень форми сигналу.

ПРИКЛАД

Розглянемо для прикладу послідовність прямокутних імпульсів тривалістю τ, амплітудою h, із періодом проходження T (рис. 3.3).

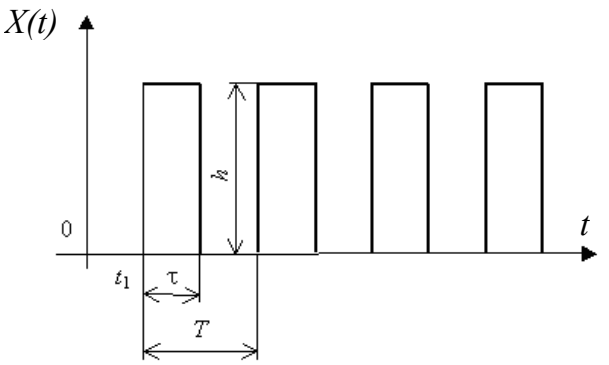


Рисунок 3.3 – Прямокутні імпульси

Розклад в ряд Фур'є періодичної послідовності прямокутних імпульсів подаєтьсяу вигляді

$$x(t) = \frac{\tau h}{T}\left[1 + 2\sum_{k=1}^{\infty} \frac{\sin\frac{K\omega_0\tau}{2}}{\frac{K\omega_0\tau}{2}} \cos K\omega_0\tau\right].$$

Спектр амплітуд такого сигналу показаний на рис. 3.4.

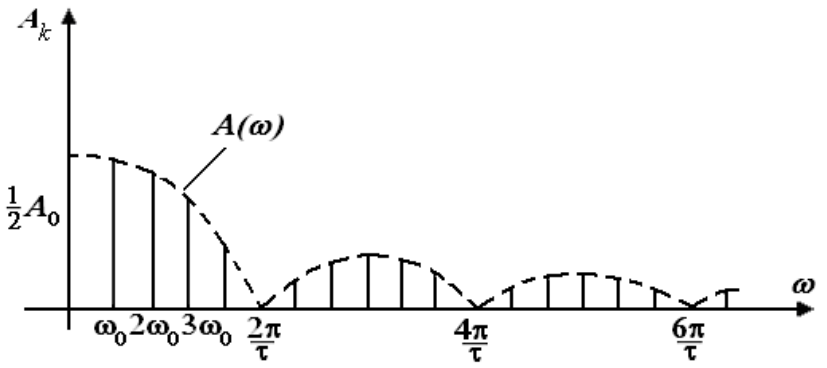


Рисунок 3.4 – Спектр амплітуд

Обвідна його визначається рівнянням

$$A(\omega) = 2\frac{\tau h}{T}\left[\frac{\sin\frac{\omega\tau}{2}}{\frac{\omega\tau}{2}}\right],$$

де $\omega = K\omega_0$ – для *K*-ої гармоніки.

Можна показати, що для періодичної послідовності імпульсів прямокутної форми тривалістю $\tau = \frac{T}{2}$ достатньо практичну ширину спектра вибрати рівною $3\omega_0 = \frac{6\pi}{T} = \frac{3\pi}{\tau}$. В цій області частот зосереджено 95% всієї потужності сигналу. Розглянемо одиничний прямокутний імпульс тривалістю *T* та величиною *h*, спектральна щільність такого сигналу визначається виразом

$$S(j\omega) = \int_{-\infty}^{\infty} x(t)\exp\{-j\omega t\}dt = \int_{\frac{-\tau}{2}}^{\frac{\tau}{2}} h\exp\{-j\omega t\}dt = \tau h\frac{\sin\frac{\omega\tau}{2}}{\frac{\omega\tau}{2}}.$$

Енергія сигналу, зосереджена в смузі частот від 0 до ω_0,

$$W_1 = \frac{1}{\pi}\int_0^{\omega_0}[S(\omega)]^2 d\omega = \frac{\tau^2 h^2}{\pi}\int_0^{\omega_1}\left[\frac{\sin\frac{\omega\tau}{2}}{\frac{\omega\tau}{2}}\right]^2 d\omega.$$

Для оцінювання впливу ширини смуги пропускання каналу зв'язку на викривлення форми сигналів розглянемо проходження прямокутного імпульсу тривалістю *T* та величиною *U* через канал зв'язку, що являє собою ідеальний фільтр низьких частот. Коефіцієнт передачі цього фільтра виражається залежністю: $K(j\omega) = K\exp\{-j\omega T_0\}$, при цьому в діапазоні частот $0 \le \omega \le \omega_b$ модуль коефіцієнта передачі $K(\omega) = K = const$ і аргументу $\varphi(\omega) = \omega T_0$; поза цим діапазоном $K(\omega) = 0$.

В теорії кіл показано, що вихідний сигнал в цьому випадку може бути поданий в аналітичному вигляді

$$a_{вих}(t) = UK\left\{\int_0^{\omega_b}\frac{\sin\omega(t-T_0)}{\omega}d\omega - \int_0^{\omega_b}\frac{\sin\omega\left[t-(T_0-\tau)\right]}{\omega}d\omega\right\}.$$

Форми переднього й заднього фронтів імпульсу спотворюються однаково. На рис. 3.5 показана форма переднього фронту вихідного сигналу.

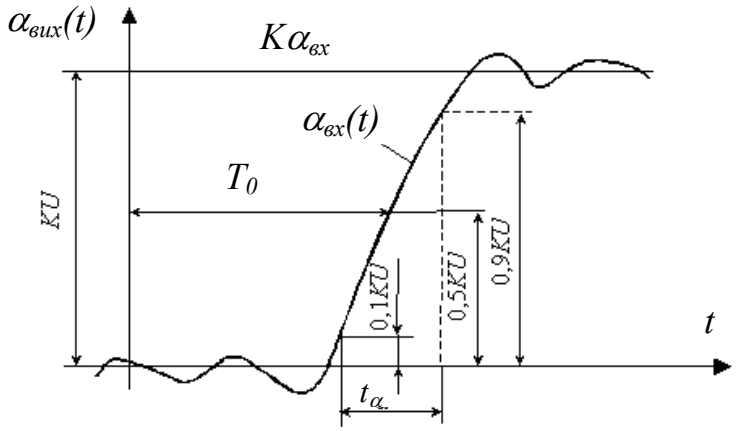


Рисунок 3.5 – Форма переднього фронту вихідного сигналу

Щоб вихідний сигнал зміг досягнути найбільшого значення, активна тривалість переднього фронту $t_{фа}$ повинна бути не більша тривалості вхідного імпульсу τ. При цьому, як показав аналіз, повинна бути справедлива умова $\tau \approx \frac{0,4}{f_b}$ або $f_b \approx \frac{0,4}{\tau}$.

Дискретне перетворення Фур'є

Припустимо, що замість функції неперервної змінної $f(x)$ задано функцію дискретної змінної на рівномірній гратці (рис. 3.6), тобто задано значення функції f_k для скінченної послідовності значень аргументу $x_k = kh$ – таблиця функції $\{f_k, kh; k = 0..N\}$. Тут за допомогою $h = \frac{L}{N}$ позначено крок гратки – відстань між сусідніми вузлами.

Перетворення Фур'є такої функції можна означити як суму

$$f_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f_n e^{-2\pi i \frac{kn}{N}}.$$

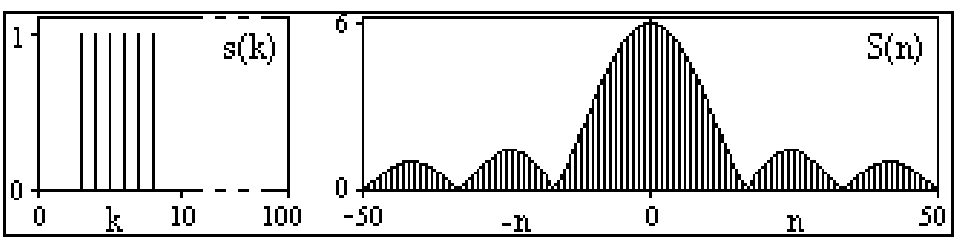


Рисунок 3.6 – Дискретний сигнал та модуль його спектра

Варто зазначити, що недоцільно використовувати суму з кількістю членів, більшою за кількість вузлів гратки.

3.2 Види модуляції сигналів

Однією з основних тенденцій розвитку мережних технологій є передача в одній мережі як дискретних, так і аналогових за своєю природою даних. Джерелами дискретних даних є комп'ютери й інші обчислювальні пристрої, а джерелами аналогових даних є такі пристрої, як телефони, відеокамери, звуко- і відеовідтворююча апаратура. На ранніх етапах розв'язання цієї проблеми в територіальних мережах усі типи даних передавалися в аналоговій формі, при цьому дискретні за своїм характером комп'ютерні дані перетворювалися в аналогову форму за допомогою модемів.

Однак в міру розвитку техніки прийому і передачі аналогових даних з'ясувалося, що передача їх в аналоговій формі не дозволяє поліпшити якість прийнятих на іншому кінці лінії даних, якщо вони істотно спотворилися при передачі. Сам аналоговий сигнал не дає ніяких вказівок ні про те, що відбулося перекручування, ні про те, як його виправити, оскільки форма сигналу може бути різною, у тому числі і такою, яку зафіксував приймач. Поліпшення ж якості ліній, особливо територіальних, потребує величезних зусиль і капіталовкладень. Тому на зміну аналоговій техніці запису і передачі звуку і зображення прийшла цифрова техніка. Ця техніка використовує так звану дискретну модуляцію вихідних неперевних у часі аналогових процесів.

Дискретні способи модуляції основані на дискретизації неперервних процесів як за амплітудою, так і за часом (рис. 3.7). Розглянемо принципи дискретної модуляції на прикладі імпульсно-кодової модуляції, ІКМ (Pulse Amplitude Modulation, PAM), що широко застосовується в цифровій телефонії.

Амплітуда початкової неперервної функції вимірюється із заданим періодом – за рахунок цього відбувається *дискретизація за часом* (англ. *diskretization on time*). Потім кожне вимірювання подається у вигляді двійкового числа певної розрядності, що означає дискретизацію за значеннями функції – неперервна множина можливих значень амплітуди заміняється дискретною множиною її значень. Пристрій, що виконує подібну функцію, називається аналого-цифровим перетворювачем (АЦП). Після цього перетворення сигнали передаються по каналах зв'язку у вигляді послідовності одиниць і нулів. При цьому застосовуються ті ж методи *кодування* (англ.

coding), що й у випадку передачі початкової дискретної інформації, тобто, наприклад, методи, основані на коді B8ZS чи 2B1Q, На прийомній стороні лінії коди перетворяться у вихідну послідовність бітів, а спеціальна апаратура, яка називається цифро-аналоговим перетворювачем (ЦАП), робить демодуляцію оцифрованих амплітуд неперевного сигналу, відновлюючи вхідну неперервну функцію часу.

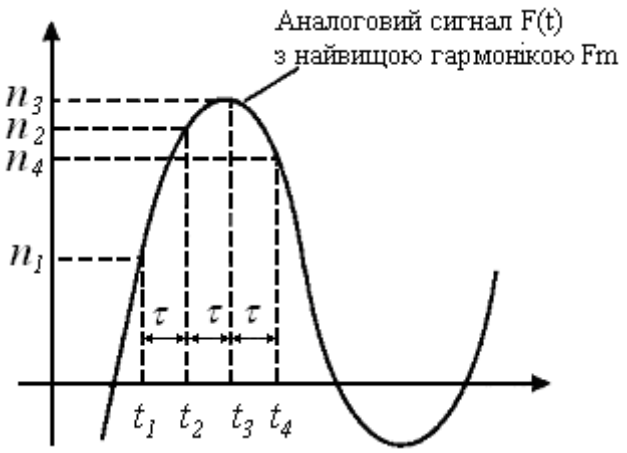


Рисунок 3.7 – Амплітудна модуляція неперервного сигналу

Дискретна модуляція основана на теорії відображення Найквіста – Котельникова. Відповідно до цієї теорії, аналогова неперервна функція, передана у вигляді послідовності її дискретних за часом значень, може бути точно відновлена, якщо частота дискретизації була в два чи більше разів вища, ніж частота найвищої гармоніки спектра вихідної функції. Якщо ця умова не дотримується, то відновлена функція буде істотно відрізнятися від вхідної. Перевагою цифрових методів запису, відтворення і передачі аналогової інформації є можливість контролю вірогідності зчитаних з носія чи отриманих по лінії зв'язку даних. Для цього можна використовувати ті ж методи, що застосовуються для комп'ютерних даних (і розглядаються більш докладно далі), – обчислення контрольної суми, повторна передача перекручених кадрів, застосування самокоректовних кодів.

Для якісної передачі голосу в методі ІКМ використовується частота квантування амплітуди звукових коливань у 8000 Гц. Це пов'язано з тим, що в аналоговій телефонії для передачі голосу був обраний діапазон від 300 до 3400 Гц, який досить якісно передає всі основні гармоніки співрозмовників. Відповідно до теореми Найквіста-Котельникова для якісної передачі голосу досить вибрати частоту дискретизації, яка у два рази перевищує найвищу гармоніку неперервного сигналу, тобто $2 \times 3400 = 6800$ Гц.

Вибрана в дійсності частота дискретизації 8000 Гц забезпечує деякий запас якості. При методі ІКМ звичайно використовується 7 чи 8 біт коду для подання амплітуди одного імпульсу. Відповідно це дає 127 чи 256 градацій звукового сигналу, що виявляється цілком достатнім для якісної передачі голосу. При використанні методу ІКМ для передачі одного голосового каналу необхідна пропускна здатність 56 чи 64 Кбіт/с в залежності від того, якою кількістю біт подається кожне вимірювання. Якщо для цих цілей використовується 7 біт, то при частоті передачі вимірювань у 8000 Гц одержуємо:

$$8000 \times 7 = 56000 \text{ біт/с чи } 56 \text{ Кбіт/с},$$

а для випадку 8-ми біт:

$$8000 \times 8 = 64000 \text{ біт/с чи } 64 \text{ Кбіт/с}.$$

Стандартним є цифровий канал 64 Кбіт/с, що також називається елементарним каналом цифрових телефонних мереж. Передача неперевного сигналу в дискретному вигляді потребує від мереж твердого дотримання часового інтервалу в 125 мкс (відповідного частоті дискретизації 8000 Гц) між сусідніми імпульсами, тобто потребує синхронної передачі даних між вузлами мережі. При недотриманні синхронності надходження імпульсів вхідний сигнал відновлюється неправильно, що приводить до перекручування голосу, зображення чи іншої мультимедійної інформації, яка передається по цифрових мережах. Так, перекручування синхронізації в 10 мс може привести до накладання слів, а зрушення між імпульсами в 200 мс приводять до втрати розпізнавання вимовлених слів. У той же час втрата одного імпульсу при дотриманні синхронності між іншими імпульсами практично не позначається на відтвореному звуці. Це відбувається за рахунок згладжувальних пристроїв у цифро-аналогових перетворювачах, що основані на властивості інерційності будь-якого фізичного сигналу – амплітуда звукових коливань не може миттєво змінитися на велику величину. На якість сигналу після ЦАП впливає не тільки синхронність надходження на його вхід імпульсів, але і похибка дискретизації амплітуд цих імпульсів. В теоремі Найквіста – Котельникова передбачається, що амплітуди функції вимірюються точно, у той же час використання для їхнього збереження двійкових чисел з обмеженою розрядністю трохи спотворює ці амплітуди. Відповідно спотворюється і відновлений неперервний сигнал. Це спотворення називається шумом дискретизації.

Існують й інші методи дискретної модуляції, що дозволяють подати вимірювання голосу в більш компактній формі, наприклад у вигляді послідо-

вності 4-бітових чи 2-бітових чисел. При цьому один голосовий канал потребує меншої пропускної здатності, наприклад 32 Кбіт/с, 16 Кбіт/с чи ще менше. З 1985 року застосовується стандарт CCITT кодування голосу, названий Adaptive Differential Pulse Code Modulation (ADPCM). Коди ADPCM основані на різниці між послідовними вимірюваннями голосу, що передаються мережею. У коді ADPCM для збереження однієї різниці використовуються 4 біт, а швидкість передачі голосу 32 Кбіт/с. Більш сучасний метод, Linear Predictive Coding (LPC), робить вимірювання вхідної функції рідше, але при цьому використовуються методи прогнозування напрямку зміни амплітуди сигналу. За допомогою цього методу можна понизити швидкість передачі голосу до 9600 біт/с.

Подані в цифровій формі непервні дані легко можна передати через комп'ютерну мережу. Для цього досить помістити кілька імпульсів у кадр будь-якої стандартної мережевої технології, постачати кадр з правильною адресою призначення і відправити адресату. Адресат повинен витягти з кадру вимірювання і подати їх з частотою квантування (для голосу – з частотою 8 000 Гц) на цифро-аналоговий перетворювач. В міру надходження наступних кадрів з вимірюваннями голосу операція повинна повторитися. Якщо кадри будуть прибувати достатньо синхронно, то якість голосу може бути досить високою. Однак, як ми вже знаємо, кадри в комп'ютерних мережах можуть затримуватися як у кінцевих вузлах, так і в проміжних комунікаційних пристроях – мостах, комутаторах і маршрутизаторах. Тому якість голосу при передачі в цифровій формі через комп'ютерні мережі звичайно буває невисокою. Для якісної передачі оцифрованих неперервних сигналів – голосу, зображення – сьогодні використовують спеціальні цифрові мережі, такі як ISDN, ATM, і мережі цифрового телебачення. Проте для передачі всередині корпоративних телефонних розмов сьогодні характерні мережі Frame relay, затримки передачі кадрів яких укладаються в припустимі рамки.

Модуляція гармонічних коливань: амплітудна, фазова, частотна

Аналогова модуляція застосовується для передачі дискретних даних по каналах з вузькою смугою частот, типовим представником яких є канал тональної частоти, наданий у розпорядження користувачам загальних телефонних мереж. Типова амплітудно-частотна характеристика каналу тональної частоти подана на рис. 3.8. Цей канал передає частоти в діапазоні

від 300 до 3400 Гц, таким чином, його смуга пропускання дорівнює 3100 Гц. Хоча людський голос має набагато більш широкий спектр – приблизно від 100 Гц до 10 кГц, – для прийнятної якості передачі мови діапазон у 3100 Гц є гарним рішенням. Строге обмеження смуги пропускання тонального каналу пов'язано з використанням апаратури ущільнення і комутації каналів у телефонних мережах. Пристрій, що виконує функції модуляції несучої синусоїди на стороні, яка передає сигнали, і демодуляції на прийомній стороні, називається модемом (модулятор-демодулятор).

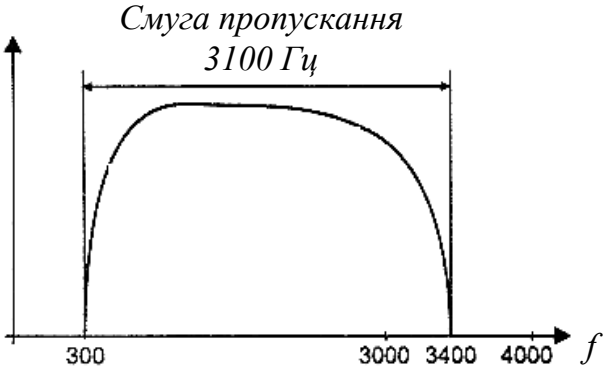


Рисунок 3.8 – Амплітудно-частотна характеристика каналу тональної частоти

Аналогова модуляція є таким способом фізичного кодування, при якому інформація кодується зміною амплітуди, фази чи частоти синусоїдального сигналу несучої частоти. Основні способи аналогової модуляції показані на рис. 3.9. На діаграмі (рис. 3.9, а) показана послідовність бітів вихідної інформації, подана потенціалами високого рівня для логічної одиниці і потенціалом нульового рівня для логічного нуля. Такий спосіб кодування називається потенційним кодом, що часто використовується при передачі даних між блоками комп'ютера.

При амплітудній модуляції (рис. 3.9, б) для логічної одиниці вибирається один рівень амплітуди синусоїди несучої частоти, а для логічного нуля – інший. Цей спосіб рідко використовується в чистому вигляді на практиці через низку завадостійкість, але часто застосовується в поєднанні з іншим видом модуляції – фазовою модуляцією.

При частотній модуляції (рис. 3.9, в) значення 0 і 1 вихідних даних передаються синусоїдами з різною частотою – f_0 і f_1. Цей спосіб модуляції не потребує складних схем у модемах і звичайно застосовується в низькошвидкісних модемах, що працюють на швидкостях 300 чи 1200 біт/с.

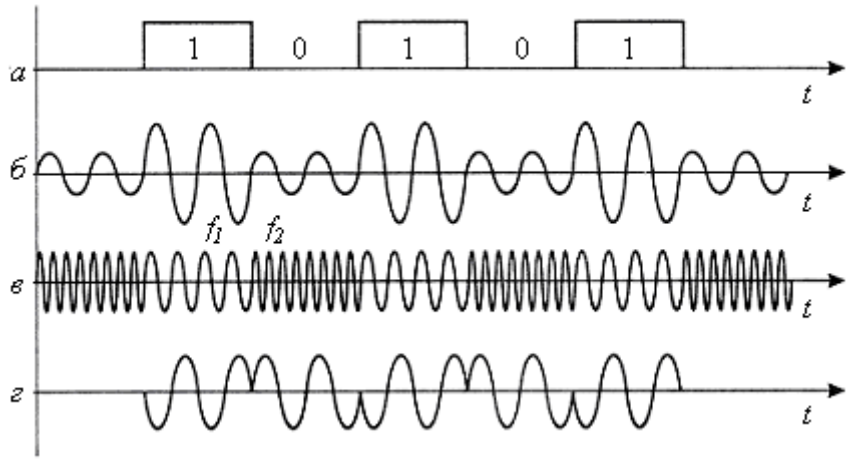


Рисунок 3.9 – Різні види модуляції

При фазовій модуляції (рис. 3.9, г) значенням даних 0 і 1 відповідають сигнали однакової частоти, але з різною фазою, наприклад 0 і 180 градусів чи 0, 90, 180 і 270 градусів.

У швидкісних модемах часто використовуються комбіновані методи модуляції, як правило, амплітудна в поєднанні із фазовою.

Порівняння спектрів сигналів з різними видами модуляції

Спектр результуючого модульованого сигналу залежить від типу модуляції і швидкості модуляції, тобто бажаної швидкості передачі бітів вихідної інформації.

Розглянемо спочатку спектр сигналу при потенційному кодуванні. Нехай логічна одиниця кодується позитивним потенціалом, а логічний нуль – негативним потенціалом такої ж величини. Для спрощення обчислень припустимо, що передається інформація, яка складається з нескінченної послідовності одиниць і нулів, що чергуються, як це і показано на рис. 3.10, а. Зауважимо, що в даному випадку величини бодів і бітів у секунду збігаються.

Для потенційного кодування спектр безпосередньо виходить з формул Фур'є для періодичної функції. Якщо дискретні дані передаються з бітовою швидкістю N біт/с, то спектр складається з постійної складової нульової частоти і нескінченного ряду гармонік з частотами f_0, $3f_0$, $5f_0$, $7f_0, \ldots$, де $f_0 = N/2$. Амплітуди цих гармонік зменшуються досить повільно – з коефіцієнтами 1/3, 1/5,1/7, ... від амплітуди гармоніки f_0 (рис. 3.2.4, а). В результаті спектр потенційного коду потребує для якісної передачі широку смугу пропущення. Крім того, потрібно врахувати, що

реально спектр сигналу постійно змінюється в залежності від того, які дані передаються по лінії зв'язку. Наприклад, передача довгої послідовності нулів чи одиниць зміщує спектр убік низьких частот, а в крайньому випадку, коли передані дані складаються тільки з одиниць (чи тільки з нулів), спектр складається з гармоніки нульової частоти. При передачі одиниць, що чергуються, і нулів постійна складова відсутня. Тому спектр результуючого сигналу потенційного коду при передачі довільних даних займає смугу від деякої величини, близької до 0 Гц, до приблизно $7f_0$ (гармоніками з частотами вище $7f_0$ можна знехтувати через їх малий внесок у результуючий сигнал).

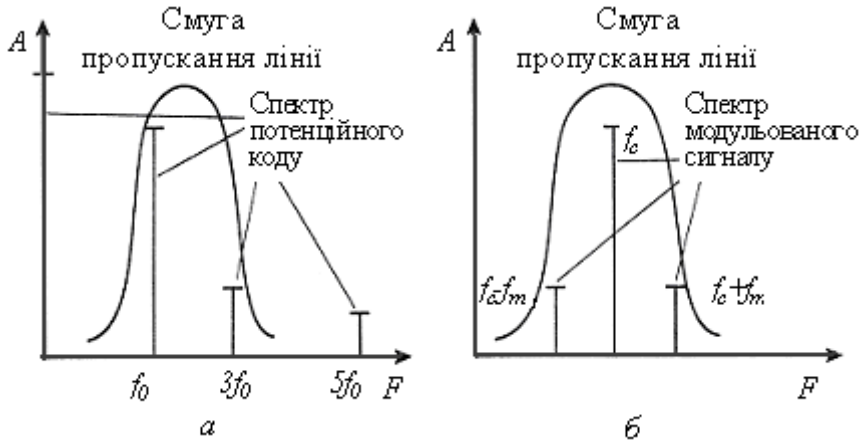


Рисунок 3.10 – Амплітуда модуляції неперервного процесу

Для каналу тональної частоти верхня границя при потенційному кодуванні досягається для швидкості передачі даних у 971 біт/с, а нижня неприйнятна для будь-яких швидкостей, тому що смуга пропускання каналу починається з 300 Гц. У результаті потенційні коди на каналах тональної частоти ніколи не використовуються.

При амплітудній модуляції спектр складається із синусоїди несучої частоти f_c і двох бічних гармонік: $(f_c + f_m)$ і $(f_c - f_m)$, де f_m – частота зміни інформаційного параметра синусоїди, що збігається зі швидкістю передачі даних при використанні двох рівнів амплітуди (рис. 3.2.4, б). Частота f_m визначає пропускну здатність лінії при даному способі кодування. При невеликій частоті модуляції ширина спектра сигналу буде також невеликою (рівною $2f_m$), тому сигнали не будуть спотворюватися лінією, якщо її смуга пропускання буде більша чи дорівнюватиме $2f_m$. Для каналу тональної частоти такий спосіб модуляції прийнятний при швидкості передачі даних не більше 3100/2=1550 біт/с. Якщо ж для подання даних вико-

ристовуються 4 рівні амплітуди, то пропускна здатність каналу підвищується до 3100 біт/с.

При фазовій і частотній модуляції спектр сигналу виходить більш складним, чим при амплітудній модуляції, тому що бічних гармонік тут утвориться більш двох, але вони також симетрично розташовані щодо основної несучої частоти, а їх амплітуди швидко зменшуються. Тому ці види модуляції також добре підходять для передачі даних по каналу тональної частоти.

Для підвищення швидкості передачі даних використовують комбіновані методи модуляції. Найбільш розповсюдженими є методи квадратурної амплітудної модуляції (Quadrature Amplitude Modulation, QAM). Ці методи основані на поєднанні фазової модуляції з 8 значеннями величин зміщення фази й амплітудної модуляції з 4 рівнями амплітуди. Однак з можливих 32 комбінацій сигналу використовуються далеко не всі. Наприклад, у кодах Трелліса припустимі всього 6, 7 чи 8 комбінацій для подання вихідних даних, а інші комбінації є забороненими. Така надмірність кодування потрібна для розпізнавання модемом помилкових сигналів, що є наслідком перекручувань через перешкоди на телефонних каналах, які комутуються дуже значними за амплітудою і тривалими за часом.

Імпульсна модуляція

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.

У залежності від вигляду модульованого параметра розрізняють такі види модуляції.

1. Амплітудно-імпульсна модуляція (АІМ), A – var. (рис. 3.11)

2. Широтно-імпульсна модуляція (ШІМ), τ_2 – var.

3. Фазово-імпульсна модуляція (ФІМ), τ_1 – var, тобто змінюється положення імпульсу на інтервалі T.

4. Частотно-імпульсна модуляція (ЧІМ), $f = 1/T$ – var.

Для прямокутних імпульсів найбільш широко використовуються амплітудно-імпульсна (АІМ) і широтно-імпульсна (ШІМ) модуляції.

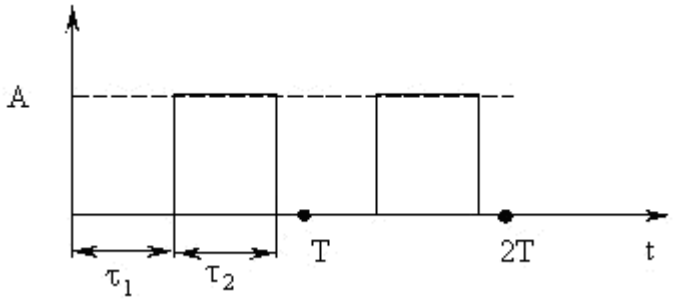


Рисунок 3.11 – Основні види імпульсної модуляції

Модуляцію розрізняють також за характером зв'язку між вхідним (модулюючим) сигналом і модульованим параметром на модуляцію І-го, ІІ-го роду (рис. 3.12).

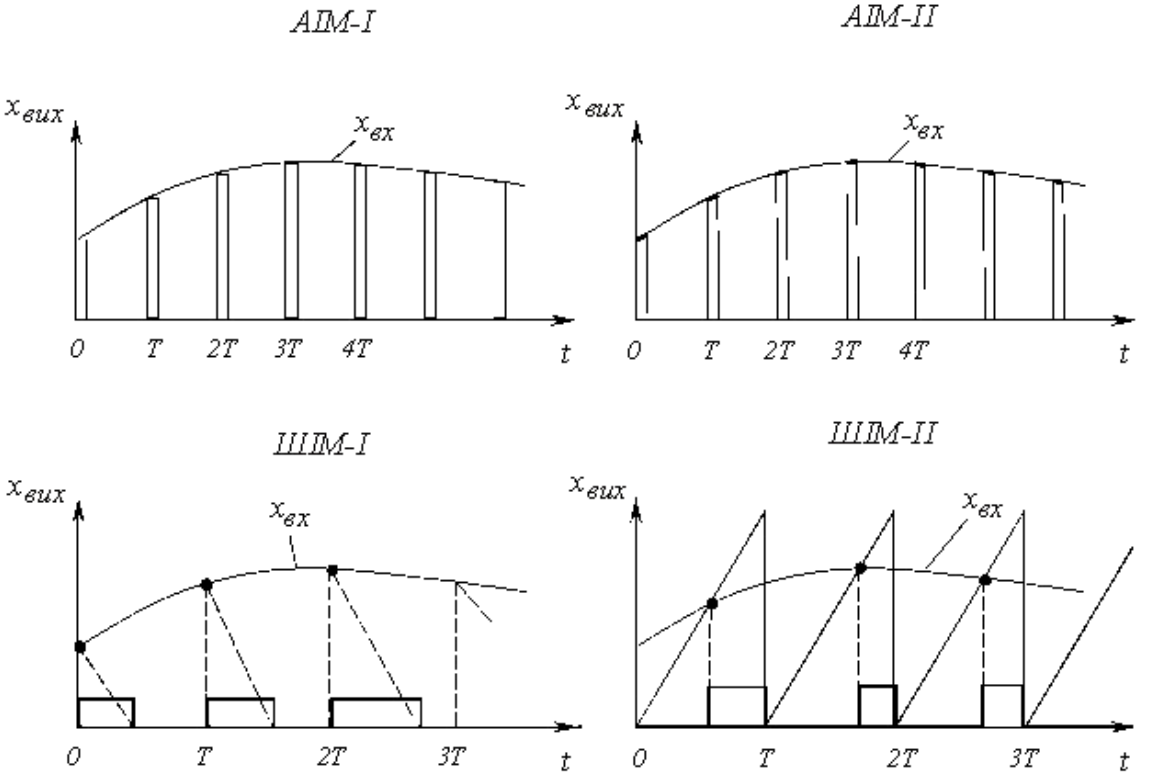


Рисунок 3.12 – Види імпульсної модуляції за характером зв'язку між вхідним сигналом і модульованим параметром

Якщо модульований параметр визначається значеннями вхідного сигналу у фіксовані, рівновіддалені один від одного моменти часу (тактові), то це модуляція першого роду (АІМ-І, ШІМ-І і т. д.) Якщо значення модульованого параметра визначаються деяким функціоналом від вхідного сигналу, або визначаються в результаті розв’язання трансцендентних рівнянь, то це модуляція другого роду (АІМ-ІІ, ШІМ-ІІ і т. д.) Іншими словами, якщо модульований параметр у процесі існування імпульсу залишаєть-

ся постійним, то це модуляція першого роду, якщо ж модульований параметр змінюється відповідно до поточного значення вхідного сигналу – то це модуляція другого роду.

Модуляційною характеристикою імпульсного елемента називається залежність величини модульованого параметра імпульсної послідовності від відповідних дискретних значень вхідної величини. Ця характеристика може бути лінійною або нелінійною. ШІМ і ЧІМ нелінійні за своєю природою. Так, для ШІМ найменша тривалість імпульсу дорівнює нулю, а найбільша – періоду T. Модуляційна характеристика може мати вигляд як на рис. 3.13.

Рисунок 3.13 – Приклад модуляційної характеристики

Як зазначалось, АІМ полягає в зміні збільшення амплітуди імпульсів пропорційно функції керуючого сигналу при постійних тривалості імпульсів і періоді їхнього проходження:

$$U(t) = U_0 + ks(t), \quad i = const, \quad T = const.$$

Спектр АІМ розглянемо на прикладі модулювання однотонального сигналу $s(t)$. Напишемо рівняння модульованого сигналу в такій формі:

$$u(t) = (1 + M\cos(\omega t))f(t),$$

де $f(t)$ – періодична послідовність прямокутних імпульсів із частотою ω, яку можна апроксимувати за Фур'є (без врахування фази):

$$f(t) = U_0 + U_n\cos(n\omega t).$$

Підставляючи значення $f(t)$ у вираз для $u(t)$, одержуємо:

$$u(t) = (1 + M\cos(\omega t))U_0 + U_n\cos(n\omega t)(1 + M\cos(\omega t)),$$

У цілому спектр нескінченний, що визначається нескінченністю спектра прямокутних імпульсів. З'являються бічні складові, що відповідають спектру моделюючої функції (при багатотональному сигналі – бічні смуги спектрів). При додатковому високочастотному заповненні імпульсів весь спектр зміщується в область високих частот на частоту заповнення.

Широтно-імпульсна модуляція (ШІМ, в англійській термінології pulse width modulation, PWM), що іноді називають модуляцією за тривалістю імпульсів, полягає в керуванні тривалістю імпульсів пропорційно функції керуючого сигналу при постійних амплітуді імпульсів і періоді проходження по фронту імпульсів:

$$t = t_0 + ks(t), \quad U = const, \quad T = const.$$

Розглянемо виконання ШІМ у найпростішому варіанті на прикладі гармонічного коливання. Передана крива 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 інтервал дискретизації даних.

Спектр сформованого сигналу ШІМ у початковій частині містить постійну складову середнього рівня сигналу й пік частоти гармоніки, закодованої в ШІМ-сигналі. Якщо виділити зі спектра ці дві складові, то відновлюється вихідний сигнал з похибкою квантування. Природно, що при малому числі рівнів квантування похибка відновлення вихідного гармонічного сигналу дуже велика.

Відмітимо, що широтно-імпульсна модуляція з наступним виділенням постійної складової може досить ефективно використовуватися (і використовується) для спостереження за середнім рівнем сигналу й автоматичного регулювання його динамічного діапазону, як, наприклад, у системах установки гучності звуку і яскравості кольорів і зображення в цілому в сучасних телевізійних установках.

Часова імпульсна модуляція є девіацією імпульсів по часовій осі за законом сигналу, що модулює, і власне кажучи аналогічна кутовій модуляції гармонічної несучої. Вона також може бути фазовою (ФІМ) або частотною (ЧІМ).

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

3.3 Випадкові сигнали та їх характеристики

Для опису випадкових процесів використовують методи теорії ймовірностей. У загальному випадку повною характеристикою випадкового процесу є його багатовимірна щільність імовірностей. Для стаціонарних гаусівських процесів одновимірна щільність імовірностей визначається дисперсією випадкового процесу. Для опису гаусівських процесів достатніми характеристиками є й кореляційна функція процесу. Однієї з характеристик випадкового сигналу є його спектральна щільність потужності, пов'язана з кореляційною функцією узагальненим перетворенням Фур'є.

Спектр випадкового процесу є суцільним. Для випадкових процесів з постійною спектральною щільністю й нескінченною смугою частот потужність нескінченна, а кореляційна функція є дельта-функцією. Такий процес має нескінченну дисперсію, є некорельованим і називається білим шумом. У випадкового процесу з постійною спектральною щільністю в обмеженій смузі частот потужність є скінченною і її можна визначити.

Отож, розглянемо дане питання більш детально. Під випадковим (стохастичним) процесом розуміють таку випадкову функцію часу $U(t)$, значення якої в кожен момент часу випадкові. Конкретний вигляд випадкового процесу, зареєстрованого у деякому досліді, називають реалізацією випадкового процесу. Дані, які характеризують всю множину можливих реалізацій називаються ансамблем.

Основними ознаками, за якими класифікують випадкові процеси є: простір станів, часовий параметр та статичні залежності між випадковими величинами $U(t_i)$ в різні моменти часу t_i.

Простором станів (англ. *space of states*) називають множину можливих значень випадкової величини $U(t_i)$. Випадковий процес у якому множини станів складають континуум, а зміна станів можлива в будь-які моменти часу, називають неперервним випадковим процесом. Якщо зміна станів допускається лише в кінцевому чи поточному числі моментів часу, то говорять про неперервну випадкову величину.

Відповідно до визначення випадковий процес $U(t)$ може бути описаний системою N звичайно залежних випадкових величин $U_1 = U(t_1),...,U_n - U(t_n)$, взятих в різні моменти часу $t_1, t_2,...,t_n$. При необмеженому збільшенні числа N така система еквівалентна випадковому процесу, що розглядається.

В більшості випадків для характеристики випадкових процесів використовують моментні функції перших двох порядків: математичне сподівання, дисперсію, а також кореляційну функцію

$$m\left[U(t_1)\right] - M\left[U(t_1)\right] = \int_{-\infty}^{\infty} U_1 P_1\left(\frac{U_1}{t_1}\right) dU_1,$$

де $P_1(U_1;t_1)$ – одновимірна щільність імовірності або одновимірна функція розподілення випадкового процесу. Фізико-математичне сподівання виражає значення сукупності вибірок випадкового процесу у визначений момент часу t_1.

Дисперсія – це математичне сподівання квадрата відхилення величини $U(t_1)$ від математичного сподівання у визначений момент часу t_1.

Дисперсія виражається формулою

$$D\left[U(t_1)\right] = M\left\{\left[U(t_1) - mU(t_1)\right]^2\right\} = \int_{-\infty}^{\infty} \left(U_1 - M\left[U(t_1)\right]^2\right) P(U_1, t_1) dU_1.$$

Вона виражає розкид значення випадкової величини навколо математичного сподівання. Корінь квадратний з дисперсії прийнято називати середнім квадратичним відхиленням випадкової величини

$$G^2\left[U(t_1)\right] = M\left[U^2(t_1)\right] = \int_{-\infty}^{\infty} U^2 P_1(U_1, t_1) dU_1.$$

Фізично початковий момент другого порядку є повною середньою потужністю випадкової величини.

Випадкові процеси можуть мати однакові математичні сподівання й дисперсію, але різко відрізняються за швидкістю зміни своїх значень у часі рис 3.14.

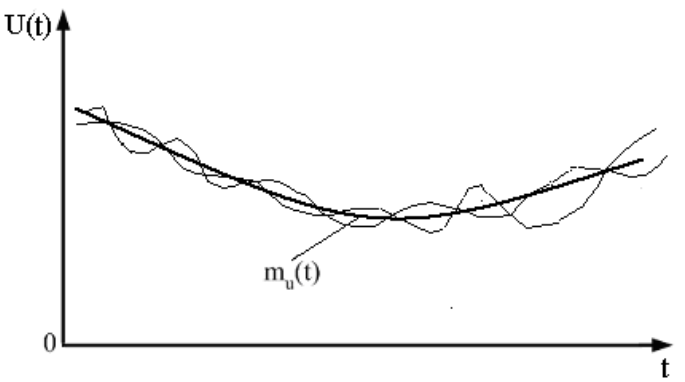


Рисунок 3.14 – Математичне сподівання для різних процесів

Тому для оцінювання ступеня статичної залежності миттєвих значень процесу $U(t)$ в будь-які моменти часу t_1 та t_2 використовується випадкова функція аргументів $R_u(t_1,t_2)$, яка називається автокореляційною або просто кореляційною функцією.

При конкретних аргументах t_1 та t_2 вона дорівнює кореляційному моменту значень процесу $U(t_1)$ та $U(t_2)$

$$R_u(t_1,t_2)=m_1\left[U(t_1),U(t_2)\right].$$

Двовимірним законом розподілу випадкової функції $X(t)$ називається закон розподілу $f(x_1,x_2,t_1,t_2)$ системи двох випадкових розмірів $X(t_1)$ та $X(t_2)$, що є значеннями випадкової функції для різних значень аргументів $t=t_1$ та $t=t_2$.

Математичним сподіванням випадкової функції $X(t)$ називається невипадкова функція $m_x(t)$, яка при кожному даному значенні аргументу дорівнює математичному сподіванню значення випадкової функції при тому ж значенні аргументу $m_x(t)=M\left[X(t)\right]$.

Кореляційною функцією випадкової функції $X(t)$ називається невипадкова функція двох аргументів $K_x(t_1,t_2)$, яка при кожній парі значень t_1 та t_2 дорівнює кореляційному моменту відповідних значень випадкової функції

$$K_x(t_1,t_2)=M\left[X(t_1),X(t_2)\right],$$

де $X(t)=X(t)-m_x(t)$ – центрована випадкова функція.

При $t_1=t_2=t$ кореляційна функція $K_x(t_1,t_2)$ перетворюється в дисперсію випадкової функції $X(t)$, тобто

$$K_x(t_1,t_2)=D_x(t)=\sigma_x^2(t).$$

Нормованою кореляційною функцією випадкової функції $X(t)$ називається функція

$$r_x(t_1,t_2)=\frac{K_x(t_1,t_2)}{\sigma_x(t_1)\sigma_y(t_2)}.$$

Взаємною кореляційною функцією двох випадкових функцій $X(t)$ та $Y(t)$ називається функція двох аргументів $R_{xy}(t_1,t_2)$, яка при кожній довільно обраній парі їх значень дорівнює кореляційному моменту відповідних значень $X(t_1)$ та $Y(t_2)$ цих випадкових функцій

$$R_{xy}(t_1,t_2) = M\left[X(t_1), Y(t_2)\right].$$

Нормованою взаємною кореляційною функцією двох випадкових функцій $X(t)$ та $Y(t)$ називається функція

$$r_{xy}(t_1,t_2) = \frac{R_{xy}(t_1,t_2)}{\sigma_x(t_1)\sigma_y(t_2)}.$$

Випадкові функції $X(t)$ та $Y(t)$ називаються некаліброваними, якщо $R_{xy}(t_1,t_2) \equiv 0$.

Канонічним розкладанням випадкової функції $X(t)$ називається подання її у вигляді

$$X(t) = m_x(t) + \sum_{k=1}^{n} X_k \varphi_k(t),$$

де X_k, $(k = 1,2,...,n)$ – центрована некорельована випадкова величина з дисперсією D_k; $\varphi_k(t)$ – невипадкова функція.

Форми сигналів

За формою сигнали поділяються на неперервні та дискретні.

Неперервні сигнали можуть приймати неперервну множину значень (континуум) в певному інтервалі (в часі і за рівнем).

Дискретні сигнали описуються за допомогою кінцевого набору чисел або дискретних значень певної функції.

Тобто, неперервні сигнали (рис. 3.15, а) зображуються функцією, безперервною в часі на відрізку спостереження $x(t)$, а дискретні (рис. 3.15, б) поступають тільки в певні моменти часу та зображаються дискретною функцією $x_g(t)$.

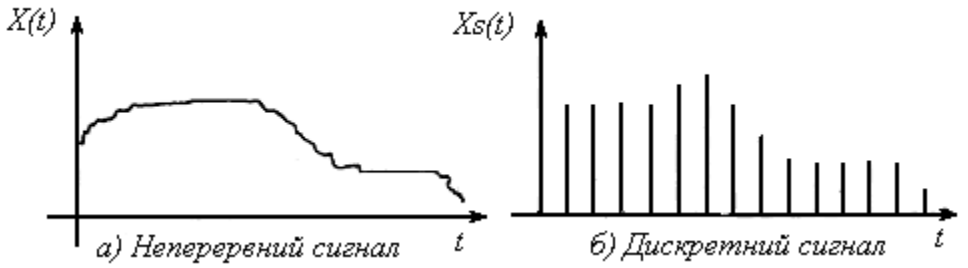


Рисунок 3.15 – Неперервні та дискретні сигнали

До елементарних детермінованих сигналів належить зокрема і одинична функція (стрибок).

При вивченні властивостей каналів передачі інформації, сигналів та завад абстрагуються від їх конкретної фізичної природи та призначення і оперують з їх моделями. *Модель* (англ. *model*) – це є вибраний спосіб опису об'єкта, процесу або явища, який відображає суттєві з погляду розв'язання даної задачі фактори. Оцінка електронних систем потребує виявлення кількісних співвідношень між основними параметрами джерела інформації і системи, тому дослідження здійснюється на математичних моделях.

Як відомо, інформація міститься в повідомленнях, зміст яких наперед невідомий і джерело інформації видає їх з певною ймовірністю. Передбачити точно зміну інформаційного параметра неможливо. Це означає, що інформаційний сигнал принципово являє собою випадкове коливання і його аналітичною моделлю може бути лише випадковий процес, який визначається імовірнісними характеристиками. Детерміноване коливання (сигнал) відображає відоме повідомлення, і немає сенсу його передавати. Йому відповідає модель у вигляді функції, що повністю визначена в часі. Разом з тим, вивчення моделей детермінованих сигналів дозволяє робити аналіз більш складних випадкових сигналів.

Сукупність детермінованих сигналів може подавати випадковий процес.

Навіть при експериментальному аналізі досить доречно вводити допоміжні функції (рис. 3.16).

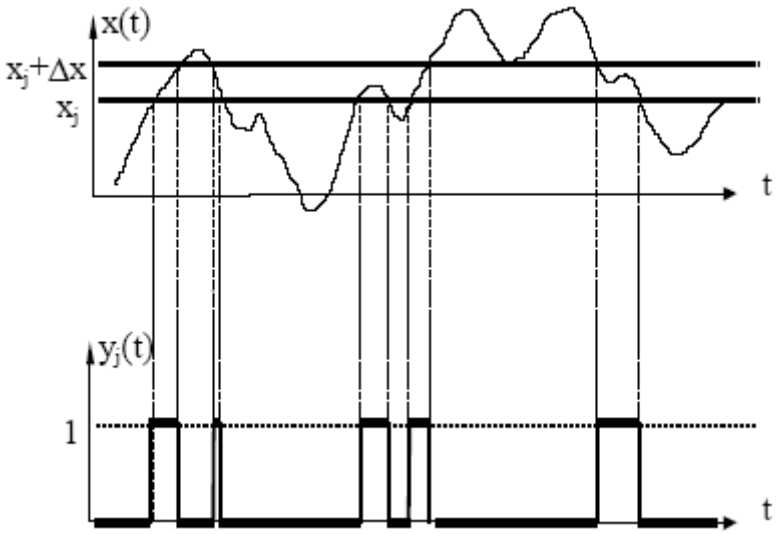


Рисунок 3.16 – Формування реалізацій $y_j(t)$ випадкових функцій $\eta_j(t)$ з реалізації $x(t)$ випадкового процесу $\xi(t)$

Існуючі аналізатори законів розподілу дозволяють визначати емпіричні функції розподілу та гістограми досліджуваних випадкових процесів. Блок-схема аналізатора наведена на рис. 3.17.

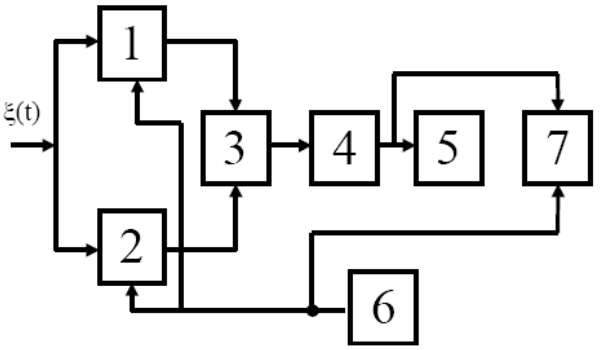


Рисунок 3.17 – Блок-схема аналізатора законів розподілу

Блок-схема складається з таких блоків:

1 – амплітудний селектор (компаратор) з порогом x_j;

2 – амплітудний селектор (компаратор) з порогом x_{j+1};

3 – пристрій віднімання;

4 – пристрій усереднення;

5 – індикатор;

6 – генератор пилкоподібної напруги;

7 – осцилограф.

Аналіз законів розподілу здійснюється, наприклад, для таких випадкових процесів як гаусів шум, синусоїдальний і пилкоподібний сигнал з випадковими фазами, адитивної суміші корисного сигналу та завади.

Спектральне подання випадкових сигналів

Розглянемо спектральне подання стаціонарних випадкових процесів. Стаціонарні випадкові процеси – це процеси, що протікають у часі однорідно, мають вигляд неперервних випадкових коливань навколо середнього значення $(-,+)$.

Якщо математичне сподівання, дисперсія, середнє квадратичне відхилення та кореляція є постійними, то такі процеси – стаціонарні.

Якщо існують випадкові процеси, що не витримують таких умов, але на деякому певному інтервалі відхиленням даних параметрів від константи можна знехтувати, то такий процес називають квазістаціонарним.

В будь-якій динамічній системі випадковий процес починається з так званого “перехідного” процесу і потім переходить в установлений режим, який з деяким наближенням можна вважати стаціонарним. Потрібно сказати, що стаціонарні випадкові процеси неперервні в часі, а значить не мають ні початку, ні кінця. Відомо два поняття: стаціонарність в обмеженому розумінні і стаціонарність у широкому.

Під стаціонарними процесами у вузькому смислі розуміють випадкові процеси, для яких функції розподілу щільності імовірності $\omega_n(x_1,t_1;x_2,t_2;...;x_n,t_n)$ вільного порядку n не змінюються при будь-якому зсуві всієї групи точок $t_1,t_2,...,t_n$ повздовж осі часу

$$\omega_n(x_1,t_1;x_2,t_2;...;x_n,t_n)=\omega_n(x_1,t_1+\tau;x_2,t_2+\tau;...;x_n,t_n+\tau).$$

З наведеного визначення можна сказати, що для стаціонарних процесів:

а) одновимірна функція розподілу щільності імовірності не залежить від часу

$$\omega_1(x,t_1)=\omega_1(x,t_1+\tau)=\omega_1(x);$$

б) двовимірна функція розподілу щільності імовірності залежить тільки від різниці часу $t_2-t_1=\tau_1$

$$\omega_2(x_1,t_1;x_2,t_2)=\omega_2(x_1,x_2;t_2-t_1)=\omega_2(x_1,x_2,\tau);$$

в) тривимірна функція розподілу щільності імовірності залежить тільки від двох різниць часу $t_2-t_1=\tau_1$ та $t_3-t_1=\tau_2$

$$\omega_3(x_1,t_1;x_2,t_2;x_3,t_3)=\omega_3(x_1,x_2,x_3;t_2-t_1;t_3-t_1)=\omega_3(x_1,x_2,x_3,\tau).$$

Оскільки математичне сподівання і дисперсія виражаються через одновимірну функцію розподілу щільності імовірності, то для стаціонарного процесу математичне сподівання й дисперсія не залежать від часу. Унаслідок залежності двовимірної функції розподілу тільки від різниці часу $t_2-t_1=\tau_1$, кореляційна функція стаціонарного процесу також залежить тільки від різниці часу τ.

Стаціонарною випадковою функцією в широкому смислі називається така випадкова функція $X(t)$, математичне сподівання якої постійне, а кореляційна функція залежить тільки від різниці аргументів, тобто

$$m_x(t)=const,\ K_x(t_1,t_2)=k_x(\tau),$$

де $\tau=t_2-t_1$.

Дисперсія стаціонарної випадкової функції постійна

$$D_x(t) = K_x(t_1, t_2) = k_x(0) = const.$$

Нормована кореляційна функція $\rho_x(\tau)$ стаціонарної випадкової функції $X(t)$ має вигляд

$$\rho_x(\tau) = \frac{k_x(\tau)}{D_x} = \frac{k_x(\tau)}{k_x(0)}.$$

Спектральне розкладання

$$X(t) = m_x + \sum_{k=0}^{\infty} (Y_k \cos \omega_k t + Z_k \sin \omega_k t),$$

де Y_k, Z_k, $(k = 0,1,2,...,n)$ – центровані некорельовані випадкові розміри.

Спектри потужності випадкових функцій визначаються аналогічно спектрам потужності детермінованих сигналів. Середня потужність випадкового процесу $X(t)$, зареєстрованого в процесі однієї реалізації на інтервалі 0 - T з використанням рівності Парсеваля може бути обчислена за формулою:

$$W_T = \int_0^T \left[X^2(t)/T\right] dt = \int_{-\infty}^{\infty} \left[|X_T(f)|^2 / T\right] df,$$

де $X_T(f)$ – спектральна густина одиничної реалізації $x(t)$.

При збільшенні інтервалу T енергія процесу на інтервалі необмежено наростає, а середня потужність прямує до певної границі:

$$W = \int_{-\infty}^{\infty} \left[\lim_{T \to \infty} \frac{1}{T} |X_T(f)|^2 \right] df,$$

де підінтегральна функція є спектральною густиною потужності даної реалізації випадкового процесу:

$$W(f) = \lim_{T \to \infty} \frac{1}{T} |X_T(f)|^2.$$

Досить часто цей вираз називають просто спектром потужності. Густина потужності є суттєвою, невід'ємною та парною функцією частоти. В загальному випадку густину потужності необхідно усереднювати за множиною реалізацій, але для ергодичних процесів допустимо усереднювати за однією тривалою в часі реалізацією.

Спектральна щільність будь-якої стаціонарної випадкової функції є невід'ємною функцією ω.

Спектральна щільність $S_x(\omega)$ і кореляційна функція $k_x(\tau)$ пов'язані перетворенням Фур'є. У дійсній формі вони мають вигляд

$$S_x(\omega) = \frac{2}{\pi}\int_0^{\infty} k_x(\tau)\cos\omega\tau d\tau,$$

$$k_x(\tau) = \int_0^{\infty} S_x(\omega)\cos\omega\tau d\omega,$$

приймаючи, що $\tau = 0,\ k_x(0) = D_x$ отримаємо $D_x = \int_0^{\infty} S_x(\omega)d\omega$.

Нормованою спектральною щільністю $\sigma_x(\omega)$ називається відношення спектральної щільності до дисперсії випадкової функції

$$\sigma_x(\omega) = \frac{S_x(\omega)}{D_x}.$$

Теорема Хінчина - Вінера

Розглянемо сигнал $q(t)$, який є однією реалізацією випадкового стаціонарного ергодичного процесу тривалістю T. Для сигналу $q(t)$ можна визначити спектр $Q(\omega)$. Якщо зсунути на τ реалізацію процесу, то отримаємо спектр $Q(\omega)\exp(j\omega\tau)$. Для дійсних сигналів $Q(\omega) = Q^*(\omega)$ рівність Парсеваля за енергією взаємодії двох сигналів

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

може бути записана в такій формі:

$$\int_{-\infty}^{\infty} q(t)q(t+\tau)dt = (1/2\pi)\int_{-\infty}^{\infty} Q(\omega)Q^*(\omega)\exp(j\omega\tau)d\omega.$$

Поділимо обидві частини рівності на T і перейдемо до границі при $T \to \infty$, при цьому в її лівій частині побачимо вираз для функції кореляції, а в правій частині – перетворення Фур'є спектра потужності сигналу:

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T q(t)\,q(t+\tau)dt = \lim_{T\to\infty}\frac{1}{2pT}\int_{-\infty}^{\infty}|Q(\omega)|^2\exp(j\omega\tau)\,d\omega,$$

$$R(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} W(\omega)\exp(j\omega\tau)d\omega.$$

Звідси випливає, що кореляційна функція випадкового стаціонарного ергодичного процесу є зворотним перетворенням Фур'є його спектра потужності. Тому для спектра потужності випадкового процесу маємо пряме перетворення Фур'є:

$$W(\omega)=\int_{-\infty}^{\infty}R(\tau)\exp(-j\omega\tau)d\tau.$$

В цьому і полягає зміст теореми Хінчина – Вінера. Функції $W(\omega)$ та $R(\tau)$ – дійсні та парні; в тригонометричній формі мають вигляд:

$$R(\tau)=2\int_{0}^{\infty}W(f)\cos(2\pi f\tau)df,$$

$$W(f)=2\int_{0}^{\infty}R(\tau)\cos(2\pi f\tau)d\tau.$$

3.4 Білий шум

Для аналізу радіотехнічних кіл та сигналів велику практичну цінність має специфічний випадковий процес, який теоретично має рівномірний спектр потужності по усій смузі частот:

$$W_x(\omega)=W_0,\quad -\infty<\omega<\infty.$$

Такий ідеалізований випадковий процес був названий білим шумом за аналогією з білим світлом, яке у видимій частині має практично рівномірний суцільний спектр.

Визначимо функцію кореляції білого шуму:

$$R_x(\tau)=\frac{W_0}{\pi}\int_{-\infty}^{\infty}e^{j\omega x}d\omega=\frac{W_0}{2\pi}\delta(\tau).$$

Функція кореляції дорівнює нулю усюди, крім $t=0$, де вона нескінченна. Отож, миттєві значення напруги (чи потужності) білого шуму в будь-які два як завгодно близькі моменти часу некорельовані, і тому інтервал кореляції $\tau_k=0$. Тому білий шум часто називають дельта Т-корельованим.

Проведемо аналітичне тлумачення поняття «білого шуму». Якщо за рівністю Парсеваля визначити його дисперсію, то отримаємо значення $\sigma_x^2=\infty$, що є абсурдним з фізичної точки зору. Цей результат є наслідком абстрактного уявлення про нескінченність протяжності спектра потужності по осі частот. Однак використання поняття «білого шуму» дуже зручне для аналізу характеристик випадкового процесу на виході радіотехнічних пристроїв, коли смуга пропускання радіотехнічного кола порівнянна з шириною спектра реального шуму.

При обробці в приймачі оптимальним фільтром ширина спектра білого шуму обмежується. Як наслідок шум на виході оптимального фільтра стає уже корельованим (рис. 3.18).

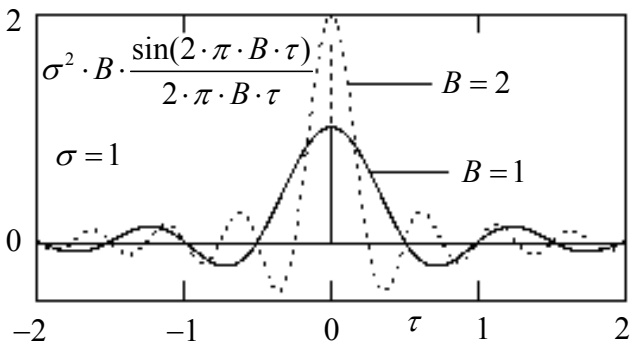


Рисунок 3.18 – Функції кореляції білого шуму

Середня потужність (дисперсія) білого шуму безмежно велика. Часова реалізація $x(t)$ білого шуму має голчату структуру з нескінченно тонкими сплесками обох знаків. Зазвичай розглядають білий шум, якщо ширина частот завади перевищує смугу частот пропускання апаратури, а амплітуди частот приблизно сталі. До таких завад відносять флуктуаційні шуми, завади в багатоканальних системах зв'язку.

3.5 Проходження випадкових сигналів через ЕС

Ефективна ширина спектра випадкового сигналу визначається як

$$B_k = G_x(f)df / G_x(f)_{\max} = S_x(f)df / S_x(f)_{\max} = K_x(0) / S_x(f)_{\max},$$

де $G_x(f)_{\max}$ – максимальне значення функції $G_x(f)$ і безпосередньо випливає з аналізу випадкових процесів.

Зауважимо, що ширина спектра є практичною характеристикою випадкового процесу і обчислюється, як правило, для реальних частот за одностороннім спектром процесу.

Нехай на вхід чотириполюсника (рис. 3.19) (ЕС) з передатною функцією $K(i\omega)$ і імпульсною характеристикою $g(t)$ діє випадковий процес $s(t)$ із заданими статичними характеристиками. Необхідно знайти його статичні характеристики $s_{вих}(t)$ на виході чотириполюсника.

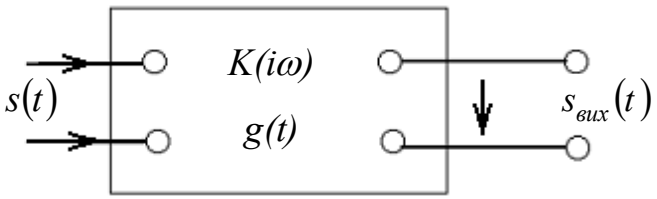


Рисунок 3.19 – Лінійний чотириполюсник з постійними параметрами

Як вже зазначалось, основними характеристиками випадкового процесу є: розподіл ймовірностей, кореляційна функція, спектральна густина потужності.

Визначити значення двох останніх досить просто. Значно складніше знайти закон розподілу випадкового процесу на виході лінійного кола. В загальному випадку при довільному розподілі процесу на виході досить складно знайти розподіл випадкового процесу на виході інерційного кола.

Процес спрощується лише при використанні нормального розподілу вхідного процесу, оскільки при будь-яких лінійних операціях з гаусовим процесом розподіл залишається нормальним і змінюються лише функції $R(\tau)$ та $W(\omega)$. Тому, якщо задано густину ймовірності вхідного процесу (з нульовим середнім)

$$p(s)=\frac{1}{\sqrt{2\pi}\sigma_s}\exp\left(-\frac{s^2}{2\sigma_s^2}\right),$$

то густина ймовірності на виході лінійного кола

$$p(s_{вих})=\frac{1}{\sqrt{2\pi}\sigma_{s\,вих}}\exp\left(-\frac{s^2{}_{вих}}{2\sigma^2_{s\,вих}}\right).$$

Дисперсія $D_{s\,вих}=\sigma^2_{s\,вих}$ легко знаходиться зі спектра чи з кореляційної функції. Таким чином, аналіз передачі гаусових процесів через лінійні кола по суті зводиться до спектрального чи кореляційного аналізу.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. У чому полягає інформаційний процес?
2. Які існують основні типи сигналів?
3. Що лежить у основі поняття спектра сигналу?
4. Для чого застосовується перетворення Фур'є?
5. Назвіть основні властивості перетворення Фур'є.
6. При виконанні яких умов є можливим подання неперіодичної функції інтегралом Фур'є?
7. Що приймають за практичну ширину спектра сигналу?
8. Для чого застосовують дискретне перетворення Фур'є?
9. Наведіть основні види модуляції сигналів.

10. Для яких сигналів і в яких випадках величини бод і біт у секунду збігаються?
11. Які види модуляції розрізняють за характером зв'язку між вхідним (модулюючим) сигналом і модульованим параметром?
12. У чому полягає ШІМ?
13. Що являє собою випадковий процес?
14. Що являє собою математичне сподівання у відношенні до сигналів?
15. Для чого призначений аналізатор закону розподілів?
16. Перелічіть основні характеристики випадкових процесів.
17. У чому сутність теореми Хінчина – Вінера?
18. Що розуміють під поняттям “білий шум”?

4 ДИСКРЕТИЗАЦІЯ, КВАНТУВАННЯ ТА КОДУВАННЯ СИГНАЛІВ В ЕЛЕКТРОННИХ СИСТЕМАХ

4.1 Цифрове подання сигналів

Під дискретизацією сигналів розуміють перетворення функцій неперервних змінних у функції дискретних змінних, по яких вихідні неперервні функції можуть бути відновлені із заданою точністю. Роль дискретних відліків виконують, як правило, квантовані значення функцій у дискретній шкалі координат.

Система дискретного часу – це алгоритм із вхідною послідовністю $s(k)$ і вихідною послідовністю $y(k)$, що може бути лінійною або нелінійною, інваріантною або змінною в часі. Система дискретного часу лінійна й інваріантна в часі, якщо вона відповідає принципу суперпозиції (відгук на кілька входів дорівнює сумі відгуків на кожен вхід окремо), а затримка вхідного сигналу викликає таку ж затримку вихідного сигналу.

Суть дискретизації аналогових сигналів полягає в тому, що неперервність у часі аналогової функції $s(t)$ замінюється послідовністю коротких імпульсів, амплітудні значення яких визначаються за допомогою вагових функцій, або безпосередньо вибірками (відліками) миттєвих значень сигналу $s(t)$ у моменти часу t_n. Подання сигналу $s(t)$ на інтервалі T сукупністю дискретних значень c_n записується у вигляді:

$$(c_1, c_2, ..., c_n) = A[s(t)],$$

де A – оператор дискретизації.

Запис операції відновлення сигналу $s(t)$:

$$s'(t) = B[(c_1, c_2, ..., c_n)].$$

Вибір операторів A та B визначається необхідною точністю відновлення сигналу. Найбільш простими є лінійні оператори. У загальному випадку:

$$c_n = q_n(t)s(t)dt,$$

де $q_n(t)$ – система вагових функцій.

Відліки в останньому виразі пов'язані з операцією інтегрування, що забезпечує високу завадостійкість дискретизації. Однак у силу складності технічної реалізації "зваженого" інтегрування, останнє використається досить рідко, при високих рівнях перешкод. Більш широке поширення одер-

жали методи, при яких сигнал $s(t)$ заміняється сукупністю його миттєвих значень $s(t_n)$ у моменти часу t_n. Роль вагових функцій у цьому випадку виконують ґратчасті функції. Відрізок часу t між сусідніми відліками називають кроком дискретизації. Дискретизація називається рівномірною із частотою $F = 1/t$, якщо значення t постійне по всьому діапазону перетворення сигналу. При нерівномірній дискретизації значення t між вибірками може змінюватися за певною програмою або залежно від зміни будь-яких параметрів сигналу.

Відтворення безперервного сигналу за вибірками може проводитися як на основі ортогональних, так і неортогональних базисних функцій. Відтворююча функція $s'(t)$ відповідно подається апроксимуючим поліномом:

$$s'(t) = c_n v_n(t),$$

де $v_n(t)$ – система базисних функцій.

Оптимальними є методи дискретизації, що забезпечують мінімальний числовий ряд при заданій похибці відтворення сигналу. При неортогональних базисних функціях використаються, в основному, статичні алгебраїчні поліноми вигляду:

$$s'(t) = c_n t_n.$$

Якщо значення апроксимуючого полінома збігаються зі значеннями вибірок у моменти їхнього відліку, то такий поліном називають інтерполяційним. Як інтерполяційний поліноми звичайно використовуються багаточлени Лагранжа. Для реалізації інтерполяційних поліномів необхідна затримка сигналу на інтервал дискретизації, що в системах реального часу потребує певних технічних рішень. Як екстраполяційні поліноми використовують, як правило, багаточлени Тейлора.

Природною вимогою до вибору частоти дискретизації є внесення мінімальних перекручувань у динаміку зміни сигнальних функцій. Логічно думати, що перекручування інформації будуть тим менші, чим вища частота дискретизації F. З іншої сторони також очевидно, що чим більше значення F, тим більшою кількістю цифрових даних будуть відображатися сигнали, і тим більший час буде затрачатися на їхню обробку. В оптимальному варіанті значення частоти дискретизації сигналу F повинно бути необхідним і достатнім для обробки інформаційного сигналу із заданою точністю, тобто забезпечувати припустиму похибку відновлення аналогової форми сигналу (середньоквадратичну в цілому за інтервалом сигналу,

або за максимальними відхиленнями від дійсної форми в характерних інформаційних точках).

Під *квантуванням* (англ. *quantization*) розуміють перетворення безперервної за значеннями величини у величину з дискретною шкалою значень із скінченної безлічі дозволених, які називають рівнями квантування. Якщо рівні квантування нумеровані, то результатом перетворення є число, що може бути виражене в будь-якій числовій системі. Округлення з певною розрядністю миттєвих значень безперервної аналогової величини з рівномірним кроком за аргументом є найпростішим випадком дискретизації й квантування сигналів при їхньому перетворенні в цифрові сигнали.

Дискретизація аналогових сигналів з перетворенням у цифрову форму пов'язана із квантуванням сигналів. Сутність квантування полягає в заміні незліченної безлічі можливих значень функції, у загальному випадку випадкових, скінченною безліччю цифрових відліків, і виконується округленням миттєвих значень вхідної функції $s(t_i)$ у моменти часу t_i до найближчих значень $s_i(t_i) = n_i$, де i – крок квантування шкали цифрових відліків. Квантування з постійним кроком називається рівномірним.

При квантуванні сигналів у великому динамічному діапазоні значень крок квантування може бути й нерівномірним, наприклад, логарифмічним, тобто пропорційним логарифму значень вхідного сигналу. Установлений діапазон шкали квантування від $s_{\min}$ до $s_{\max}$ і крок квантування визначають число розподілів шкали $N = (s_{\max} - s_{\min})/i$ і відповідно цифрову розрядність квантування. У результаті дискретизації й квантування безперервна функція $s(t)$ заміняється числовою послідовністю $\{s(kt)\}$. Похибка округлення $\eta = s(t_i) - s_i(kt)$ називається шумом квантування. Необхідна точність квантування оцінюється за впливом виниклого шуму квантування на наступну обробку сигналів.

При досить малому кроці квантування будь-яке значення в його межах можна вважати рівноймовірним, при цьому значення розподілені за рівномірним законом.

Вхідний сигнал містить, як правило, адитивну суміш корисного сигналу $s(t)$ і перешкоди $q(t)$.

На практиці крок квантування вибирають звичайно таким, щоб не відбувалося помітної зміни відношення сигнал/шум.

Тепер перейдемо до подання сигналів у неперервній формі. Поштовхом до подання неперервних (аналогових) сигналів у цифровій формі

стала необхідність засекречування мовних сигналів під час другої світової війни. Ще більшим стимулом до цифрового перетворення аналогових сигналів було створення комп'ютера, який можна використовувати як кінцевий пристрій системи передачі цифрової інформації, доручаючи йому виконання логічних операцій з прийому та обробки сигналів.

У наш час у цифровій формі можна передавати будь-який вид інформації, забезпечуючи необхідну надійність за значної швидкості передачі. Сильним поштовхом у розвитку цифрових систем передачі інформації послужило створення інформаційно-обчислювальних мереж і наявність технологічної бази для виробництва високошвидкісних перемикачів, широкосмугових модемів, мультиплексорів, підсилювачів та ін.

Кодуванням є подання за певними правилами дискретних повідомлень у деякі комбінації, складені з певного числа елементів-символів. Ці елементи називаються елементами коду, а число різноманітних елементів, із яких складаються комбінації, – основою коду. Елементи коду утворюють кодові комбінації. Наприклад, якщо ми складаємо комбінації із різноманітних сполучень (0 і 1), то це – код із основою два або двійковий код. Якщо всі комбінації мають однакове число знаків, код називається рівномірним. Широко відомий код Морзе – нерівномірний код. Правило кодування звичайно виражається кодовою таблицею, в якій кожному символу повідомлення ставиться у відповідність певна кодова комбінація.

Кодове подання дискретних значень сигналу здійснюється за допомогою цифр, але не обов'язково десяткових. Нагадаємо, що в десятковій системі, називаючи число, ми вказуємо, скільки одиниць від нуля до дев'яти є в розряді одиниць, у розряді десятків, сотень, тисяч і т. д. Те саме відбувається в будь-якій іншій системі числення з іншою основою. У десятковій системі користуємося десятьма цифрами: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. У двійковій системі числення в нашому розпорядженні тільки дві цифри: 0 та 1. Тому десяткова цифра 23, наприклад, у двійковій системі виглядатиме так:

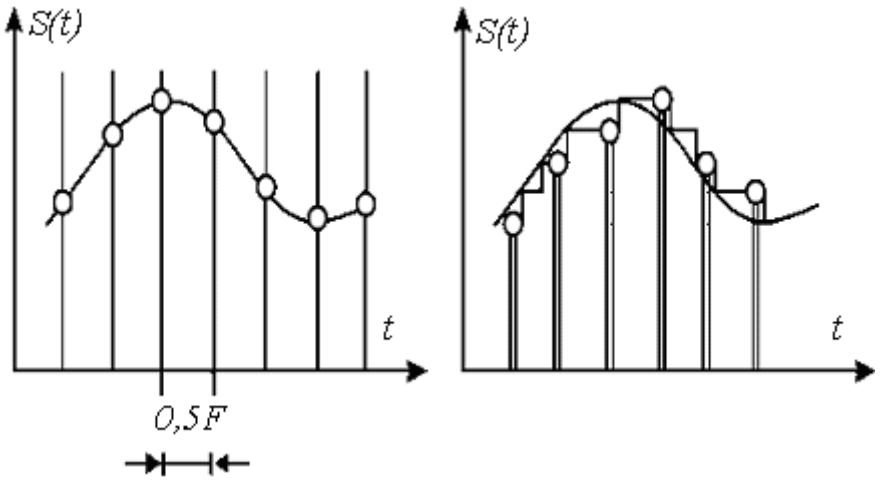
$$1 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 23,$$

і це число запишеться в послідовності розрядів як 10111.

На рис. 4.1. показаний процес цифрового подання інформації квантуванням за: а) частотою і б) амплітудою.

Двійкова система числення широко використовується в обчислювальній техніці та зв'язку внаслідок використання електричних і електронних приладів із двома елементами стану (реле в замкнутому або розімкнутому

стані, магнітний матеріал намагнічений або розмагнічений, два можливих стани тригера). Дуже зручним з погляду створення і розпізнавання є сигнал, що має два стани: "так" або "ні" (1 або 0).



а) подання сигналу, узяте через інтервал 1/2F

б) сигнал, квантований за рівнем

Рисунок 4.1 – Подання аналогового сигналу у вигляді дискретних відліків (вибірок)

Отже, однією з головних переваг передачі інформації в цифровій формі є можливість використання кодованих сигналів і оптимального в заданих умовах способу їх прийому. Важливо, що за цифрової передачі усі типи сигналів, такі як мова, музика, телебачення, дані, можуть об'єднуватися в один загальний потік інформації, передача якого формалізована. Крім того, ущільнення при одночасному використанні комп'ютера дозволяє більш ефективно використовувати спектр і час, захистити канал від несанкціонованого доступу, об'єднати в єдиний процес передачу цифрової інформації і цифрової комутації каналів та повідомлень.

Звичайно, за деякі переваги цифрових видів передачі необхідно сплачувати, зокрема, більшою шириною спектра сигналу, що випромінюється, наявністю міжсимвольних перешкод, необхідністю синхронізації систем передачі та інше. Темпи впровадження цифрової техніки говорять про те, що всі ці труднощі можливо подолати.

А тепер розглянемо відновлення сигналів. Відновлення неперервного сигналу за вибірками може проводитися як на основі ортогональних, так і не ортогональних базисних функцій. Відновлююча функція $s'(t)$ відповідно подається апроксимуючим поліномом:

$$s'(t) = \sum_n c_n v_n(t),$$

де $v_n(t)$ – система базисних функцій.

Ортогональні базисні функції забезпечують збіжність ряду до $s(t)$ при $n \to \infty$. Оптимальними є методи дискретизації, які забезпечують мінімальний числовий ряд при заданій похибці відтворення сигналу. При неортогональних базисних функціях переважно використовуються степеневі алгебраїчні поліноми вигляду:

$$s'(t) = \sum_{n=0}^{N} c_n t^n.$$

Якщо значення апроксимуючого полінома збігаються зі значеннями вибірок в моменти їх відліку, то такий поліном називають інтерполювальним. Як інтерполювальні поліноми зазвичай використовуються багаточлени Лагранжа. Для реалізації інтерполювальних поліномів необхідна затримка сигналу на інтервал дискретизації, що в системах реального часу потребує певних технічних рішень. Як екстраполювальні поліноми використовують, як правило, багаточлени Тейлора.

Зрозумілою вимогою до вибору частоти дискретизації є внесення мінімальних спотворень в динаміку змін сигнальних функцій. Логічно вважати, що спотворення інформації будуть тим менші, чим вища частота дискретизації F. З другого боку, також очевидно, що чим більше значення F, тим більшою кількістю цифрових даних будуть відображатися сигнали, і тим більший час буде затрачено на їх обробку. В оптимальному варіанті значення частоти дискретизації сигналу F повинне бути необхідним і достатнім для обробки інформаційного сигналу з заданою точністю, тобто забезпечувати допустиму похибку відновлення аналогової форми сигналу (середньоквадратичну в цілому за інтервалом сигналу, або за максимальними відхиленнями від істинної форми в характерних інформаційних точках сигналів).

Теорема Котельникова

Частотний критерій Котельникова базується на теоремі Котельникова: якщо неперервна функція $x(t)$ відповідає умовам Діріхле (тобто, обмежена, кусково-неперервна і має кінцеву кількість екстремумів) та її спектр обмежений деякою частотою f_c – то вона повністю визначається

послідовністю своїх значень в точках, які розташовані одна від одної на відстані

$$T_k = \frac{1}{2f_c}.$$

Аналітично теорема Котельникова записується інтерполяційним рядом

$$x(t) = \sum_{k=-\infty}^{\infty} x(k\Delta t)\frac{\sin\left[\omega_c(t-k\Delta t)\right]}{\omega_c(t-k\Delta t)},$$

де $\Delta t = \frac{\pi}{\omega_c} = \frac{1}{2f_c}$.

Очевидно, що неперервна функція з обмеженим спектром може бути подана у вигляді суми нескінченно великої кількості членів, кожний з яких є множником функції вигляду $\frac{\sin y}{y}$ (функції відліку) та коефіцієнта $x(k\Delta t)$, який визначає значення функції $x(t)$ в момент відліку.

Функція вигляду $\frac{\sin\omega_c\tau}{\omega_c\tau}$ являє собою реакцію ідеального фільтра нижніх частот з граничною частотою ω_c на дельта-функцію. Тоді, якщо через такий фільтр пропустити квантований в часі сигнал із частотою квантування $f_c = 2f_c = \frac{\omega_c}{\pi}$, то, знаходячи добуток вихідних сигналів фільтра, можна отримати вихідний неперервний сигнал функції $x(t)$.

При використанні теореми Котельникова для квантування сигналів реальний спектр сигналу умовно обмежують деяким діапазоном частот від нуля до ω_c, в якому зосереджена основна частина енергії спектра сигналу.

Енергія сигналу, що відсікається, призводить до появи похибки, яка в загальному вигляді виражається відношенням енергії відсіченої частини сигналу до загальної енергії сигналу. Але похибку зручніше виразити як дисперсію зведеної похибки

$$D_\omega = \frac{\gamma\omega_c P_T}{(x_{\max} - x_{\min})^2},$$

де похибка дорівнює $\gamma\omega_c = \frac{\omega}{\int_0^x S(j\omega)^2 d\omega} \Big/ \int_0^x S(j\omega)^2 d\omega$.

Середня потужність сигналу може бути обчислена за формулою

$$P_T = \frac{1}{\pi T_c} \int_0^{\infty} |S(j\omega)|^2 \, d\omega = \int_{x_{\min}}^{x_{\max}} x^2 \omega(x) dx .$$

Якщо відомі $D_{\omega_c}, x_{\max}, x_{\min}$ та спектри функції $x(t)$ можна визначити частоту, яка обмежує спектр сигналу.

4.2 Дискретизація сигналів

Задача абсолютно точного відновлення сигналу на практиці зазвичай не ставиться, на відміну від задачі мінімального фізичного об'єму інформації, при якому зберігається можливість її відновлення в неперервній формі з визначеним допустимим значенням похибки. Така задача актуальна завжди, а особливо при дистанційних методах реєстрації та обробки інформації, передачі сигналів по каналах зв'язку і при підготовці інформації до тривалого зберігання. Одним з методів розв'язання цієї задачі є дискретизація сигналів за критерієм найбільшого відхилення.

В процесі дискретизації за критерієм найбільшого відхилення задається допустиме значення похибки відновлення сигналу σ. При відновленні сигналу неперервна функція $s(t)$ апроксимується, як правило, степеневими поліномами n-го порядку. Похибка відновлення функції $s(t)$ поліномом $s_a(t)$ визначається залишковим членом $L(t)$:

$$L(t) = s(t) - s_a(t).$$

Крок дискретизації вибирається з умови забезпечення $L(t) < \sigma$ по всьому інтервалу визначення функції $s(t)$. Як правило, динаміка функції $s(t)$ може суттєво змінюватися в різні моменти часу за інтервалом реєстрації, при цьому крок дискретизації також може змінюватися, за умови не перевищення заданої похибки на кожному кроці. При установленому значенні σ зменшення числа відліків забезпечується підвищенням степені апроксимувального багаточлена. На практиці зазвичай обмежуються сходинковим, лінійним та параболічним апроксимаційними поліномами відповідно нульового, першого і другого степенів.

Як було сказано вище, як інтерполювальні багаточлени використовують багаточлени Лагранжа, а як екстраполювальні багаточлени використовують багаточлени Тейлора. Для багаточлена Тейлора нульового степеня умови відновлення сигналу практично не відрізняються від багаточлена Лагранжа, за винятком напряму (від поточного зареєстрованого відліку і

вперед до t). Для багаточленів Тейлора більш високих степенів при відновленні сигналу крім відліку $s(t_i)$ використовуються також відповідні значення похідних в точці відліку. Відновлення сигналу багаточленами Тейлора відбувається без затримки в часі. Однак при використанні багаточленів вище нульового степеня для точного відновлення сигналу порівняно з інтерполяційними методами потрібно в два рази більш висока частота дискретизації.

Для багаточленів Лагранжа нульового степеня значення $s_a(t)$ в момент часу t на інтервалі $t_i < t < t_{i+1}$ між двома послідовними відліками функції приймається рівним відліку $s(t_{i+1})$. Якщо відновлення сигналу $s(t)$ проводити за двома відліками: $s_a(t_i) = [s(t_{i+1}) - s(t_i)]/2$, то при цьому ж кроці дискретизації похибка відновлення сигналу зменшується вдвічі. Але при використанні двох послідовних відліків краще застосовувати багаточлени Лагранжа першого степеня, тобто з'єднання двох послідовних відліків прямою лінією, що дає ще більше зменшення похибки відновлення аналогової форми сигналу.

Адаптивна дискретизація

Частота рівномірної дискретизації інформації розраховується за граничними значеннями частотних характеристик сигналів. Адаптивна дискретизація орієнтована на динамічні характеристики сигналу, що дозволяє забезпечити його відновлення при мінімальній кількості вибірок. В основі принципів адаптивної дискретизації лежить спостереження за поточною похибкою відновлення сигналу. Найбільш широке застосування отримали алгоритми дискретизації з адаптацією по довжині інтервалу апроксимації. Суть дискретизації полягає в послідовному нарощенні інтервалу апроксимації з неперервним порівнянням сигналу $s(t)$ з відтворювальною функцією $s_a(t)$. При досягненні заданого значення нарощення інтервалу зупиняють, і проводиться відлік значень $s(t_i)$, тобто дискретизація є нерівномірною. Для відтворення сигналів нерегулярної дискретизації зазвичай використовують степеневі алгебраїчні поліноми нульового чи першого степеня в інтерполяційному чи в екстраполяційному варіантах.

Найбільш простою є техніка адаптивної дискретизації з використанням багаточлена нульового степеня. На момент t_j початку кожного інтервалу апроксимувальний поліном $s_a(t)$ приймається рівним $s(t_i)$, обчис-

люється поточна різниця $L(t)=s(t)-s_a(t)$ і проводиться порівняння її значення із заданим значенням. При фіксуванні рівності $L(t)$ та заданого значення проводиться черговий відлік і починається наступний інтервал.

При використанні апроксимувального багаточлена першого степеня обчислюється значення $s_a(t)=s(t_j)+s'(t_j)$, де $s'(t)$ – похідна сигналу. Момент чергового відліку визначається виконанням рівності $s(t)-s(t_j)-s'(t_j)=\sigma$. Варто мати на увазі, що даний алгоритм неефективний при наявності високочастотних завад, до яких досить чутлива операція диференціювання.

Дискретизація випадкових сигналів належить до класу непростих задач аналізу випадкових сигналів. Тому обмежимось лише згадуванням основних особливостей такого аналізу на прикладі дискретизації квантованих за рівнем випадкових сигналів.

Отож, для обчислення оптимального кроку рівномірної дискретизації достатньо знати лише величину інтервалу квантування та кореляційну функцію вимірюваного процесу.

При оперативних змінах із запам'ятовуванням результату лише останньої відлікової операції оптимальна тривалість наступного кроку практично не залежить від цього результату і дорівнює оптимальному кроку рівномірної дискретизації вимірюваного процесу.

В околі точки оптимуму існує деякий досить широкий інтервал приблизно рівнозначних значень кроку дискретності Δt. Так в межах $\pm 10\%$-вих відхилень Δt від оптимального значення Δt_0 зміни імовірнісного показника якості $p(\Delta t)$ не перевищують 1%.

Оптимальний крок дискретизації визначається так

$$R(\Delta t_0)=\sqrt{1-0{,}9\left(\frac{\Delta x}{\sigma}\right)^2},$$

де Δx – величина кванта;

σ – середньоквадратичне відхилення процесу.

Дана формула дозволяє визначити оптимальний крок дискретизації безпосередньо на графіку кореляційної функції. На практиці переважно використовують інше співвідношення:

$$\Delta t_0 = 0{,}95\frac{\Delta x}{\omega\sigma},$$

де $\omega = \sqrt{-\frac{d^2R(t)}{dt^2}}\Big|_{t=0}$.

Усереднена за усіма номерам k похибка визначається як

$$\bar{\delta} = 0{,}01\left(\frac{\Delta x}{\sigma}\right)^4.$$

Варто зазначити, що при $\frac{\Delta x}{\sigma} < 0{,}4 - 0{,}5$ розрахункову похибку оптимального значення критерію якості вимірювальної системи можна оцінити однією похибкою, яка складає приблизно 0,5%.

4.3 Квантування сигналів

Дискретизація аналогових сигналів з перетворенням в цифрову форму пов'язана з квантуванням сигналів. Суть квантування полягає в заміні нескінченної множини можливих значень функції, в загальному випадку випадкових, скінченною множиною цифрових відліків, і виконується округленням миттєвих значень вхідної функції $s(t_i)$ в моменти часу t_i до найближчих значень $s_i(t_i) = n_i r$, де r – крок квантування шкали цифрових відліків. Квантування з постійним кроком r називають рівномірним. Математично операція квантування може бути виражена формулою:

$$s_i(t_i) = \left[\frac{s(t_i)}{r} + \frac{1}{2}\right] r,$$

де результат обчислення в дужках округляється до цілого значення.

Квантування за амплітудою – це процес заміни реальних (виміряних) значень амплітуди сигналу значеннями, наближеними з деякою точністю. Кожен з $2N$ можливих рівнів називається рівнем квантування, а відстань між двома найближчими рівнями квантування називається кроком квантування. У випадку лінійного розбиття амплітудної шкали на рівні, квантування називають лінійним (однорідним). На рис. 4.2. наведено приклад такого квантування.

Як бачимо, результатом такого оцифрування став ступінчатий сигнал, складений із прямокутників, кожен з яких має ширину, рівну величині кроку дискретизації, і висоту, рівну виміряному значенню амплітуди сигналу. Очевидно, що точність округлення залежить від вибраної кількості ($2N$) рівнів квантування, яка, в свою чергу, залежить від кількості бітів (N), відведених для запису значення амплітуди. Число N називають роз-

рядністю квантування (маючи на увазі кількість розрядів, тобто бітів, в кожному слові), а отримані в результаті округлення значень амплітуди числа – відліками чи семплами (от англ. «sample» – «замір»). Вважається, що квантування з розрядністю 16 біт залишається для слухача майже непомітним.

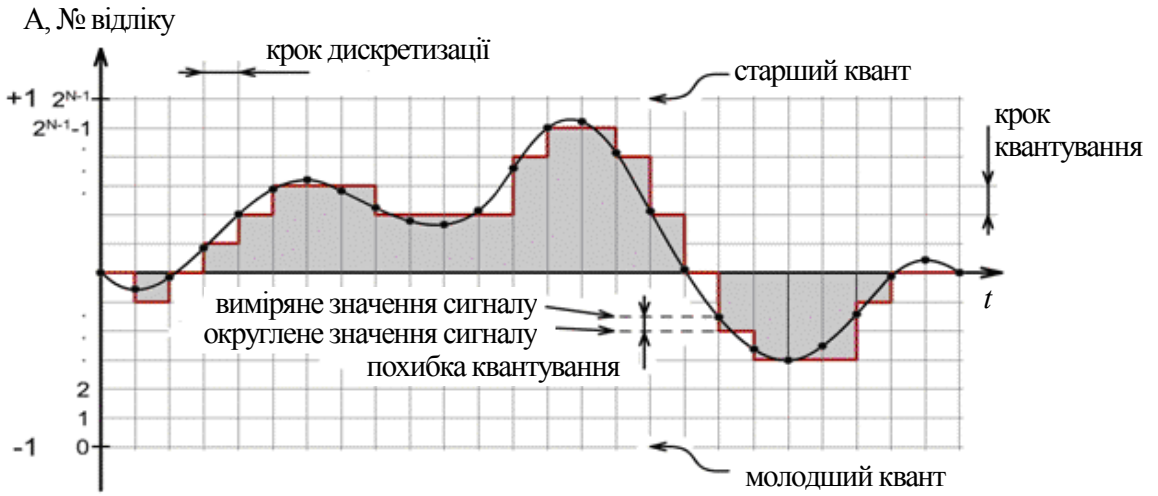


Рисунок 4.2 – Квантування неперервних сигналів по амплітуді

При квантуванні сигналів у великому динамічному діапазоні значень крок квантування може бути й нерівномірним, наприклад, логарифмічним, тобто пропорційним логарифму значень вхідного сигналу. Установлений діапазон шкали квантування від $s_{\min}$ до $s_{\max}$ і крок квантування r визначають число поділок шкали $N = \dfrac{(s_{\min} - s_{\max})}{r}$ і відповідно цифрову розрядність квантування. В результаті дискретизації й квантування неперервна функція $s(t)$ замінюється числовою послідовністю $\{s(kt)\}$. Похибка округлення $\varepsilon_i = s(t_i) - s_i(kt)$ знаходиться в межах $-kt/2 < \varepsilon < kt/2$ і називається шумом квантування. Необхідна точність квантування оцінюється за впливом на подальшу обробку сигналів виниклого шуму квантування.

При достатньо малому кроці квантування будь-яке значення в його межах можна вважати рівноймовірним, при цьому значення ε розподілені за рівномірним законом.

Відповідно дисперсія та середнє квадратичне значення шуму квантування:

$$\sigma^2 = kt^2/12, \qquad \bar{\varepsilon} \approx 0{,}3kt.$$

При заданому рівні шуму квантування з використанням попередньої формули неважко визначити допустиме значення кроку квантування.

Вхідний сигнал містить, як правило, адитивну суміш власне сигналу $s(t)$ та завади $q(t)$ з дисперсією відповідно σ_q^2. Якщо завади не корельовані із сигналом, то після квантування сумарна дисперсія шумів:

$$\sigma_{\sum}^2 = \sigma_q^2 + \sigma^2.$$

На практиці крок квантування вибирають зазвичай таким, щоб не відбувалось помітної зміни відношення сигнал/шум, тобто $\sigma_q^2 >> \sigma^2$.

Цифрове подання інформації можна коротко зробити таким чином. Якщо сигнал, який був дискретизований за часом і рівнем, потім подається у цифровому вигляді, то такий процес називається аналого-цифровим перетворенням. Оскільки в порівнянні із, наприклад, записаним на носії (касеті чи пластинці) аналоговим сигналом, цифровий запис характеризується високим відношенням сигнал-шум і широким динамічним діапазоном (відношенням мінімального сигналу до максимального неспотвореного сигналу) і забезпечує високу якість відтворення, наприклад, звуку. Але для відтворення записаного звуку його слід знову перевести в аналогову форму з допомогою цифро-аналогового перетворення, про що буде йти мова далі.

Перейдемо до розгляду позиційних систем числення. Позиційна система числення з основою P має P цифр $C_0, C_1, ..., C_{p-1}$, що звичайно позначають натуральні числа від 0 до $P-1$. Ці записи та позначені ними числа – значення цих записів – називаються однорозрядними.

Цифри десяткової системи 0, 1, 2, … , 9 називаються арабськими, хоча й були запозичені арабами в індусів.

У програмуванні широко застосовується шістнадцяткова система, в якій перші 10 цифр арабські, а наступні шість, A, B, C, D, E, F, позначають числа, десятковий запис яких 10, 11, 12, 13, 14, 15, відповідно.

Число P у P-ковій системі позначається дворозрядним записом C_0C_1, число $P+1$ – записом C_1C_1 тощо. Наприклад, 10, 11, ... , 99 у десятковій системі, 10, 11 у двійковій, 10, 11, … , 1F, 20, … , FF у 16-ковій. Наприклад, 100, 101, … , 999 у десятковій системі, 100, 101, 110, 111 у двійковій, 100, 101, … , FFF у 16-ковій.

Наприклад, двійковий запис (10011)2 позначає число, яке в десятковому записі має вигляд $1\times2^4+0\times2^3+0\times2^2+1\times2^1+1\times2^0=19$. 16-ковий запис (1BC)16 позначає десяткове $1\times162+11\times16+12=444$.

Права цифра в записі числа позначає кількість одиниць і називається молодшою, ліва – кількість чисел P^k і називається старшою.

Ми звикли до десяткового подання чисел, і саме воно, головним чином, використовується в програмах, але в комп'ютері числа, як правило, подаються в двійковій системі. Таким чином, виникає необхідність створювати двійкове подання числа за його десятковим записом і навпаки.

За P-ковим записом натурального числа N можна побудувати десяткове подання, обчисливши значення полінома за допомогою операцій множення та додавання в десятковій системі.

P-кове подання чисел, менших 1, має вигляд $0, a_{-1}, 0, a_{-2}, ..., 0, a_{-i}$, де a_{-i} – P-кові цифри. Цей запис позначає дійсне число – значення виразу

$$a_{-1}P^{-1} + a_{-2}P^{-2} + ... + a_{-i}P^{-i}.$$

Нагадаємо, що якщо основа P має прості дільники, відмінні від 2 і 5, то число зі скінченним P-ковим записом зображується нескінченним, але періодичним десятковим дробом. Якщо ж простими дільниками P є тільки 2 і 5, то й десятковий дріб скінченний.

Перейдемо до цифрового кодування. При цифровому кодуванні дискретної інформації застосовують потенціальні й імпульсні коди.

У потенціальних кодах для подання логічних одиниць і нулів використовується тільки значення потенціалу сигналу, а його перепади, що формують закінчені імпульси, до уваги не беруться. Імпульсні коди дозволяють подати двійкові дані або імпульсами певної полярності, або частиною імпульсу – перепадом потенціалу певного напрямку.

Вимоги до методів цифрового кодування

При використанні прямокутних імпульсів для передачі дискретної інформації необхідно вибрати такий спосіб кодування, який одночасно досягав би декількох цілей:

а) мав при одній і тій же бітовій швидкості найменшу ширину спектра результуючого сигналу;

б) забезпечував синхронізацію між передавачем і приймачем;

в) мав здатність розпізнавати помилки;

г) мав низьку вартість реалізації.

Більш вузький спектр сигналів дозволяє на одній і тій же лінії (з однієї і тією же смугою пропускання) домагатися більш високої швидкості передачі даних. Крім того, часто до спектра сигналу висувається вимога відсутності постійної складової, тобто наявності постійного струму між передавачем і приймачем. Зокрема застосування різних трансформаторних

схем гальванічної розв'язки перешкоджає проходженню постійного струму.

Синхронізація передавача і приймача потрібна для того, щоб приймач точно знав, у який момент часу необхідно зчитувати нову інформацію з лінії зв'язку. Ця проблема в мережах важить більше, ніж при обміні даними між близько розташованими пристроями, наприклад між блоками усередині комп'ютера чи між комп'ютером і принтером. На невеликих відстанях добре працює схема, що базується на окремій тактовій (синхронізуючій) лінії зв'язку (рис. 4.3), тому інформація знімається з лінії даних тільки в момент приходу тактового імпульсу.

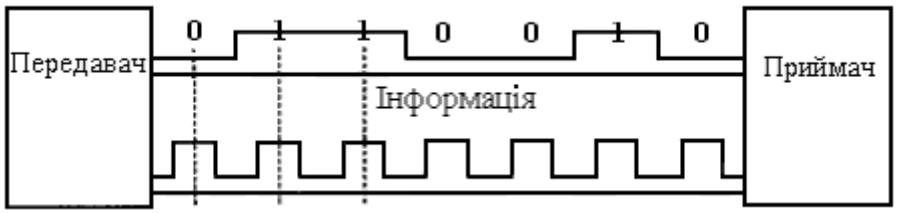


Рисунок 4.3 – Синхронізація приймача і передавача на невеликих відстанях

У мережах використання цієї схеми спричиняє труднощі через неоднорідність характеристик провідників у кабелях. На великих відстанях нерівномірність швидкості поширення сигналу може привести до того, що тактовий імпульс прийде настільки пізніше чи раніше відповідного сигналу даних, що біт даних буде пропущений чи врахований повторно. Іншою причиною, через яку у мережах відмовляються від використання тактуючих імпульсів, є економія провідників у дорогих кабелях.

Тому в мережах застосовуються самосинхронізуючі сигнали, які несуть для передавача інформацію про те, у який момент часу потрібно здійснювати розпізнавання чергового біта (чи декількох біт, якщо код орієнтований більш ніж на два стани сигналу). Будь-який різкий перепад сигналу – що називається фронтом – може нести інформацію для синхронізації приймача з передавачем.

При використанні синусоїд як несучого сигналу результуючий код має властивість самосинхронізації, тому що зміна амплітуди несучої частоти дає можливість приймачу визначити момент появи вхідного коду.

Розпізнавання і корекцію перекручених даних складно здійснити засобами фізичного рівня, тому найчастіше цю роботу беруть на себе протоколи, що лежать вище: канальний, мережний, транспортний чи прикладний. З іншого боку, розпізнавання помилок на фізичному рівні заощаджує

час, тому що приймач не чекає повного переміщення кадру в буфер, а відбраковує його відразу при розпізнаванні помилкових бітів усередині кадру.

4.4 Аналого-цифрові та цифро-аналогові перетворювачі

Аналого-цифрові перетворювачі (АЦП) це пристрої, які приймають вхідні аналогові сигнали та генерують відповідні до них цифрові сигнали, які придатні для обробки мікропроцесорами та іншими цифровими пристроями.

Принципово не виключена можливість безпосереднього перетворення різних фізичних величин в цифрову форму, однак це завдання вдається розв'язати тільки досить рідко через складність таких перетворювачів. Тому зараз найраціональнішим вважається спосіб перетворення різних за фізичною природою величин спочатку в функціонально пов'язані з ними електричні, а потім уже за допомогою перетворювачів напруга - код – в цифрові. Саме ці перетворювачі і мають на увазі, коли говорять про АЦП.

Процедура аналого-цифрового перетворення неперервних сигналів, яку реалізовують за допомогою АЦП, це перетворення неперервної функції часу $U(t)$, яка описує вхідний сигнал, у послідовність чисел $\{U'(t_j)\}$, j=0,1,2,…, що віднесені до деяких фіксованих моментів часу. Цю процедуру можна розділити на дві самостійні операції: дискретизацію і квантування.

Найпоширенішою формою дискретизації, як зазначалось, є рівномірна дискретизація, в основі якої лежить теорема відліків. Згідно з цією теоремою як коефіцієнти a_j потрібно використовувати миттєві значення сигналу $U(t_j)$ в дискретні моменти часу $t_j = j\omega t$, а період дискретизації вибирати з умови:

$$t = 1/2F_m,$$

де F_m – максимальна частота спектра сигналу, що перетворюється.

Тоді отримаємо відомий вираз теореми відліків

$$U(t) = \sum_{j=-\infty}^{\infty} U(j\omega\Delta t)\frac{\sin[2\pi F_m(t - j\Delta t)]}{2\pi F_m(t - j\Delta t)}.$$

Для сигналів зі строго обмеженим спектром цей вираз є тотожністю. Однак спектри реальних сигналів прямують до нуля тільки асимптотично. Застосування рівномірної дискретизації до таких сигналів викликає виникнення в системах обробки інформації специфічних високочастотних

спотворень, які зумовлені вибіркою. Для зменшення цих спотворень необхідно або збільшувати частоту дискретизації, або використовувати перед АЦП додатковий фільтр нижніх частот, який обмежуватиме спектр вхідного сигналу перед його аналого-цифровим перетворенням.

У загальному випадку вибір частоти дискретизації буде залежати також від вигляду функції $f_j(t)$, що використовується в першій формулі розділу та допустимого рівня похибок, які виникають при відновленні початкового сигналу за його відліками. Усе це необхідно враховувати при виборі частоти дискретизації, яка визначає необхідну швидкодію АЦП. Часто цей параметр задають розробнику АЦП.

Розглянемо докладніше місце АЦП при виконанні операції дискретизації.

Для достатньо вузькосмугових сигналів операцію дискретизації можна виконувати за допомогою самих АЦП і суміщати таким чином з операцією квантування. Основною закономірністю такої дискретизації є те, що за рахунок скінченного часу одного перетворення та невизначеності моменту його закінчення, який, у загальному випадку, залежить від параметрів вхідного сигналу, не вдається отримати однозначної відповідності між значеннями відліків та моментами часу, до яких їх потрібно віднести. В результаті при роботі із сигналами, які змінюються в часі, виникають специфічні похибки, динамічні за своєю природою, для оцінки яких вводять поняття апертурної невизначеності, яка переважно характеризується апертурним часом.

Апертурним часом t_a називають час, протягом якого зберігається невизначеність між значенням вибірки та часом, до якого вона відноситься. Ефект апертурної невизначеності проявляється або як похибка миттєвого значення сигналу при заданих моментах вимірювання, або як похибка моменту часу, в який проводиться вимірювання при заданому миттєвому значенні сигналу. При рівномірній дискретизації наслідком апертурної невизначеності є виникнення амплітудних похибок, які називаються апертурними та чисельно рівні приростові сигналу протягом апертурного часу.

Якщо використовувати іншу інтерпретацію ефекту апертурної невизначеності, то її наявність викликає "тремтіння" істинних моментів часу, в які беруться відліки сигналу, відносно моментів, які рівновіддалені на осі часу. В результаті замість рівномірної дискретизації зі строго постійним періодом проводиться дискретизація с флюктуючим періодом повторення.

Це викликає порушення умов теореми відліків та появи уже розглянутих апертурних похибок в системах цифрової обробки інформації.

Таке значення апертурної похибки можна визначити, розклавши вираз для вхідного сигналу в ряд Тейлора в околі точок відліку, який для j-ї точки має вигляд:

$$U(t) = U(t_j) + t_a U'(t_j) + \frac{t_a^2}{2} U''(t_a) + \ldots$$

та в першому наближенні дає апертурну похибку:

$$\Delta U_a(t_j) \approx t_a U'(t_j),$$

де t_a – апертурний час, який для розглянутого випадку в першому наближенні є часом перетворення АЦП.

Зазвичай для оцінки апертурних похибок використовують синусоїдальний випробувальний сигнал $U(t) = U_m \sin \omega t$.

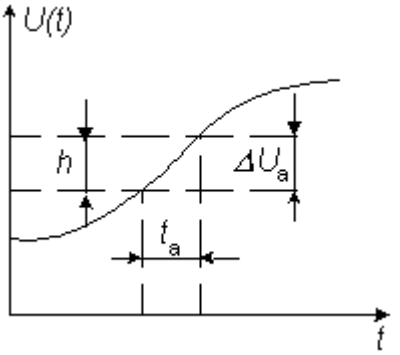


Рисунок 4.4 – Утворення апертурної похибки для випадку, коли вона дорівнює крокові квантування

Якщо прийняти, що для N-розрядного АЦП з роздільною здатністю 2^{-N} апертурна похибка не повинна перевищувати кроку квантування (рис. 4.4), то залежність між частотою сигналу ω, апертурним часом t_a та відносною апертурною похибкою буде такою: $1/2^N = \omega t_a$.

Для забезпечення дискретизації синусоїдального сигналу частота якого 100 кГц з похибкою 1% час перетворення АЦП повинен бути рівним 25 нс. У той же час за допомогою такого швидкодіючого АЦП принципово можна дискретизувати сигнали, які мають ширину спектра біля 20 МГц. Таким чином, дискретизація за допомогою самого АЦП викликає суттєве розходження вимог між швидкодією АЦП та періодом дискретизації. Це розходження досягає 2...3 порядків та дуже ускладнює і здорожує процес дискретизації, оскільки навіть для порівняно вузькополосних сигналів пот-

ребує досить швидкодіючих АЦП. Для достатньо широкого класу сигналів, які швидко змінюються, цю проблему вирішують за допомогою пристроїв вибірки-зберігання, що мають малий апертурний час.

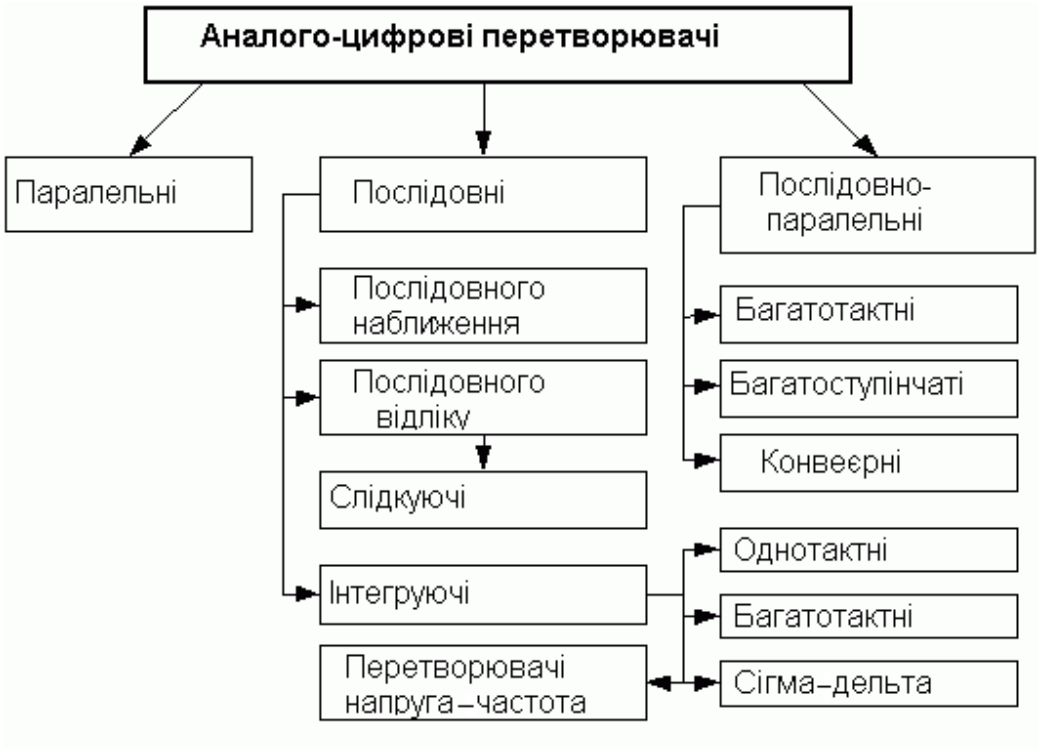


Рисунок 4.5 – Класифікація аналогово-цифрових перетворювачів

Зараз відома велика кількість методів перетворення напруга – код. Ці методи суттєво відрізняються один від одного потенційною точністю, швидкістю перетворення та складністю апаратної реалізації. На рис. 4.5 наведена класифікація АЦП за методами перетворення.

В основу класифікації АЦП покладено ознаку, яка вказує на те, як в часі розгортається процес перетворення аналогової величини в цифрову. В основі перетворення вибіркових значень сигналу в цифрові еквіваленти лежать операції квантування та кодування. Вони можуть проводитись за допомогою або послідовної, або паралельної, або послідовно-паралельної процедур наближення цифрового еквівалента до перетворюваної величини.

Розглянемо детальніше найбільш поширені типи АЦП.

Функціонування аналого-цифрового перетворення за методом послідовного підрахунку можна проілюструвати за допомогою структурної схеми на рис. 4.6.

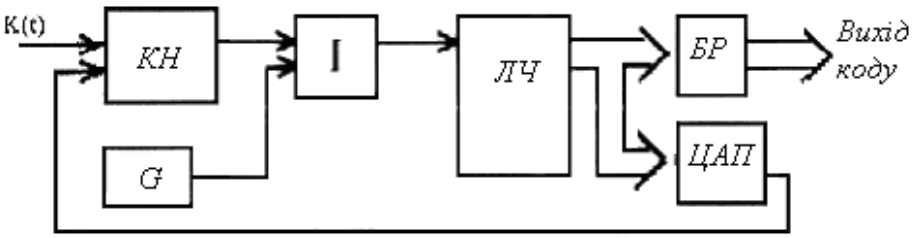


Рисунок 4.6 – АЦП послідовного підрахунку

До складу схеми входять: генератор тактових сигналів (G), компаратор напруги (КН), схема І, лічильник (ЛЧ), буферний регістр (БР), цифро-аналоговий перетворювач (ЦАП). Схема працює наступним чином. На вхід перетворювача подається аналоговий сигнал $x(t)$, який підключається до одного з входів компаратора напруги КН. На другий вхід компаратора подається еталонна напруга (U_{em}), яка формується на виході ЦАП під управлінням колового слова на виході ЛЧ. Компаратор формує на своєму виході сигнал або логічної одиниці, або логічного нуля в залежності від того, яке значення більше. Якщо $U_{em} < x(t)$, то на виході компаратора формується одиниця, яка дозволяє проходження імпульсів з тактового генератора через схему І на лічильний вхід лічильника ЛЧ. На виході лічильника йде процес підрахунку цих імпульсів в двійковому коді від 2^0 до 2^{n-1}. Двійковий код з ЛЧ подається на вхід ЦАП, на виході якого формується ступінчатий сигнал U_{em}. Кожна сходинка цього сигналу відповідає за рівнем інтервалу дискретизації q. Сигнал U_{em} порівнюється із сигналом $x(t)$ і в момент, коли $x(t)$ стає меншим за U_{em}, на виході компаратора формується сигнал логічного нуля. Схема І закривається, лічильник зупиняє підрахунок і набраний двійковий код переписується у вихідний буферний регістр БР для видачі користувачу.

Метод безпосереднього зчитування реалізовується за допомогою так званого АЦП паралельної дії. Такий перетворювач має лінійку 2^{n-1} компараторів напруги, перші входи яких запаралелені і на них подається сигнал $x(t)$. На інші входи під'єднуються виходи подільника еталонної напруги. Виходи компараторів під'єднані до перетворювача одиничного коду в двійковий. Процес перетворення здійснюється за один такт, причому на виході лінійки компараторів до компаратора, який зафіксує $x(t) < U$, буде хвиля одиниць, а далі хвиля нулів одиничного коду. Структурно-

функціональна схема перетворення зображена на рис. 4.7, а часова діаграма аналогічна рис. 4.9.

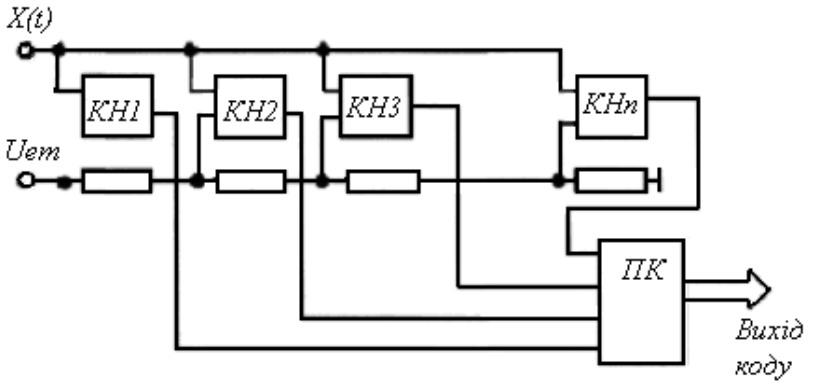


Рисунок 4.7 – АЦП безпосереднього зчитування

Найбільше поширення отримав метод порозрядного зрівноваження, який забезпечує час перетворення від 1 мкс до 1 мс. Структурно-функціональна схема перетворення зображена на рис. 4.8, а часова діаграма – на рис. 4.9.

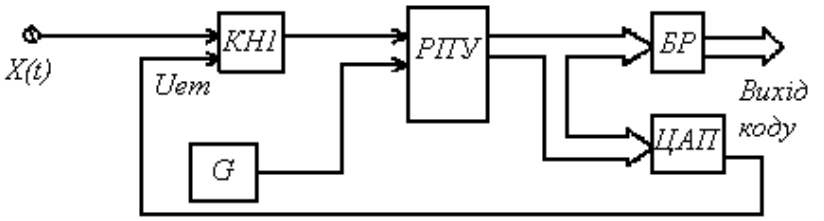


Рисунок 4.8 – АЦП порозрядного зрівноваження

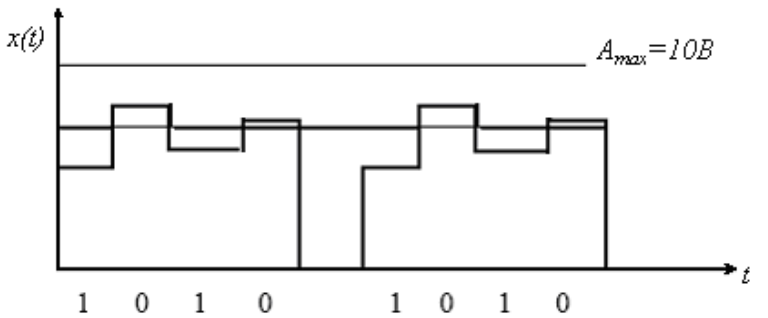


Рисунок 4.9 – Часова діаграма перетворення АЦП за методом порозрядного зрівноваження при $n = 4$

Схема працює так. На вхід АЦП подається вхідний сигнал $x(t)$, який порівнюється з еталонним сигналом U_{em}, що формується на виході ЦАП. ЦАП складається із сукупності $3n$ еталонних джерел сигналів, які управляються за допомогою спеціального регістра порозрядного зрівнова-

ження (РПУ). Перетворення проходить за n часових тактових інтервалів. Причому на першому такті РПУ примусово вмикає в роботу перший розряд ЦАП. Значення першого розряду еталонних величин на виході ЦАП дорівнює половині діапазону перетворення сигналу. Потім в кінці першого тактового інтервалу компаратор проводить порівняння $x(t)$ з U_{em}. Якщо $x(t) < U_{em}$, то примусово увімкнений старший розряд ЦАП залишається ввімкненим до закінчення процесу перетворення. Це забезпечується під управлінням певного сигналу на виході компаратора (1чи 0). Якщо ж $x(t) > U_{em}$, то перший розряд вимикається на початку другого такту. На початку другого такту в роботу примусово вмикається другий розряд ЦАП і знову проводиться порівняння $x(t)$ з U_{em}. Процедура повторюється доти, поки всі розряди ЦАП не візьмуть участі у процесі зрівноваження. В результаті на виході АЦП формується код, що відповідає вхідному сигналу.

Інтегруючі АЦП

Відомо, що недоліком послідовних АЦП є низька завадостійкість результатів перетворення. Дійсно, вибірка миттєвого значення вхідної напруги, переважно включає доданок у вигляді миттєвого значення завади. Згодом при цифровій обробці послідовності вибірок ця складова може бути подавлена, однак на це потрібен час та обчислювальні ресурси. Переважно у АЦП вхідний сигнал інтегрується або неперервно, або у певному часовому діапазоні, тривалість якого зазвичай вибирається кратною періодові завади. Це дозволяє в багатьох випадках приглушити заваду ще на етапі перетворення. Платою за це є понижена швидкодія інтегруючих АЦП.

Спрощена схема АЦП, який працює в два основних такти (АЦП двотактного інтегрування), наведена на рис. 4.10.

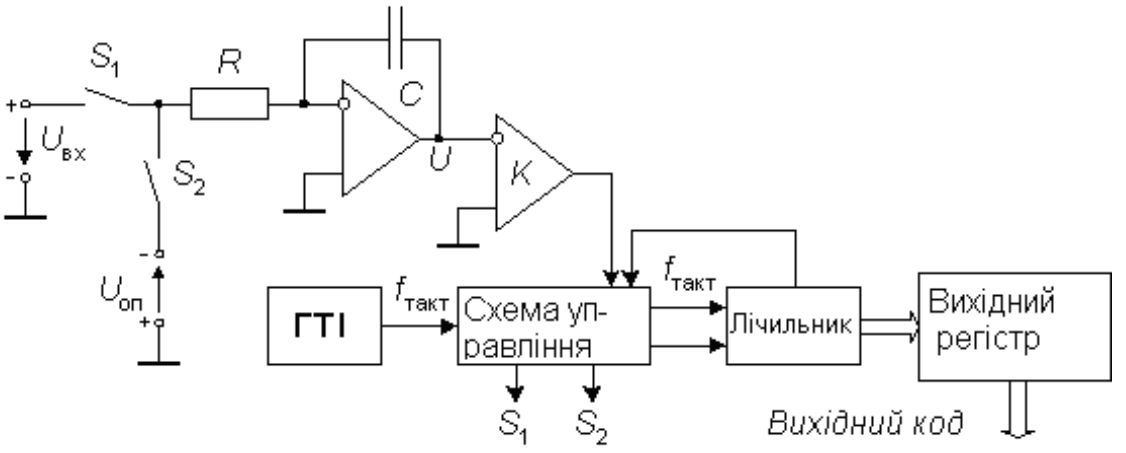


Рисунок 4.10 – Спрощена схема АЦП двотактного інтегрування

Перетворення проходить протягом двох стадій: стадії інтегрування та стадії підрахунку. На початку першої стадії ключ S_1 замкнутий, а ключ S_2 розімкнутий. Інтегратор I інтегрує вхідну напругу $U_{вх}$. Час інтегрування вхідної напруги t_1 постійний; як таймер використовується лічильник з коефіцієнтом підрахунку $K_л$, так, що

$$t_1 = K_л / f_{такт}.$$

До моменту закінчення інтегрування вихідна напруга інтегратора складає

$$U_{вих}(t_1) = -\frac{1}{RC}\int_0^{t_1} U_{вх}(t)dt = -\frac{U_{вх.сер}K_л}{f_{такт}RC},$$

де $U_{вх.сер}$ – середнє за час t_1 значення вхідної напруги.

Після закінчення стадії інтегрування ключ S_1 розмикається, а ключ S_2 замикається та опорна напруга $U_{оп}$ надходить на вхід інтегратора. При цьому вибирається опорна напруга, протилежна за знаком вхідній напрузі. На стадії підрахунку вихідна напруга інтегратора лінійно зменшується за абсолютною величиною, як показано на рис. 4.11.

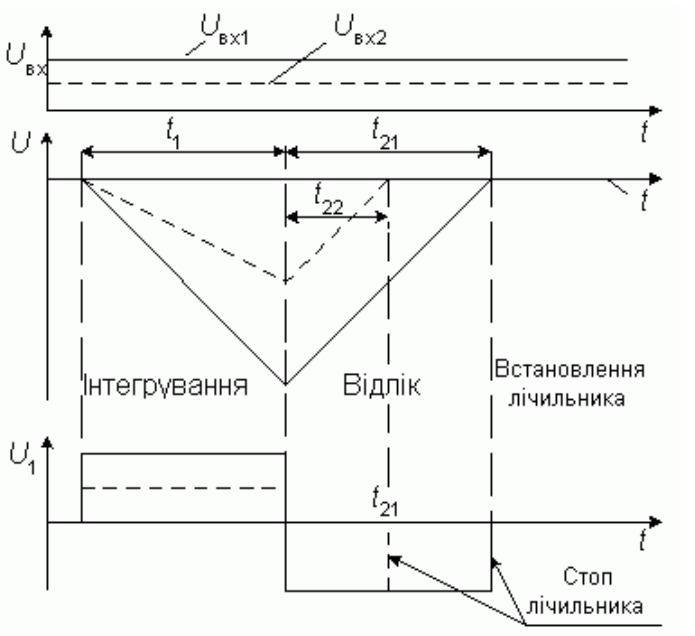


Рисунок 4.11 – Часові діаграми АЦП двотактного інтегрування

Стадія підрахунку закінчується, коли вихідна напруга інтегратора переходить через нуль. При цьому компаратор К переключається та підра-

хунок зупиняється. Діапазон часу, у якому проходить стадія підрахунку, визначається рівнянням

$$U_{вих}(t_1)+\frac{1}{RC}\int_0^{t_1+t_2}U_{оп}dt=0.$$

Далі, виконавши прості математичні дії і врахувавши, що:

$$t_2=\frac{n_2}{f_{такт}},$$

де n_2 – вміст лічильника після закінчення стадії підрахунку,

отримаємо результат

$$n_2=\frac{U_{вх.сер}K_л}{U_{оп}}.$$

З цієї формули випливає, що відмітною рисою методу багатотактного інтегрування є те, що ні тактова частота, ні постійна інтегрування RC не впливають на результат. Необхідно тільки, щоб тактова частота протягом часу t_1+t_2 залишалася постійною. Це можна забезпечити при використанні простого тактового генератора, оскільки істотні часові чи температурні дрейфи частоти відбуваються за час який більший, ніж час перетворення.

При виведенні попередніх виразів ми бачили, що в остаточний результат входять не миттєві значення перетворюваної напруги, а тільки значення, усереднені за час t_1. Тому змінна напруга послабляється тим сильніше, чим вища її частота.

Визначимо коефіцієнт передачі завади $K_П$ для АЦП двотактного інтегрування. Нехай на вхід інтегратора надходить гармонічний сигнал одиничної амплітуди частотою f з довільною початковою фазою φ. Середнє значення цього сигналу за час інтегрування t_1 дорівнює

$$U_{ñð}=\frac{1}{t_1}\int_0^{t_1}\sin(2\pi ft+\varphi)\,dt=\frac{\sin(\pi ft_1+\varphi)\sin\pi ft_1}{\pi ft_1}.$$

Коли ця величина досягає максимуму за модулем, то

$$K_П=\left|\frac{\sin^2\pi ft_1}{\pi ft_1}\right|.$$

Частотна характеристика коефіцієнта приглушення завад АЦП двотактного інтегрування наведена на рис. 4.12.

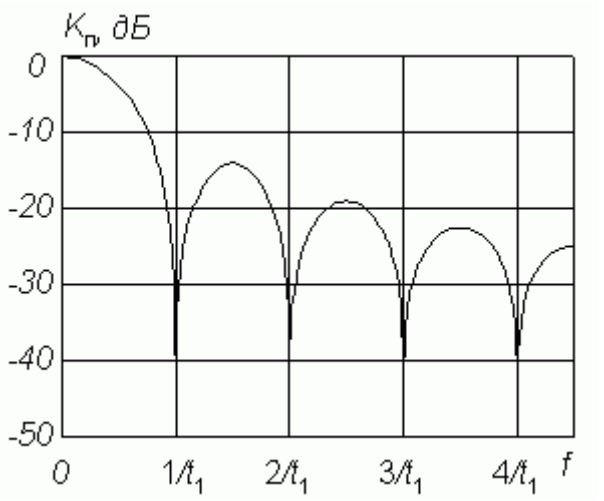


Рисунок 4.12 – Частотна характеристика коефіцієнта приглушення завад АЦП двотактного інтегрування

Як бачимо змінна напруга, період якої в ціле число раз менший t_1, приглушується зовсім Тому доцільно вибрати тактову частоту такою, щоб добуток $K_л \cdot f_{такт}$ був рівним чи кратним періоду напруги промислової мережі.

Багатоканальні АЦП

Багатоканальні АЦП на сьогодні досить поширені, особливо там, де потрібно об'єднати інформацію, отриману від кількох її джерел, тобто, наприклад, від різних сенсорів. Такі АЦП можна застосовувати, наприклад, для моніторингу напруги на входах, контролю крайніх значень, реєстрації показів, управління виходами (навантаженням) тощо. Схема багатоканального АЦП УМ-АЦП1 на основі мікроконтролера PIC16F876A наведена на рис. 4.13.

Комерційна версія такого пристрою має по 40 входів та виходів, але їх кількість може бути й іншою.

Структурні методи покращення характеристик АЦП

Необхідно відзначити, що висока точність досягається за рахунок як вдосконалення елементної бази, процесу виготовлення, так і застосовуваними матеріалами. Певні відхилення ваг розрядів від необхідних значень зазвичай корегуються шляхом лазерної підгонки в процесі виготовлення резисторів АЦП. При цьому вимагається збільшення площі внутрішньо кристальних компонентів і кристала в цілому, а також виникає проблема вилучення матеріалів кристала в ході пригонки. Ці процеси порушують

структуру матеріалів компонентів, зменшують часову і температурну стабільність схеми.

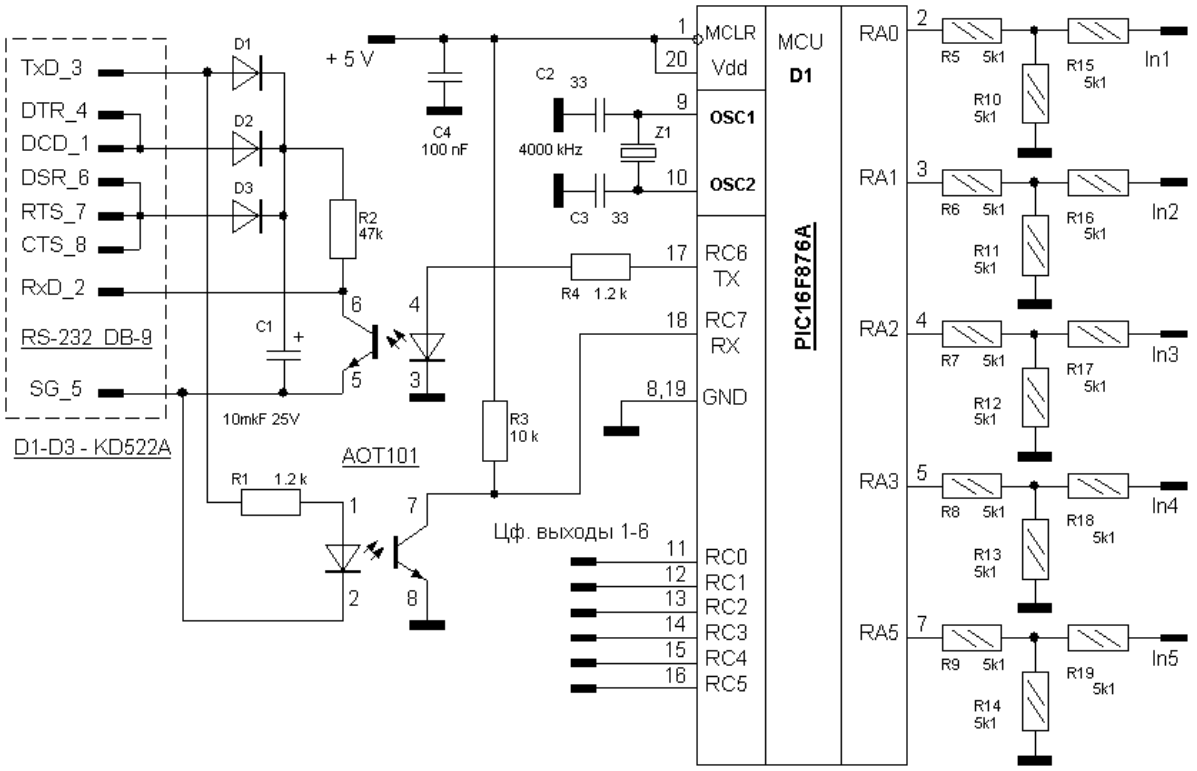


Рисунок 4.13 – Схема багатоканального АЦП УМ-АЦП1

Більш перспективним є підхід, який уникає фізичного впливу на елементи схеми. Наприклад, у випадку використання ЦАП на основі двійкової системи числення, зменшення статичних похибок досягається корекцією вихідної величини шляхом введення поправки в аналоговій формі, що формується додатковим корегувальним ЦАП. В цьому випадку перетворюваний код $K_{вх}$ подається, як показано на рис. 4.14, на вхід основного ЦАП і в цифровий обчислювальний пристрій (ЦОП).

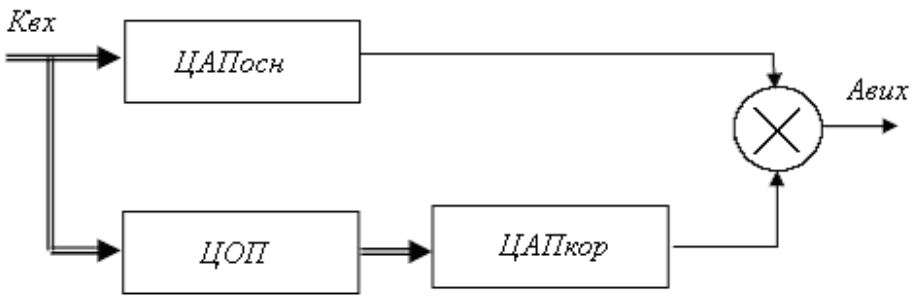


Рисунок 4.14 – Корекція вихідної величини двійкового ЦАП

В ЦОП вираховується код поправки, що надходить на вхід корегувального ЦАП. Результат перетворення $A_{вих}$ формується за допомогою суматора $\otimes$ аналогових величин. Але застосування такого принципу при аналого-цифровому перетворенні дає деяке зниження швидкості пристроїв.

Також використовують введення в прилади при розробці інформаційної надмірності у вигляді надлишкових позиційних систем числення (НПСЧ), що комплексно вдосконалює водночас кілька характеристик аналого-цифрового перетворення. Збільшення розрядної сітки пристрою (а відповідно і збільшення кількості тактів при порозрядному зрівноваженні) підвищує точність АЦП середньої і високої швидкості, реалізованих на грубих аналогових вузлах, а з іншого боку підвищує швидкодію високоточних АЦП на елементній базі середньої швидкості.

Підвищення швидкодії багаторозрядних АЦП досягається двома шляхами. Перший орієнтується на використання більш досконалої елементної бази, що не є досить простим шляхом. Другий шлях пов'язаний із введенням надмірності, переважно структурної, при проектуванні.

Цифро-аналогові перетворювачі

Необхідність здійснення операції відновлення вихідного сигналу з дискретних відліків, а також необхідність здійснення операцій формування еталонних сигналів при аналого-цифровому перетворенні висуває задачу цифро-аналогового перетворення. Суть операції цифро-аналогового перетворення полягає у формуванні аналогових сигналів, що відповідають кодовим словам дискретного сигналу. Технічно це формування здійснюється цифро-аналоговим перетворювачем (ЦАП).

Аналоговий сигнал на виході ЦАП може бути сформований шляхом множення опорної напруги $E_{оп} = q$ на вагові розрядні коефіцієнти кодового слова $a_i = 2^i$, таким чином, що $U_{вих} = q\left(a_0 2^0 + a_1 2^1 + ... + a_{n-1} 2^{n-1}\right)$.

Технічно найпростіше ЦАП реалізовується на принципі підсумовування розрядних струмів $U_{вих} = \sum I_1 R_{оп} = R_{оп}\left(a_0 I_1 + a_1 I_2 + ... + a_n I_n\right)$ (рис. 4.15).

Схема реалізації ЦАП для підсумовування струму містить джерело стабільної напруги E_0, матрицю двійково-зважених резисторів $\left(R \cdot 2^i\right)$, набір ключів $КЕ_i$, що реалізовують розрядні коефіцієнти a_i і перетворювач струму в напругу на операційному підсилювачі ОП.

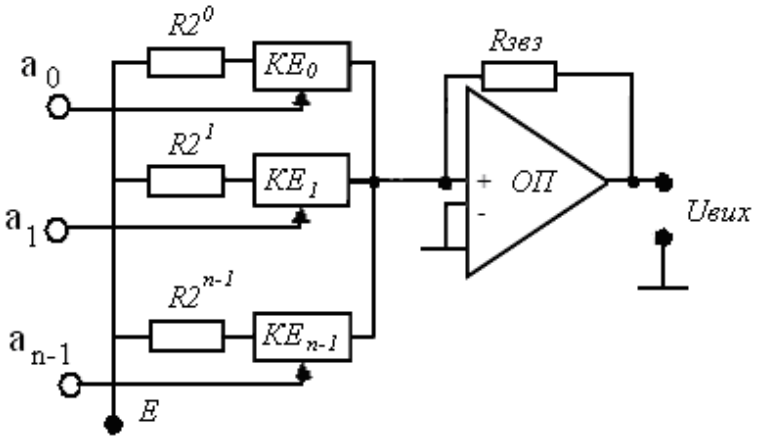


Рисунок 4.15 – ЦАП для підсумовування струму

Часова діаграма класичного процесу цифро-аналогового перетворення має вигляд (рис. 4.16).

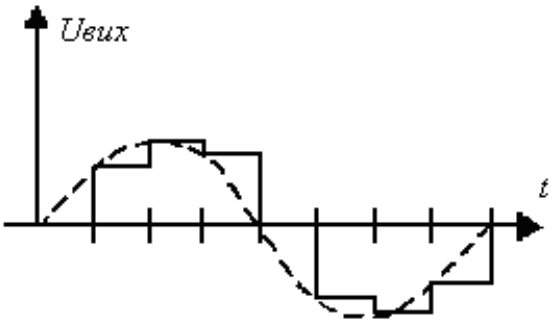


Рисунок 4.16 – Часова діаграма процесу ЦА перетворення

При малій кількості дискретних вибірок миттєвих значень сигналу, цей сигнал мало нагадує вихідний, однак може бути наближеним до нього шляхом аналогової фільтрації або інтерполяції.

Основні типи електронних ЦАП

1. Широтно-імпульсний модулятор – найпростіший тип ЦАП. Стабільне джерело струму чи напруги періодично вмикається на час, пропорційний перетворюваному цифровому коду, далі отримана імпульсна послідовність фільтрується аналоговим фільтром низьких частот. Такий спосіб часто використовується для керування швидкістю електромоторів, а також стає популярним в Hi-Fi аудіотехніці.

2. ЦАП передискретизації, такі, як дельта-сигма ЦАП, основані на змінюваній густоті імпульсів. Передискретизація дозволяє використовувати ЦАП з меншою розрядністю для досягнення більшої розрядності кінцевого перетворення; часто дельта-сигма ЦАП будується на основі найпростішого однобітового ЦАП, який є практично лінійним. На ЦАП малої розрядності надходить імпульсний сигнал з модульованою густотою імпульсів (з постійною тривалістю імпульсу, але зі змінною шпаруватістю), створений з використанням негативного зворотного зв'язку. Негативний зворотний зв'язок виступає в ролі фільтра високих частот для шуму квантування. Більшість ЦАП більшої розрядності (більше 16 біт) побудовані на цьому принципі внаслідок його високої лінійності і низької вартості. Швидкодія дельта-сигма ЦАП сягає сотень тисяч відліків в секунду, розрядність – до 24 біт. Для генерації сигналу з модульованою густотою імпульсів можна використати простий дельта-сигма модулятор першого порядку чи більш високого порядку як MASH (англ. Multi stage noise SHaping). Зі збільшенням частоти передискретизації знижуються вимоги до вихідного фільтра низьких частот і поліпшується приглушення шуму квантування.

3. ЦАП зважування, в якому кожному біту перетворюваного двійкового коду відповідає резистор чи джерело струму, підключене до спільної точки додавання. Сила струму джерела (провідність резистора) пропорційна вазі біта, якому він відповідає. Таким чином, всі ненульові біти коду додаються з вагою. Метод зважування – один з найшвидших, але йому властива низька точність через необхідність наявності набору множини різних прецизійних джерел чи резисторів. Через цю причину ЦАП зважування мають розрядність не більше восьми біт.

4. Ланцюгова R-2R схема є варіацією ЦАП зважування. В R-2R ЦАП зважені значення створюються в спеціальній схемі, яка складається з резисторів опорами R і 2R. Це дозволяє суттєво збільшити точність порівняно зі звичайним ЦАП зважування, оскільки порівняно просто виготовити набір прецизійних елементів з однаковими параметрами. Недоліком методу є більш низька швидкість внаслідок паразитної ємності.

5. Сегментний ЦАП містить по одному джерелу струму чи резистору на кожне можливе значення вихідного сигналу. Так, наприклад, восьмибітовий ЦАП цього типу містить 255 сегментів, а 16-бітовий – 65535. Теоретично, сегментні ЦАП мають найбільшу швидкодію, оскільки для перетворення достатньо замкнути один ключ, який відповідає вхідному коду.

6. Гібридні ЦАП використовують комбінацію перерахованих вище способів. Більшість мікросхем ЦАП належать до цього типу, вибір конкретного набору способів є компромісом між швидкодією, точністю і вартістю ЦАП.

Цифрові вимірювачі неелектричних величин

Сприйняття інформації про об'єкти чи процеси здійснюється за допомогою пристроїв, які називаються первинними перетворювачами чи сенсорами. В більшості випадків сенсори відображають вхідну інформацію у вигляді еквівалентного електричного параметра. Тобто, сенсором називається елемент, який приймає контрольований параметр і перетворює його до вигляду, зручного для подальшої обробки (вимірювання, передачі, контролю).

Відповідно до схеми вмикання сенсорів можна визначити дві групи узгоджувально-нормувальних пристроїв. До першої групи належать пристрої, в яких сенсори є елементами подільників напруги (рис. 4.17, а), до другої – пристрої, в яких сенсори є елементами коливальних систем ВЧ генераторів (рис. 4.17, б).

В пристроях першої групи сенсори найчастіше вмикаються за диференціальною чи мостовою схемою (рис. 4.17, в, г, відповідно).

а) з подільниками напруги; б) з елементами ВЧ генераторів; в) за диференціальною схемою увімкнення; г) за мостовою схемою увімкнення

Диференціальні схеми відрізняються високою стабільністю, оскільки дестабілізуючі фактори одночасно діють на обидва елементи диференціального сенсора, що компенсує цей вплив.

Сенсори в мостовій схемі входять до складу моста, який врівноважений при деякому (звичайно нульовому чи початковому) значенні контрольованого параметра.

При вимірюваннях деяких неелектричних величин не завжди вдається перетворити їх безпосередньо в електричну величину. В цих випадках здійснюють подвійне перетворення первинної вимірюваної величини в проміжну неелектричну величину, яку перетворюють потім у вихідну електричну величину.

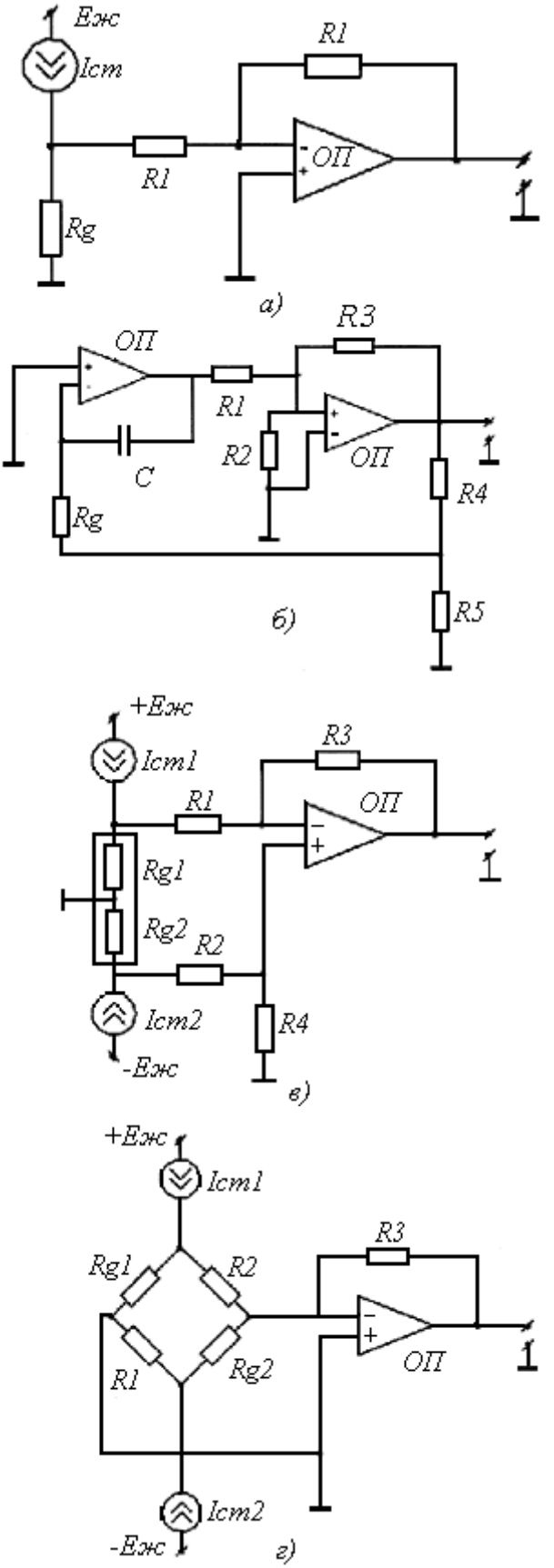


Рисунок 4.17 – Схеми узгодження

Сукупність двох відповідних вимірювальних перетворювачів утворюють комбінований сенсор (рис. 4.18).

Рисунок 4.18 – Блок-схема комбінованого сенсора

Подібні перетворювачі зручні для вимірювання неелектричних (механічних) величин, які викликають в первинному перетворювачі деформацію або переміщення вихідного елемента, до яких чутливий вторинний перетворювач.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Що розуміють під системою дискретного часу?
2. Яка головна вимога оптимальності методу дискретизації?
3. У чому полягає суть квантування?
4. У чому відмінність дискретизації від квантування?
5. У чому полягає сутність кодування інформації?
6. Які переваги та недоліки цифрового подання інформації?
7. Як відбувається відновлення сигналів?
8. У яких випадках використовуються поліноми Лагранжа та Тейлора?
9. Поясніть принципи адаптивної дискретизації.
10. У чому сутність теореми Котельникова?
11. Перелічіть основні вимоги до методів цифрового кодування.
12. У чому відмінність між потенціальними та імпульсними кодами?
13. Що розуміють під терміном “апертурний час”?
14. Назвіть та наведіть схеми найбільш поширених типів АЦП.
15. Які основні переваги та недоліки інтегруючих АЦП?
16. Для чого використовуються багатоканальні АЦП?
17. Назвіть та наведіть схеми найбільш поширених типів ЦАП.
18. Наведіть основні схеми цифрових вимірювачів неелектричних величин.

5 ПЕРЕДАЧА ІНФОРМАЦІЇ В ЕЛЕКТРОННИХ СИСТЕМАХ

5.1 Типова система передачі даних

Повідомлення передаються від об’єкта до адресата за допомогою сукупності технічних засобів, які утворюють систему передачі даних.

Будь-яка система передачі даних (СПД) може бути описана через три основні свої компоненти. Такими компонентами є передавач (джерело передачі інформації), канал передачі даних і приймач (одержувач інформації). При двосторонньому обміні (дуплексній передачі) джерело й одержувач можуть бути об'єднані так, що їхнє устаткування може передавати і приймати дані одночасно. У найпростішому випадку СПД між пунктами А і В (рис. 5.1) складається з таких основних семи частин: кінцевого устаткування даних у пункті А; інтерфейсу (або стику) між кінцевим устаткуванням даних і апаратурою каналу даних; апаратури каналу даних у пункті А; каналу передачі між пунктами А і В; апаратури каналу даних у пункті В; інтерфейсу (або стику) апаратури каналу даних; кінцевого устаткування даних у пункті В.

Кінцеве устаткування даних (КУД) – це узагальнене поняття, використовуване для опису кінцевого приладу користувача або його частини.

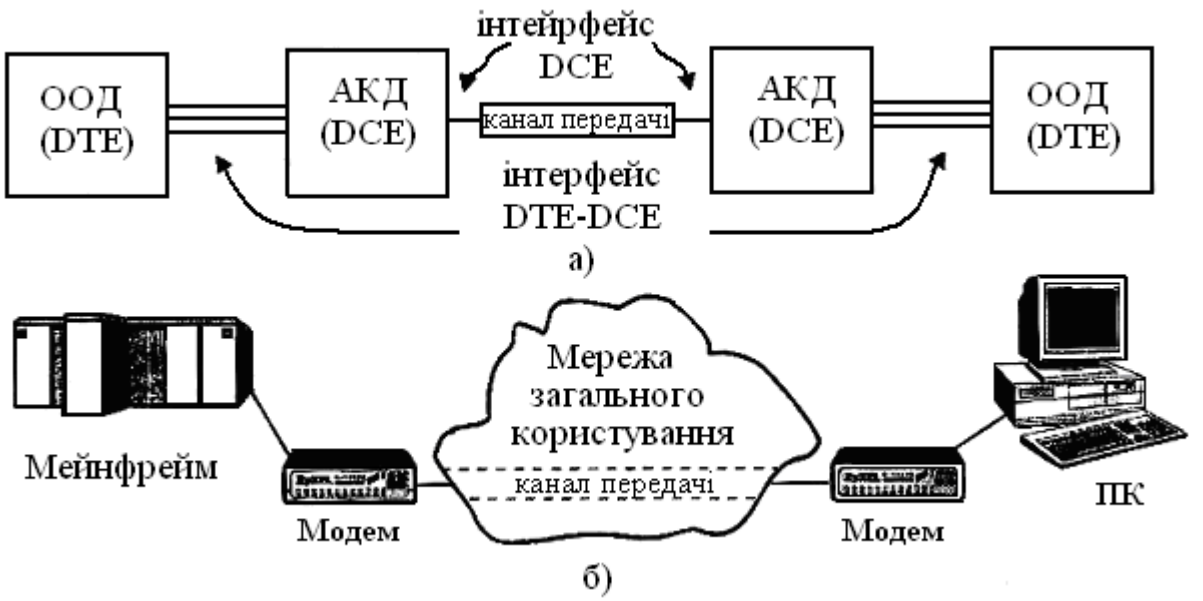


Рисунок 5.1 – Типова система передачі даних: а – блок-схема системи передачі даних; б – реальна система передачі даних

КУД може бути джерелом інформації, її одержувачем або тим і іншим одночасно, передає і (або) приймає дані за допомогою використання апаратури каналу даних і каналу передачі. У літературі часто вживається відповідний міжнародний термін – DTE (Data Terminal Equipment). Як DTE може виступати персональний комп'ютер, велика ЕОМ (Mainframe Computer), термінал, касовий апарат, приймач сигналів глобальної навігаційної системи або будь-яке інше устаткування, спроможне передавати або приймати дані.

Апаратуру каналу даних також називають апаратурою передачі даних. Широко використовується міжнародний термін DCE (Data Communications Equipment), що і будемо вживати надалі. Функція DCE полягає в забезпеченні можливості передачі інформації між двома або великим числом DTE по каналу певного типу, наприклад по телефонному. Для цього DCE повинен забезпечити з'єднання з DTE з одного боку, і з каналом передачі – з іншого. На рис. 5.1, а DCE може бути аналоговим модемом, якщо використовується аналоговий канал, або, наприклад, пристроєм обслуговування каналу/даних (CSU/DSU – Channel Servis Unit/Data Service Unit), якщо використовується цифровий канал типу El / T1 або ISDN.

Модеми, розроблені в 60-70-х роках, являли собою пристрої винятково з функціями перетворення сигналів. Проте в останні роки модеми набули значну кількість складних функцій, що будуть розглянуті нижче. Слово модем є скороченою назвою пристрою, що здійснює процес МОдуляції/ ДЕМодуляції.

Модуляцією називається процес зміни одного або декількох параметрів вихідного сигналу за законом вхідного сигналу. При цьому вхідний сигнал є, як правило, цифровим і називається модулюючим. Вихідний сигнал – звичайно аналоговий і часто називається модульованим сигналом.

Дискретні або цифрові сигнали як засіб передачі інформації знайшли більш широке застосування, ніж аналогові. Це пояснюється тим, що дискретні сигнали меншою мірою схильні до впливу завад у каналах зв'язку, перекручування дискретного сигналу легше виявити, ніж перекручування аналогового, і, головне, дискретні сигнали легко опрацьовуються на електронно-обчислювальних машинах.

Важливу роль у взаємодії DTE і DCE відіграє інтерфейс, що складається з вхідних/вихідних кіл у DTE і DCE, роз’єднувачів і з'єднувальних кабелів. У вітчизняній літературі і стандартах також часто застосовується термін стик.

З'єднання DTE із DCE відбувається по одному зі стиків типу С2. При підключенні DCE до каналу зв'язку або середовища поширення застосовується один із стиків типу С1.

5.2 Канали зв'язку. Аналогові і цифрові канали

Під каналом зв'язку розуміють сукупність середовища поширення і технічних засобів передачі між двома канальними інтерфейсами або стиками типу С1. З цієї причини стик С1 часто називається канальним стиком.

В залежності від типу сигналів, що передаються, розрізняють два великих класи каналів зв'язку – цифрові та аналогові.

Цифровий канал є бітовим трактом із цифровим (імпульсним) сигналом на вході і виході каналу. На вхід аналогового каналу надходить неперервний сигнал, і з його виходу також знімається неперервний сигнал (рис 5.2). Як відомо, сигнали характеризуються формою свого подання.

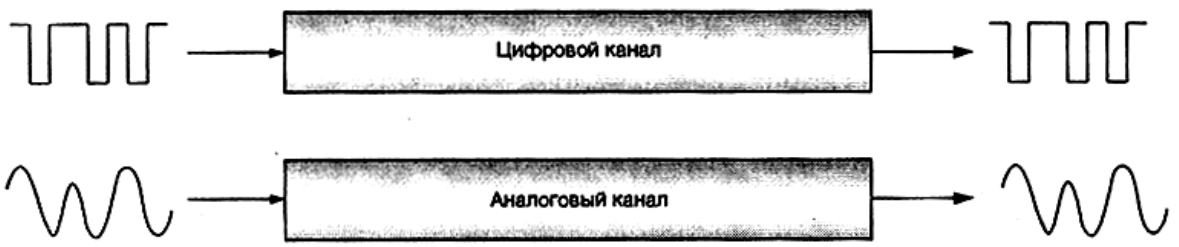


Рисунок 5.2 – Цифрові й аналогові канали передачі

Цифровими є канали систем ІКМ, ISDN, канали типу Т1/ Е1 і більшість інших. Нові СПД будують на основі цифрових каналів, що мають ряд переваг перед аналоговими.

Аналогові канали є найбільш поширеними через тривалий історичний шлях їхнього розвитку і простоти реалізації. Типовим прикладом аналогового каналу є канал тональної частоти, а також групові тракти на 12, 60 і більше каналів тональної частоти. Комутуюча телефонна мережа загального користування, як правило, включає численні комутатори, розподільні пристрої, групові модулятори і демодулятори. Для такої мережі канал передачі (його фізичний маршрут і ряд параметрів) змінюється при кожному черговому виклику.

При передачі даних на вході аналогового каналу повинен знаходитись пристрій, що перетворював би цифрові дані, які приходять від DTE, в

аналогові сигнали, що посилаються в канал. Приймач повинен містити пристрій, що перетворював би назад прийняті неперервні сигнали в цифрові дані. Цими пристроями є модеми. Аналогічно, при передачі по цифрових каналах дані доводиться приводити до вигляду, прийнятого для даного конкретного каналу. Цим перетворенням займаються цифрові модеми, яких часто називають адаптери ISDN, адаптери каналів E1/T1, лінійні драйвери, і так далі (у залежності від конкретного типу каналу або середовища передачі).

Термін модем використовується широко. При цьому необов'язково повідомляється, яка модуляція використовується, а просто вказується на певні операції перетворення сигналів, що надходять від DTE для їх подальшої передачі по каналу. Таким чином, у широкому змісті поняття модем і апаратура каналу даних (DCE) є синонімами.

Комутовані канали надаються споживачам на час з'єднання за вимогою (дзвоником). Такі канали принципово містять у своєму складі комутаційне устаткування телефонних станцій (АТС). Звичні телефонні апарати використовують комутовані канали.

Виділені (орендовані) канали орендуються у телефонних компаній або (дуже рідко) прокладаються зацікавленою організацією. Такі канали є принципово двоточковими. Їх якість у загальному випадку вище якості комутованих каналів через відсутності впливу комутаційної апаратури.

Як правило, канали мають двопровідне або чотирипровідне закінчення. Їх називають, відповідно, двопровідними і чотирипровідними.

Чотирипровідні канали надають два провідники для передачі сигналу і ще два провідники для прийому. Перевагою таких каналів є, практично, повна відсутність впливу сигналів, що передаються у зустрічному напрямку.

Двопровідні канали дозволяють використовувати два провідники як для передачі, так і для прийому сигналів. Такі канали дозволяють заощаджувати на вартості кабелів, але потребують ускладнення каналоутворювальної апаратури й апаратури користувача. Двопровідні канали потребують вирішення задачі поділу сигналів, що приймаються і передаються. Така розв'язка реалізовується за допомогою диференціальних систем, що забезпечують необхідне загасання на зустрічних напрямках передачі. Але неідеальність диференціальних систем (а ідеального нічого не буває) приводить до перекручувань амплітудно-частотних і фазочастотних характеристик каналу.

Теоретичну основу сучасних інформаційних систем визначає Базова еталонна модель взаємодії відкритих систем (OSI – Open Systems Interconnection) Міжнародної організації стандартів. Вона описана стандартом ISO 7498. Модель є міжнародним стандартом для передачі даних. Згідно з еталонною моделлю виділяють сім рівнів, які утворюють область взаємодії відкритих систем. До них відноситься фізичний, канальний, мережевий, транспортний, сеансовий, представницький і прикладний рівні. Кожний рівень визначається групою стандартів, які включають в себе дві специфікації: протокол і сервіс для вищого рівня.

Найближчим до користувача є прикладний рівень. Його головна задача – подати уже перероблену інформацію. Для цього, звичайно, використовується системне програмне забезпечення та програмне забезпечення користувача.

Найнижчий фізичний рівень описує процедури передачі сигналів в канал і отримання їх із каналу. Як фізичне середовище передачі можуть виступати канал тональної частоти, з'єднувальна провідна лінія, радіоканал і т. д.

Фізичний рівень виконує три основні функції: встановлення зв'язку, перетворення сигналів, реалізацію інтерфейсу.

5.3 Кодування інформації

При передачі по каналах зв'язку завжди виникають помилки. Причини їх можуть бути різні, але результат видається один – дані спотворюються і не можуть бути використані на прийомній стороні для подальшого опрацювання. Як правило, можливість перекручування біта в потоку переданих даних на рівні фізичного каналу знаходиться в межах $10^{2}...10^{-6}$. У той же час із боку користувачів і багатьох прикладних процесів часто висовується вимога до можливості помилок у прийнятих даних не гірше $10^{-6}... 10^{-12}$. Боротьба з виникаючими помилками ведеться на різних рівнях семирівневої моделі OSI (в основному на перших чотирьох). Для боротьби з виникаючими помилками відомо багато різноманітних способів.

В одному із способів на передавальній стороні передані дані кодуються одним із відомих кодів із виправленням помилок. На приймальній стороні, відповідно, проводиться декодування прийнятої інформації і виправлення виявлених помилок. Можливість застосовуваного коду з виправленням помилок залежить від числа надлишкових бітів, що генеруються кодером. Якщо внесена надмірність невелика, тобто існує небезпека того,

що прийняті дані будуть містити незнайдені помилки, це може призвести до помилок у роботі прикладного процесу. Якщо ж використовувати код із високою виправлювальною здатністю, то це приводить до низької швидкості передачі даних. Таким чином, знання теорії завадостійкого кодування дозволяє визначити оптимальні параметри завадостійкого коду в залежності від поставленої задачі.

5.3.1 Завадостійке кодування

Завадостійкі коди – один з найбільш ефективних засобів забезпечення високої вірності як при зберіганні, так і при передачі дискретної інформації. Створено спеціальну теорію завадостійкості кодування, що швидко розвивається останнім часом.

К. Шеннон сформулював теорему для випадку передачі дискретної інформації з каналу із завадами, яка стверджує, що ймовірність помилкового декодування прийнятих сигналів може бути забезпечена як завгодно малою шляхом вибору відповідного способу кодування сигналів.

Під завадостійкими кодами розуміють коди, що дозволяють виявляти або виявляти і виправляти помилки, які виникають у результаті впливу завад.

Завадостійкість кодування забезпечується за рахунок введення надмірності в кодові комбінації, тобто за рахунок того, що не всі символи в кодових комбінаціях використовуються для передачі інформації.

Всі завадостійкі коди можна розділити на два основних класи: блокові і неперервні (рекуррентні або ланцюгові).

У блокових кодах кожному повідомленню (або елементу повідомлення) відповідає кодова комбінація (блок) із певної кількості сигналів. Блоки кодують і декодують окремо. Блокові коди можуть бути рівномірними, коли довжина кодових комбінацій n постійна, або нерівномірними, коли n мінлива.

Нерівномірні завадостійкі коди не одержали практичного застосування через складність їх технічної реалізації.

Як блокові, так і неперервні коди в залежності від методів внесення надмірності розділяються на роздільні і нероздільні. У роздільних кодах чітко розмежована роль окремих символів. Одні символи є інформаційними, інші є перевірними і служать для виявлення і виправлення помилок. Роздільні блокові коди називаються звичайно n-кодами, де n – довжина кодових комбінацій, k – число інформаційних символів у комбінаціях. Не-

роздільні коди не мають чіткого розділення кодової комбінації на інформаційні і перевірні символи. Цей клас кодів поки нечисленний. Роздільні блокові коди розділяються, у свою чергу, на несистематичні і систематичні.

Більшість відомих роздільних кодів складають систематичні коди. У цих кодів перевірні символи визначаються в результаті проведення лінійних операцій над певними інформаційними символами. Для випадку двійкових кодів кожний перевірний символ вибирається таким, щоб його сума за модулем два з певними інформаційними символами стала рівною нулю. Декодування зводиться до перевірки на парність певних груп символів. У результаті таких перевірок дається інформація про наявність помилок, а в разі потреби - про позицію символів, де є помилки.

Основні принципи завадостійкого кодування

Для з'ясування ідеї завадостійкого кодування розглянемо двійковий код, що знайшов на практиці найбільш широке застосування. Нагадаємо, що двійковий код – це код із основою $m = 2$. Кількість розрядів n у кодовій комбінації прийнято називати довжиною або значністю коду. Символи кожного розряду можуть приймати значення 0 і 1. Кількість одиниць у кодовій комбінації називають вагою кодової комбінації і позначають w.

Ступінь відмінності будь-яких двох кодових комбінацій даного коду характеризується так званою відстанню між кодами d. Вона виражається числом позицій або символів, у яких комбінації відрізняються одна від одної, і визначається як вага суми за модулем два цих кодових комбінацій.

Помилки, внаслідок впливу завад, виявляються в тому, що в однім або декількох розрядах кодової комбінації нулі переходять в одиниці і, навпаки, одиниці переходять у нулі. В результаті створюється нова - помилкова кодова комбінація.

Якщо помилки відбуваються тільки в одному розряді кодової комбінації, то їх називають однократними. При наявності помилок у двох, трьох і т. д. розрядах помилки називають дворазовими, триразовими і т. д.

Експериментальні дослідження каналів зв'язку показали, що помилки символів при передачі по каналу зв'язку, як правило, групуються в пачки різної тривалості. Під пачкою помилок розуміють ділянку послідовності, що починається і закінчується помилково прийнятими символами. Всередині пачки можуть бути і правильно прийняті елементи.

Для вказання місць у кодовій комбінації, де є перекручування символів, використовується вектор помилки e. Вектор помилки n-розрядного коду – це n-розрядна комбінація, одиниці в якій указують положення перекручених символів кодової комбінації.

Вага вектора помилки *we* характеризує кратність помилки. Сума за модулем два для перекрученої кодової комбінації і вектори помилки дають вихідну невикривлену комбінацію.

Як уже відзначалося, завадостійкість кодування забезпечується за рахунок внесення надмірності в кодові комбінації. Це значить, що з n символів кодової комбінації для передачі інформації використовується $k < n$ символів. Отже, із загального числа $N_o=2^n$ можливих кодових комбінацій для передачі інформації використовується тільки $N = 2^k$ комбінацій. Відповідно до цього вся множина $N_o=2^n$ можливих кодових комбінацій поділяється на дві групи. У першу групу входить множина $N = 2^k$ дозволених комбінацій, друга група містить у собі множину $(N_o - N) = 2^n-2^k$ заборонених комбінацій.

Якщо на стороні приймання встановлено, що прийнята комбінація відноситься до групи дозволених, то вважається, що сигнал прийшов без перекручувань. В іншому випадку робиться висновок, що прийнята комбінація перекручена. Однак це справедливо лише для таких перешкод, коли усунута можливість переходу одних дозволених комбінацій в інші.

У загальному випадку кожна з N дозволених комбінацій може трансформуватися в будь-яку з N_o можливих комбінацій, тобто усього є N·No можливих варіантів передачі, із них:

N варіантів безпомилкової передачі;

$N(N-1)$ варіантів переходу в інші дозволені комбінації;

$N(No-N)$ варіантів переходу в заборонені комбінації.

Таким чином, не всі перекручування можуть бути виявлені. Частка помилкових комбінацій, що виявляються, складає

$$\frac{N(N_0 - N)}{N \cdot N_0} = 1 - \frac{N}{N_0}.$$

Для використання даного коду як виправного множина заборонених кодових комбінацій розбивається на N підмножин, що не перетинаються. Кожна з підмножин ставиться у відповідність одній із дозволених комбінацій. Помилка виправляється в $(N_o - N)$ випадках, рівних кількості заборонених комбінацій. Частка помилкових комбінацій, що виправляються, від загального числа помилкових комбінацій, що виявляються, складає

$$\frac{N_0 - N}{N\left((N_0 - N\right)} = \frac{1}{N}.$$

Спосіб розбиття на підмножини залежить від того, які помилки повинні виправлятися даним кодом.

ПРИКЛАД

Нехай необхідно побудувати код, що виявляє всі помилки кратністю t і нижче.

Побудувати такий код – це означає із множини N_o можливих вибрати N дозволених комбінацій так, щоб будь-яка з них у сумі за модулем два з будь-яким вектором помилок із вагою $W_l \le t$ не дала б у результаті ніякої іншої дозволеної комбінації. Для цього необхідно, щоб найменша кодова відстань задовольняла умову

$$d_{\min} \ge t + 1.$$

У загальному випадку для усунення помилок кратності σ кодова відстань повинна задовольняти умову

$$d_{\min} \ge 2\sigma + 1.$$

Аналогічно міркуючи, можна встановити, що для виправлення всіх помилок кратності не більше σ і одночасно виявлення всіх помилок кратності не більше t і (при $t \ge \sigma$) кодова відстань повинна задовільняти умову

$$d_{\min} \ge t + \sigma + 1.$$

5.3.2 Коди із заданою виправною здатністю

Підвищення коригуючої здатності коду досягається при зберіганні n за рахунок зменшення множини N дозволених комбінацій (або зменшення кількості k інформаційних символів). Звичайно ж на практикці коди будуються в зворотному порядку: спочатку вибирається кількість інформаційних символів k, виходячи з об'єму алфавіту джерела, а потім забезпечується необхідна коригуюча спроможність коду за рахунок додавання надлишкових символів.

Нехай відомий об'єм алфавіту джерела N. Необхідна кількість інформаційних символів

$$k = \log_2 N.$$

Нехай також відомо повне число помилок E, що необхідно виправити.

Завдання полягає в тому, щоб при заданих N і E визначити значність коду n, що має необхідні коригуючі можливості.

Повне число помилкових комбінацій, які підлягають виправленню, дорівнює $E \cdot 2^k = E \cdot N$. Тому, що кількість помилкових комбінацій дорівнює $N_o - N$, то код забезпечує виправлення не більше $N_o - N$ комбінацій. Отже, необхідну умову для можливості виправлення помилок можна записати у вигляді

$$\boldsymbol{NE} \le \boldsymbol{N}_0 - \boldsymbol{N} ,$$

і отримаємо

$$\boldsymbol{N}_0 \ge (1 + \boldsymbol{E})\boldsymbol{N} ,$$

або

$$\boldsymbol{N} \le \frac{2^{\boldsymbol{n}}}{1 + \boldsymbol{E}} . \qquad (5.1)$$

Формула (5.1) виражає умову для вибору значності коду n. Розглянемо окремі випадки. Якщо є помилки різної кратності, то насамперед необхідно забезпечити усунення однократних помилок, ймовірність появи яких найбільша. Можлива кількість векторів однократних помилок

$$\boldsymbol{E} = \boldsymbol{C}_{\boldsymbol{n}}^{1} = \boldsymbol{n} .$$

У цьому випадку залежність (5.1) набуде вигляду

$$\boldsymbol{N} = 2 \le \frac{2^{\boldsymbol{n}}}{1 + \boldsymbol{n}} . \qquad (5.2)$$

При побудові коду доцільно користуватися табл. 5.1. Потрібно при цьому мати на увазі, що код повинен також задовільняти умову $d_{min} \ge 3$.

Таблиця 5.1 – Виявлення значності коригуючого коду

n	2	3	4	5	6	7	8	9
2n/(1+n)	1,88	2	3,2	5,33	9,2	16	28,4	51,2

Якщо необхідно забезпечити усунення всіх помилок кратності від 1 до l, то потрібно врахувати, що

кількість можливих однократних помилок $E_1 = C^l_n$,

кількість можливих дворазових помилок $E_2 = C^2_n$

кількість можливих l-кратних помилок $E_l = C^l_n$.

Загальна кількість помилок $E = \sum_{i=1}^{l} C_n^i$.

При цьому залежність (5.1) набуде вигляду

$$N \le \frac{2^n}{1+\sum_{i=1}^{l} C_n^i}. \qquad (5.3)$$

Умова (5.3) є нижньою оцінкою для довжини коригуючоого коду, тобто вона визначає необхідну мінімальну довжину коду n, що забезпечує виправлення помилок заданої кратності при відомому числі дозволених комбінацій N або числі інформаційних символів $k = log_2 N$.

Це ж умова є верхньою оцінкою для N або k, тобто визначає максимально можливе число дозволених комбінацій або інформаційних символів для коду довжини n, що забезпечує виправлення помилок заданої кратності.

Будь-який коригуючий код характеризується рядом показників: довжиною n, кількістю інформаційних символів k або надлишкових символів $\rho = n - k$, повним числом усіх можливих кодових комбінацій $N_o = m^n$ (для двійкових кодів $N = 2^n$), числом дозволених кодових комбінацій (потужністю коду) N, вагою кодової комбінації w, вагою вектора помилки w_1, кодовою відстанню d і ін.

Однак основним показником якості коригуючого коду є його спроможність забезпечити правильне прийняття кодових комбінацій при наявності перекручувань під впливом перешкод, тобто завадостійкість коду.

Коригуюча можливість коду забезпечується за рахунок надмірності, тобто подовження кодових комбінацій. При подовженні кодових комбінацій ускладнюється апаратура, збільшується час передачі й опрацювання інформації. Тому надмірність також є важливою характеристикою коду.

Для оцінки надмірності коду користуються поняттям коефіцієнта надмірності

$$K_{над} = \frac{\rho}{n} = \frac{n-k}{n},$$

де ρ – кількість надлишкових символів у кодовій комбінації.

ПРИКЛАД

Визначити коригуючу здатність коду, що має такі дозволені комбінації: 00000; 01110; 10101; 11011.

Розв'язання. Коригуюча здатність коду визначається мінімальною кодовою відстанню. Складемо матрицю відстаней між кодовими комбінаціями (табл. 5.2).

Таблиця 5.2 – Матриця відстаней

	00000	01110	10101	11011
00000	0	3	3	4
01110	3	0	4	3
10101	3	3	0	3
11011	4	3	4	0

Як видно з матриці, мінімальна кодова відстань $d_{min}=3$. Отже, даний код здатний:

а) виявляти дворазові помилки;

б) усувати однократні помилки;

в) усувати і виявляти однократні помилки.

5.3.3 Систематичні коди

В наш час найбільш широкий клас коригуючих кодів складають систематичні коди. Ці коди відносяться до групи роздільних блокових кодів. Для систематичного коду сума за модулем для двох дозволених комбінацій також дає дозволену комбінацію.

В теорії кодування широко використовується матричне подання кодів. Всі дозволені кодові комбінації систематичного (n, k)-коду можна одержати, маючи k вихідних дозволених кодових комбінацій. Вихідні кодові комбінації повинні задовольняти такі умови:

1. У число вихідних комбінацій не повинна входити нульова.

2. Кодова відстань між будь-якими парами вихідних комбінацій повинна бути не менше d_{min}.

3. Кожна вихідна комбінація, як і будь-яка ненульова дозволена комбінація, повинна містити кількість одиниць не менше d_{min}.

4. Всі вихідні комбінації повинні бути лінійно незалежні, тобто, жодна з них не може бути отримана шляхом підсумовування інших.

Вихідні комбінації можуть бути отримані з матриці, що складається з k рядків і n стовпців:

$$\boldsymbol{P}_{n,k} = \begin{vmatrix} \boldsymbol{a}_{11}\boldsymbol{a}_{12}\ \ldots\boldsymbol{a}_{1k}\,\boldsymbol{b}_{11}\boldsymbol{b}_{12}\ \ldots\boldsymbol{b}_{1\rho} \\ \boldsymbol{a}_{21}\boldsymbol{a}_{22}\ \ldots\boldsymbol{a}_{2k}\,\boldsymbol{b}_{21}\boldsymbol{b}_{22}\ \ldots\boldsymbol{b}_{2\rho} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \boldsymbol{a}_{k1}\boldsymbol{a}_{k2}\ \ldots\boldsymbol{a}_{kk}\,\boldsymbol{b}_{k1}\boldsymbol{b}_{k2}\ \ldots\boldsymbol{b}_{k\rho} \end{vmatrix}.$$

Тут символи перших k стовпців є інформаційними й останні ρ стовпців кількість перевірними. Матрицю $P_{n,k}$ називають породжувальною. Породжувальна матриця $P_{n,k}$ може бути подана двома підматрицями – інформаційною U_k і перевірною H_p:

$$\boldsymbol{P}_{n,k} = \left|\begin{matrix} \boldsymbol{a}_{11}\boldsymbol{a}_{12}\ \ldots\boldsymbol{a}_{1k} \\ \boldsymbol{a}_{21}\boldsymbol{a}_{22}\ \ldots\boldsymbol{a}_{2k} \\ \ldots\ldots\ldots\ldots\ldots \\ \boldsymbol{a}_{k1}\boldsymbol{a}_{k2}\ \ldots\boldsymbol{a}_{kk} \end{matrix}\right| \left|\begin{matrix} \boldsymbol{b}_{11}\boldsymbol{b}_{12}\ \ldots\boldsymbol{b}_{1\rho} \\ \boldsymbol{b}_{21}\boldsymbol{b}_{22}\ \ldots\boldsymbol{b}_{2\rho} \\ \ldots\ldots\ldots\ldots\ldots \\ \boldsymbol{b}_{k1}\boldsymbol{b}_{k2}\ \ldots\boldsymbol{b}_{k\rho} \end{matrix}\right|,$$

де

$$\boldsymbol{U}_k = \begin{vmatrix} \boldsymbol{a}_{11}\boldsymbol{a}_{12}\ \ldots\boldsymbol{a}_{1k} \\ \boldsymbol{a}_{21}\boldsymbol{a}_{22}\ \ldots\boldsymbol{a}_{2k} \\ \ldots\ldots\ldots\ldots\ldots \\ \boldsymbol{a}_{k1}\boldsymbol{a}_{k2}\ \ldots\boldsymbol{a}_{kk} \end{vmatrix};\ \boldsymbol{H}_\rho = \begin{vmatrix} \boldsymbol{b}_{11}\boldsymbol{b}_{12}\ \ldots\boldsymbol{b}_{1\rho} \\ \boldsymbol{b}_{21}\boldsymbol{b}_{22}\ \ldots\boldsymbol{b}_{2\rho} \\ \ldots\ldots\ldots\ldots\ldots \\ \boldsymbol{b}_{k1}\boldsymbol{b}_{k2}\ \ldots\boldsymbol{b}_{k\rho} \end{vmatrix}.$$

Для побудови породжувальної матриці зручно інформаційну матрицю брати у вигляді квадратної одиничної матриці

$$\boldsymbol{U}_k = \begin{vmatrix} 1\ 0\ 0\ \ldots\ 0 \\ 0\ 1\ 0\ \ldots\ 0 \\ 0\ 0\ 1\ \ldots\ 0 \\ \ldots\ldots\ldots\ldots \\ 0\ 0\ 0\ \ldots\ 1 \end{vmatrix}.$$

При цьому перевірна підматриця H_p повинна будуватися з дотриманням таких умов:

а) кількість одиниць у рядку повинна бути не менше d_{min-1};

б) сума за модулем два двох будь-яких рядків повинна містити не менше d_{min-2} одиниць.

Перевірні символи утворюються за рахунок лінійних операцій над інформаційними символами. Для кожної кодової комбінації повинно бути складено ρ незалежних сум за модулем два. Вибір інформаційних символів, що беруть участь у формуванні того або іншого перевірного символу, залежить від способу декодування коду. Дуже зручно перевірні суми складати за допомогою перевірної матриці H, що будується в такий спосіб.

Спочатку будується підматриця H^1, що є транспонованою стосовно підматриці H_ρ:

$$\boldsymbol{H}^1 = \begin{vmatrix} \boldsymbol{b}_{11}\,\boldsymbol{b}_{21}\ \ldots \boldsymbol{b}_{k1} \\ \boldsymbol{b}_{12}\,\boldsymbol{b}_{22}\ \ldots \boldsymbol{b}_{k2} \\ \ldots\ldots\ldots\ldots\ldots \\ \boldsymbol{b}_{1\rho}\,\boldsymbol{b}_{2\rho}\ \ldots \boldsymbol{b}_{k\rho} \end{vmatrix}.$$

Потім до неї справа приписується одинична матриця

$$\boldsymbol{H} = \begin{vmatrix} \boldsymbol{b}_{11}\,\boldsymbol{b}_{21}\ \ldots \boldsymbol{b}_{k1}\ 1\ 0\ \ \ldots\ 0 \\ \boldsymbol{b}_{12}\,\boldsymbol{b}_{22}\ \ldots \boldsymbol{b}_{k2}\ 0\ 1\ \ \ldots\ 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \boldsymbol{b}_{1\rho}\,\boldsymbol{b}_{2\rho}\ \ldots \boldsymbol{b}_{k\rho}\,0\ 0\ \ \ldots\ 1 \end{vmatrix}. \qquad (5.4)$$

Алгоритм визначення перевірних символів за інформаційними за допомогою матриці (5.4) такий. Позиції, які займаються одиницями в першому рядку підматриці H^1, визначають інформаційні розряди, що повинні брати участь у формуванні першого перевірного розряду кодової комбінації. Позиції одиниць у другому рядку підматриці Н1 визначають інформаційні розряди, що беруть участь у формуванні другого перевірного розряду і т. д.

Перевірна матриця H дуже зручна для визначення місця помилки в кодовій комбінації, а, отже, виправлення помилок. Перевірка кодових комбінацій при цьому виконується шляхом підсумовування за модулем два перевірних символів кодових комбінацій і перевірних символів, обчислених за прийнятими інформаційними. У результаті буде отримана сукупність контрольних рівностей, кожна з яких являє суму за модулем два одного з контрольних символів і певної кількості інформаційних.

Склад контрольних рівностей легко визначається з перевірної матриці H. До складу першої контрольної рівності повинні входити символи, позиції яких зайняті одиницями в першому рядку матриці H. До складу другої контрольної рівності повинні входити символи, позиції яких зайняті одиницями в другому рядку матриці Н, і т. д.

В результаті ρ таких перевірок буде отримане ρ-розрядне двійкове число (синдром), що буде рівне нулю при відсутності помилок і відмінним від нуля у випадку наявності помилок.

Якщо для виправлення однократних помилок у кодових комбінаціях одержати синдроми досить просто, то для виправлення дворазових, триразових і т. д. помилок, а також для виправлення пачок помилок побу-

дова синдромів досить важка й у цих випадках звертаються звичайно за допомогою до ЕОМ.

ПРИКЛАД

Побудувати породжувальну матрицю систематичного коду, здатного виправляти одиничну помилку (σ= 1) при передачі 16 повідомлень.

Розв'язання. Тому, що число дозволених кодових комбінацій $N = 16$, то число інформаційних розрядів у кодових комбінаціях

$$\boldsymbol{k} = \log_2 16 = 4 .$$

Отже, число рядків породжувальної матриці $P_{n,k}$ повинно дорівнювати чотирьом.

Користуючись умовою (5.2) і табл. 5.1, знаходимо довжину коду $n = 7$. Отже, число стовпців матриці $P_{n,k}$ дорівнює сім.

Оскільки число перевірних розрядів $p = n - k = 3$, то число стовпців перевірної підматриці H_p дорівнює трьом. Кількість одиниць у кожному рядку підматриці H_p повинна бути не менша $d_{min} - 1$. Для $\sigma=1$ $d_{min} = 2\sigma + 1 = 3$. Вибираємо для підматриці H_p такі комбінації: 111, 110, 101, 011. Отже, перевірна підматриця може мати вигляд:

$$\boldsymbol{H}_p = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix}, \text{ або } \boldsymbol{H}_p = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix}, \text{ або } \boldsymbol{H}_p = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix}.$$

Інформаційну підматрицю U_k беремо у вигляді одиничної матриці з кількістю рядків і стовпців, рівною $k = 4$:

$$\boldsymbol{U}_k = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}.$$

Остаточний вид породжувальної матриці:

$$\boldsymbol{P}_{7,4} = \begin{vmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{vmatrix}, \text{ або } \boldsymbol{P}_{7,4} = \begin{vmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{vmatrix}, \text{ або}$$

$$\boldsymbol{P}_{7,4} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{vmatrix}.$$

ПРИКЛАД

Побудувати перевірну матрицю H систематичного коду (7,4), утворююча матриця якого має вигляд

$$\boldsymbol{P}_{7,4} = \begin{vmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{vmatrix}.$$

Скласти рівняння для визначення перевірних символів.

Розв'язання. Знаходимо підматрицю H^1, що є транспонованою стосовно підматриці H_p:

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

Приписуючи до неї справа одиничну матрицю U_k, отримаємо перевірну матрицу

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

Перевірні символи визначають шляхом підсумовування за модулем два певних інформаційних символів. Алгоритм визначення перевірних символів визначається підматрицею H^1. Позиції, що займаються одиницями в першому рядку H^1 визначають інформаційні розряди, що повинні брати участь у формуванні першого перевірного розряду. Позиції одиниць у другому рядку визначають інформаційні розряди, що беруть участь у формуванні другого перевірного розряду. І, нарешті, позиції одиниць у третьому рядку визначають інформаційні розряди, що беруть участь у формуванні третього перевірного розряду. Отже, перевірні символи визначають рівностями

$$\boldsymbol{b}_1 = \boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3;\ \ \boldsymbol{b}_2 = \boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_4;\ \ \boldsymbol{b}_3 = \boldsymbol{a}_1 + \boldsymbol{a}_3 + \boldsymbol{a}_4.$$

ПРИКЛАД

Побудувати комбінацію систематичного коду (7, 4) для випадку, коли проста ненадмірна комбінація має вигляд 0110.

Розв'язання. У даному випадку значення інформаційних символів такі: a_1=0; a_2= 1; a_3 = 1; a_4 = 0.

Тоді, використовуючи отримані в попередньому прикладі рівності для перевірних символів, визначимо:

$$\boldsymbol{b}_1 = 0+1+1=0;\ \ \boldsymbol{b}_2 = 0+1+0=1;\ \ \boldsymbol{b}_3 = 0+1+0=1.$$

Отже, повідомлення 0110, закодоване систематичним кодом (7, 4), прийме вигляд: 0110011.

ПРИКЛАД

Утворююча матриця коду (7, 4) має вигляд

$$\boldsymbol{P}_{11,7} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{vmatrix}.$$

Побудувати комбінацію систематичного коду (11, 7), що відповідає ненадмірній комбінації 1101001.

Розв'язання. Оскільки для одержання комбінації вигляду 1101001 необхідно скласти перший, другий, четвертий і сьомий рядки одиничної матриці U_k, для побудови перевірних розрядів варто скласти ці ж рядки підматриці H_p:

$$\begin{array}{r} 1111 \\ +\,1110 \\ 1011 \\ 0110 \\ \hline 1100 \end{array}.$$

Таким чином, комбінація коду (11, 7) має вигляд 11010011100.

5.4 Коди з виявленням помилок

До найбільш поширених кодів з виявленням помилок відносять код з перевіркою на парність, коди з прямим та інверсним повторенням та кореляційний код. Розглянемо їх детальніше.

5.4.1 Код із перевіркою на парність

Код містить лише один надлишковий символ. Вибирається надлишковий символ таким чином, щоб загальна кількість одиниць у кодовій комбінації була парною. Перевірка кодової комбінації робиться шляхом підсумовування за модулем два всіх його символів. Надмірність коду дорівнює

$$K_{над} = \frac{\rho}{n} = \frac{1}{n},$$

де n – довжина кодової комбінації;

$\rho = 1$ – число перевірних символів.

Код дозволяє виявляти однократні помилки і всі помилки непарної кратності, тому що тільки в цих випадках кількість одиниць у комбінації стане непарною. Не виявляються помилки парної кратності.

Кодер коду з перевіркою на парність

Кодером коду з перевіркою на парність є електронний пристрій, який при надходженні на його входи п інформаційних символів двійкового коду автоматично добавляє до інформаційних символів один контрольний символ «0» чи «1», значення якого залежить від кількості одиниць в двійковій кодовій комбінації. Якщо кількість одиниць в двійковому коді парна, то контрольний символ дорівнює «0», якщо непарна - «1».

Таким чином у всіх кодових комбінаціях коду з перевіркою на парність кількість одиниць завжди повинна бути парною. Виявити спотворену внаслідок завад комбінацію можна на основі аналізу парності кількості одиниць в комбінації. Якщо парність порушена, кодова комбінація не приймається.

Схема кодера коду з перевіркою на парність залежить від того, як вводиться звичайний двійковий код – паралельно чи послідовно. На рис. 5.3 наведена схема кодера при паралельному введенні двійкового коду в регістр.

При подачі сигналу з рівнем логічної «1» на вхід L регістра, двійкова комбінація, яка подається на інформаційні входи D регістра DD1, запи-

сується на виходи Q1-Q4. Елементом, який виробляє п'ятий контрольний символ, є суматор за модулем 2 m2 - DD2.

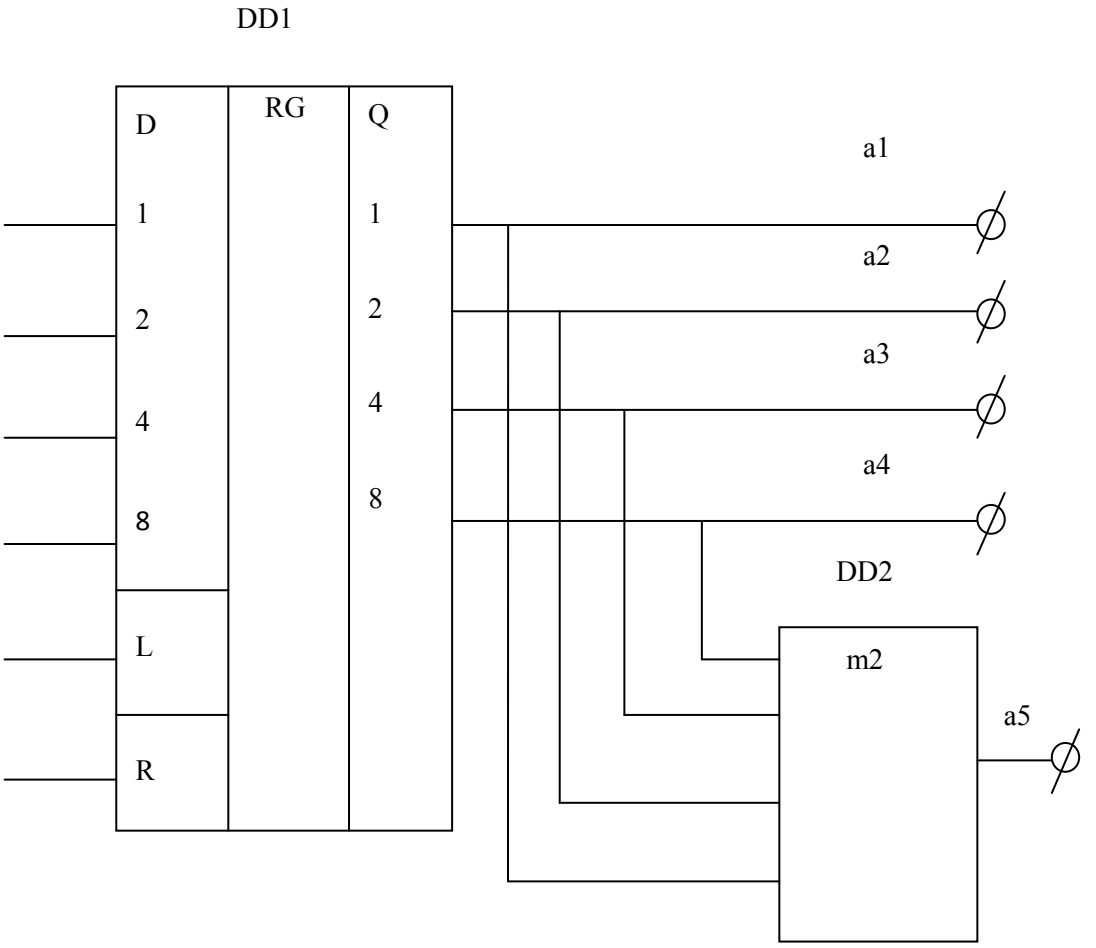


Рисунок 5.3 – Кодер коду з перевіркою на парність при паралельному введенні в регістр двійкової кодової комбінації

Якщо кількість одиниць на його вході парна, то на його виході з'явиться нуль, якщо непарна – одиниця. Тому в деяких літературних джерелах суматор за модулем два називається *генератором парності*.

Таким чином, вхідна чотирирозрядна комбінація на вході кодера автоматично перетворилась в п'ятирозрядну. П'ятий символ – контрольний, який в приймальному пристрої дозволяє визначити чи є помилка в прийнятій комбінації.

Якщо інформаційна комбінація звичайного двійкового коду вводиться в кодер послідовно, схема ускладнюється (рис. 5.4).

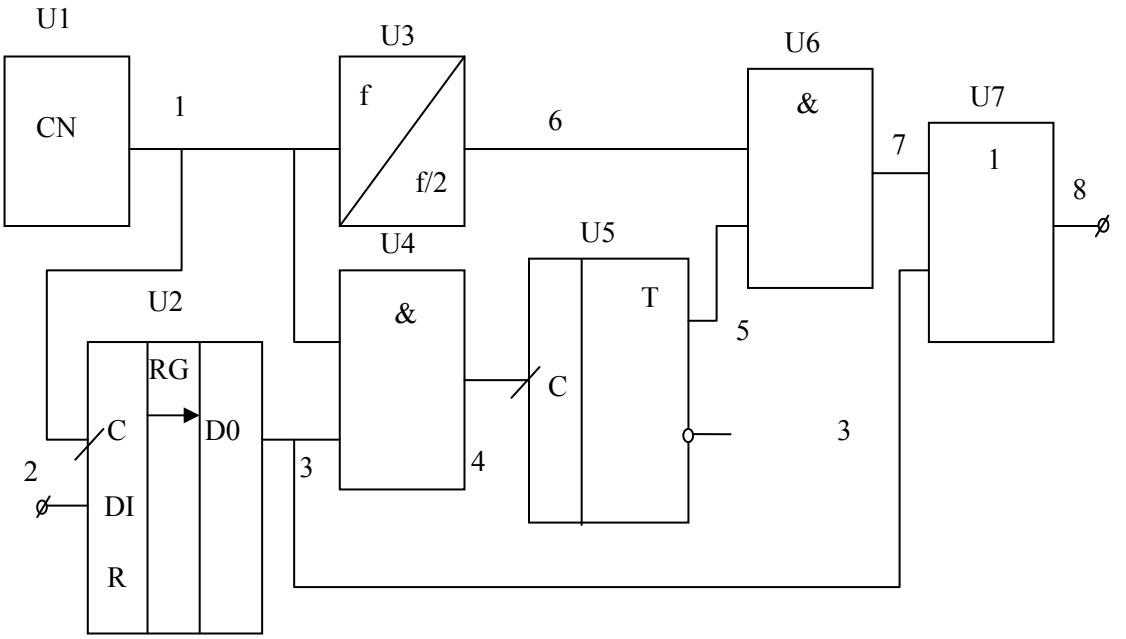


Рисунок 5.4 – Кодер коду з перевіркою на парність при послідовному введенні інформації в регістр

ПРИКЛАД

Для пояснення роботи схеми наведемо часові діаграми її роботи на прикладі чотирирозрядного двійкового коду 1101 і 1001 (рис. 5.5).

Вхідна інформаційна двійкова комбінація 1101 надходить на послідовний вхід послідовного регістра U2 і виводиться в послідовному коді через елемент U7 на зовнішні пристрої відповідно до тактових імпульсів генератора U1 (рис. 5.5, а). Елементом, що визначає парність одиниць вхідної комбінації, є лічильний тригер U5. При парній кількості імпульсів, що надходять на його лічильний вхід, на прямому виході тригера встановиться рівень логічного нуля, при непарній – рівень логічної одиниці. Якщо діє вхідна комбінація 1101, то протягом перших чотирьох тактів вона пройде на зовнішні пристрої. П'ятий імпульс від генератора виділяється з послідовності імпульсів схемою U3, яка є подільником частоти на 5. Цим імпульсом знімається інформація про стан прямого входу тригера U5, передається через елемент U7 і додається до основної кодової комбінації.

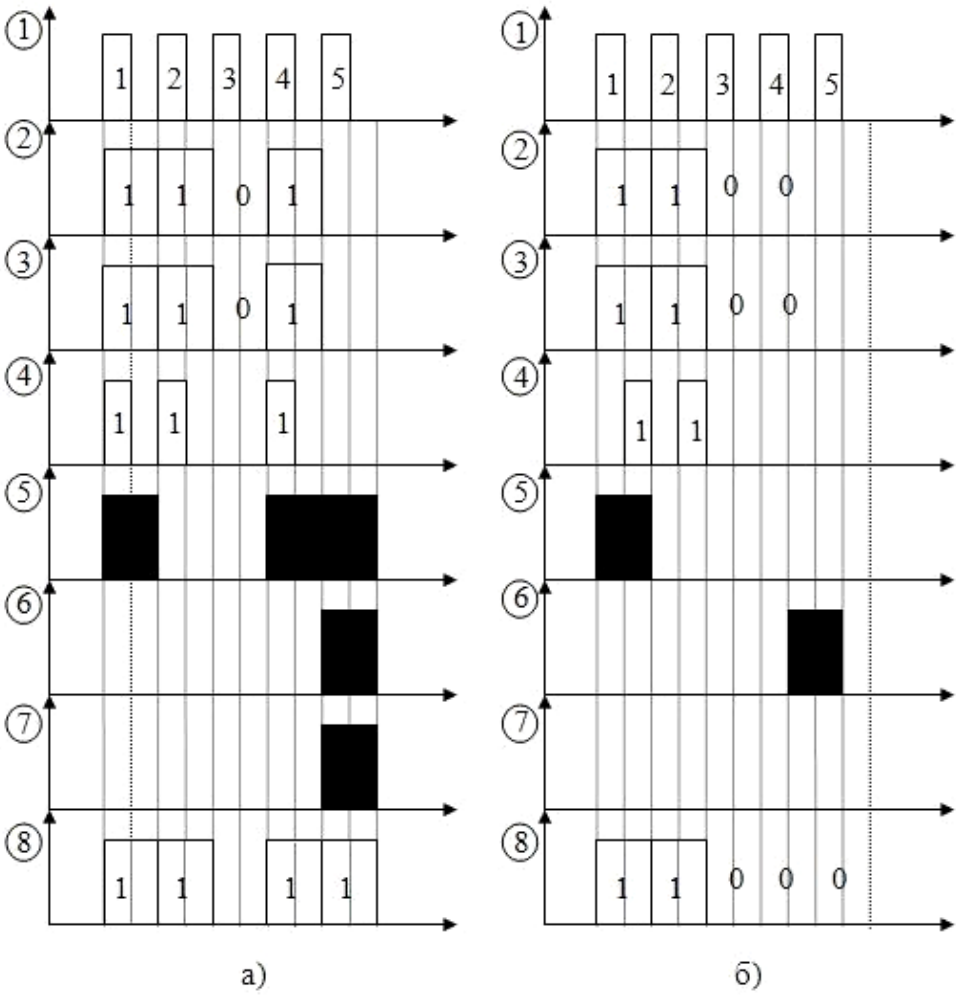


Рисунок 5.5 – Часові діаграми роботи кодера з перевіркою на парність при послідовному введенні інформації в регістр
а) для комбінації 1101; б) для комбінації 1001

По суті лічильний тригер виконує таку ж функцію як суматор за модулем два в схемі рис. 5.3.

Елемент U4 розділяє інформаційні одиниці вхідного коду паузою, якщо вони слідують одна за одною.

Декодер коду з перевіркою на парність

Декодером коду з перевіркою на парність називається електронний пристрій, який аналізує прийняту комбінацію на парність і, якщо вона порушена, забороняє її подальше перетворення. Якщо прийнята комбінація записана в послідовно-паралельному регістрі, то схема декодера має такий вигляд (рис. 5.6).

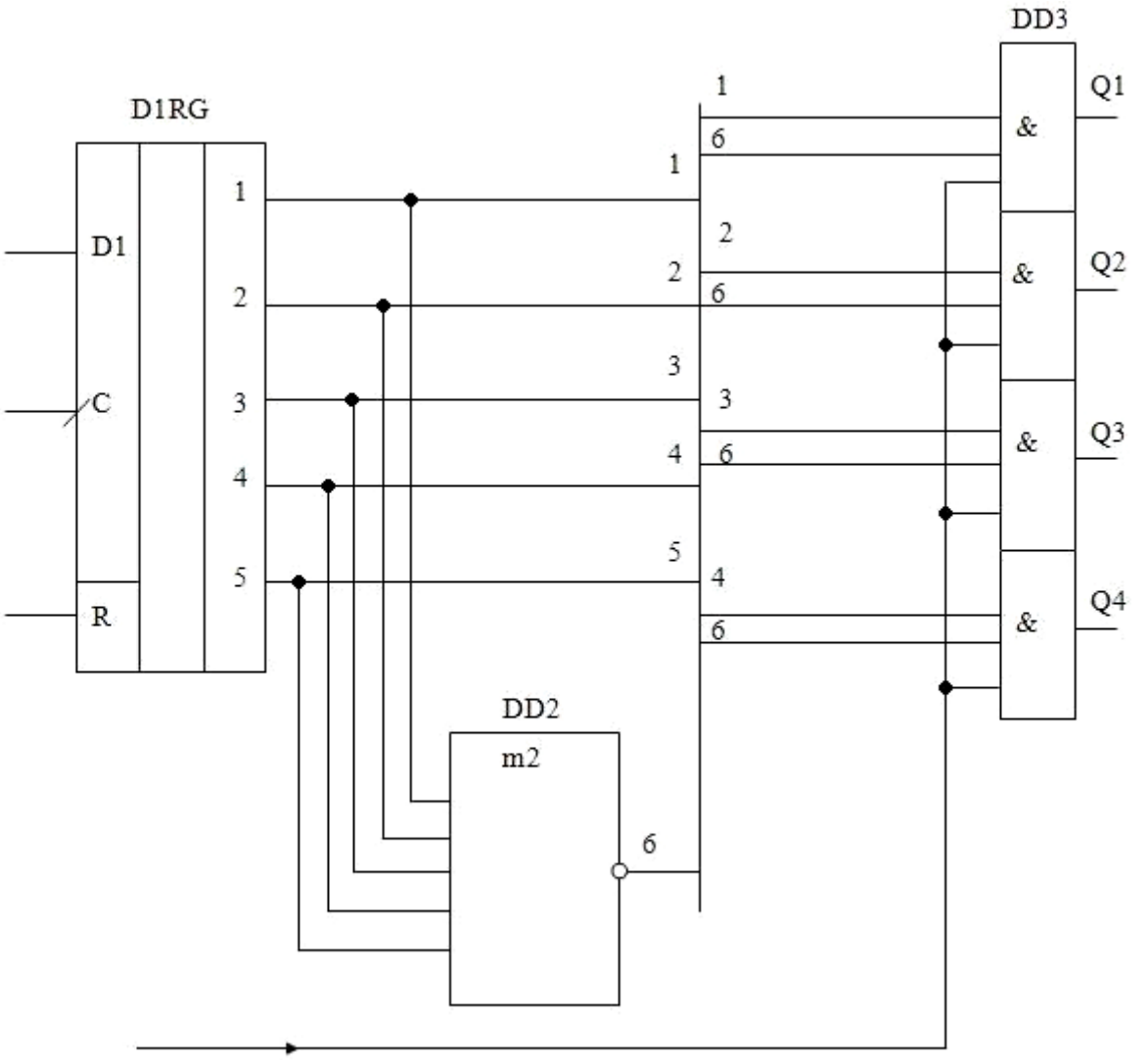


Рисунок 5.6 – Декодер коду з перевіркою на парність

Прийнята в послідовно-паралельний регістр комбінація коду з перевіркою на парність, надходить на суматор за модулем 2 DD2, аналізується, і в разі, якщо кількість одиниць парна, на інверсному виході суматора з'явиться рівень логічної одиниці, який поступає на всі елементи схеми DD3.

При наявності роздільного імпульсу, який надходить одночасно на всі елементи схеми DD3, прийнята комбінація з'являється на виходах Q1 - Q4 і проходить далі. Якщо парність порушена, на виході інверсного сумматора за модулем два m2 з'явиться рівень логічного нуля, який закриє всі елементи DD3 і прийнята комбінація далі не пройде.

5.4.2 Код з прямим повторенням

Існує ряд кодів, в яких для збільшення завадозахищенності до інформаційної кодової комбінації, яка складається з n_0 інформаційних символів двійкового безнадлишкового коду, додається k контрольних символів, причому $k = n_0$. Таким чином загальна кількість елементів в кодовій комбінації дорівнює $2n_0$. Тому такі коди називають *кодами з подвоєнням числа елементів*.

В залежності від способу утворення контрольних символів такі коди класифікують на:

а) коди з прямим подвоєнням;

б) коди з інверсним подвоєнням;

в) кореляційні коди.

Розглянемо більш детально схеми, що реалізовують вказані вище коди.

Кодер коду з прямим повторенням

Кодер коду з прямим повторенням - це пристрій, який автоматично добавляє до інформаційної кодової комбінації двійкового коду таку ж саму (повторює її), збільшуючи загальну кількість двійкових символів вдвічі.

На рис. 5.7 зображена схема такого кодера. Припустимо, що кількість інформаційних символів двійкового коду n_0=4.

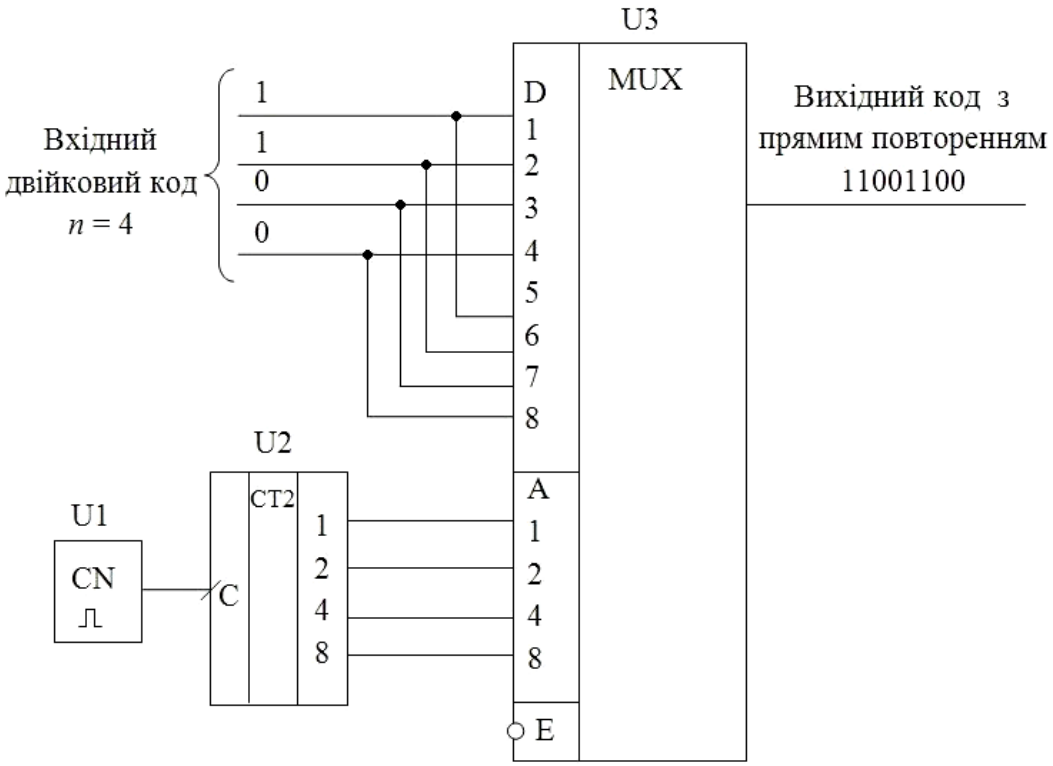


Рисунок 5.7 – Кодер коду з прямим повторенням

Вхідний код передається на входи D1-D4 мультиплексора U3. П’ятий - восьмий входи цього ж мультиплексора паралельно з'єднуються відповідно з першим - четвертим його входами. При надходженні імпульсів від генератора U1 на вхід двійкового лічильника U2 на його виходах утворюється двійковий (адресний код), який управляє перетворенням паралельного коду на входах мультиплексора у послідовний на його виході. Таким чином, якщо на входах мультиплексора подано код 1101, то за 8 тактів генератора на подальше перетворення зовнішніми пристроями надійде код 11011101.

Декодер коду з прямим повторенням

Декодер коду з прямим повторенням (рис. 5.8) порівнює інформаційні та контрольні символи комбінації, що надійшли на приймач (вони повинні повторювати одна одну) і дає команду на її подальше перетворення, якщо дія завади не виявлена.

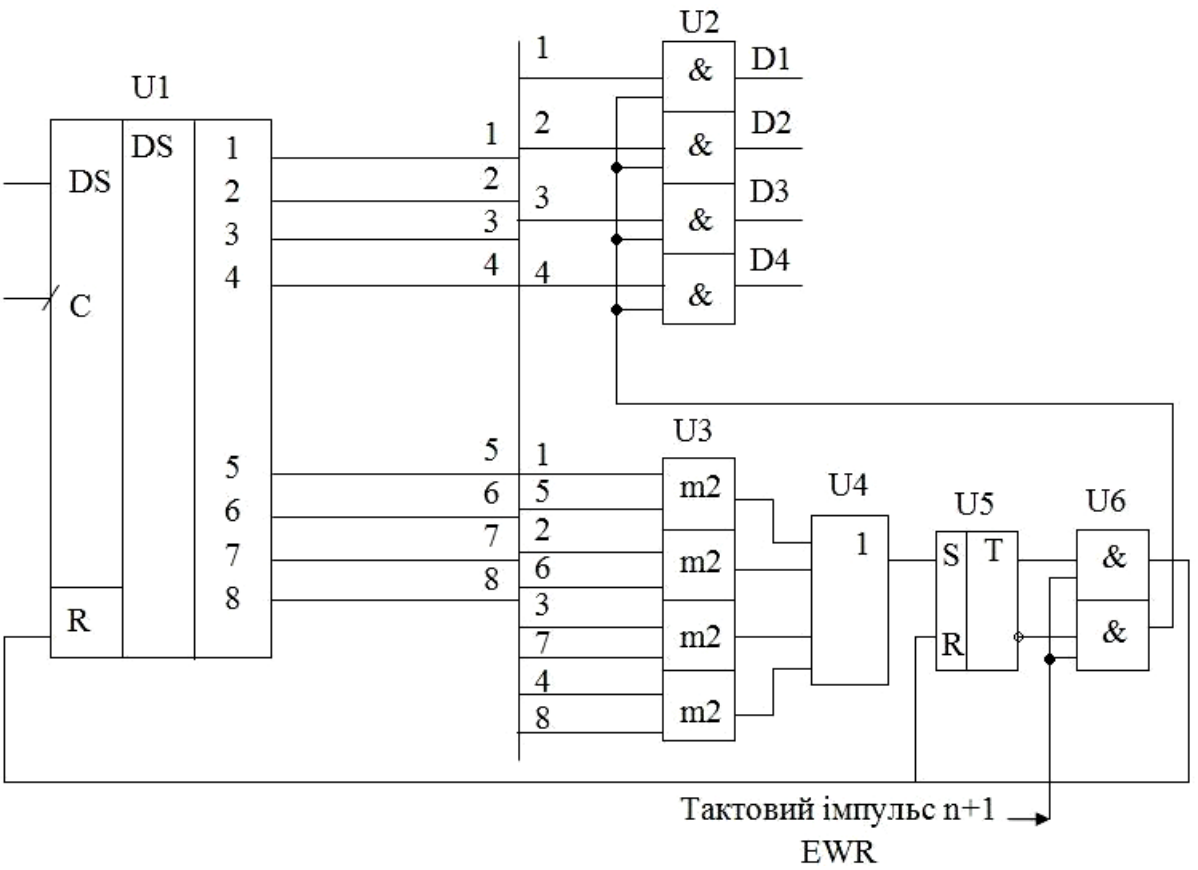


Рисунок 5.8 – Декодер коду з перевіркою з прямим повторенням

Після запису в послідовно-паралельний регістр U1 комбінація коду з прямим повторенням з'явиться на виходах 1-8 регістра. Комбінації на ви-

ходах 1-4 і 5-8 повторюють одна одну. Відповідні символи комбінацій n_0 (на виходах 1-4 регістра) порівнюються із символами $k = n_0$ (на виходах 5-8) за допомогою блока суматорів за модулем 2 – U3.

Якщо комбінація коду з прямим повторенням прийнята правильно на виходах всіх суматорів за модулем 2 буде рівень логічного нуля. На виході схеми «АБО» U4 також буде рівень логічного «0» і тригер U5 знаходиться у початковому стані (прямий вихід тригера має потенціал логічного «0». На виходах блока елементів U6 також зберігається рівень логічного «0», бо вони закриті нульовим потенціалом входу, який дозволяє зчитування інформації EWR. Імпульс на цьому вході з'являється після закінчення записування всієї комбінації в регістр. Якщо комбінація коду має *n* розрядів, то роздільний імпульс з'являється на *n+1* такті (в нашому прикладі на дев'ятому такті).

В момент надходження цього імпульсу логічні елементи блока U6 відкриваються, на вході нижнього елемента цього блока з'являється рівень логічної одиниці, який відкриває всі елементи «І» блока U2. Вихідна комбінація D1-D4 з виходів регістра 1-4 через елементи «І» блока U2 подається для подальшого перетворення.

Якщо внаслідок дії завад комбінація буде спотворена, на одному (чи декількох) виходах суматорів блока U3 з'явиться рівень логічної одиниці, який приведе тригер U5 в одиничний стан. 9-й тактовий імпульс на вході EWR відкриває верхній логічний елемент блока U6, на його виході з'явиться рівень логічної одиниці, який надійде на входи R регістра Ul і тригера U5 і скине їх у початковий стан (комбінація стирається). Елементи блока U2 при цьому будуть закриті і спотворена завадами комбінація на його виходах не з'явиться.

Звернемо увагу на те, що вихідний блок декодера U2 містить тільки чотири інформаційні розряди. Контрольні розряди виконали свою роль при аналізі правильності прийнятої комбінації і в подальшому не використовуються.

5.4.3 Коди з інверсним повторенням (інверсні коди)

Коди з інверсним повторенням інформаційної двійкової кодової комбінації утворюються за таким правилом: якщо кількість одиниць в інформаційній комбінації парна, то до неї автоматично додається така сама комбінація (пряме повторення) і якщо кількість одиниць в інформаційній

комбінації непарна, до неї додається інвертована інформаційна комбінація. Наприклад, якщо інформаційна кодова комбінація має вигляд 1100, то до неї додається така сама комбінація 1100 і в цілому для подальшого перетворення надійде комбінація 11001100.

Якщо інформаційна комбінація має вигляд 1110, то до неї буде додана комбінація 0001 і комбінація інверсного коду буде мати вигляд 11100001.

Кодер коду з інверсним повторенням

Кодером коду з інверсним повторенням називається електронний пристрій, який до інформаційної n_0-розрядної комбінації двійкового коду автоматично додає контрольну комбінацію з $k=n_0$ символів за таким алгоритмом: якщо кількість одиниць інформаційного коду парна, то до неї додається така ж сама комбінація, якщо непарна, до неї додається інвертована інформаційна комбінація. Схема такого кодера наведена на рис. 5.9.

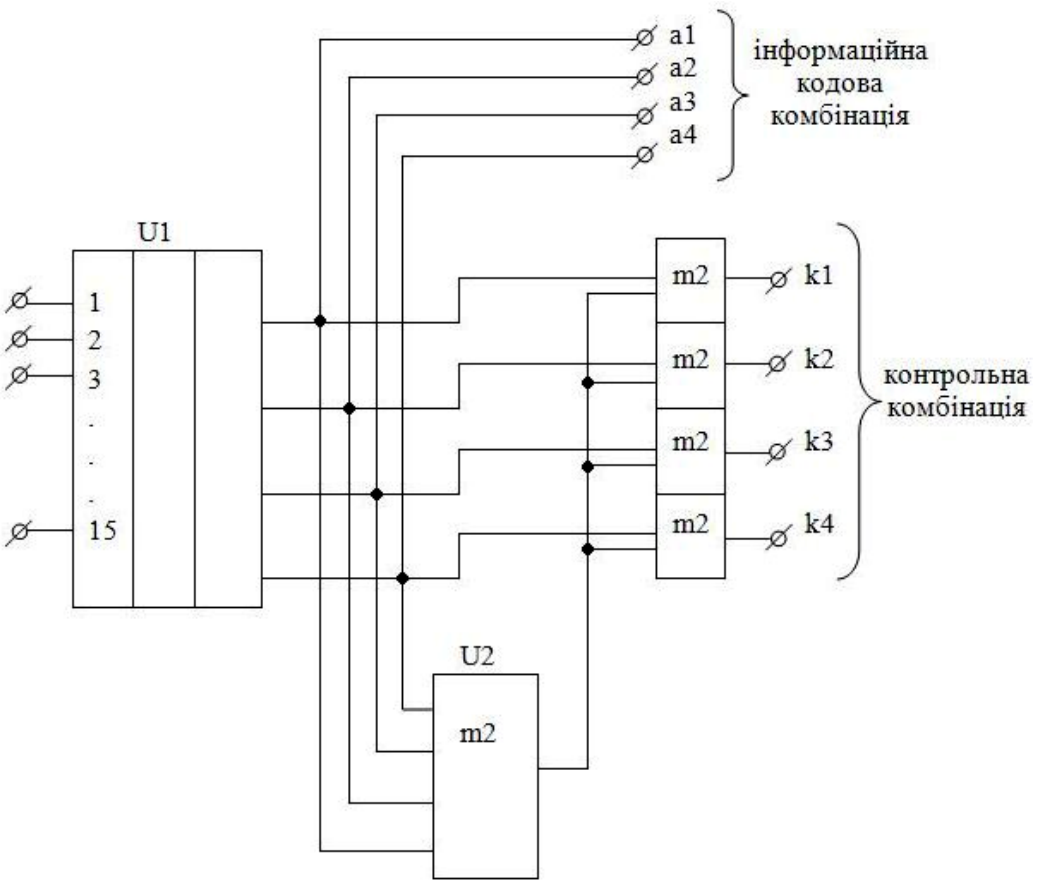


Рисунок 5.9 – Схема коду з інверсним повторенням

Інформаційна кодова комбінація з'являється на виходах кодера U1 при подачі на один з його 16 входів рівня логічної «1». Суматор за моду-

лем 2 U2 аналізує парність кількості одиниць в інформаційній кодовій комбінації.

Якщо ця кількість парна, на його виході з'явиться рівень логічного нуля і на виходи блока суматорів за модулем 2 U3 пройдуть всі символи, які встановились на виходах кодера U1. Якщо кількість одиниць інформаційної комбінації непарна, на виході елемента U2 з'явиться рівень логічної одиниці і на виходи блока U3 пройдуть інвертовані символи інформаційної кодової комбінації, яка встановилась на виходах кодера U1, тобто схема працює за встановленим алгоритмом.

Декодер коду з інверсним повторенням

Декодер коду з інверсним повторенням – пристрій, який за відомим алгоритмом утворення порівнює інформаційні та контрольні символи, що надійшли на приймальний пристрій, і дає команду на її подальше перетворення, якщо дія завади не виявлена.

Схема декодера інверсного коду подана на рис.5.10.

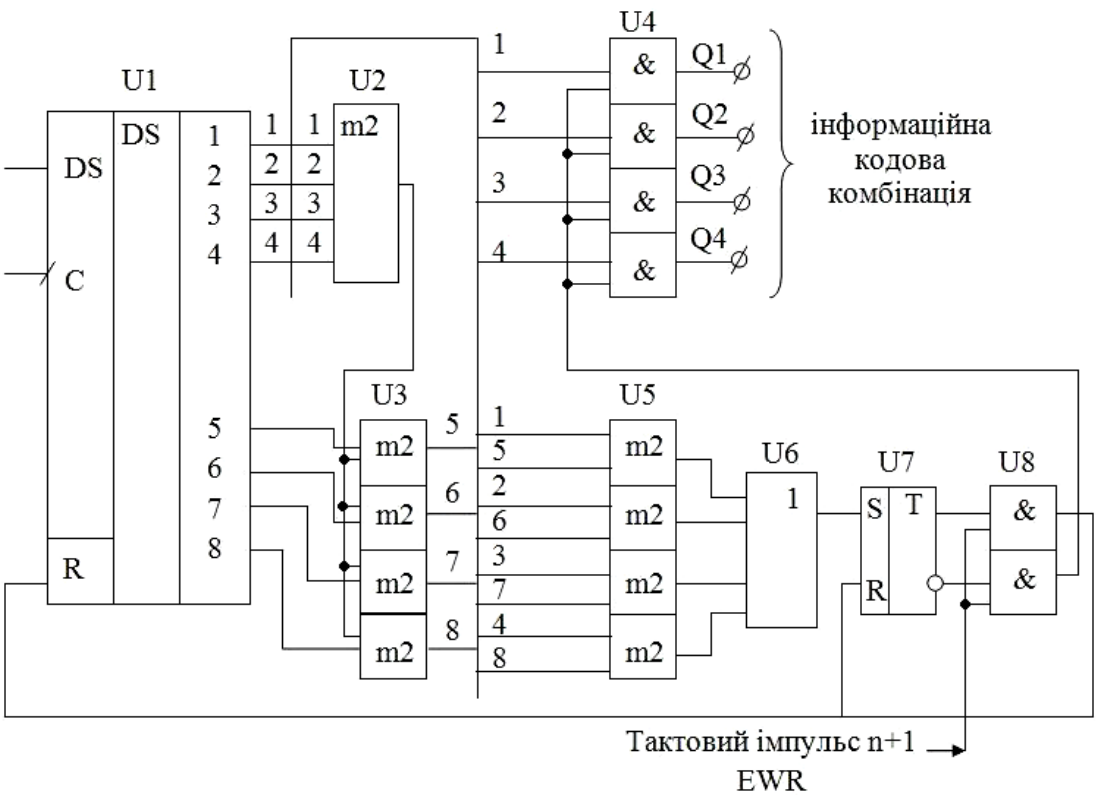


Рисунок 5.10 – Декодер інверсного коду

Декодер працює таким чином. Прийнята комбінація надходить на вхід DS послідовно-паралельного регістра U1 і після надходження на вхід

«С» *n* тактів фіксується на виходах 1-8 регістра. (На виходах 1-4 – інформаційна частина кодової комбінації, на виходах 5-8 - контрольна комбінація).

Суматор за модулем 2 U2 аналізує кількість одиниць в інформаційні частині кодової комбнації і дає команду на інвертувальний блок U3, в залежності від парності кількості одиниць в інформаційній комбінації, символи контрольної комбінації на виходах 5-8.

В результаті при правильному надходженні комбінації інверсного коду інформаційна комбінація з виходів регістра 1-4 і комбінація з виходів блока U3 повинні бути однаковими. Блок суматорів за модулем 2 U5 порівнює відповідні символи цих комбінацій і при правильному прийомі на виходах повинні бути рівні логічного «0». При наявності помилки в прийнятій комбінації на одному (чи декількох) виходах блока U5 з'явиться логічна «1», тригер U7 перейде в одиничний стан і після надходження на вхід EWR *n+1* тактового імпульсу через верхній логічний елемент блока U8 скидає (стирає) дані, які були записані в регістр Ul; і повертає тригер U7 в початковий стан. При правильному прийомі після надходження на вхід EWR *n+1* імпульсу (в нашому прикладі – 9-го) сигнал з нижнього елемента «І» блока U8 дає дозвіл на проходження інформаційної частини кодової комбінації інверсного коду через елементи «І» блока U4 для подальшого перетворення.

5.4.4 Кореляційний код

Кодером кореляційного коду є електронний пристрій, який при надходженні на його входи n_0 інформаційних символів двійкового коду автоматично додає до кожного інформаційного символу по одному контрольному "0" чи "1" і значення якого утворюється за правилом: інформаційний нуль перетворюється в 01, інформаційна одиниця перетворюється в 10. Таким чином загальна кількість символів кодової комбінації кореляційного коду $n = n_0 + k = 2n_0$. Тому цей код відносять до коду з подвоєнням числа елементів. Схема кодера кореляційного коду наведена на рис. 5.11, а часові діаграми його роботи на рис. 5.12.

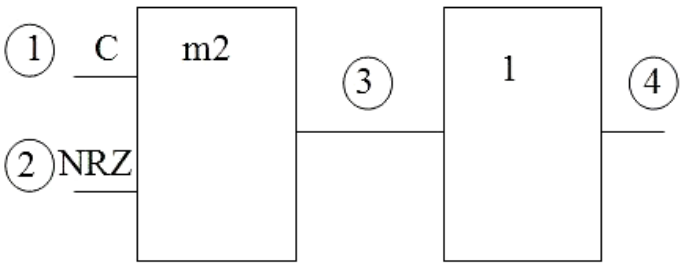


Рисунок 5.11 – Кодер кореляційного коду

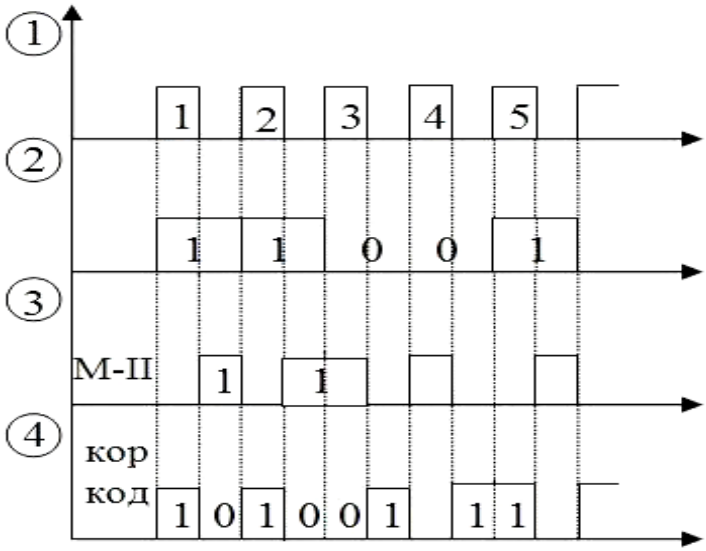


Рисунок 5.12 – Часові діаграми роботи кодера кореляційного коду

Декодер кореляційного коду

Декодером кореляційного коду є пристрій, який аналізує прийняту кодову комбінацію. На прийомі помилка визначається в тому випадку, коли в парних елементах містяться однакові символи, тобто 11 чи 00 (замість 10 та 01). При правильному прийомі парні елементи відкидаються і залишається початкова комбінація.

Схема декодера кореляційного коду наведена на рис. 5.13.

Для виявлення помилки прийнята кодова комбінація розбивається на групи по 2 символи в кожній і аналізується парність кожної групи.

В разі правильного прийому в кожній групі мають бути різні символи (01 чи 10). Якщо в будь-якій групі прийняті однакові символи (00 чи 11) – комбінація вважається спотвореною завадами і стирається, тому що в такому випадку помилка не виявляється.

Означений алгоритм виявлення помилки реалізовується таким чином.

Протягом N в послідовно-паралельний регістр U1 записується комбінація прийнятого кореляційного коду. Кожна пара символів (1-2, 3-4,... , [(N-1)- N]) порівнюється за допомогою суматорів за модулем 2 (блок U2) і в разі правильного прийому на всіх виходах суматорів будуть рівні логічної "1", які надходять на входи схеми "І" U3. З надходженням (N +1) тактового імпульсу від схеми управління на виході елемента U3 виникне імпульс, який дасть дозвіл на подальше перетворення інформаційної частини комбінації (символи Ul, U3, U5,..... U_{N-1}).

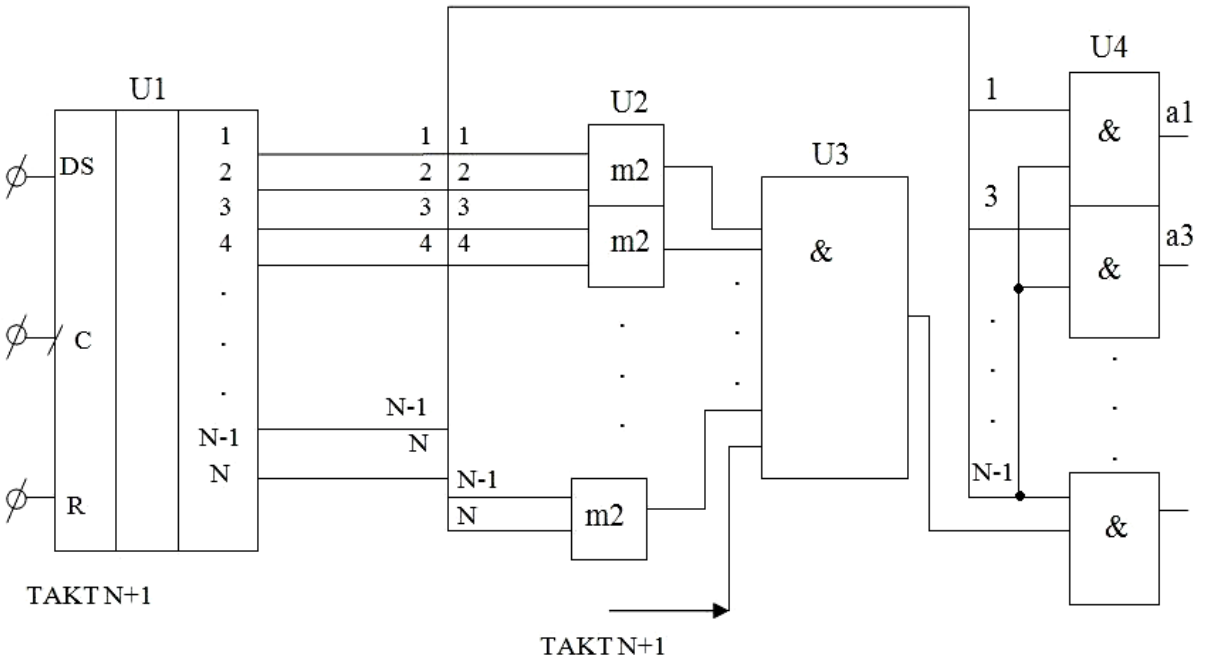


Рисунок 5.13 – Декодер кореляційного коду

В разі наявної помилки в будь-якій парі символів на виході елемента U3 значення логічного "0" не змінюється з приходом (N+l)-го такту і комбінація через блок елементів U4 не пройде.

З приходом (*N+2*) такту записана в регістр U1 комбнація кореляційного коду скидається.

На рис. 5.14 подана ще одна схема декодера кореляційного коду.

В цій схемі кожна пара символів кореляційного коду порівнюється між собою одним суматором за модулем 2 безпосередньо під час послідовного надходження символів коду на вхід послідовно-паралельного регістра U1, який має тільки два виходи. Елемент U3 -суматор за модулем 2 з інверсією порівнює кожну пару символів через кожних 2 такти.

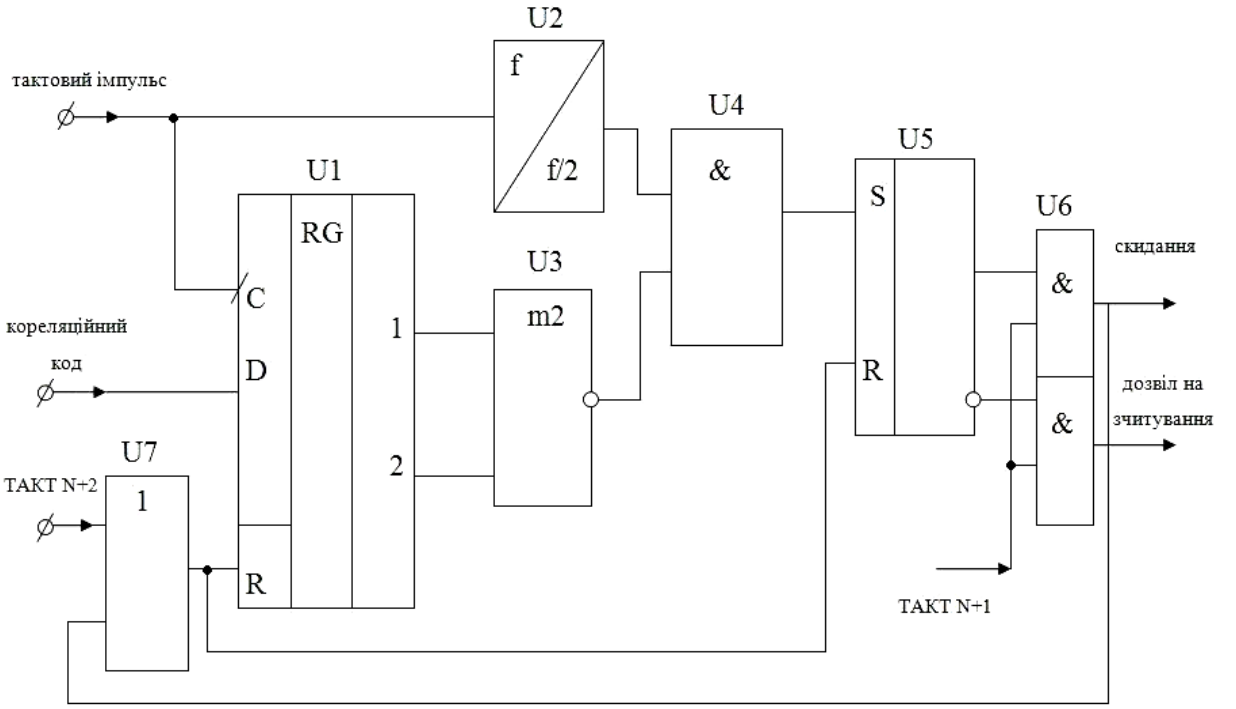


Рисунок 5.14 – Схема декодера кореляційного коду

Для того, щоб результат порівняння кожної пари символів здійснювався тільки на парних тактах вхідних імпульсів, дозвіл на передачу результату порівняння подається через подільник частоти тактових імпульсів U2 з коефіцієнтом підрахунку 2. В разі, коли у відповідний парі символів з'являться однакові символи, на виході елемента U3 з'явиться рівень логічної "1", який буде записаний в RS-тригер U5. Після закінчення прийому кодової комбінації кореляційного коду на (N +1) такті елементи схеми U6 дають дозвіл на зчитування інформації в разі правильного прийому або на скидання інформації при наявності помилки. З надходженням (N+2) такту регістр Ul і тригер U5 скидаються в обох випадках і починається аналіз нової кодової комбінації.

5.5 Коди з виявленням і виправленням помилок

До найбільш поширених кодів з виявленням та виправленням помилок відносять код Хемінга, циклічні та рекурентні коди. Розглянемо їх більш детально.

5.5.1 Код Хемінга

Код Хемінга відноситься до систематичних кодів, в яких з п символів, які утворюють комбінацію, n_0 символів є інформаційними, а останні

$k = n - n_0$ с надлишковими (контрольними), призначеними для перевірки (контрольні символи у всіх комбінаціях займають однакові позиції). Коди Хемінга дозволяють виправити всі одиничні помилки (при кодовій відстані d=3) і визначити всі подвійні помилки (при *d=4*), але не виправляти їх.

Зв'язок між кількістю інформаційних та контрольних символів в коді Хемінга знаходять на основі таких міркувань. При передачі комбінації по каналу з шумами може бути спотворений довільний з п символів коду, або комбінація може бути передана без спотворень. Таким чином може бути *n + 1* варіантів спотворення (включаючи передачу без спотворення). Використовуючи контрольні символи, необхідно перевірити всі *n+1* варіантів. За допомогою контрольних символів *k* можна описати *2k* подій. Для цього повинна бути використана умова:

$$2^k \geq n + 1 = n_0 + k + 1.$$

В таблиці 5.3 подана залежність між *k* і n_0, яка отримана з цієї невірності, де *k* - число контрольних символів в коді Хемінга, n_0 - інформаційних символів.

Таблиця 5.3 – Розміщення контрольних символів в комбінаціях коду Хемінга

n_0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
k	2	3	3	3	4	4	4	4	4	4	4	5	5	5	5

В коді Хемінга контрольні символи розташовують на місцях, кратних степеню числа 2, тобто на позиціях 1, 2, 4, 8 і т. п. Інформаційні символи розташовують на місцях, що залишилися. Наприклад, для семиелементної закодованої комбінації можна записати

k_1	k_2	a_{04}	k_3	a_{03}	a_{02}	a_{01}
a_1	a_2	a_3	a_4	a_5	a_6	a_7

Символи коду Хемінга, які обведені прямокутниками, є *контрольними*, останні – *інформаційні*, де a_3 – старший (четвертий) розряд вихідної кодової комбінації двійкового коду, який необхідно кодувати, a_7 – молодший (перший) розряд. Після розташування на відповідних місцях кодової комбінації контрольних і інформаційних символів в коді Хемінга склада-

ють спеціальні перевірні рівняння, які використовують для визначення наявності спотворень і їх виправлення. З перевірних рівнянь і отримують контрольні символи при кодуванні вихідної кодової комбінації двійкового коду. Для визначення контрольних символів необхідно використати такий алгоритм.

1. Всі символи коду Хемінга з номерами розрядів розташовують в порядку збільшення номерів і під ними записують номери розрядів в двійковому коді

a_1	a_2	a_3	a_4	a_5	a_6	a_7
0001	0010	0011	0100	0101	0110	0111

2. Перше перевірне рівняння складають як суму за mod 2 всіх розрядів, в номерах яких в молодшому розряді 2^0 стоїть одиниця: $S_1 = a_1 \oplus a_2 \oplus a_5 \oplus a_7$.

Друге перевірне рівняння складають як суму за mod 2 всіх розрядів, в номерах яких стоїть одиниця на другому місці відповідного двійкового еквівалента (2^1): $S_2 = a_2 \oplus a_3 \oplus a_6 \oplus a_7$.

Третє перевірне рівняння складають як суму за mod 2 всіх розрядів, в номерах яких стоїть одиниця на третьому місці (2^2): $S_3 = a_4 \oplus a_5 \oplus a_6 \oplus a_7$.

Аналогічно утворюються і інші перевірні суми (при більшій кількості інформаційних і контрольних символів, відповідно).

Як видно з наведених рівнянь, в кожну перевірну суму входить тільки один невизначений контрольний символ k_i (a_1, a_2, a_4, відповідно), а всі інші інформаційні символи відомі

Всі перевірні рівняння за умовою Хемінга повинні дорівнювати 0 при підсумовуванні за mod 2. З цієї умови і знаходять контрольні символи.

ПРИКЛАД

Наприклад, необхідно передати інформаційну кодову комбінацію:

a1	a2	a3	a4
1	1	0	0

з числом розрядів $n_0 = 4$.

З формули $2^k \geq n + 1 = n_0 + k + 1$ визначимо, що $k = 3$. Запишемо перевірні суми:

$$S_1 = a_1 + a_3 + a_5 + a_7;$$

$$S_2 = a_2 + a_3 + a_6 + a_7;$$

$$S_3 = a_4 + a_5 + a_6 + a_7.$$

Підставимо у рівняння значення відомих інформаційних символів. З умови рівності нулю всіх перевірних сум визначаємо відповідно контрольні символи. З першого рівняння $a_0 = 0$, з другого $a_1 = 1$, з третього $a_4 = 1$.

Відповідно буде передана така комбінація:

a_1	a_2	a_3	a_4	a_5	a_6	a_7
0	1	1	1	1	0	0,

яка є комбінацією коду Хемінга.

Алгоритм декодування коду Хемінга

На приймальному пристрої комбінація коду Хемінга декодується - визначається наявність спотворення прийнятої кодової комбінації і, якщо один символ комбінації спотворений, він автоматично спочатку визначається, а потім виправляється. Визначення і виправлення спотвореного символу здійснюється на основі перевірних рівнянь. При правильному прийомі всі суми повинні дорівнювати нулю. В разі спотворення двійкове число, яке є результатом цих сум (синдром помилки), переводять в десяткове число, яке вказує номер спотвореного символу, який в подальшому виправляється шляхом інвертування.

ПРИКЛАД

Наприклад, при прийомі закодованої вище неспотвореної кодової комбінації Хемінга:

a_1	a_2	a_3	a_4	a_5	a_6	a_7
0	1	1	1	1	0	0,

значення всіх символів комбінації підставляють на відповідні місця у перевірних сумах:

$$S_1 = a_1 + a_3 + a_5 + a_7 = 0;$$

$$S_2 = a_2 + a_3 + a_6 + a_7 = 0;$$

$$S_3 = a_4 + a_5 + a_6 + a_7 = 0.$$

Ознака правильно прийнятої комбінації – рівність нулю всіх сум.

Припустимо, що шостий символ кодової комбінації спотворений, тобто замість комбінації 0111100 буде прийнята:

a_1	a_2	a_3	a_4	a_6	a_7	a_8
0	1	1	1	1	1	0

Підставимо значення прийнятих символів у перевірні рівняння.

Одержимо: $S_1 = a_1 + a_3 + a_5 + a_7 = 0$;

$S_2 = a_2 + a_3 + a_6 + a_7 = 0$;

$S_3 = a_4 + a_5 + a_6 + a_7 = 0$.

Переведемо двійкове число (синдром) $1^{2^2}1^{2^1}1^{2^0}$ в десяткове. Одержимо десяткове число 6, яке і вказує на номер спотвореного символу.

Кодер коду Хемінга

На основі наведених вище правил будується кодер коду Хемінга (рис. 5.15). Він автоматично визначає значення контрольних символів коду при відомих інформаційних, які складаються з безнадлишкових комбінацій звичайного двійкового коду.

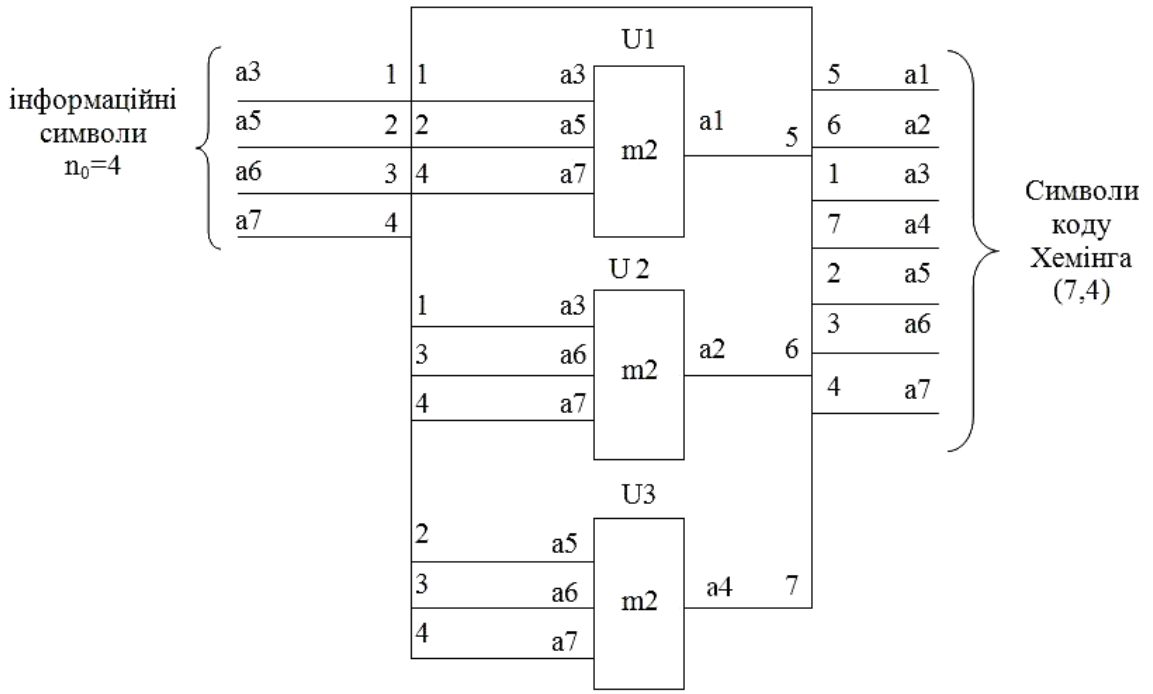


Рисунок 5.15 – Кодер коду Хемінга

На входи суматорів за mod 2 Ul, U2, U3 подаються інформаційні символи відповідно до перевірних сум. На вході суматорів одержуються контрольні символи a_1, a_2, a_4. Інформаційні і контрольні символи розташо-

вуються в необхідному порядку і через шину виводяться для подальшого перетворення.

Декодер коду Хемінга

Декодер коду Хемінга (рис. 5.16) аналізує прийняту кодову комбінацію і в разі спотворення будь-якого одного символу (інформаційного чи контрольного) автоматично виправляє спотворений символ.

Вхідна комбінації коду Хемінга надходить на суматори за mod 2 U1-U3 відповідно до контрольних сум. На виходах суматорів утворюється результат контрольних сум S1-S3. При правильному прийомі суми SI, S2, S3 повинні дорівнювати нулю.

В разі спотворення одного символу на виходах елементів U1-U3 з'явиться двійкова кодова комбінація (синдром помилки), яка декодується декодером U4. На одному з виходів декодера з'являється рівень логічної одиниці, який відповідає номеру спотвореного символу.

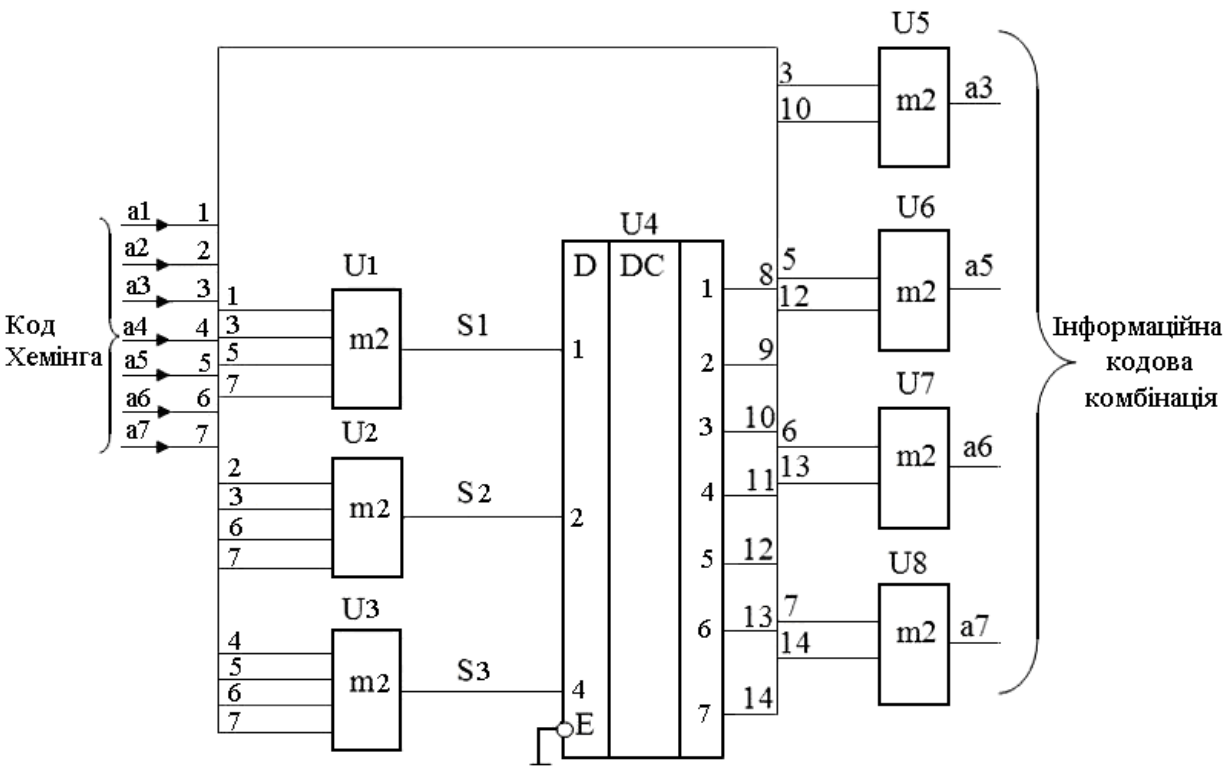


Рисунок 5.16 – Декодер коду Хемінга для семиелементної комбінації

На суматори за mod 2 U5-U8 надходять відповідні сигнали з входів a_3, a_5, a_6, a_7 (інформаційні символи) і відповідний вихід декодера.

При правильному прийомі на виходах 1-7 декодера - логічні нулі і на виходах суматорів U5-U8 з'являться інформаційні символи а3, а5 , а6 , а7 без змін.

В разі спотворення одного з символів, наприклад, третього, на третьому виході декодера з'явиться рівень логічної одиниці і на виході сумматора за mod2 U5 спотворений символ автоматично інвертується, перетворюючись у правильний.

5.5.2 Циклічні коди

Циклічні коди одержали досить широке застосування завдяки їхній ефективності при виявленні і виправленні помилок. Схеми кодувальних і декодувальних пристроїв для цих кодів надзвичайно прості і будуються на основі звичайних регістрів зсуву.

Назва кодів пішла від їх властивості, яка полягає в тому, що кожна кодова комбінація може бути отримана шляхом циклічної перестановки символів комбінації, що належить до цього ж коду. Це значить, що якщо, наприклад, комбінація $a^0a^1a^2...a^{n-1}$ є дозволеною комбінацією циклічного коду, то комбінація $a^{n-1}a^0a^1a^2...a^{n-2}$ також належить цьому кодові.

Циклічні коди зручно розглядати, подаючи комбінацію двійкового коду не у вигляді послідовностей нулів і одиниць, а у вигляді полінома від фіктивної змінної x

$$G(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + ... + a_1x + a_0, \qquad (5.5)$$

де a_i – цифри даної системи числення (у двійковій системі 0 і 1).

Найбільший степінь x з ненульовим коефіцієнтом a називається степенем полінома.

Подання кодових комбінацій у формі (5.5) дозволяє звести дії над комбінаціями до дії над многочленами. При цьому додавання двійкових многочленів зводиться до додавання за модулем два коефіцієнтів при рівних степенях змінної х; множення та ділення здійснюється за звичайними правилами множення та ділення логічних функцій, отримані при цьому коефіцієнти при рівних степенях змінної х додаються за модулем два.

Подання комбінацій за формулою (5.5) зручно ще і тим, що згадана раніше циклічна перестановка є результат простого множення даного полінома на х. Дійсно, якщо одна з кодових комбінацій виражається поліномом

$$V(x) = = a_0 + a_1x + a_2x^2 + ... a_{n-1}x^{n-1},$$

то нова комбінація за рахунок циклічного зсуву буде

$$x \cdot V(x) == a_0x + a_1x^2 + a_2x^3 + \ldots + a_{n-2}x^{n-1} + a_{n-1}x^n.$$

Однак в останньому члені необхідно замінити xn на 1. Отже, нова комбінація буде

$$\boldsymbol{V}^1(\boldsymbol{x}) = \boldsymbol{a}_{\boldsymbol{n}-1} + \boldsymbol{a}_0\boldsymbol{x} + \boldsymbol{a}_1\boldsymbol{x}^2 + \ldots + \boldsymbol{a}_{\boldsymbol{n}-2}\boldsymbol{x}^{\boldsymbol{n}-1}.$$

Відповідно до визначення циклічного коду для побудови породжувальної матриці $P_{n,\,k}$ досить вибрати тільки одну вихідну n-розрядну комбінацію $V_i(x)$. Циклічним зсувом можна одержати (n - 1) різних комбінацій, із котрих будь-які k комбінацій можуть бути взяті як вихідні. Підсумовуючи рядки породжувальної матриці у всіх можливих комбінаціях, можна одержати інші кодові комбінації. Можна показати, що кодові комбінації, одержані з деякої комбінації $V_i\,(x)$ циклічним зсувом, задовольняють умови, запропоновані до сукупності вихідних комбінацій.

Циклічний зсув комбінації з одиницею в старшому n-му розряді тотожний множенню відповідного многочлена на x з одночасним відніманням від результату многочлена (x^n- 1) або (x^n + 1), тому що операції здійснюються за модулем два. Отже, якщо як вихідний взяти деякий поліном $P(x)$, то процес одержання базових поліномів можна подати в такому вигляді:

$$\begin{aligned}
&\boldsymbol{U}_1(\boldsymbol{x}) = \boldsymbol{P}(\boldsymbol{x});\\
&\boldsymbol{U}_2(\boldsymbol{x}) = \boldsymbol{P}(\boldsymbol{x})\boldsymbol{x} - \boldsymbol{C}_2(\boldsymbol{x}^{\boldsymbol{n}} + 1);\\
&\boldsymbol{U}_3(\boldsymbol{x}) = \boldsymbol{P}(\boldsymbol{x})\boldsymbol{x}^2 - \boldsymbol{C}_3(\boldsymbol{x}^{\boldsymbol{n}} + 1);\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\boldsymbol{U}_{\boldsymbol{n}}(\boldsymbol{x}) = \boldsymbol{P}(\boldsymbol{x})\boldsymbol{x}^{\boldsymbol{n}-1} - \boldsymbol{C}_{\boldsymbol{n}}(\boldsymbol{x}^{\boldsymbol{n}} + 1),
\end{aligned}$$

де C_2, C_3, ..., C_n – коефіцієнти, що приймають значення 1 при $P(x)\cdot x^i \geq (x^n\text{-}1)$ і значення 0 при $P(x)\cdot x^i < (x^n\text{-}1)$.

При такому способі побудови базових поліномів поліном $P(x)$ називають утворюючим.

Якщо прийняти умову, що поліном $P(x)$ є дільником двочлена (x^n + 1), то базові комбінації, а разом із ними і всі дозволені комбінації коду, набувають властивості подільності на $P(x)$. З цього випливає, що належність кодової комбінації до групи дозволених можна легко перевірити розподілом її полінома на утворюючий поліном $P(x)$. Якщо залишок від розподілу дорівнює нулю, то комбінація є дозволеною.

Ця властивість циклічного коду використовується для виявлення або виправлення помилок. Дійсно, якщо під впливом завад дозволена ко-

дова комбінація трансформується в заборонену, то помилка може бути виявлена за наявністю залишку при розподілі комбінації на утворюючий поліном *P(x)*.

Таким чином утворюючий поліном *P(x)* повинен задовольняти вимогу – він повинен бути дільником двочлена (x^n+1). Вибір *P(x)* однозначно визначає циклічний код і його коригувальні властивості.

Циклічний (*n*, А)-код може бути отриманий шляхом множення простого А-значного коду, вираженого у вигляді полінома степеня (*k* - 1), на деякий утворюючий поліном *P(x)* степеня (*n* - *k*).

Можлива й інша процедура одержання циклічного коду. Для цього кодова комбінація простого *k*-значного коду *G(x)* збільшується на одночлен x^{n-k}, а потім ділиться на утворюючий поліном *P(x)* степеня (*n*-*k*). У результаті множення *G(x)* на x^{n-k} степінь кожного одночлена, що входить у *G(x)*, підвищиться на (*n* - *k*). При діленні добутку $x^{n-k}G(x)$ на утворюючий поліном *P(x)* утвориться частка *Q(x)* такого ж степеня, як і *G(x)*.

Результат множення і ділення можна подати в такому вигляді:

$$\frac{\boldsymbol{x}^{n-k}\boldsymbol{G}(\boldsymbol{x})}{\boldsymbol{P}(\boldsymbol{x})} = \boldsymbol{Q}(\boldsymbol{x}) + \frac{\boldsymbol{R}(\boldsymbol{x})}{\boldsymbol{P}(\boldsymbol{x})},$$

де *R(x)* – залишок від ділення $x^{n-k}G(x)$ на *P(x)*.

Оскільки частка *Q(x)* має такий ж степінь, як і кодова комбінація *G(x)*, то *Q(x)* також є комбінацією простого *k*-значного коду.

Таким чином, кодова комбінація циклічного (*n*, *k*)-коду може бути отримана двома способами:

1) шляхом множення простої кодової комбінації степеня (*k* - 1) на одночлен x^{n-k} і додавання до цього добутку залишку, отриманого від ділення отриманого добутку на утворюючий поліном *P(x)* степеня (*n* - *k*),

2) шляхом множення простої кодової комбінації степеня (*k* - 1) на утворюючий поліном *P(x)* степеня (*n* – *k*).

Згідно з першим способом кодування перших *k* символів отриманої кодової комбінації збігаються з відповідними символами вихідної простої кодової комбінації.

Згідно з другим способом в отриманій кодовій комбінації інформаційні символи не завжди збігаються із символами вихідної простої комбінації. Такий спосіб легко реалізовується, але внаслідок того, що в отриманих кодових комбінаціях не містяться інформаційні символи в явному вигляді, ускладнюється процес декодування.

У зв'язку з вищевикладеним на практиці звичайно використовується перший спосіб одержання циклічного коду.

Матричне подання циклічних кодів

Для формування рядків породжувальної матриці за першим способом утворення циклічного коду беруть комбінації простого k-розрядного коду *G(x)*, що містять одиницю в одному розряді. Ці комбінації збільшуються на x^{n-k}, визначається залишок *R(x)* від ділення отриманого добутку $x^{n-k}G(x)$ на утворюючий поліном і записується відповідний рядок матриці у вигляді суми добутку $x^{n-k}G(x)$ і залишку *R(x)*. При цьому породжувальна матриця $P_{n,k}$ подається двома підматрицями – інформаційною U_k і додатковою H_p:

$$\boldsymbol{P}_{n,k} = \left|\boldsymbol{U}_k, \boldsymbol{H}_\rho\right|.$$

Інформаційна підматриця U_k, являє собою квадратну одиничну матрицю з кількістю рядків і стовпців, рівною *k*. Додаткова підматриця H_p містить $p = n - k$ стовпців і *k* рядків і утворена залишками *R(x)*.

Породжувальна матриця дозволяє одержати *k* комбінацій коду. Інші комбінації утворюються підсумовуванням за модулем два рядків породжувальної матриці у всіх можливих поєднаннях.

При другому способі утворення циклічного коду породжувальна матриця $P_{n,k}$ формується шляхом множення утворюючого полінома *P(x)* степеня $\rho = n - k$ на одночлен x^{k-1} наступних *k-1* зсувів отриманої комбінації.

Вибір утворюючого полінома

При побудові циклічного коду спочатку визначається число інформаційних розрядів *k* за заданим об’ємом коду. Потім знаходиться найменша довжина кодових комбінацій *n*, що забезпечує виявлення або виправлення помилок заданої кратності. Ця проблема зводиться до перебування потрібного утворюючого полінома *P(x)*.

Як уже відзначалося раніше, степінь утворюючого полінома повинен дорівнювати числу перевірних розрядів ρ.

Оскільки в циклічному коді розпізнювачами помилок є залишки від ділення многочлена прийнятої комбінації на утворюючий коректувальний многочлен, то спроможність коду буде тим вища, чим більше залишків може бути утворено в результаті цього ділення.

Найбільше число залишків, що дорівнює 2^{ρ}-1 (крім нульового), може забезпечити тільки многочлен порядка степеня ρ, що не ділиться ні на який інший многочлен.

Відомо, що двочлен типу $(x^n + 1) = (x^{2z-1} + 1)$, у розкладанні якого як співмножник повинен входити утворюючий многочлен, має ту властивість, що він є спільним кратним для усіх без винятку незвідних поліномів степеня n і розкладається на множники з усіх незвідних поліномів степеня 2, які діляться без залишку на число z.

Найпростішим циклічним кодом є код, що забезпечує виявлення однократних помилок. Вектору однократної помилки відповідає одночлен x^i, степінь котрого i може приймати значення від 1 до n. Для того щоб могла бути виявлена помилка, одночлен x^i не повинен ділитися на утворюючий поліном $P(x)$. Серед незвідних многочленів, що входять у розкладання двочлена $x^n + 1$, є багаточлен найменшого степеня $x + 1$. Таким чином утворюючим поліномом даного коду є двочлен $P(x) = x + 1$. Залишок від ділення будь-якого многочлена на $x + 1$ може приймати тільки два значення: 0 і 1. Отже, при будь-якому числі інформаційних розрядів необхідний тільки один перевірний розряд. Значення символу цього розряду забезпечує парність числа одиниць у кодовій комбінації.

Даний циклічний код із перевіркою на парність забезпечує виявлення не тільки однократних помилок, але і всіх помилок непарної кратності.

Для побудови циклічного коду, що виправляє однократні або виявляє дворазові помилки, необхідно, щоб кожній одиничній помилці відповідав свій розпізнавач, тобто залишок від ділення многочлена прийнятої комбінації на утворюючий многочлен. Оскільки кількість можливих однократних помилок дорівнює n, а незвідний многочлен степеня ρ може дати 2^{ρ} - 1 ненульових залишків, то необхідною умовою виправлення будь-якої одиничної помилки є виконання нерівності

$$2^{\rho} - 1 \geq n .$$

Звідси знаходять степінь утворюючого полінома

$$\rho = n - k \geq \log_2 (n + 1) \qquad (5.6)$$

і загальну довжину n кодової комбінації.

Оскільки утворюючий многочлен $P(x)$ повинен входити як співмножник в розкладання двочлена $(x^n + 1) = (x^{2z-1} + 1)$, то використовуючи

відзначені раніше властивості цього двочлена, а також умову (5.6), можна вибрати утворюючий поліном.

Однак не всякий мнгочлен степеня ρ, що входить у розкладання двочлена $x^n + 1$, може бути використаний як утворюючий поліном. Необхідно, щоб для кожної із n однократних помилок забезпечувався свій, відмінний від інших, залишок від розподілу прийнятої кодової комбінації на утворюючий поліном. Це буде мати місце за умови, якщо обраний незвідний многочлен степеня ρ, будучи дільником двочлена $x^n + 1$, не входить у розкладання ніякого іншого двочлена $x^i + 1$, степінь якого $i < n$.

Утворюючі поліноми кодів, які здатні виправляти помилки будь-якої кратності, можна визначати, користуючись таким правилом Хеммінга.

1. За заданим числом інформаційних розрядів k визначається число перевірних розрядів ρ, необхідне для виправлення однократних помилок, і знаходиться утворюючий поліном.

2. Розглядаючи отриманий (n, k)-код і некоректуючий n-розрядний код визначають додаткові розряди для забезпечення виправлення однієї помилки в цьому коді і знаходять відповідний утворюючий поліном.

3. Повторюється дана процедура стільки разів, поки не буде отриманий код, що виправляє незалежні помилки до даної кратності включно.

Закодувати просту інформаційну групу $G(x) = 1011$ циклічним кодом, що забезпечує виявлення дворазових або усунення однократних помилок.

ПРИКЛАД

Розв'язання. За заданою кількістю інформаційних символів $k = 4$ визначаємо значність коду. Користуючись співвідношенням (5.6) і табл. 5.1, отримаємо $n = 3{,}2$.

Для побудови циклічного коду необхідно вибрати утворюючий поліном $P(x)$ степеня $\rho = n - k = 3$, що повинен входити як співмножник, а для розкладання двочлена $x^n + 1 = x^{2z-1} - 1$. В нашому випадку цей двочлен має вигляд x^7+1. Складові його співмножники повинні бути незвідними поліномами, степені яких є дільниками числа $z = 3$. Отже, співмножниками двочлена $x^7 + 1$ повинні бути незвідні поліноми першого і третього степенів. Користуючись таблицями незвідних поліномів отримаємо

$$x^7 + 1 = (x + 1)(X^3 + x + 1)(x^3 + x^2 + 1).$$

Вибираємо як утворюючий поліном співмножник

$$\boldsymbol{P(x) = x^3 + x^2 + 1}.$$

Кодування здійснюємо першим способом. Для цього вихідну кодову комбінацію *G(x)* множимо на $x^{n-k}=x^3$

$$\boldsymbol{x^{n-k}G(x) = x^3(x^3 + x + 1) = x^6 + x^4 + x^3}.$$

Визначаємо залишок *R(x)* від розподілу x^{n-k} *G(x)* на утворюючий багаточлен *P(x)*:

$$\begin{array}{r|l} x^6 + x^4 + x^3 & x^3 + x^2 + 1 \\ \hline + \; x^6 + x^5 + x^3 & x^3 + x^2 \\ \hline x^5 + x^4 & \\ + \; x^5 + x^4 + x^2 & \\ \hline x^2 & \end{array}.$$

Залишок $R(x) = x^2$.

Отже, поліном *F(x)* циклічного коду буде мати вигляд

$$\boldsymbol{F(x) = x^{n-k}G(x) + R(x) = x^6 + x^4 + x^3 + x^2 = 1011100}.$$

ПРИКЛАД

Отримане повідомлення циклічним кодом $P^*(x) = x^6 + x^4 + x^3 + x^2$. Перевірити декодуванням наявність помилок у прийнятій комбінації, якщо утворюючий поліном $P(x) == x^3 + x^2 + 1$.

Розв'язання. Декодування здійснюється діленням полінома отриманої комбінації на утворюючий поліном

$$\begin{array}{r|l} x^6 + x^4 + x^3 + x^2 & x^3 + x^2 + 1 \\ \hline + \; x^6 + x^5 + x^3 & x^3 + x^2 \\ \hline x^5 + x^4 + x^2 & \\ + \; x^5 + x^4 + x^2 & \\ \hline 0 & \end{array}.$$

Залишок від розподілу $R(x) = 0$. Отже, комбінація прийнята без перекручувань.

ПРИКЛАД

Отримана комбінація $P^*(x) = x^6 + x^4 + x^2 + 1$, закодована циклічним кодом. Утворюючий поліном $P(x) = x3 + x2 + 1$. Перевірити наявність помилок у кодовій комбінації.

Розв'язання. Ділимо поліном отриманої комбінації на утворюючий поліном

$$
\begin{array}{r|l}
+\begin{array}{l} x^6 + x^4 + x^2 + 1 \\ x^6 + x^5 + x^3 \end{array} & \begin{array}{l} x^3 + x^2 + 1 \\ \hline x^3 + x^2 + 1 \end{array} \\
\hline
+\begin{array}{l} x^5 + x^4 + x^3 + x^2 + 1 \\ x^5 + x^4 + x^2 \end{array} & \\
\hline
+\begin{array}{l} x^3 + 1 \\ x^3 + x^2 + 1 \end{array} & \\
\hline
x^2 &
\end{array}
.
$$

залишок – $R(x) = x^2 \neq 0$. Отже, комбінація прийнята з помилками.

ПРИКЛАД

Побудувати породжувальну матрицю циклічного коду при $n = 7$; $k = 4$. Утворюючий поліном $P(x) = x^2 + x + 1$.

Розв'язання. Оскільки $k = 4$, інформаційна підматриця має вигляд

$$
\boldsymbol{U}_k = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix}.
$$

Для одержання першого рядка додаткової підматриці множимо перший рядок інформаційної підматриці на x^3. Ділимо на утворюючий поліном, тобто виконуємо операції $\dfrac{0001 \cdot 1000}{1011}$. Залишок від цих операцій 011 складає перший рядок додаткової підматриці. Аналогічно визначаємо інші рядки додаткової підматриці. Додаткова підматриця має вигляд:

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

Остаточно одержимо таку породжувальну матрицю:

$$
\boldsymbol{P}_{n,k} = \left| \boldsymbol{U}_k, \boldsymbol{H}_\rho \right| = \begin{vmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{vmatrix}.
$$

Для побудови породжувальної матриці другим способом перший рядок матриці одержуємо шляхом множення утворюючого полінома на

x^{k-1}: $P(x)\cdot x^{k-1}$=1011·1000 = 1011000. Наступні рядки матриці отримаємо шляхом циклічного зміщення отриманої комбінації. Таким чином породжувальна матриця

$$\boldsymbol{P}_{n,k} = \begin{vmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{vmatrix}.$$

5.5.3 Рекурентні коди

Циклічні коди дозволяють виявляти і виправляти як одиночні і подвійні помилки, так і пачки помилок. Однак практичне застосування цих кодів для виправлення пачок помилок є не складним тим, що при не дуже великій надмірності довжина кодових комбінацій значно перевищує довжину пачок. У цьому відношенні більш зручними є рекурентні коди.

Рекурентні (неперервні) коди відрізняються тим, що в них операції кодування і декодування робляться неперервно над послідовністю символів без розподілу їх на блоки.

Ці коди призначені в основному для виправлення пачок помилок. Формування перевірних символів здійснюється шляхом додавання двох або більше інформаційних символів, зміщених один відносно одного на визначену відстань *t*, звану кроком додавання.

Довжина, що виправляється рекурентним кодом пачки помилок *l* залежить від кроку додавання *t* і визначається з умови

$$\boldsymbol{l} = 2\boldsymbol{t}\,.$$

Мінімальна необхідна відстань між пачками помилок, при якій забезпечується виправлення всіх помилок у пачці довжиною *l*, дорівнює

$$\boldsymbol{a} = 6\boldsymbol{t} + 1\,.$$

З усіх рекурентних кодів найбільше поширення одержав так званий ланцюговий код, що відрізняється гранично простими методами кодування і декодування.

У ланцюговому коді кожний перевірний символ формується шляхом додавання за модулем два двох інформаційних символів, що віддалені один від одного на крок додавання *t*.

Позначаючи послідовність інформаційних символів через $a_0a_1a_2...a_{2t}a_{2t+1}...$, отримаємо таку послідовність перевірних символів для ланцюгового коду: $b_0=a_0+a_t$; $b_1=a_1+a_{t+1}$; .. $b_t=a_t+a_{2t}$; $b_{t+1}=a_{t+1}+a_{2t+1}$... У загально-

му потоці символів ланцюгового коду між кожними двома інформаційними символами міститься один перевірний

$$a_0b_0a_1b_1a_2b_2...a_tb_ta_{t+1}b_{t+1}...a_{2t}b_{2t}a_{2t+1}b_{2t+1}...$$

Тому, що число перевірних символів, сформованих за деякий час, дорівнює числу інформаційних символів, що надійшли за той же час, то надмірність ланцюгового коду дорівнює 0,5.

На прийомі інформаційні і перевірні символи розділяються і реєструються незалежно один від одного. З прийнятої послідовності інформаційних символів формуються контрольні символи так само, як формувалися перевірні символи при передачі. Після затримки на величину (3t + 1) кожний сформований контрольний символ звірюється з відповідним прийнятим перевірним символом. У випадку перекручування одного з перевірних символів буде розбіжність відповідних контрольних і перевірних символів. Якщо ж перекручений один з інформаційних символів, то буде розбіжність двох контрольних і відповідних перевірних символів, у формуванні яких бере участь даний інформаційний символ.

З розглянутого принципу виправлення помилок у коді випливає, що правильне виправлення помилок можливо тільки в тому випадку, якщо два з трьох символів, охоплених перевіркою, прийняті правильно. Це виконується за умови, що при кроці додавання t, довжина пачки помилок не більше *2t* і перевірні символи передаються в канал із затримкою на (3t+ 1).

Таким чином, код виявляє і виправляє пачки помилок порівняно просто, але ціною великої надмірності.

5.6 Кодування із стисненням інформації

Як відомо, застосування стиснення даних дозволяє більш ефективно використовувати ємність дискової пам'яті. Не менш корисне застосування стиснення при передача інформації в будь-яких системах зв'язку. У останньому випадку з'являється можливість передавати значно меншу (як правило, у декілька разів) кількість даних і, отже, необхідні значно менші ресурси пропускної спроможності каналів для передачі тієї ж самої інформації. Виграш може виражатися в скороченні часу занятості каналу і, відповідно, у значній економії орендної плати.

Науковою передумовою можливості стиснення даних виступає відома з теорії інформації теорема кодування для каналу без перешкод, опублікована наприкінці 40-х років у статті Клода Шеннона "Математична

теорія зв'язку". Теорема підтверджує, що в каналі зв'язку без перешкод можна так перетворити послідовність символів джерела у послідовність символів коду, що середня довжина символів коду може бути як завгодно близька до ентропії джерела повідомлень *H(X)*, обумовленої як:

$$H(X) = -\sum_{i=1}^{N} p(x_i) \times \log_2 p(x_i)$$

де $p(x_i)$ – можливість появи конкретного повідомлення x_i із N можливих символів алфавіту джерела.

Число N називають об'ємом алфавіту джерела.

5.6.1 Код Шеннона-Фано

При передачі повідомлень, закодованих двійковим рівномірним кодом, не враховують статистичну структуру повідомлень, які незалежно від імовірності їх появи є кодовими комбінаціями однакової довжини, тобто кількість двійкових символів на одне постійне повідомлення. Такі коди мають надлишковість.

Найбільш ефективним способом зменшення надлишковості повідомлення є побудова оптимальних кодів, які мають мінімальну довжину кодових слів.

Враховуючи статистичні властивості джерела повідомлення, можна мінімізувати середнє число символів, які потрібні для вираження одного знаку повідомлення, що при відсутності шуму дозволяє зменшити час передачі або об'єм запам'ятовувального пристрою.

Ефективне кодування повідомлення для передачі їх по дискретному каналу без перешкод базується на теоремі Шеннона, яку можна сформулювати так: повідомлення джерела з ентропією $H(Z)$ завжди можна закодувати послідовностями символів з об'ємом алфавіту так, що середнє число символів на знак повідомлення $L_{сер}$ буде якомога ближче до величини $\frac{H(Z)}{\log(m)}$, але менше за неї.

Для одержання оптимального коду, що має мінімальну довжину кодової комбінації, необхідно домагатися найменшої надмірності кожного з кодів слова, що у свою чергу повинні будуватися з рівноймовірнісних і взаємонезалежних символів. При цьому кожен кодовий елемент повинен приймати значення 0 чи 1 по можливості з рівними ймовірностями, а вибір наступного елемента повинен бути незалежним від попереднього. Алго-

ритм побудови такого коду вперше був запропонований К. Шенноном у 1948 році і трохи пізніше модифікований Р. Фано, у зв'язку з чим він одержав назву Шеннона-Фано.

Відповідно до алгоритму на початку процедури кодування всі символи алфавіту джерела заносяться в таблицю в порядку зменшення ймовірностей. На першому етапі кодування символи розбиваються на дві групи таким чином, щоб суми ймовірностей символів у кожній з них були по можливості однакові. Усім символам верхньої групи присвоюється елемент кодової комбінації 0, а всім нижнім – 1. На другому етапі кодування кожна з груп знову розбивається на дві рівноймовірнісні підгрупи. Другому елементу кодових комбінацій для верхньої підгрупи присвоюється значення 0, а нижній – 1. Процес кодування продовжується до тих пір, поки в кожній підгрупі не залишиться по одному символу. Аналогічно може бути побудований альтернативний варіант оптимального префіксного коду Шеннона-Фано, у якому в процесі кодування верхнім підгрупам символів присвоюється кодовий елемент 1, а нижнім – 0.

Цей код відрізняється від попереднього тим, що його кодові комбінації для відповідних символів будуть інверсними.

ПРИКЛАД

Розглянемо алгоритм Шеннона-Фано на прикладі кодування джерела, алфавіт якого складається з 8 символів a_i, $(i=1,2,...,8)$, а імовірності появи символів у повідомленні дорівнюють від'ємним степеням двійки, тобто $P(a)=(1/2)^i$. Щоб ансамбль повідомлень джерела подав повну групу подій, у прикладі прийнято $P(a_8)=P(a_7)=(1/2)^7$. Процедура розбивання символів на групи й підгрупи та утворення кодових слів показані у таблиці 5.4.

Середнє число бітів на символ при кодуванні в даному випадку кодом Шеннона-Фано дорівнює

$$L_{сер}=\sum_{i=1}^{8}P(a_i)l(a_i)=\frac{1}{2}+\frac{1}{4}\cdot 2+\frac{1}{8}\cdot 3+\frac{1}{16}\cdot 4+\frac{1}{32}\cdot 5+\frac{1}{64}\cdot 6+\frac{1}{128}\cdot 7+\frac{1}{128}\cdot 8=1\frac{63}{64},$$

а ентропія джерела

$$H(A)=\sum_{i=1}^{8}P(a_i)\log(a_i)=\frac{1}{2}\cdot\log\left(\frac{1}{4}\right)+\frac{1}{4}\cdot\log\left(\frac{1}{4}\right)+\frac{1}{8}\cdot\log\left(\frac{1}{8}\right)+\frac{1}{16}\cdot\log\left(\frac{1}{16}\right)+$$

$$\frac{1}{32}\cdot\log\left(\frac{1}{32}\right)+\frac{1}{64}\cdot\log\left(\frac{1}{64}\right)+\frac{1}{128}\cdot\log\left(\frac{1}{128}\right)+\frac{1}{128}\cdot\log\left(\frac{1}{128}\right)=1\frac{63}{64}.$$

Таблиця 5.4 – Розбивання символів на групи й підгрупи

Символи a_i	Імовірності $P(a_i)$	Етапи кодування							Кодові комбінації
		I	II	III	IV	V	VI	VII	
a_1	$\frac{1}{2}$	0							0
a_2	$\frac{1}{4}$	1	0						10
a_3	$\frac{1}{8}$	1	1	0					110
a_4	$\frac{1}{16}$	1	1	1	0				1110
a_5	$\frac{1}{32}$	1	1	1	1	0			11110
a_6	$\frac{1}{64}$	1	1	1	1	1	0		111110
a_7	$\frac{1}{128}$	1	1	1	1	1	1	0	1111110
a_8	$\frac{1}{128}$	1	1	1	1	1	1	1	1111111

Таким чином, код Шеннона-Фано для заданого розподілу ймовірностей символів є оптимальним, оскільки середнє число бітів на символ у точності дорівнює ентропії джерела. При звичайному кодуванні для подання кожного символу необхідно 3 біти. Отже, коефіцієнт стиснення повідомлення за рахунок нерівномірного кодування Шеннона-Фано дорівнює

$$\frac{1\frac{63}{64}}{3}=0,66=66\%.$$

Приклад кодування за алгоритмом Шеннона-Фано для ансамбля символів з довільним розподілом імовірностей наведено у таблиці 5.5.

Таблиця 5.5 – Кодування за алгоритмом Шеннона-Фано

Символи a_i	Імовірності $P(a_i)$	Етапи кодування				Кодові комбінації
		I	II	III	IV	
a_1	0,22	0	0			00
a_2	0,20	0	1	0		010
a_3	0,16	0	1	1		011
a_4	0,16	1	0	0		100
a_5	0,10	1	0	1		101
a_6	0,10	1	1	0		110
a_7	0,04	1	1	1	0	1110
a_8	0,02	1	1	1	1	1111

5.6.2 Код Хаффмена

Алгоритм Шеннона-Фано не завжди приводить до однозначної побудови коду. Це викликано тим, що при розбитті m символів джерела на підгрупи можна зробити більшими за імовірністю як верхню, так і нижню підгрупи. При різному розбитті на підгрупи середнє число бітів на символ може бути різним. Більш ефективний є алгоритм, запропонований Хаффменом у 1952 р., що дозволяє побудувати оптимальний код із найменшим для даного розподілу імовірностей середнім числом бітів на символ. Тобто

$$l_{cp} = \min\left[\sum_{i=1}^{m} P(a_i) l(a_i)\right].$$

Для двійкового коду алгоритм Хаффмена зводиться до такого. Символи повідомлення упорядковуються за зменшенням ймовірностей і розташовуються в основний стовпець таблиці таким чином, що $P(a_i) \geq P(a_j)$ для усіх $i<y$ (табл. 5.6).

ПРИКЛАД

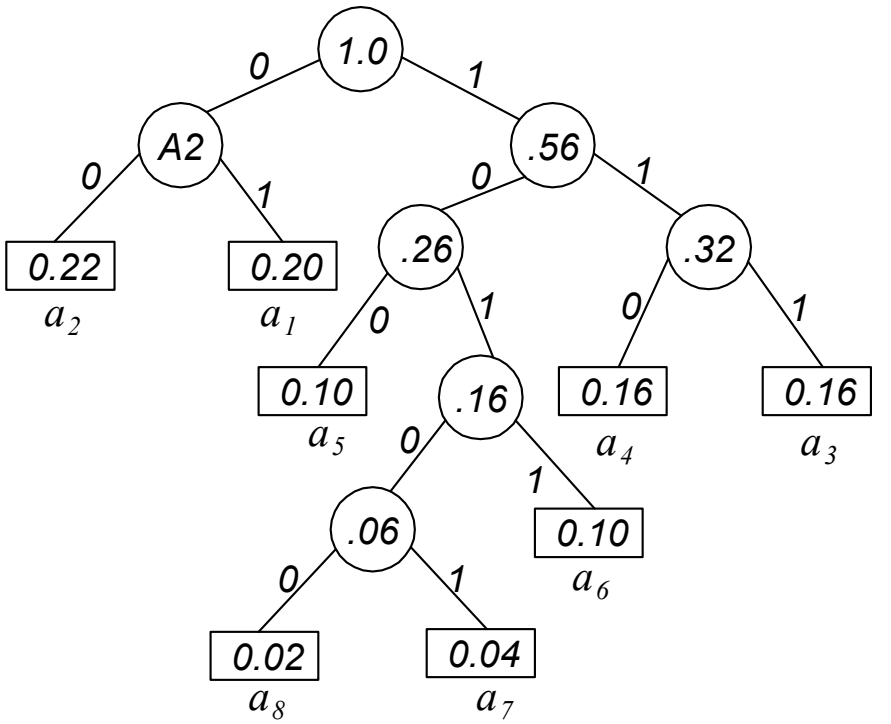
Таблиця 5.6 – Алгоритм Хаффмена

Символи a_i	Імовірності $P(a_j)$	Додаткові стовпці						
a_1	0,22	0,22	0.22	0,26	0,32	0,42	0,58	1,0
a_2	0,20	0,20	0,20	0,22	0,26	0,32	0,42	
a_3	0,16	0,16	0,16	0,20	0,22	0,26		
a_4	0,16	0,16	0,16	0,16	0,20			
a_5	0,10	0,10	0,16	0,16				
a_6	0,10	0,10	0,10					
a_7	0,04	0,06						
a_8	0,02							

Два останніх символи поєднуються в один допоміжний, імовірність якого дорівнює сумарній імовірності складових його символів. Усі символи, що залишилися, разом з утвореним допоміжним символом, знову розташовуються за зменшенням імовірностей у додатковому стовпці. Два останніх елементи стовпця поєднуються в другий допоміжний символ, і утворюється наступний додатковий стовпець, у якому всі елементи, розташовані в порядку спадання ймовірностей. Процедура продовжується доти, поки не вийде єдиний допоміжний символ, що має імовірність, рівну одиниці. Для формування кодових комбінацій, що відповідають символам даного повідомлення, необхідно простежити шлях переходу символів по рядках і стовпцях таблиці. При побудові кодів Хаффмена найчастіше використовуються кодові дерева. З однієї сторони: це дозволяє більш наочно відобразити процедури кодування і декодування, а, з іншого боку – полегшити програмну реалізацію цих процедур (рис. 5.17).

Побудова дерева (рис. 5.17, а) починається з кореневого вузла, імовірність якого дорівнює 1. З кореня проводяться дві вітки, причому вітці з більшою імовірністю присвоюється значення (*біт*) 1, а з меншою імовірністю – 0. Знову утворені вузли можуть відображати одиничний чи допоміжний символи. В останньому випадку вузол є проміжним і з кожного з них, у свою чергу, знову проводяться по дві вітки. Таке послідовне розга-

луження продовжується доти, поки не буде побудований вузол, що відповідає імовірності символу алфавіту (*вузол листа*). Рухаючись по кодовому дереву від кореня зверху вниз, можна записати для кожного символу відповідну йому кодову комбінацію (рис. 5.17, б).



$a_1 = 01$
$a_2 = 00$
$a_3 = 111$
$a_4 = 110$
$a_5 = 100$
$a_6 = 1011$
$a_7 = 10101$
$a_8 = 10100$

а) б)

Рисунок 5.17 – Кодове дерево (а) і таблиця коду Хаффмена (б)

Середнє число бітів на символ при такій побудові коду складає

$$L_{ср} = \sum P(a_i) l(a_i) = 0.22 \times 2 \times 0.2 \times 2 \times 0.16 \times 3 \times 0.16 \times 3 \times 0.1 \times$$
$$\times 3 \times 4 \times 0.04 \times 5 = 2.8 \text{ біта}$$

Ентропія джерела повідомлення дорівнює:

$H(A) = -\sum P(a_i) \log P(a_i) =$ - (0,22*log*0,22 + 0,2*log*0,2 + 0,16*log*0,16 + +0,16*log*0,16 + 0,l*log*0,l + 0,1*log*0,1 + 0,04*log*0,04 + 0,02*log*0,02 = 2,754 біта

Як видно з розглянутого прикладу, середня довжина кодової комбінації й ентропія джерела практично збігаються, тобто отриманий код є оптимальним.

Алгоритм є двопрохідним, тому що при його реалізації потрібно двічі переглядати кодоване повідомлення. При першому проході обчислюються імовірності (*частоти*) появи символів у повідомленні і будується хаффменівське дерево.

При другому проході здійснюється кодування символів, що надійшли від джерела. У цьому випадку визначається значення віток дерева при русі від *листка*, що відповідає кодованому символу, до *кореня*. Очевидно, що для прискорення процедури декодування біти кодової комбінації на вихід кодера повинні видаватися, починаючи зі старшого розряду, тобто з вітки дерева, що виходить від кореня. На практиці замість імовірностей символів використовують абсолютні значення кількості (*вага*) символів у переданому повідомленні, тому що кількість символів пропорційна імовірності їхньої появи. Для однозначного декодування таблиця імовірностей символів повідомляється декодеру.

5.6.3 Арифметичне кодування

При арифметичному стисненні повідомлень алфавіту джерела ставиться у відповідність числовий, відкритий справа, інтервал [0,1), а кожен символ алфавіту зіставляється з різними ділянками цієї числової осі. Ширина інтервалу (*діапазон*) кожної ділянки залежить від імовірності (*частоти*) появи символу в повідомленні.

ПРИКЛАД

Розглянемо, як приклад, алфавіт, що складається із шести символів A, B, C, D, E, ! з імовірностями появи відповідно 0,1; 0,1; 0,3; 0,2; 0,2 та 0,1. Тоді інтервал [0,1) може бути розділений на ділянки таким чином:

A: [0, 0,1), B: [0,1, 0,2), C: [0,2, 0,5),

D: [0,5, 0,7), E: [0,7, 0,9), !: [0,9, 1,0).

Розподіл числової осі ілюструється рис. 5.18. Як видно з прикладу, поділ одиничного інтервалу зручно здійснюється підсумовуванням імовірності кожного з границею попереднього.

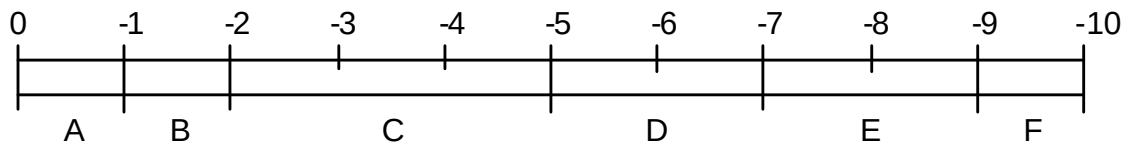


Рисунок 5.18 – Розподіл відрізків числової осі між символами при арифметичному кодуванні

У процесі стиснення рядок текстового повідомлення подається

двома дробовими числами, що відображають деякий інтервал числової осі [0, 1). До початку кодування допустимою областю для повідомлення є інтервал [0, 1). При надходженні від джерела першого символу йому виділяється інтервал, що відповідає розташуванню цього символу на осі [0, 1) відповідно до імовірності появи його в повідомленні. Надходження наступного символу звужує виділений інтервал. При цьому границі нового інтервалу визначаються числовим діапазоном, що відповідає черговому символу, тобто ширина діапазону пропорційна імовірності символу. З приходом наступного символу відбувається подальше звуження інтервалу. Таким чином, із збільшенням довжини рядка, що кодується, відбувається постійне звуження інтервалу. Теоретично його можна звужувати як завгодно довго, а на практиці цей інтервал обмежується технічними можливостями кодувальних пристроїв (розрядністю слова).

Число бітів, необхідне для подання інтервалу шириною s, дорівнює $-\log(s)$. Основа логарифму тут і в наступних виразах дорівнює 2. Ширина s останнього інтервалу рядка повідомлення, що складається з N_c символів визначається добутком ймовірностей символів P_{jC} повідомлення

$$S = \prod_{i=1}^{N_C} P_{jC} .$$

В зв'язку з цим можна записати, що

$$-\log(s) = -\sum \log P_i = -\sum_{i=1}^{m} P(a_i) \log P(a_i),$$

де m – кількість різних символів повідомлення a_i, $(i = 1,2,...,m)$.

Таким чином, число бітів, затрачуване на кодування повідомлення, при арифметичному кодуванні в точності дорівнює ентропії джерела.

Розглянемо процедуру арифметичного стиснення на прикладі кодування рядка повідомлення (CADA!).

Тут символ ! є ознакою закінчення рядка. Імовірності появи символів і відповідні їм числові діапазони візьмемо з попереднього прикладу. Початковий числовий діапазон, який виділяється для кодера, дорівнює одиниці [0,1). Після надходження від джерела першого символу С числовий інтервал звужується до величини [0,2, 0,3). З приходом наступної букви А границі діапазону стають рівними [0,2, 0,23).

Процедура кодування наочно ілюструється на рис. 5.19. Вертикальні лінії відображають початкову числову вісь [0, 1) і її наступні ділянки в

збільшеному розмірі (промасштабовані на одиничну вісь) на різних етапах кодування. Який би не був довгий рядок, кінцевий діапазон завжди буде розташований в інтервалі першого кодованого символу. Підрахунок інтервалів здійснюється за рекурентною формулою

$$(X_{\max} - X_{\min})X'_{\min} + X_{\min}$$

$$(X_{\max} - X_{\min})X'_{\max} + X_{\min},$$

де $X_{\max}$ і $X_{\min}$ – нижня і верхня границі для попередньої ітерації;

$X'_{\min}$ і $X'_{\max}$ – границі інтервалу нового символу.

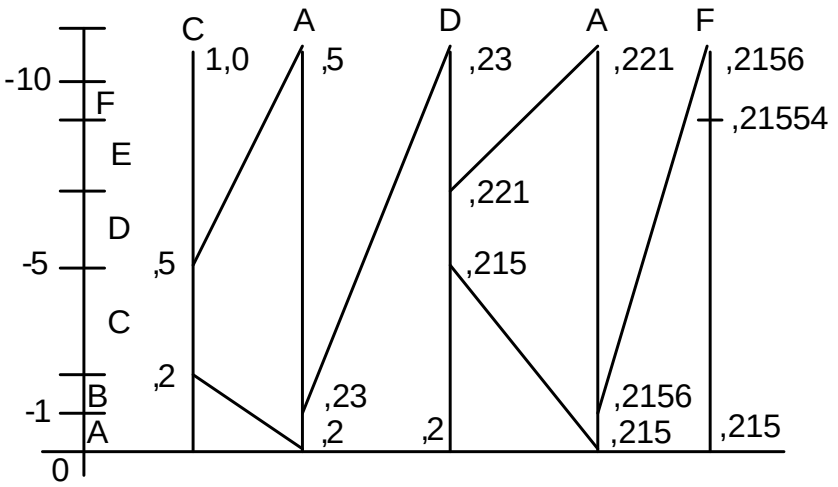


Рисунок 5.19 – Ілюстрація процедури арифметичного кодування

Межі інтервалів для всіх етапів кодування наведені в табл.5.7. Таким чином, числовий інтервал з межами [0,21554-0,2156) відображає рядок повідомлення CADA! Для кодування цього рядка може бути використано будь-яке число з отриманого діапазону. Нехай таким числом є нижня межа інтервалу, тобто X = = 0,21554.

Таблиця 5.7 – Результати арифметичного кодування

Символи	Нижня межа інтервалу	Верхня межа інтервалу
Вихідний стан	0,0	1.0
C	0,2	0.5
A	0.2	0,23
D	0,215	0.221
A	0,215	0,2156
!	0,21554	0,2156

У процесі декодування здійснюється зворотна процедура ототожнення прийнятого числа (інтервалу) із закодованим рядком символів. Очевидно, що на приймальній стороні повинні бути відомі ймовірності декодованих символів і відповідні їм інтервали числової осі [0,1). Нехай на вхід декодера надійшло число X = 0,21554. Оскільки воно перебуває в діапазоні 0,2 - 0,5, тобто $0,2 \le X < 0,5$, то на вихід декомпресора видається символ C. Потім декодер визначає, в яку відносну частку відрізка [0,2 – 0,5) потрапляє число 0,21554, для цього він робить обчислення за формулою

$$X' = \frac{(X - X_{\min})}{(X_{\max} - X_{\min})} \quad , \tag{5.7}$$

де X' – відносне значення числа X на черговій ділянці числової осі.

У нашому випадку X' = (0.21554 - 0,2) / 0,3 = = 0,0518. Оскыльки $0 < X < 0,1$, то на вихід декомпресора надходить символ A. Неважко помітити, що процедура визначення вихідного символу рекурентна, при якій після знаходження першого діапазону для заданого числа X на одиничній числовій осі визначається відносна частка (діапазон) цієї частини відрізка, куди потрапляє декодоване повідомлення.

Обчислення на наступному кроці проводиться за формулою (5.7) з припущенням, що $X = X'$ де X' – відносне значення числа, знайденого в попередній ітерації. Нові значення X_{min} і X_{max} визначаються межами інтервалу, ідентифікованого на попередній ітерації. На цьому кроці X = 0,0518, $X_{min} = 0$, $X_{max} = 0,1$, а $X' = 0,0518/0,1 = 0,518$. Нове X' відповідає символу D, оскільки $0,5 < X' < 0,7$.

Далі, вважаючи, $X = 0,518$, $X_{\min} = 0,5$, $X_{\max} = 0,7$, знайдемо аналогічно $X' = (0,518 - 0,5) / 0,2 = 0,09$, що відповідає символу A. Оскільки на наступній ітерації $X' = 0,09 / 0,1 = 0,9$, то декодер виводить символ !.

У зв'язку з тим, що символ ! є ознакою закінчення рядка повідомлення, то декодування на цьому закінчується.

ПРИКЛАД

Провести компресію і декомпресію повідомлення ”ІНФОРМАЦІЯ”, використовуючи метод арифметичного кодування.

Визначимо ймовірності появи символів у заданому повідомленні і розподілимо символи на одиничному відрізку (табл. 5.8). На підставі алго-

ритму кодування розрахуємо нижні і верхні границі кодового числа, що відображає стисле повідомлення, в залежності від кількості закодованих символів.

Таблиця 5.8 – Розподіл символів на одиничному відрізку

Символи	Вірогідність появи	Розташування на відрізку
А	0,1	$0,0 \le X < 0,1$
І	0,2	$0,1 \le X < 0,3$
М	0,1	$0,3 \le X < 0,4$
Н	0,1	$0,4 \le X < 0,5$
О	0,1	$0,5 \le X < 0,6$
Р	0,1	$0,6 \le X < 0,7$
Ф	0,1	$0,7 \le X < 0,8$
Ц	0,1	$0,8 \le X < 0,9$
Я	0,1	$0,9 \le X < 1,0$

Дані розрахунку зведемо в табл. 5.9. Як кодове число, що відображає стисле повідомлення, приймемо нижню межу цього числа: 0.1951261656.

Таблиця 5.9 – Дані розрахунку

Символи	Нижня границя інтервалу	Верхня границя інтервалу
Початковий стан	0,0	1,0
І	0,1	0,3
Н	0,18	0,2
Ф	0,194	0,196
О	0,1950	0,1952
Р	0,19512	0,19514
М	0,195126	0,195128
А	0,195126	0,1951262
Ц	0.19512616	0,19512618
І	0,195126162	0,195126166
Я	0,1951261656	0,195126166

Використовуючи алгоритм декодування, зробимо процедуру деко-

мпресії, етапи якої наведено в табл. 5.10.

Таблиця 5.10 – Результати процедури декомпресії

Кодове число	Вихідний символ	Нижня границя	Верхня границя	Інтервал
0,1951261656	І	0,1	0,3	0,2
0,4756300828	Н	0,4	0,5	0,1
0.75630828	Ф	0,7	0,8	0,1
0,5630828	О	0,5	0,6	0,1
0 630828	Р	0,6	0,7	0,1
0,30828	М	0,3	0,4	0,1
0,0828	А	0,0	0,1	0,1
0,828	Ц	0,8	0,9	0,1
0.28	І	0,1	0,3	0.2
0.9	Я	0,9	1,0	0,1
0,0 – кінець декодування				

5.6.4 Динамічне кодування методом Хаффмена

Класичний метод кодування Хаффмена визначає від початку перетворення вірогідності появи символів на виході джерела інформації, причому символи впорядковується за спаданням вірогідностей їх виникнення. Для впорядкованого списку складається кодова таблиця, у якій довжина вихідної комбінації визначається вірогідністю вихідного символу. Природно, що на передавальній та приймальній сторонах повинні бути відомі таблиці (кодові дерева) для кожного повідомлення, що стискається. Цей метод має два вагомих недоліки. Перший полягає в тому, що для його реалізації потрібно два проходи кодованого масиву. При першому перегляді обчислюються імовірності появи кожного знака в повідомленні і складається таблиця коду Хаффмена. На наступному етапі здійснюється кодування на підставі статичної структури дерева Хаффмена і передача символів у стиснутому вигляді. Таким чином, виграш, отриманий за рахунок стискання даних, може помітно знижуватись, особливо при передачі щодо короткого повідомлення, у зв’язку з необхідністю передавати декодеру додаткову інформацію про кодове дерево. Другий недолік – наявність затри-

мки від моменту надходження даних від джерела до видачі відповідних кодових комбінацій, що обмежує використання нерівномірного кодування в синхронних мережах передачі інформації, а також у системах, що функціонують у реальному часі.

На початку 70-х років були розроблені однопрохідні методи стискання інформації, основані на класичній процедурі кодування Хаффмена. Усі ці методи незначно відрізняються один від одного і їхня суть полягає в тому, що передавач будує дерево Хаффмена в темпі надходження даних від джерела, тобто "*на ходу*". У процесі кодування відбувається "*навчання*" кодера на основі статичних характеристик джерела повідомлень, у ході якого обчислюється оцінка вихідних імовірностей повідомлення і виробляється відповідна модифікація кодового дерева, цей процес отримав назву *динамічного кодування Хаффмена*. Очевидно, що для правильного відновлення стиснутих даних, декодер також повинен постійно "*вчитися*" поряд із кодером, здійснюючи синхронну зміну кодової таблиці на приймальній стороні. Для забезпечення синхронності процесів кодування і декодування кодер видає символ у нестиснутому вигляді, якщо він уперше з'явився на виході джерела, і визначає його на кодовому дереві. З повторною появою символу на вході кодера він передається нерівномірною кодовою комбінацією, обумовленою позицією символу на поточному кодовому дереві. Кодер корегує дерево Хаффмена збільшенням частоти передачі символів, що уже введені в дерево, чи нарощує дерево, додаючи в нього нові вузли.

Найважливішою умовою, якої потрібно дотримуватись при модифікації кодового дерева, є збереження властивостей хаффменівського дерева. Для формулювання цих властивостей звернемося ще раз до алгоритму побудови оптимального коду Хаффмена. При статичному кодуванні символи розміщуються в списку в порядку спадання ваг (ймовірностей). Потім проводиться об'єднання двох вузлів найменшої ваги W_i, W_j і заміна їх внутрішнім вузлом з вагою, рівною сумі вихідних ваг $W_i + W_j$. Знову утворений вузол розміщується в списку таким чином, щоб не порушувався порядок розташування вузлів за вагами. Цей процес повторюється доти, поки в списку не залишиться один, так званий *кореневий* вузол.

Розглянемо приклади побудови дерев оптимального коду Хаффмена для букв алфавіту від A до H. Ваги цих символів відповідно мають: A=18; B=10; C=2; D=2; E=1; F=1; G=1; H=1. Проаналізуємо їх властивості. Як видно з рисунків, вузли дерева розташовані в порядку зростання їхніх

ваг при обході дерева від крайнього нижнього вузла до кореня зліва, направо і знизу вверх. У зв'язку з тим, що вузли з вагами W_i, W_j поєднуються парно, то на одному рівні не може бути менше двох вузлів, причому пари вузлів є дочірні в загальному батьківському вузлі, вага якого дорівнює сумі ваг дочірніх вузлів. Неважко переконатися, якщо при побудові дерева припустити, що дочірній вузол із великою вагою з'єднаний нульовою віткою, а з меншою вагою – одиничною, то хаффменівське дерево залишається упорядкованим за зростанням ваг при русі від нижнього вузла до кореня по рівнях справа наліво.

При побудові кодових дерев і їхньої модифікації проводиться нумерація вузлів із першого до $(2m-1)$-го в порядку збільшення їхніх ваг. Перший номер присвоюється вузлу з мінімальною вагою. Тут m - число символів алфавіту джерела. На рис. 5.20 показано дерево Хаффмена, побудовене для вищенаведеного прикладу 18A, 10B, 3C, 2D, 1E, 1F, 1G, 1H, відповідно. Зробимо нумерацію вузлів дерева, починаючи з листка з мінімальною вагою (нижнім крайнім лівим вузлом). Як видно з рисунка, властивості хаффменівського дерева, побудованого для статичного коду, зберігаються, хоча воно й набуло іншу конфігурацію.

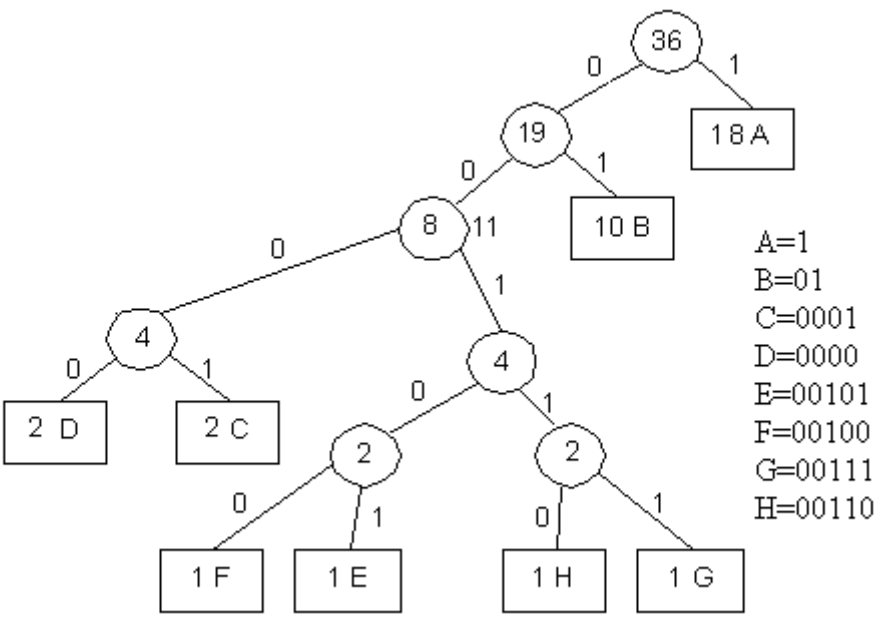


Рисунок 5.20 – Дерево Хаффмена для повідомлення, що складається зі символів A; B; C; D; E; F; G; H

При динамічному кодуванні Хаффмена після одержання від джерела наступного символу (наприклад, “C”) повинен збільшитись на 1 вагу відповідного йому листка (рис. 5.21). При цьому ваги усіх вузлів на шляху від листка із символом C до кореня збільшиться на 1. Але це приведе до порушення порядку розташування вузлів за вагами, властивому оптимальному кодовому дереву. Тому при одержанні чергового символу від джерела необхідно здійснити модифікацію кодового дерева, щоб воно залишалося оптимальним, так званим хаффменівським деревом.

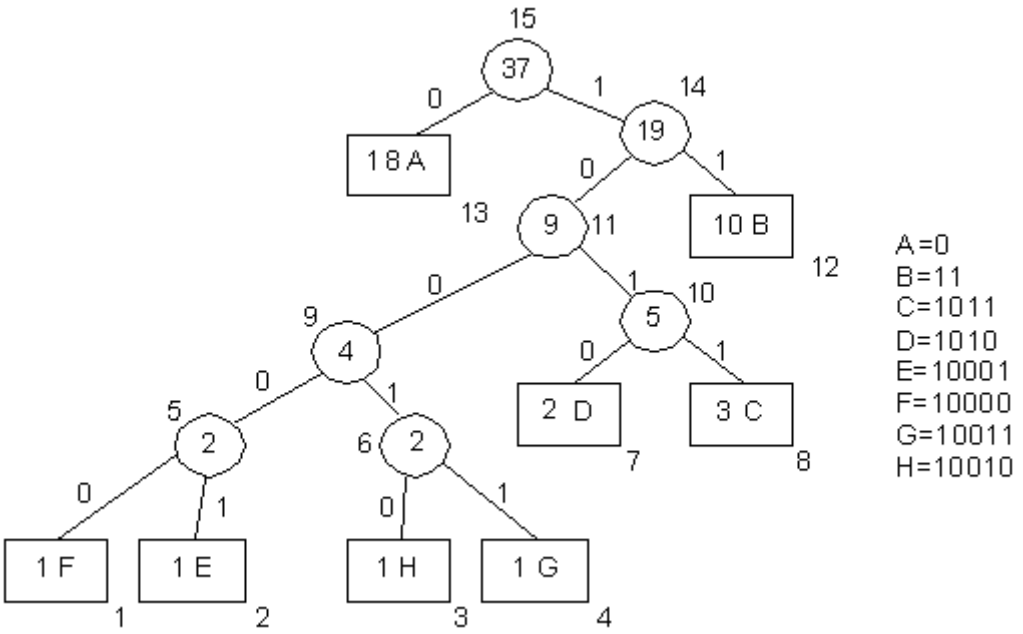


Рисунок 5.21 – Модифіковане дерево Хаффмена після надходження чергового символу “C”

Таким чином, на підставі аналізу властивостей оптимальних кодових дерев можна зробити висновок, що кодове дерево, яке має m зовнішніх вузлів, є хаффменівським, якщо воно має такі властивості:

а) зовнішні вузли (листки) хаффменівського кодового дерева мають вагу $W > 0$; кожен внутрішній (*батьківський*) вузол має підпорядковані (дочірні) вузли, а його вага дорівнює сумі підпорядкованих (*дочірніх*) ваг;

б) на кожному рівні дерева (за винятком кореневого) повинно бути не менше однієї пари вузлів, що мають загальний батьківський вузол;

в) усі вузли нумеруються в зростаючому порядку таким чином, що вузли з номерами (*2j - 1*) та *2j* є вузлами одного рівня для $l < j < m - l$, а їх загальний батьківський вузол має більш високий рівень;

г) нумерація вузлів відповідає тому порядку, у якому вузли поєднуються відповідно до статичного алгоритму Хаффмена.

5.6.5 Динамічне кодування методом FGK

Уперше алгоритм синтезу динамічного коду Хаффмена був запропонований Н. Феллером у 1973 році, а потім модифікований Р. Галлагером і Д. Батогом. У зв'язку з цим він одержав назву "*алгоритм FGK*".

Введемо деякі позначення, що будуть використовуватись при аналізі й синтезі динамічного кодового дерева Хаффмена за алгоритмом FGK:

m – розмір алфавіту джерела повідомлень;

z_i – j – й символ алфавіту;

$M(k) = z(1),\ z(2),...,\ z(k)$ – перші k символів у повідомленні;

k – число символів у повідомленні, оброблених до поточного моменту часу;

$z(k)$ – k-й символ у повідомленні;

K – кількість різних символів, оброблених на даний момент часу;

W_j – число (вага) символів z_j, що надійшли на момент обробки повідомлення;

l_j – відстань від кореня дерева до z_j-го листка.

Суть алгоритму синтезу динамічного кодового дерева Хаффмена полягає в процедурі обчислення листків і побудові бінарного дерева з мінімальною вагою шляху $\sum_j W_j \cdot l_j$. Процедуру можна умовно розбити на два етапи, хоча при реалізації алгоритму вони можуть легко бути об'єднані в один. На першому етапі дерево Хаффмена, побудоване після обробки повідомлення $M(k)$, перетвориться в інше, еквівалентне вихідному, яке потім простим збільшенням ваг може бути перетворене в хаффменівське дерево для $M(k+l)$.

Перший етап починається після одержання від джерела символу $z(k+l)$ із присвоєння статусу *поточного* вузла листка $z(k+l)$. Потім проходить обмін поточного вузла (включаючи утворене ним піддерево), із вузлом, що має *найбільший порядковий номер та вагу*. Після цього як новий поточний вузол ініціюється батьківський вузол останнього поточного вузла. Обмін вузлами в разі потреби багаторазово повторюється, поки не буде досягнутий корінь дерева. Неважко переконатися, що максимальна кіль-

кість перестановок, які можуть знадобитися при модифікації кодового дерева, дорівнює висоті дерева $l_{\max}$.

На другому етапі формується листок дерева, що відповідає обробленому символу і наступні проміжні вузли, розташовані на шляху руху від листка до кореня дерева.

Оригінальний спосіб був запропонований Д. Кнутом. Його суть полягає в тому, що усі символи використовуваного алфавіту, які ще не з'явилися на виході джерела, відмічаються на дереві *нульовим вузлом*. Тому, коли символ, що підлягає стисненню, генерується вперше, кодер "відзначає" його, виробляючи *префіксну комбінацію*. Слідом за нею на вихід надходить код нестиснутого символу. Як префіксна комбінація використовується код, що відповідає листку *нульової ваги*. Для кодування символів, що надійшли від джерела вперше, використовується не 8-ми розрядна комбінація ASCII-коду, а мінімальний префіксний код, побудований на підставі таких міркувань: якщо є повідомлення, що складається із m символів $a_i,..., a_m$, то число m можна подати у вигляді цілого степеня двійки і цілого числа m

$$m = 2^e + r,$$

де $0 < m < 2^e$.

Тоді a_k-й символ кодується $(e+1)$-ю бітовою комбінацією числа $k-1$, якщо $1 \le k \le 2r$.

У протилежному випадку – e-ю бітовою комбінацією числа $k-r-1$. Наприклад, якщо $m=6$, то $e=2$ та $r=2$, а символи джерела будуть відображатися комбінаціями:

$$a_1 \Rightarrow 000;\ a_2 \Rightarrow 001;\ a_3 \Rightarrow 010;\ a_4 \Rightarrow 001;\ a_5 \Rightarrow 10;\ a_6 \Rightarrow 11.$$

Неважко помітити, що отриманий код має властивість префікса. Цей код є оптимальним, якщо символи мають однакову імовірність.

5.6.6 Динамічне кодування методом Віттера

Подальше удосконалення алгоритму динамічного кодування даних нерівномірними кодами *FGK* було запропоноване Д. Віттером. Цей алгоритм одержав назву *алгоритму V*. При пошуку шляхів оптимізації процедури кодування даних кодом Хаффмена автор виходив із того, що в новому алгоритмі число обмінів вузлів у процесі модифікації кодового дерева повинна обмежуватись деяким малим числом (у кращому випадку одини-

цею), а динамічне хаффменівське дерево повинна будуватися таким чином, щоб мінімізувати не тільки сумарну довжину зовнішнього шляху $\sum W_j \cdot l_j$, але і величину $\sum l_j$, $\max\{l_j\}$. Мінімізація висоти дерева $h = \max\{l_j\}$ дозволить запобігати утворенню довгих кодових комбінацій при кодуванні чергового символу в повідомленні. Віттеру значною мірою вдалося вирішити поставлену задачу. Розроблений ним алгоритм має в порівнянні з алгоритмом FGK такі дві переваги.

1. Кількість обмінів вузлами, при яких поточний вузол переміщується вгору по кодовому дереві в процесі його модифікації, обмежується одиницею. В алгоритмі FGK верхня границя числа обмінів складає $\frac{l_j}{2}$, де l_j - довжина кодового слова для $Z_j(k+l)$-го символу до початку процедури модифікації.

2. Алгоритм *V* мінімізує довжину зовнішнього шляху дерева l_j і гарантує дерево мінімальної висоти $h = \max\{l_j\}$ за умови мінімізації сумарної довжини зовнішнього шляху дерева $\sum W_j \cdot l_j$.

Суть удосконалення алгоритму *V* полягає у введені нової системи нумерації вузлів кодового дерева, що одержала назву *неявної нумерації (Implicit numbering)*.

При неявній нумерації вузли хаффменівського дерева нумеруються в порядку збільшення по рівнях зліва направо і знизу вверх, тобто, вузли більш низького рівня мають номери менші ніж вузли наступного рівня. Найважливішою особливістю неявної нумерації є дотримання необхідної умови побудови дерева.

Для кожної ваги *W* всі зовнішні вузли (листи) дерева з вагою *W* повинні передавати всім внутрішнім вузлам ваги *W*.

Неважко помітити, що ця умова є однією з відмінних рис неявної нумерації у порівнянні з іншими алгоритмами.

Тому, якщо передача продовжиться із символів *d, c, g, e, f* чи з ще не використаних символів, то формовані кодові слова в алгоритмі *V* будуть коротші, ніж у *FGK*, тобто стратегія мінімізації зовнішнього шляху й висоти дерева оптимальна при припущенні, що будь-який символ, який з'являється, рівноймовірний.

Другою відмінною рисою алгоритму *V* є введення поняття блока еквівалентних вузлів. При цьому вузли *v* та *в* еквівалентні тільки тоді, якщо вони мають однакову вагу і обоє є або внутрішніми, або зовнішніми вузлами. Вузол блока, що має (при неявній нумерації) найвищий номер, називається *лідером блока.* Блоки впорядковуються за збільшенням ваги, причому блок листків ваги *W* повинен дорівнювати блоку внутрішніх вузлів однієї і тієї ж самої ваги.

Третя відмінна риса алгоритму *V* полягає в способі модифікації дерева після одержання чергового символу. Головною операцією алгоритму за підтримкою умови неявної нумерації (*) є ковзання й збільшення (*Slide and Increment*). Суть цієї операції полягає в тому, що вузол, оголошений поточним, обмінюється з лідером свого блока і потім зісковзує в напрямку кореня дерева по сусідньому блоці, що безпосередньо примикає до блоку поточного вузла. Ковзання продовжується доти, поки поточний вузол не пройде весь блок і буде установлений у початок цього блока. Потім здійснюється збільшення ваги поточного вузла і новим поточним вузлом призначається батько старого поточного вузла. Операція ковзання із збільшенням продовжується до досягнення кореня дерева. При цьому вибір батьківського вузла залежить від того чи був поточний вузол листком, чи внутрішнім вузлом. Якщо поточний вузол був листком, то новим поточним вузлом стає батько, із яким виявився зв'язаний поточний вузол після завершення ковзання. А у випадку, якщо поточним вузлом був внутрішній вузол, то новим поточним вузлом призначається його батьківський вузол, із яким був пов'язаний поточний до початку ковзання.

5.7 Методи стиснення зображень

В залежності від оброблюваного зображення всі методи стиснення зображень можна умовно поділити на методи стиснення бінарних зображень, півтонових зображень та кольорових зображень. Також стиснення може відбуватись із втратами інформації та без втрат. Розглянемо більш детально ці методи.

Методи стиснення без втрат

Факсовий метод

Факсовий (використовуваний у факсах) метод кодування полягає в тому, що кодуються одиницями перепад кольорів сусідніх пікселів в рядку,

а нулем – відсутність перепаду (рис. 5.22). Для зображень, що містять однорідні ділянки великої площі, цей метод є надлишковим, легко кодується методами оптимізації.

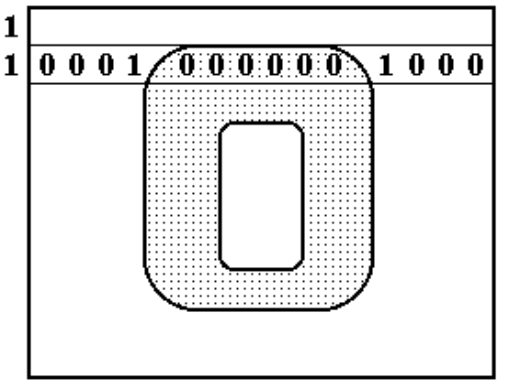


Рисунок 5.22 – Приклад кодування

До недоліків методу слід віднести: при збої виходять перешкоди у вигляді чорних або білих смуг; не використовується двовимірна природа зображення.

RLE метод

RLE код відрізняється тим, що порядково кодується колір пікселя і кількість пікселів цього кольору, що йдуть підряд у цьому рядку поки не зустрінеться зміна кольору (рис. 5.23).

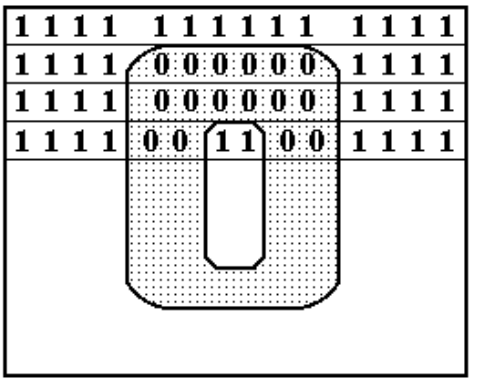


Рисунок 5.23 – Приклад кодування RLE методом

Метод досить простий у реалізації і застосовується при стисненні bmp-зображень.

Недолік RLE методу полягає в тому, що він при стисненні зображення, у якому плавно змінюються кольори, приводить не до стиснення, а навпаки до збільшення розміру місця, займаного зображенням.

Кодування кутами Шлезенгера

Для простоти розглянемо стиснення чорно-білого зображення. Наприклад, маємо чорно-біле зображення, записане растром.

Для реалізації методу необхідно:

а) дописати рядок і стовпець нулів;

б) як запропонував Шлезенгер, надалі необхідно кодувати не кожен піксел, а четвірки пікселів. Таких четвірок може бути всього шістнадцять видів (рис. 5.24).

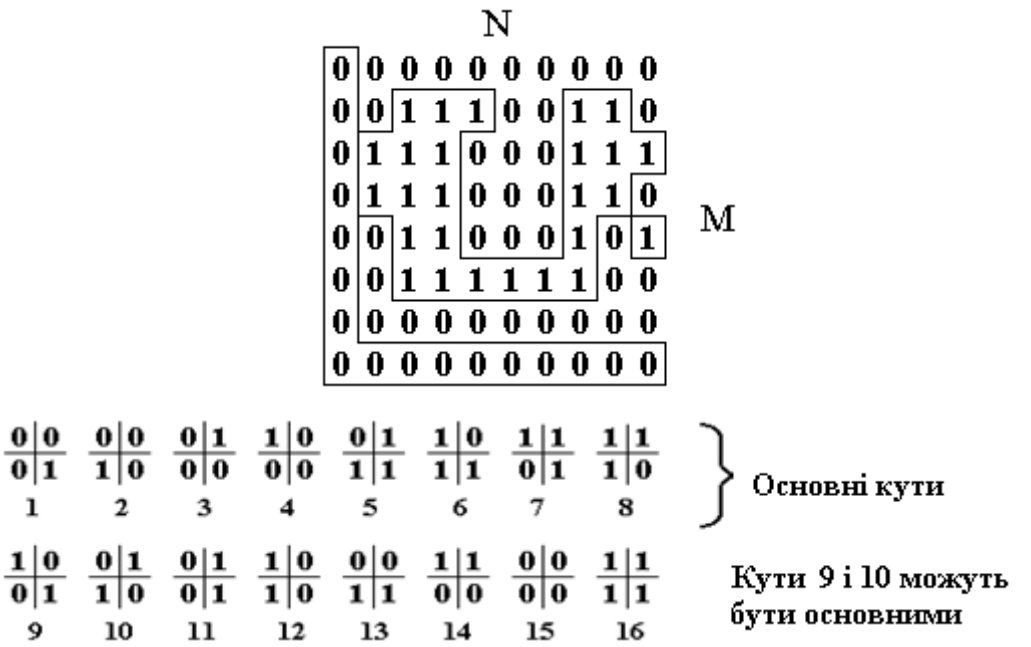
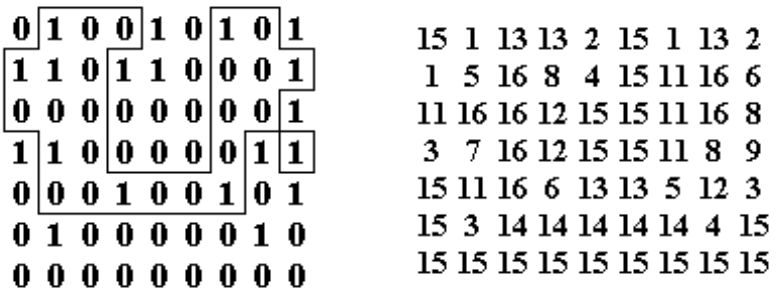


Рисунок 5.24 – Типи кутів Шлезенгера

З вихідної растрової (N +1) на (M +1) картинки ми отримаємо масив розміром N на M біт (рис. 5.25, а).



а) б)

Рисунок 5.25 – Результати кодування методом кутів Шлезенгера

Далі проводимо виділення кутів та записуємо їх у вигляді послідовності (рис. 5.25, б).

Наступним кроком є аналіз цієї послідовності. Якщо сканувати зліва направо, зверху вниз, і якщо є кут № 1, то його обов'язково замикає кут № 2. Тому для кодування використовуємо основні кути. Із зображення розміром N на M, кожен піксел якого є 1 біт, замінюємо зображенням основних кутів розміром 1 біт, вказуючи 1 – якщо кут дуальний (основний), і вказуючи 0 – якщо кут недуальний (неосновний).

Можна зробити такі висновки.

1. Отриманий растр менш надлишковий.

2. Перетворення є стискальним для зображень з великими площами рівної зафарбованості (креслення).

3. Дане кодування зворотне, тобто зображення відновлюється без втрат.

4. Перетворення Шлезенгера знижує ентропію, отже підвищує ступінь стиснення.

Wavelet – хвильове перетворення

Таке перетворення полягає в тому, що вихідний сигнал пропускається через систему лінійних фільтрів (як правило ФНЧ і ФВЧ), в результаті сигнали, отримані на виході цих фільтрів, пропускаються через аналогічні фільтри (рис. 5.26).

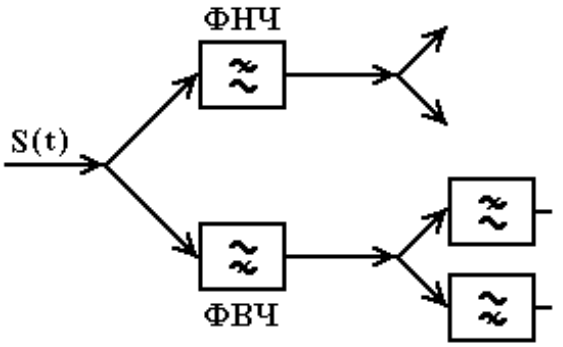


Рисунок 5.26 – Схема Wawelet-перетворення

У результаті отримуємо розкладання вихідного сигналу на складові, пропорційні ортогональним функціям імпульсних характеристик фільтрів, які називаються вейвлетами (Wavelet).

Фільтри, що зазвичай використовуються в Wavelet-парі повинні бути ортогональними, тобто, згортка імпульсних характеристик повинна дорівнювати нулю.

S-перетворення

Розглянемо принцип методу S-перетворення на основі певного прикладу.

Виберемо як сигнал певну послідовність відліків фіксованої величини S(i): $S_{MIN} \le S(i) \le S_{MAX}$.

Далі необхідно здійснити перетворення за схемою, що зображена на рис. 5.27.

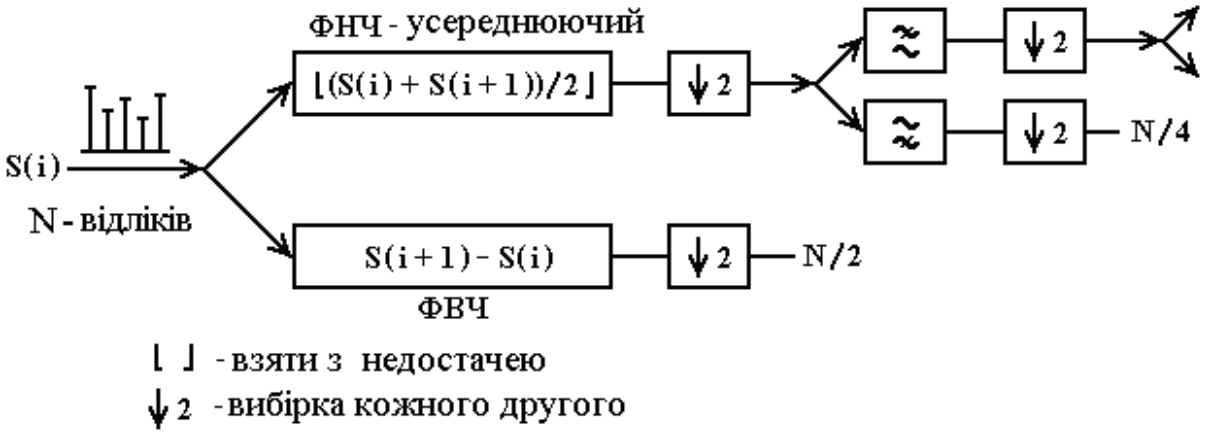


Рисунок 5.27 – Схема S-перетворень

Таких перетворень необхідно здійснити (Log2N) раз, це буде кількість кроків.

$$H = S_{i+1} - S_i$$
$$L = \lfloor (S_{i+1} - S_i)/2 \rfloor .$$

У результаті S-перетворення отримаємо зображення з ентропією меншою, ніж у вихідного зображення за умови, що вихідне зображення було природного походження, тобто, поряд розташовані пікселі мали схожі характеристики (рис. 5.28).

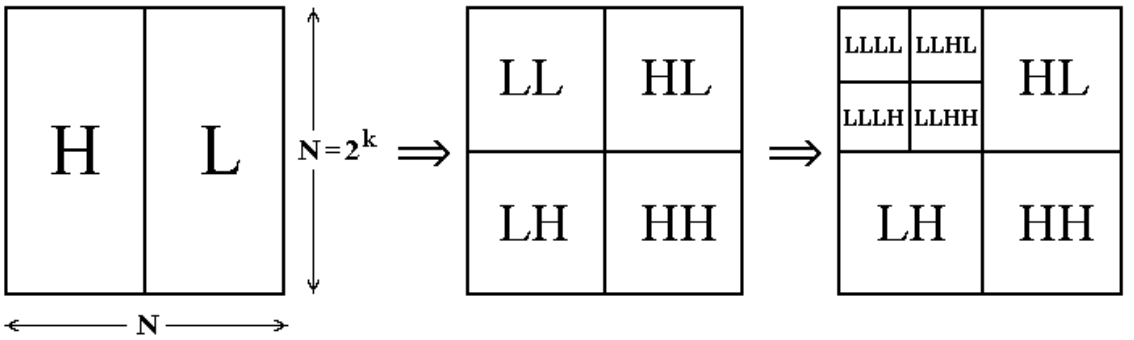


Рисунок 5.28 – Результати S-перетворення

Це дає можливість стискати перетворені за допомогою цього методу зображення без втрат більш ефективно, ніж без використання S-перетворення. Операція зворотна стисненню і абсолютно симетрична.

Методи стиснення із втратами

При стисненні зображень необов'язковим є стиснення без втрат. Це пояснюється особливостями зорового сприйняття людини. Тому незначні зміни кольорів зображення або його відтінків зазвичай залишаються непоміченими. Для характеристики втрат при стисненні використовують такі параметри.

1. Середня похибка:

$$\delta_{CP} = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{\Delta ij}{Xij}.$$

де Δij – різниця в числових характеристиках піксела, що стоїть в і рядку і в j стовпці вихідного зображення і стисненого;

Xij – значення піксела вихідного зображення в точці (i, j).

Оцінити середню помилку на зображенні можна за формулою:

$$\delta_{CP} = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{\Delta ij}{X_{MAX}}.$$

2. Максимальна помилка на зображенні:

$$\Delta_{MAX} = \max_{\substack{i=1...N \\ j=1...M}} \left(\frac{\Delta ij}{X_{MAX}} \right).$$

Перевага методів стиснення із втратами у тому, що коефіцієнт стиснення на порядок вищий. Тому ці методи використовуються в мережі Інтернет, телекомунікаціях, супутниковому зондуванні тощо.

Метод JPEG

У цьому методі зображення розкладається за двовимірним косинус-перетворенням. Розглянемо приклад.

Нехай є функція ні парна, ні непарна, тоді перетворення Фур'є має складові і косинуси, і синуси (парна має тільки косинуси, непарна синуси). Доповнимо її симетрично і вона стане парною.

Косинусний спектр вихідної функції буде мати вигляд як на рис. 5.29.

Рисунок 5.29 – Вихідна функція стиснення

Стиснення відбувається за рахунок відкидання ВЧ частини складової спектра, що має малі амплітуди.

У результаті з'являється розмитість (рис. 5.30).

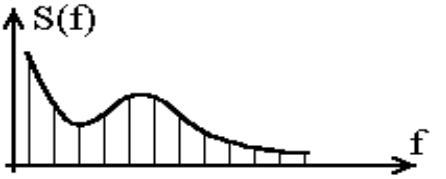


Рисунок 5.30 - Результат обробки

Даний метод стиснення є найбільш розповсюдженим внаслідок простоти реалізації. При цьому коефіцієнт стиснення методом JPEG – від десяти до декількох сотень. Метод зазвичай застосовується для зображень природного походження.

Фрактальний метод стиснення

Згідно з фрактальним методом зображення поділяється на певні подібні ділянки (фрактали), частина яких в подальшому відкидається (рис. 5.31).

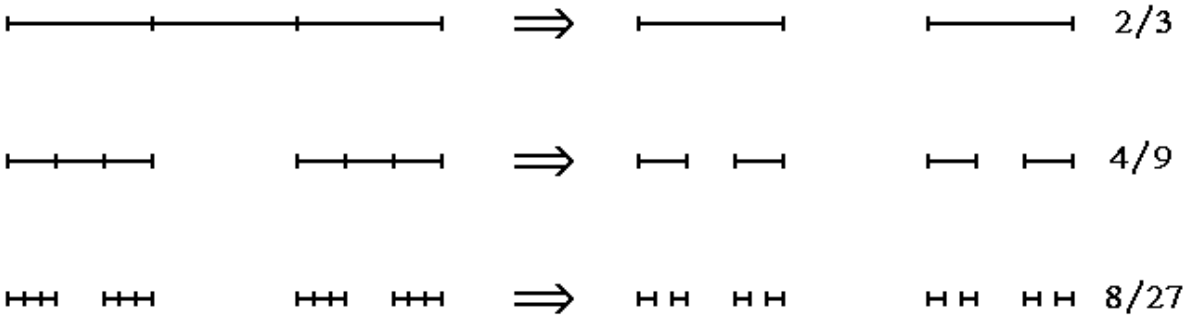


Рисунок 5.31 – Поділ зображення на фрактали

У результаті отримано множину, що має нульову довжину – це множина кінців відрізків – канторова множина (2^n / 3^n).

Фрактал – множина точок, результат канторової множини. Але не всякий фрактал є канторовою множиною. Фрактал має властивість самоподібності.

На початковому зображенні відшукуються самоподібні ділянки і вони беруться як базові елементи фрактального зображення. Останнє апроксимується за допомогою фрактальних перетворень, потім отримуємо зображення у вигляді формули, що відображає фрактальні перетворення.

Цей метод не отримав широкого поширення через ряд недоліків:

а) відсутність автоматичного перетворення для множини випадків.

б) похибка методу залежить від часу обчислення, щоб отримати гарне стиснення інформації необхідно витратити на стиснення декілька годин.

в) вузьке застосування – там, де необхідно створити за великий час компактний приклад сигналу.

SPITH-перетворення

Це перетворення базується на Wavelet-перетворенні. Відбувається нормування зображення за певною схемою і вводиться норма на пікселі, які множаться на вибрані коефіцієнти (рис. 5.32).

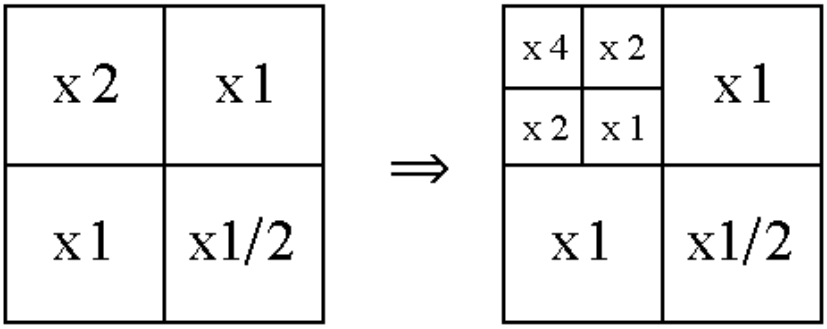


Рисунок 5.32 – Принцип SPITH-перетворення

Далі відкидаються ті коефіцієнти, які менші наперед заданого порогового значення. Процес повторюється доки зображення не стиснеться до необхідних меж або явно погіршиться якість.

До переваг даного методу можна віднести те, що при одних і тих самих коефіцієнтах стиснення втрат виходить менше, ніж при JPEG стисненні.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Які складові входять до типової системи передачі?
2. Які типи каналів зв’язку бувають?
3. Сформулюйте основні принципи завадостійкого кодування.
4. Які недоліки завадостійкого кодування?
5. Які ви знаєте коди з виявленням помилок? Дайте їх загальну характеристику.
6. Що являють собою систематичні коди?
7. Як працює кодер коду з перевіркою на парність?
8. Як працює декодер коду з перевіркою на парність?
9. Які коди відносяться до кодів з повторенням числа елементів?
10. Як працює кодер і декодер коду з прямим повторенням?
11. Кодер коду з інверсним повторенням. Алгоритм побудови.
12. Як працює декодер коду з інверсним повторенням?
13. Як працює кодер і декодер кореляційного коду?
14. Як кодується код Хемінга?
15. Алгоритм декодування коду Хемінга.
16. Побудуйте кодер коду Хемінга на 5 інформаційних розрядів.
17. Побудуйте декодер коду Хемінга (9, 5).
18. Яка залежність між кількістю інформаційних і контрольних символів коду Хемінга?
19. У чому сутність циклічного кодування?
20. Наведіть переваги та недоліки рекурентних кодів.
21. Розкрийте сутність алгоритму коду Шеннона-Фано.
22. Розкрийте сутність алгоритму коду Хаффмена.
23. Що розуміють під арифметичним кодуванням?
24. У чому відмінність динамічного кодування Хаффмена від звичайного?
25. Які модифікації має алгоритм динамічного кодування Хаффмена? Розкрийте їх суть.
26. Наведіть приклади методів стиснення зображень без втрат.
27. Які переваги та недоліки методів стиснення зображень із втратами?

6 ПРИЙМАННЯ ІНФОРМАЦІЇ В ЕЛЕКТРОННИХ СИСТЕМАХ

6.1 Оптимальне приймання та завадостійкість

Одною з важливих задач радіотехніки є задача знаходження сигналу на фоні завади. Нехай на вхід системи, яка розв'язує цю задачу, надходить коливання $u(t)$, яке при відсутності сигналу є шумом $n(t)$, тобто $u(t)=n(t)$, а при наявності сигналу в найпростішому випадку воно є сумою сигналу і шуму: $u(t)=n(t)+s(t)$.

Таким чином сигнал на вході приймача можна записати в такому вигляді

$$u(t)=n(t)+\lambda s(t),$$

де $\lambda=0$, якщо сигналу немає (позначимо цей випадок A_0) і $\lambda=1$, якщо він є (випадок A_1).

Величина λ наперед невідома і її необхідно визначити. Через випадковий характер шуму чи завади при розв'язанні задачі знаходження можливі помилки. Якщо приймач визначає, що $\lambda=0$, будемо вважати, – це випадок B_0, а якщо $\lambda=1$, то випадок B_1. При цьому можливі чотири варіанти розв'язування.

1. Сигналу немає і приймач визначає, що його немає – правильне знаходження сигналу (ситуація A_0B_0);

2. Сигналу немає, а приймач визначає, що він є – помилкове знаходження сигналу, хибна тревога (ситуація A_0B_1);

3. Сигнал є і приймач визначає, що він є – правильне знаходження сигналу (ситуація A_1B_1);

4. Сигнал є, а приймач визначає, що його немає – пропуск сигналу (ситуація A_1B_0).

Найбільш просто розв'язується задача знаходження повністю невідомого сигналу. Задача ускладнюється зі збільшенням випадкових невідомих параметрів (чи характеристик) знайденого сигналу. В цьому випадку сигнал описується не просто функцією часу, але й функцією випадкових параметрів $s(t,\alpha)$, де $\alpha=\alpha(\alpha_1,\alpha_2,...,\alpha_n)$ – матриця параметрів сигналу. Ще більш складною є задача розв'язання кількох сигналів. В цьому випадку на

вході приймача діє сигнал $u(t)$, який є сумою завади $n(t)$ і двох сигналів $s_1(t)$ та $s_2(t)$, які можливо накладаються, тобто

$$u(t) = \lambda_1 s_1(t) + \lambda_2 s_2(t) + n(t),$$

де λ_1 і λ_2 – випадкові величини, які можуть приймати значення 0 чи 1. Якщо сигнали не повністю відомі, то постає задача розв'язання сигналів з випадковими параметрами: $s_1(t,\alpha_1), s_1(t,\alpha_2)$.

В поняття "розв'язати два сигнали" можна вкладати різний зміст. При можливості одночасної наявності в прийнятій реалізації $u(t)$ двох сигналів, можна мати на увазі лише окреме знаходження сигналів із заданими показниками, а можна не тільки окреме знаходження, але й визначення будь-якого параметра α_i. Так, якщо показники знаходження другого сигналу залишаються вище допустимих в присутності випадкового першого сигналу, то можна говорити, що другий сигнал розв'язується по суті знаходження (оцінки параметра α_i). Якщо в доповнення до цього розв'язується і перший сигнал при наявності другого, то можна говорити, що сигнали взаємно розв'язується по суті знаходження (оцінки параметра α_i). Математичною базою для розв'язання всіх задач прийому сигналів, пов'язаних з ризиком помилок прийняття рішення про наявність сигналу при його відсутності, і пропуск сигналу, є теорія ймовірності і математичної статистики, при цьому забезпечується можливість врахування не тільки інформації про прийняте коливання, але й апріорні дані відносно сигналу й завад.

Характеристика завад

Завадами зазвичай називають сторонні електричні збурення, які накладаються на сигнал, що передається, й ускладнюють його прийом. При великій інтенсивності завад прийом стає практично неможливим.

Класифікація завад:

а) завади від сусідніх радіопередавачів (станцій);

б) завади від промислових установок;

в) атмосферні завади (грози, опади);

г) завади, обумовлені проходженням електромагнітних хвиль через шари атмосфери: тропосферу, іоносферу;

д) теплові і дробові шуми в елементах радіокіл, обумовлені тепловим рухом електронів.

В математичному описі завади є випадковими функціями часу. Випадкову функцію неперервного часу зазвичай називають випадковим процесом, її дискретний аналог – випадковою послідовністю. Як правило, завади належать до класу стаціонарних випадкових процесів і характеризуються як своїми розподіленнями й моментами розподілень, так і їх числовими параметрами. Основне розподілення більшості шумових сигналів – нормальне (гаусів процес). Це пояснюється тим, що розподілення сум незалежних випадкових величин, з яких складаються випадкові завади, сходиться до нормального, не залежно від характеру розподілення доданків (теорема Ляпунова).

Момент першого порядку виражає середнє значення (постійну складову) випадкового процесу:

$$M\{q\} = \bar{q} = \int_{-\infty}^{\infty} q \cdot p(q)\, dq,$$

де $p(q)$ – густина ймовірностей значень q.

Центральний момент другого порядку визначає дисперсію процесу:

$$D\{q\} = \sigma^2 = \int_{-\infty}^{\infty} (q - \bar{q})^2\, p(q) dq = \overline{q^2} - (\bar{q})^2.$$

Дисперсія виражає потужність змінної складової процесу. Корінь квадратний із значення дисперсії, тобто значення σ, є середнім квадратичним значенням розкиду випадкових значень q відносно середнього значення $\bar{q}$.

Змішаний момент другого порядку називається функцією автокореляції випадкового процесу $q(t)$:

$$M\{q(t)q(t+\tau)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2\, p(x_1, x_2) dx_1 dx_2 = B(\tau).$$

Величина $B(\tau)$ при $\tau = 0$ дорівнює загальній потужності випадкового процесу $q(t)$.

На практиці більшість випадкових процесів має властивість ергодичності. Вона полягає в тому, що середні значення за множиною реалізацій (математичні очікування, обчислені за густинами розподілів) збігаються з середніми значеннями за часом T однієї реалізації процесу при $T \to \infty$. Це дозволяє проводити оцінку числових значень параметрів завад безпосередньо за довільними інтервалами $[a, b]$ задання сигналів:

$$\bar{q} = \lim_{T\to\infty} \frac{1}{T}\int_0^T q(t)dt \approx \frac{1}{b-a}\int_a^b q(t)dt.,$$

$$\sigma^2 = \lim_{T\to\infty} \frac{1}{T}\int_0^T (q(t)-\bar{q})^2 dt \approx \frac{1}{b-a}\int_a^b q(t)-\bar{q})^2 dt.,$$

$$B(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T q(t)q(t+\tau)dt \approx \frac{1}{b-a}\int_a^b q(t)q(t+\tau)dt.$$

Спектральна густина потужності випадкового процесу (розподілення потужності завад і шумів за частотою) пов'язана з функцією автокореляції перетворенням Фур'є. В односторонньому (фізичному) поданні спектра:

$$B(f) = 4\int_0^\infty B(\tau)\cos 2\pi f\tau\, d\tau,$$

$$B(\tau) = \int_0^\infty B(f)\cos 2\pi f\tau\, d\tau.$$

Адитивну заваду з рівномірним спектром $B(f) = B_0 = const$ називають білим шумом. Потужність білого шуму в смузі частот $0 - F$ пропорційна ширині смуги:

$$W_F = \int_0^F B(f)df = B_0 F.$$

Сигнал с адитивною завадою зазвичай характеризують не абсолютною потужністю завади, а відношенням середніх потужностей сигналу і завади, яке коротко називають відношенням сигнал/завада:

$$\rho = W_c / W_q.$$

Значення випадкових процесів є некорельованими тільки при необмеженій смузі частот. Будь-яке обмеження частотної смуги вносить певну кореляцію в процес і незалежними один від одного можна вважати тільки значення процесу, які віддалені одне від одного як мінімум на інтервал кореляції τ_0:

$$\tau_0 = \frac{2}{W_F}\int_0^\infty B(t)dt = 1/2F.$$

Завадостійкість оптимального прийому безперервних повідомлень

Наявність завад утруднює прийом сигналів, при великій інтенсивності завад розпізнавання сигналу може стати практично неможливим.

Здатність системи протистояти негативному пливу завади носить назву завадостійкості.

Мірою завадостійкості оптимального прийому безперервних повідомлень є середньоквадратичне відхилення чи середній квадрат похибки.

Нехай $\varepsilon(t)$ – завада на виході приймача. Величина $M\{\varepsilon^2(t)\} = P_\varepsilon$ є потужністю завади. Потужність первинного сигналу (повідомлення) $b(t)$ дорівнює $P_b = M[b^2(t)]$. Замість середньоквадратичного відхилення часто використовують інший показник – відношення потужності сигналу до потужності шуму (на виході приймача):

$$\rho_{вих} = P_b / P_\varepsilon .$$

Якщо відношення сигнал/шум на вході приймача дорівнює $\rho_{вх} = P_c / P_ш$, то відношення

$$g = \frac{\rho_{вих}}{\rho_{вх}} = \frac{P_b}{P_\varepsilon} / \frac{P_c}{P_ш} = \frac{P_b}{P_c} / \frac{P_\varepsilon}{P_ш}$$

характеризує виграш (чи програш) в завадостійкості за рахунок використання того чи іншого способу обробки сигналу.

Корисним показником також є узагальнений виграш, що дорівнює:

$$g' = \frac{\rho'_{вих}}{\rho'_{вх}} = \frac{g}{\alpha},$$

де $\rho'_{вих} = \dfrac{P_b}{P_\varepsilon / F_c}$;

F_c – ширина спектра повідомлення;

$\rho'_{вх} = \dfrac{P_c}{P_ш / F}$;

F – ширина спектра сигналу;

$\alpha = F / F_c$.

Потенційна завадостійкість сигналів з амплітудною, частотною та фазовою модуляцією

Для передачі елементів двійкового коду (0 і 1) зазвичай використовують сигнали з дискретною амплітудною модуляцією (АМ), частотною модуляцією (ЧМ) і фазовою модуляцією (ФМ).

В процесі передачі елементи коду спотворюються завадами, причому спостерігаються помилки двох видів.

1. При передачі елемента 0 може бути прийнятий елемент 1, ймовірність такої події (переходу $0 \to 1$) позначимо через $p(1/0)$.

2. При передачі елемента 1 може бути прийнятий елемент 0, ймовірність такої події (переходу $1 \to 0$) позначимо через $p(0/1)$.

Середня ймовірність помилки визначається за формулою:

$$P_{пом} = P(0)p(1/0) + P(1)p(0/1).$$

В подальшому будемо вважати, що апріорні ймовірності передачі елементів коду рівні, тобто $P(0) = P(1) = 0{,}5,$ при цьому:

$$P_{пом} = 0{,}5\left[p(1/0) + p(0/1)\right].$$

Заваду в каналі зв'язку будемо вважати флуктуаційною з нормальним законом розподілу миттєвих значень:

$$W(\xi) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{\xi^2}{\sigma^2}}.$$

Завадостійкість прийому сигналів АМ, ЧМ, ФМ при вказаних вище умовах можна визначити, обчислюючи середню ймовірність помилки таким чином.

Амплітудна модуляція

Елементами сигналу АМ є посилка (кодовий елемент "1") і пауза (кодовий елемент "0"):

$$\left.\begin{array}{l} S_1(t) = a \cdot \sin_0 t, \\ S_2(t) = 0 \end{array}\right\}, 0 \le t \le T,$$

де T – тривалість елементу сигналу.

Прийом сигналу АМ в випадку некогерентного прийому здійснюється шляхом порівняння рівня сигналу після амплітудного детектора (детектора обвідної) з деяким пороговим рівнем U_n. Помилки виникають у випадках.

1. При передачі посилки обвідна суми сигналу і завади ($E_{сп}$) виявляється менша порогового рівня U_n (перехід $1 \to 0$).

2. При передачі паузи обвідна завади E виявляється більша U_n (перехід $0 \to 1$).

Ймовірність першої події рівна

$$\left.\begin{aligned} p(0/1) &= \int_{\infty}^{U_n} \omega(E_{cn})dE_{cn}, \\ p(1/0) &= \int_{U_n}^{0} \omega(E_n)dE_{cn} \end{aligned}\right\},$$

ймовірність другої

$$W(E_{cn}) = \frac{E_{cn}}{\sigma^2}\exp\left[\frac{-E_{cn}{}^2 + a^2}{2\sigma^2}\right] I_0\left[\frac{E_{cn}a}{\sigma^2}\right],$$

де $W(E_{cn})$ – густина розподілу обвідної суми сигналу і завади, яка, як відомо, визначається узагальненим законом Релея;

$W(E_c) = \frac{E_c{}^2}{\sigma^2}\exp\left[\frac{-E_n{}^2}{2\sigma^2}\right]$ –густина розподілу обвідної завади (простий закон Релея).

Середня ймовірність похибки рівна

$$P_{помАМ} = 0{,}5\left[\int_0^{U_n} \omega(E_{cn})dE_{cn} + \int_{U_n}^{\infty} \omega(E_n)dE_n\right].$$

Значення $P_{помАМ}$ залежить від порогового рівня U_n. Можна показати, що ймовірність помилки мінімальна, коли $U_n = \frac{1}{2}a$ (якщо $a^2 >> \sigma^2$).

Після обчислення інтеграла отримаємо

$$P_{помАМ} = 0{,}5(0{,}5\left[1 - \Phi\left(\frac{h}{\sqrt{2}}\right)\right] + \exp\left(-\frac{h^2}{4}\right)),$$

де $h^2 = \frac{a^2}{2\sigma^2}$ – відношення потужностей сигналу і завади (відношення сигнал/шум);

$\Phi(z) = \frac{2}{\sqrt{2\pi}}\int_0^Z e^{\frac{-X^2}{2}}dx$ – табульований інтеграл ймовірностей.

Якщо $h^2 >> 1$, то $P_{помАМ} \approx 0{,}5\exp\left(-\frac{h^2}{4}\right)$.

Максимальна завадостійкість при прийомі сигналів АМ спостерігається в тому випадку, якщо перед детектором застосовується оптимальна фільтрація сигналів. При цьому забезпечується максимальне відношення сигнал/шум, рівне h^2.

При когерентному прийманні застосовується синхронний детектор, який усуває ортогональну складову вектора завади. Складова $x = E_n \cos\Theta$ має нормальний закон розподілу і модальність $\overline{x^2} = \sigma^2$. Тому ймовірність спотворення посилки $p(0/1)$ і ймовірність спотворення паузи $p(1/0)$ будуть рівні:

$$p(0/1) = \int_0^{U_n} W(x/a)dx \text{ і } p(1/0) = \int_{U_n}^{\infty} W(x)dx,$$

де $W(x/a) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[\frac{(x-a)^2}{2\sigma^2}\right]$ та $W(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{x^2}{2\sigma^2}\right]$ – густини розподілу ймовірностей сигналів на виході детектора при прийманні посилки та паузи.

Середня ймовірність помилки

$$P_{помАМ} = 0{,}5\left[\int_0^{U_n} W(x/a)dx + \int_{U_n}^{\infty} W(x)dx\right].$$

Якщо $U_n = 0{,}5a$, то ймовірність помилки мінімальна і дорівнює

$$P_{помАМ} = 0{,}5\left[1 - \Phi\left(\frac{h}{\sqrt{2}}\right)\right],$$

де $h^2 = \frac{a^2}{2\sigma^2}$ – відношення сигнал/шум.

Максимальна (потенційна) завадостійкість прийому сигналів АМ має місце, коли після детектора здійснюється оптимальна фільтрація сигналу. При цьому досягається максимальне відношення сигнал/шум.

Частотна модуляція

Елементами сигналу при ЧМ є

$$\left.\begin{aligned} S_1(t) &= a \cdot \sin\omega_1 t, \\ S_2(t) &= a \cdot \sin\omega_2 t \end{aligned}\right\}, 0 \le t \le T.$$

В приймачі сигнали розділяються з допомогою фільтрів, налаштованих на частоти ω_1 і ω_2 з подальшим детектуванням.

При некогерентному прийомі сигналів ЧМ в одному із фільтрів завжди присутня сума сигналу і завади, а в другому лише завада. Помилка при реєстрації сигналу, очевидно, буде в тому випадку, коли обвідна завади в фільтрі без сигналу перевищить обвідну суми сигналу і завади в фільтрі із сигналом.

Враховуючи, що потужності сигналів і завади в кожному із фільтрів однакові (канал симетричний), ймовірності спотворення "1" і "0" будуть однакові, тобто $p(0/1) = p(1/0)$.

Ймовірність того, що обвідна завади в фільтрі буде більша обвідної суми сигналу і завади в другому фільтрі, рівна

$$p(E_n > E_{cn}) = \int_{E_{cn}}^{\infty} W(E_n) dE_n.$$

Обвідна суми сигналу і завади є випадковою величиною, яка має узагальнений закон розподілу Релея. Тому для визначення ймовірності помилки необхідно усереднити ймовірність $p(E_n > E_{cn})$ за всіма значеннями E_{cn}:

$$p(0/1) = p(1/0) = \int_0^{\infty} W(E_{cn}) p(E_n > E_{cn}) dE_{cn} = \int_0^{\infty} W(E_{cn}) \left[\int_{E_{cn}}^{\infty} W(E_n) dE_n \right] dE_{cn}.$$

Підставляючи замість $W(E_{cn})$ і $W(E_n)$ їх значення і обчислюючи інтеграл, отримаємо

$$p(0/1) = p(1/0) = \frac{1}{2} e^{-\frac{h^2}{2}},$$

де h^2 – відношення сигнал/шум на виході фільтра з сигналом.

Середня ймовірність помилки рівна:

$$P_{помАМ} = 0{,}5[p(1/0) + p(0/1)] = \frac{1}{2} e^{-\frac{h^2}{2}}.$$

Максимальна завадостійкість досягається при прийманні сигналів ЧМ в тому випадку, коли здійснюється оптимальна фільтрація сигналу. При цьому $h^2 = h_0^2$.

При когерентному прийманні сигналів ЧМ на завадостійкість впливають тільки синфазні із сигналом складові завад x_1 в фільтрі ω_1 і x_2 в фільтрі ω_2. Ці складові мають нормальний закон розподілу амплітуд

$$W(x_1) = W(x_2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{x^2}{2\sigma^2} \right]$$

і однакові дисперсії $\overline{x_1^2} = \overline{x_2^2} = \sigma^2$.

Ймовірність перевищення значення обвідної завади в фільтрі без сигналу x_2 обвідною суми сигналу і завади в фільтрі з сигналом $a + x_1$ дорівнює ∞.

Для визначення ймовірності помилки необхідно усереднити ймовірність $p(x_2 > a + x_1)$ за всіма значеннями випадкової величини:

$$p(0/1) = p(1/0) = \int_{-\infty}^{\infty} W(a + x_1) \left[\int_{a+x_1}^{\infty} W(x_2) dx_2 \right] dx_1 = 0{,}5[1 - \Phi(h)],$$

де h^2 – відношення сигнал/шум.

Середня ймовірність помилки рівна

$$P_{помАМ} = 0{,}5[p(1/0) + p(0/1)] = 0{,}5[1 - \Phi(h)].$$

Максимальна завадостійкість при когерентному прийманні сигналів ЧМ досягається при оптимальній фільтрації сигналів, коли $h^2 = h_0^2$.

Фазова модуляція

Елементами сигналу при ФМ є:

$$\left.\begin{aligned} S_1(t) &= a \cdot \sin \omega_1 t, \\ S_2(t) &= -a \cdot \sin \omega_1 t \end{aligned}\right\}, \; 0 \le t \le T.$$

Приймання сигналів фазової модуляції можливе лише з допомогою синхронного (когерентного) детектора, який розрізняє фази сигналів.

Ймовірності переходів $p(0/1)$ і $p(1/0)$ при флуктуаційній заваді в каналі зв'язку однакові і рівні:

$$p(0/1) = p(1/0) = \frac{1}{\sqrt{2\pi}\sigma} \int_a^{\infty} \exp\left(-\frac{x^2}{2\sigma^2} \right) dx = 0{,}5\left[1 - \Phi(\sqrt{2}h)\right]$$

Середня ймовірність помилки

$$P_{помАМ} = 0{,}5[p(1/0) + p(0/1)] = 0{,}5\left[1 - \Phi\left(\sqrt{2}h\right)\right].$$

Максимальна завадостійкість сигналу ФМ досягається при оптимальній фільтрації сигналів, коли $h^2 = h_0^2$.

Розподіл каналів при багатоканальній передачі безперервних повідомлень

Використовувані методи розподілу каналів можна розподілити на лінійні і нелінійні (комбінаційні).

В більшості випадків розподілу каналів кожному джерелу повідомлень виділяється спеціальний сигнал, названий канальним. Промодульовані повідомленнями канальні сигнали об'єднуються, в результаті чого утворюється груповий сигнал. Якщо операція об'єднання лінійна, то результуючий сигнал називають лінійним груповим сигналом.

Для уніфікації багатоканальних систем зв'язку за основний чи стандартний канал приймають канал тональної частоти (канал ТЧ), який забезпечує передачу повідомлень зі смугою частот 300…3400 Гц, яка ефективно передається і відповідає основному спектру телефонного сигналу.

Багатоканальні системи утворюються шляхом об'єднання каналів ТЧ в групи, зазвичай кратні 12 каналам. В свою чергу, часто використовують "вторинне ущільнення" каналів ТЧ телеграфними каналами і каналами передачі даних.

На рис. 6.1. наведена узагальнена структурна схема системи багатоканального зв'язку.

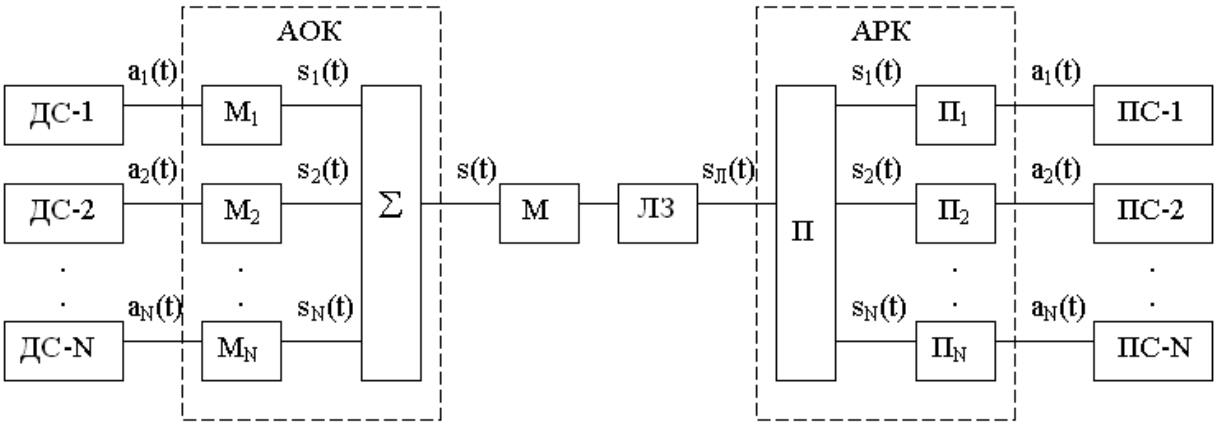


Рисунок 6.1 – Узагальнена структурна схема системи багатоканального зв'язку

Реалізація повідомлень кожного джерела $a_1(t), a_2(t),..., a_N(t)$ з допомогою індивідуальних передатчиків (модуляторів) $M_1, M_2,..., M_N$ перетворюється у відповідні канальні сигнали $s_1(t), s_2(t),..., s_N(t)$. Сукупність канальних сигналів на виході апаратури об'єднання каналів утворює груповий сигнал $s(t)$. В груповому передавачі M сигнал $s(t)$ перетворюється в лінійний сигнал $s_л(t)$, який і надходить в лінію зв'язку (ЛЗ). Припустимо, що лінія пропускає сигнал практично без спотворень і не вносить шумів. Тоді на приймальному кінці лінії зв'язку лінійний сигнал $s_л(t)$ з допомогою апаратури розділення каналів може бути знову перетворений в груповий сигнал $s(t)$. Канальними чи індивідуальними приймачами $П_1, П_2,..., П_N$ з групового сигналу $s(t)$ виділяються відповідні канальні сигнали $s_1(t), s_2(t),..., s_N(t)$ і потім перетворюються в призначені отримувачам повідомлення $a_1(t), a_2(t),..., a_N(t)$.

Канальні передавачі разом з пристроєм додавання утворюють апаратуру об'єднання. Груповий передавач M, лінія зв'язку ЛЗ і груповий приймач $П$ утворюють груповий канал зв'язку (тракт передачі), який разом с апаратурою об'єднання і індивідуальними приймачами утворюють систему багатоканального зв'язку.

Індивідуальні приймачі системи багатоканального зв'язку $П_k$ поряд з виконанням звичайної операції перетворення сигналів $s_k(t)$ в відповідні повідомлення $a_k(t)$ повинні забезпечити виділення сигналів $s_k(t)$ із групового сигналу $s(t)$. Інакше кажучи, в складі технічних засобів на передавальній стороні багатоканальної системи повинна бути передбачена апаратура об'єднання, а на приймальній стороні – апаратура розділення.

Щоб пристрої розділення були здатні розрізняти сигнали окремих каналів, повинні існувати певні ознаки, властиві тільки даному сигналу. Такими ознаками можуть бути параметри перенесення, наприклад, амплітуда, частота чи фаза у випадку неперервної модуляції гармонічного перенесення. При дискретних видах модуляції ознакою може служити і форма сигналів. Відповідно розрізняються і способи розділення сигналів: частотний, часовий, фазовий та інші. Ознайомимось з деякими із них.

Розглянемо частотний розділення каналів.

Функціональна схема простої системи багатоканального зв'язку з розділенням каналів за частотою подана на рис. 6.2. Спочатку відповідно до переданих повідомлень первинні (індивідуальні) сигнали, з енергетичними спектрами $G_1(\omega), G_2(\omega),..., G_N(\omega)$ модулюють піднесучі частоти ω_k кожного каналу, відповідно. Цю операцію виконують модулятори $M_1, M_2,..., M_N$ канальних передавачів.

Модулятори – це чотириполюсники з нелінійною амплітудною характеристикою, яка в загальному випадку апроксимується поліномом n-го степеня:

$$U_{вих} = a_1 U_{вх} + a_2 U^2{}_{вх} + ... + a_n U^n{}_{вх},$$

де $a_1, a_2,..., a_n$ – коефіцієнти апроксимації.

Для простоти візьмемо поліном другого степеня, тобто:

$$U_{вих} = a_1 U_{вх} + a_2 U^2{}_{вх}.$$

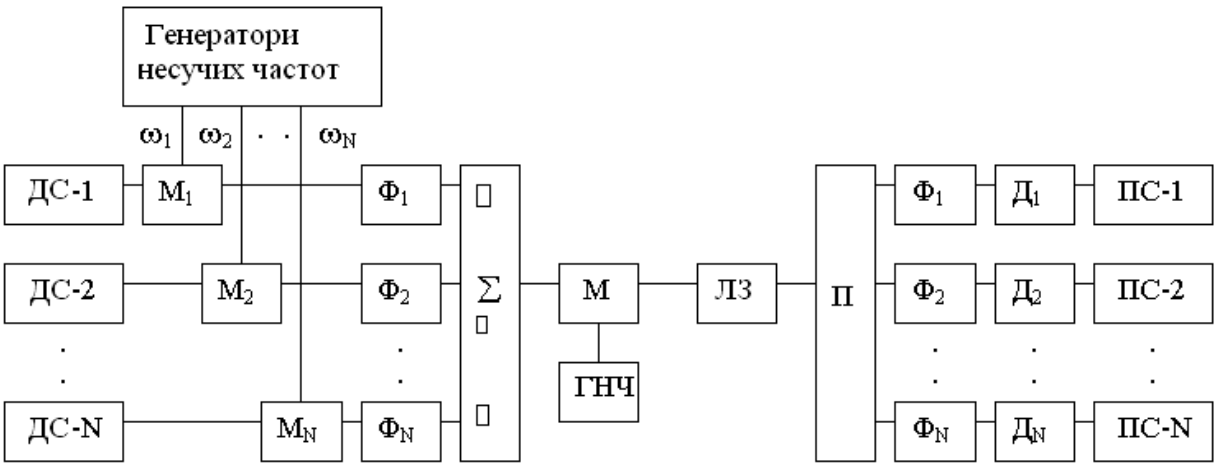


Рисунок 6.2 – Функціональна схема простої системи багатоканального зв'язку з розподілом каналів за частотою

Подамо на такий чотириполюсник сигнали двох частот, тобто

$$U_{вх} = U_{m1}\cos\omega t + U_{m2}\cos\Omega t,$$

де $\omega > \Omega$.

Тоді

$$U_{вих}(t) = a_1(U_{m1}\cos\omega t + U_{m2}\cos\Omega t) + a_2(U_{m1}\cos\omega t + U_{m2}\cos\Omega t)^2 = a_1U_{m1}\cos\omega t +$$
$$+ a_1U_{m2}\cos\Omega t + a_2U^2{}_{m1}\cos^2\omega t + a_2U^2{}_{m2}\cos^2\Omega t + 2a_2U_{m1}U_{m2}\cos\omega t\cos\Omega t.$$

Після відповідних перетворень отримаємо:

$$U_{вих}(t) = \frac{a_2}{2}(U^2{}_{m1} + U^2{}_{m2}) + a_1U_{m1}\cos\omega t + a_1U_{m2}\cos\Omega t + \frac{a_2U^2{}_{m2}}{2}\cos 2\omega t +$$
$$+ \frac{a_2U^2{}_{m1}}{2}\cos 2\Omega t + a_2U_{m1}U_{m2}\cos(\omega+\Omega)t + a_2U_{m1}U_{m2}\cos(\omega-\Omega)t.$$

Спектр сигналу на виході чотириполюсника буде мати вигляд (рис. 6.3):

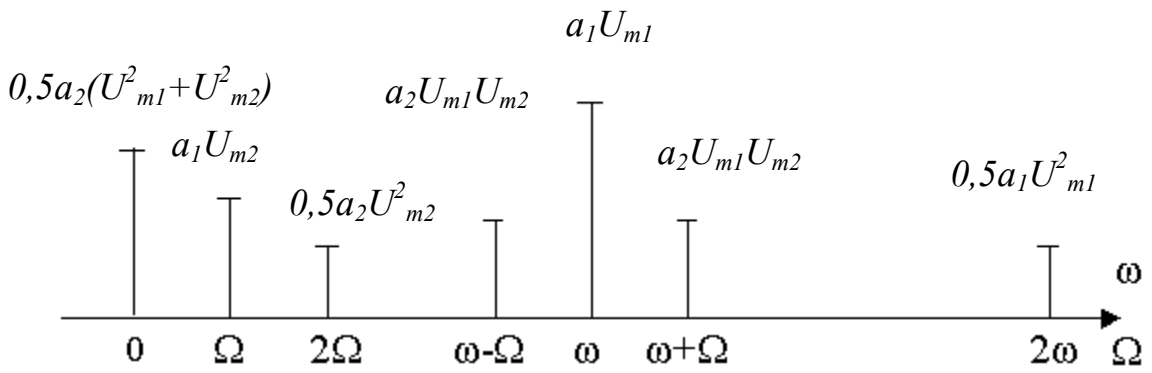


Рисунок 6.3 – Спектр сигналу на виході чотириполюсника

Таким чином, на виході чотириполюсника поряд з частотами вхідних сигналів (ω,Ω) з'явились постійна складова $\frac{a_2U^2{}_{m2}}{2}+\frac{a_2U^2{}_{m1}}{2}$, другі гармоніки вхідних сигналів 2ω і 2Ω та складові сумарної $(\omega+\Omega)$ та різницевої частот $(\omega-\Omega)$.

Якщо на чотириполюсник подати сигнал несучої частоти $U_1=U_m\cos\omega_i t$ і сигнал тональної частоти в смузі $\Omega_н...\Omega_в$ (де $\Omega_н$ = 0,3 кГц, $\Omega_в$ = 3,4 кГц), то спектр сигналу на виході чотириполюсника буде мати вигляд (рис. 6.4):

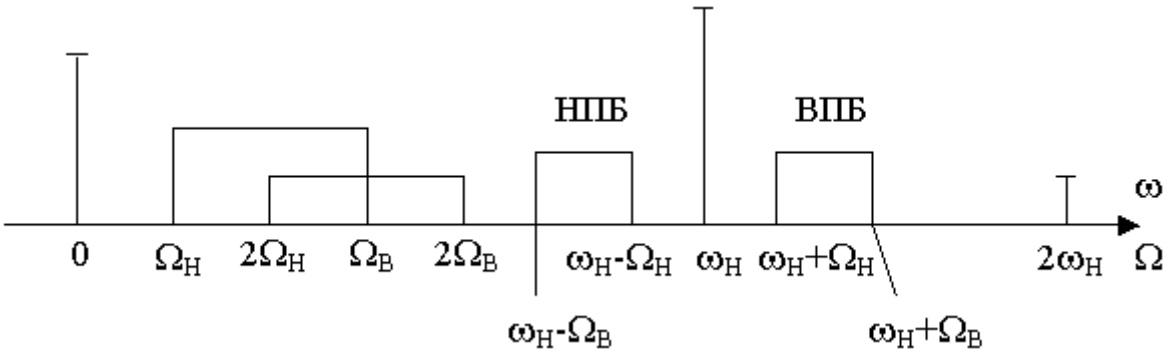


Рисунок 6.4 – Спектр сигналу на виході чотириполюсника

Корисними продуктами перетворення (модуляції) є верхня і нижня бокові смуги. Для відтворення сигналу на прийомі на вхід демодулятора достатньо подати несучу частоту $(\Omega_н)$ і одну із бокових частот.

В багатоканальних системах передачі з частотним розділенням каналів по каналу передається тільки сигнал однієї бокової смуги, а несуча частота береться від місцевого генератора. Таким образом, на виході кожного канального модулятора включається смуговий фільтр зі смугою пропускання $\Delta\omega=\Omega_н-\Omega_в=3{,}1$кГц. Спектри $G_1(\omega), G_2(\omega),..., G_N(\omega)$ після транспонування (переносу) на різні частотні інтервали і інверсування складаються і утворюють груповий спектр $G_{гр}(\omega)$.

З метою зменшення впливу сусідніх каналів (зменшення перехідних завад), обумовленого неідеальністю АЧХ фільтрів, між спектрами сигнальних повідомлень вводяться захисні інтервали. Для каналів ТЧ вони рівні 0,9 кГц. Таким чином, ширина смуги каналу ТЧ з врахуванням захисного інтервалу дорівнює 4 кГц (рис. 6.5).

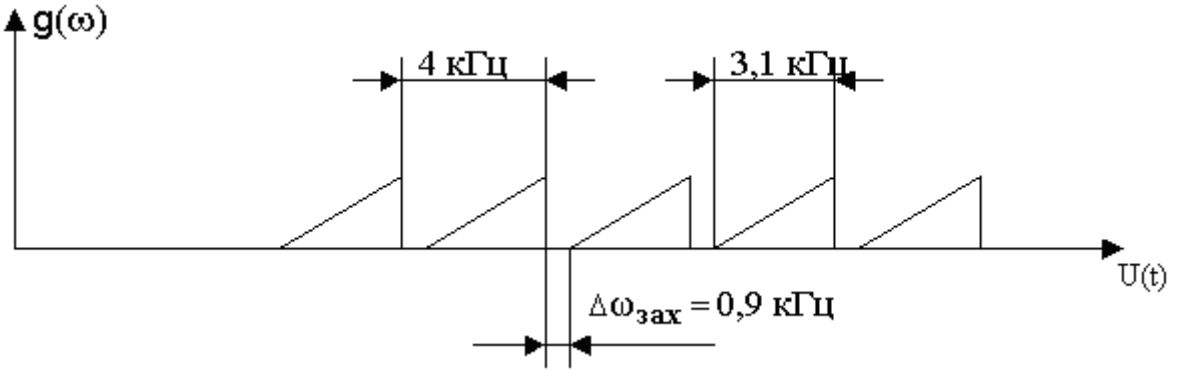


Рисунок 6.5 – Спектр групового сигналу із захисними інтервалами

Розглянемо часовий розподіл каналів.

При формуванні сигналу лінійного тракту систем передачі при ЧРК на передавальній стороні неперервні сигнали від абонентів передаються почергово (рис. 6.6)

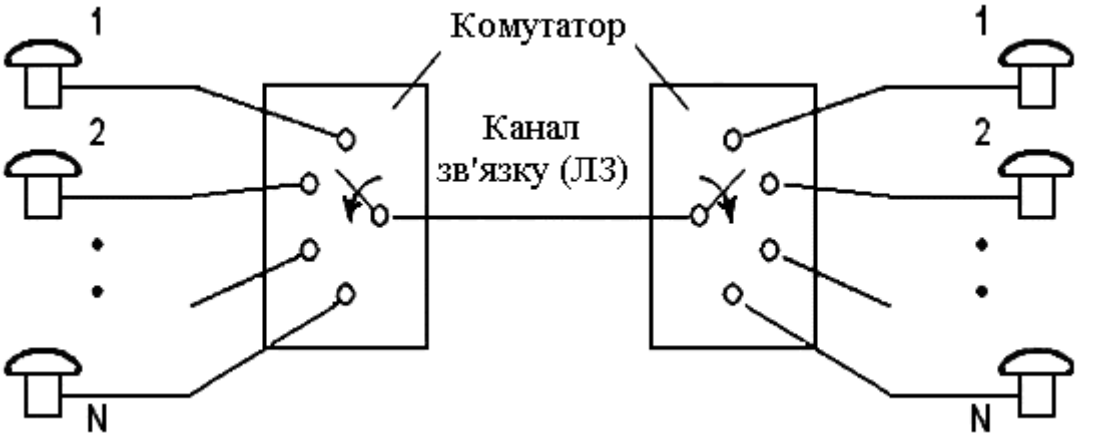


Рисунок 6.6 – Принцип часового розподілу каналів

Для цього ці сигнали перетворюються в ряд дискретних значень, періодично повторюваних через визначені інтервали часу $T_Д$, які називаються періодом дискретизації. Згідно з теоремою Найквіста-Котельникова період дискретизації неперервного, обмеженого за спектром сигналу з верхньою частотою $F_н << F_в$ повинен дорівнювати

$$T_Д = 1/F_Д, \quad F_Д \geq 2F_в,$$

де $F_Д$ – частота дискретизації.

Інтервал часу між найближчими імпульсами групового сигналу T_K називається канальним інтервалом чи тайм-слотом (Time Slot).

Із принципу часового об'єднання сигналів випливає, що передача в таких системах здійснюється циклами, тобто періодично у вигляді груп із $N_{гр} = N + n$ імпульсів, де N – кількість інформаційних сигналів, n – кількість службових сигналів (імпульсів синхронізації – ІС, службові зв'язки, управління і виклики). Тоді величина канального інтервалу:

$$\Delta t_K = T_Д / N_{гр}.$$

Таким чином, при ЧРК повідомлення від N абонентів і додаткових пристроїв передаються по спільному каналу зв'язку у вигляді послідовності імпульсів, тривалість кожного з яких $\Delta t_И = \Delta t_K$ (рис. 6.7, 6.8).

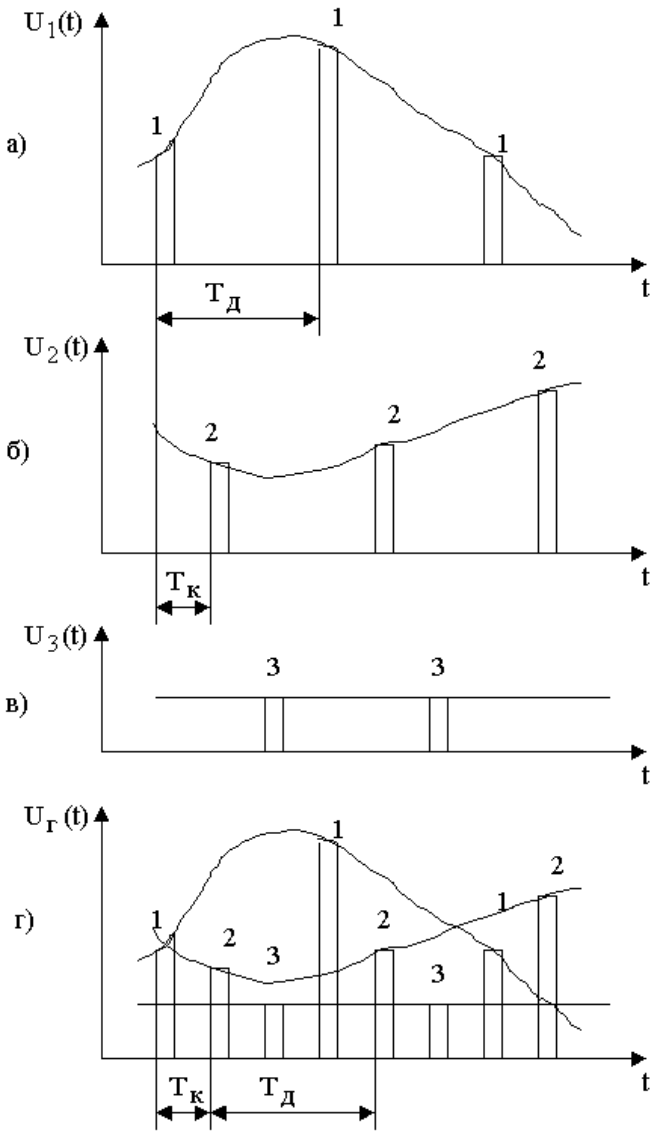


Рисунок 6.7 – Перетворення сигналів при ЧРК

При часовому розподілі каналів можливі такі види імпульсної модуляції (рис. 6.9): АІМ – амплітудноімпульсна модуляція; ШІМ – широтноімпульсна модуляція; ФІМ – фазоімпульсна модуляція.

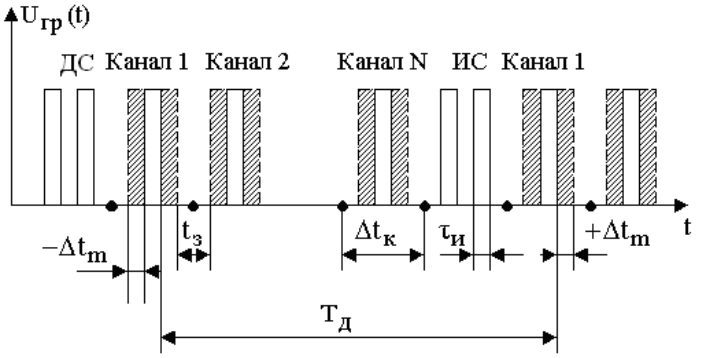


Рисунок 6.8 – Груповий сигнал при ЧРК з ФІМ

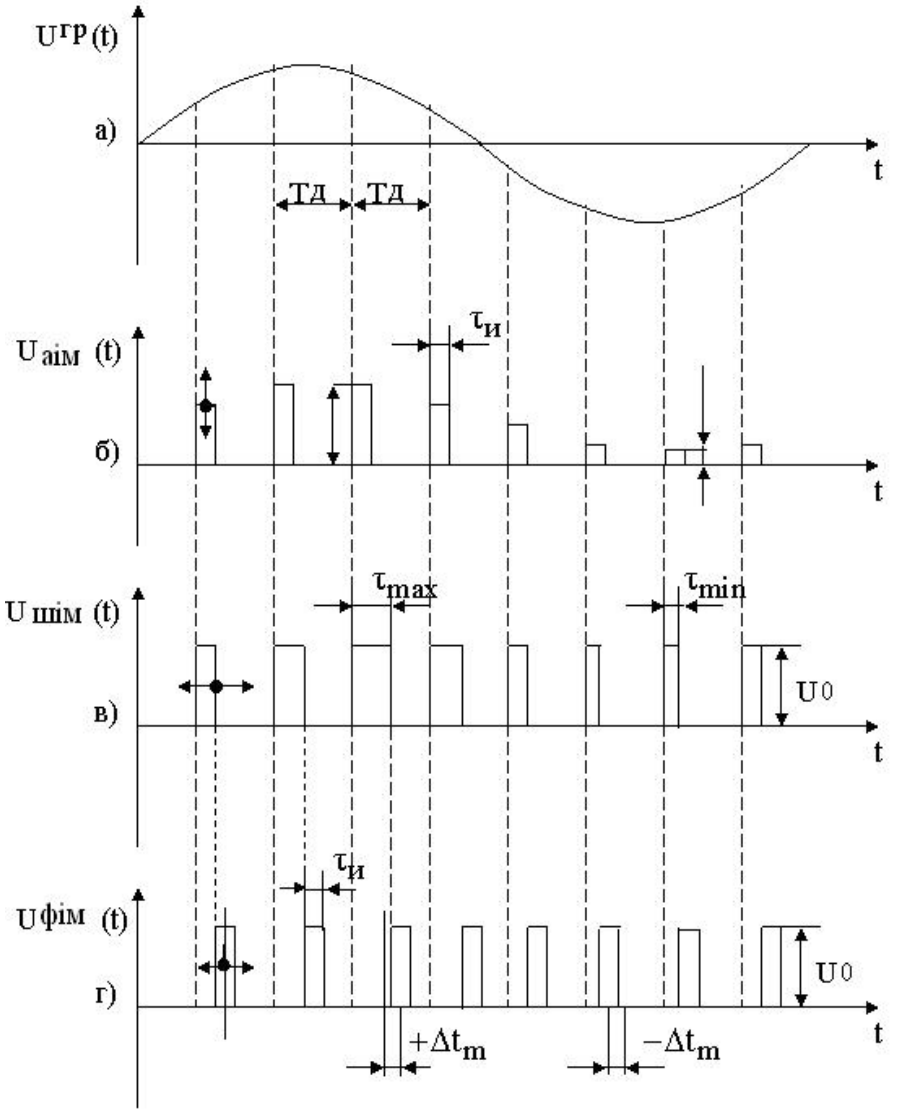


Рисунок 6.9 – Модуляція канальних імпульсів при ЧРК: а) неперервне повідомлення; б) АІМ; в) ШІМ; г) ФІМ

Зазвичай приймають $\omega_Д = (2{,}3\ \ldots\ 2{,}4)\Omega_B$ і при дискретизації телефонного повідомлення зі смугою частот 0,3…3,4 кГц частоту дискретизації $F_Д = \omega_Д / 2\pi$ вибирають рівною 8 кГц.

6.2 Виявлення та приймання інформації

Розпізнавання дискретних повідомлень при наявності завад

Когерентне приймання сигналів здійснюється за таких умов:

1) передаванні сигнали $s_i(x), i = 1,..., m$ повністю відомі;

2) канал зв'язку має відомі параметри;

3) завада $n(x)$ носить адитивний характер, має гауссівську густину ймовірності і відому спектральну густину G_n;

4) синхронізація сигналів – ідеальна.

Уявимо реальний сигнал моделлю

$$\xi(x) = s_i(x) + n(x),$$

де $\xi(x)$ – m-вимірний вектор, який враховує всі можливі передаванні сигнали $s_i(x)$ (для конкретного сигналу $s_i(x)$ (m-1) компонентів вектора $\xi(x)$ є нульовими).

Необхідно забезпечити оптимальну різницю сигналів $s_i(x)$.

Сигнали s_i є детермінованими і відомими, тому густина ймовірності прийнятого сигналу $\xi(x_k), (k = 1,..., K)$ повністю визначається K-вимірною густиною ймовірності значень завади $p_n(n)$, тобто, функція правдоподібності є

$$p(\xi / s_i) = p_n(n).$$

У випадку гауссівської завади

$$p(\xi / s_i) = \left[1/\left(2\pi\sigma^2\right)^{K/2}\right] \exp\left[-1/\left(2\sigma^2\right)\sum_{k=1}^{K} n_k^2\right],$$

де $\sigma^2 = G_n \Delta u$;

Δu – ширина смуги частот;

n_k – відліки завади.

Енергія завади на інтервалі L рівна

$$\int_0^L n^2(x)dx = (1/2\Delta u)\sum_{k=1}^{K} n_k^2,$$

чи

$$(1/2\sigma^2)\sum_{k=1}^{K} n_k^2 = (2\Delta u/2\sigma^2)\int_0^L n^2(x)dx = (1/G_n)\int_0^L n^2(x)dx.$$

Оскільки $n(x) = \xi(x) - s_i(x)$, то

$$p(\xi/s_i) = \left[1/(2\pi\sigma^2)^{K/2}\right]\exp\left[-(1/2G_n)\int_0^L [\xi(x) - s_i(x)]^2 dx\right].$$

Алгоритм оптимального когерентного прийому за критерієм максимальної правдоподібності полягає в отриманні максимального за i значення в останньому виразі. Ця умова забезпечується при $\int_0^L [\xi(x) - s_i(x)]^2 dx \xrightarrow{i} \min$.

Можна показати, що алгоритм забезпечує у вказаних умовах також мінімум повної ймовірності помилки і відповідає критерію ідеального спостерігача.

Для двійкових сигналів (імпульсів) є два варіанти $s_1(x)$ чи $s_2(x)$. При цьому для їх розпізнавання перевіряється умова

$$\int_0^L [\xi(x) - s_1(x)]^2 dx < \int_0^L [\xi(x) - s_2(x)]^2 dx.$$

Після розкриття дужок отримуємо нерівність

$$\int_0^L \xi(x)s_1(x)dx - \int_0^L \xi(x)s_2(x)dx > \frac{1}{2}[\int_0^L s_1{}^2(x)dx - \int_0^L s_2{}^2(x)dx],$$

де в лівій частині записана різниця кореляційних інтегралів, а в правій – різниця енергій сигналів.

Блок-схема розпізнавання двох сигналів показана на рис. 6.10.

Для розпізнавання сигналів $s_i(x), (i = 1,...,m)$ одним із найбільш ефективних методів є використання узгоджених фільтрів. Відомо, що такі фільтри забезпечують найбільше відношення сигналу до шуму на виході фільтра.

Узгоджений фільтр дає сигнал на виході:

$$\int_0^L \xi(x-\chi)\,h_i(\chi)d\chi\big|_{x=L} = \int_0^L \xi(L-\chi)\,s_i(L-\chi)d\chi,$$

де імпульсний відгук фільтра $h_i(x)$ є оберненою копією сигналу $s_i(x)$.

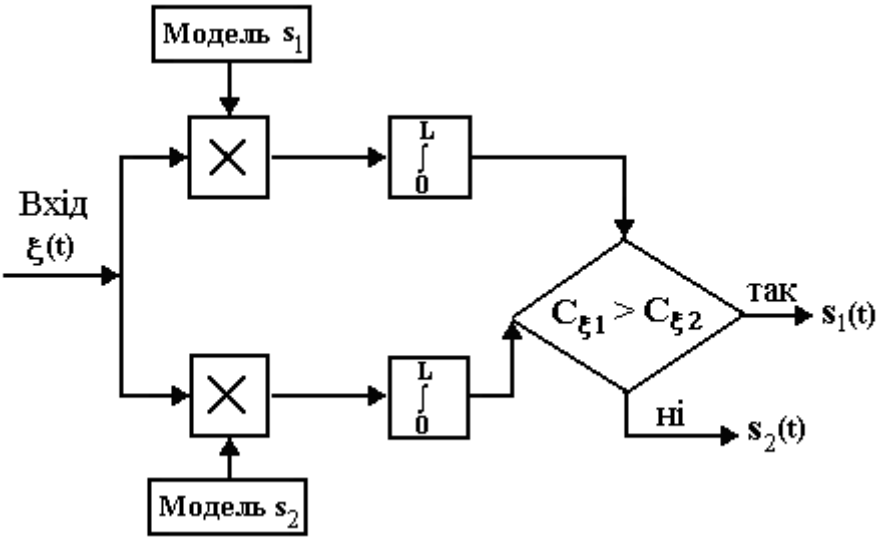


Рисунок 6.10 – Структурна схема алгоритму розпізнавання двох сигналів

Оскільки сигнали, що приймаються, $\xi(x)$ є реалізаціями стаціонарного випадкового процесу, значення інтеграла в останньому виразі не залежить від зсуву аргументів підінтегральних функцій і дорівнює кореляційному інтегралу

$$\int_0^L \xi(t)\, s_i(t) dt = C_{ki}.$$

Для обчислення різниці кореляційних інтегралів $C_{k1} - C_{k2}$ в схемі на рис. 6.10 можна задати імпульсний відгук узгодженого фільтра в формі

$$h(x) = s_1(T - \chi) + s_2(T - \chi)$$

і використовувати тільки один фільтр замість двох помножувачів. Запуск фільтра забезпечується схемою синхронізації когерентного приймача в момент $x = 0$ і зняття показів в кінці інтервалу L.

Якщо енергія різниці двох розпізнаваних сигналів W_d не перевищує енергію завади на виході узгодженого фільтра, тобто

$$\int_0^L \left[s_2(x) - s_1(x)\right]^2 dx < 2\int_0^L n(x)\left[s_1(x) - s_2(x)\right] dx,$$

то має місце помилковий прийом сигналу. Ймовірність помилок P_0, яка виникає через вплив гауссівського шуму, визначається інтегруванням гауссівської густини ймовірності $p_n(n)$ шуму в формі

$$P_0 = \int p_n(n)\, dn = 1 - \Phi\left(\sqrt{W_d / 2G_n}\right),$$

де $\Phi(y) = \left(2/\sqrt{2\pi}\right)\int_0^y \exp\left(-z^2/2\right) dz$.

Величина $S/N = \dfrac{W_d}{2G_n}$ характеризує відношення сигналу до шуму. При заданому відношенні S/N ймовірність випадкового прийому можна знайти із останньої формули, використовуючи табульовані значення функції $\Phi(y)$.

При некогерентному прийомі моменти появи випадкового за формою сигналу (його фаза ε) розглядаються як значення випадкової змінної. При цьому математичне сподівання функції правдоподібності можна виразити в формі

$$< p(\xi / s_i) >= \int_{-x}^{x} \left(1/2\pi\sigma^2\right)^{K/2} \exp\left\{-(1/G_n)\int [\xi(x) - s_i(x)]^2 dx\right\} p_z(\varepsilon)\, d\varepsilon.$$

Згідно з критерієм максимальної правдоподібності, необхідно забезпечити виконання умови $< p(\xi / s_i) > \xrightarrow{i} \max$.

Також можна показати, що оптимальний некогерентний приймач виділяє обвідну взаємної кореляційної функції.

Оптимальний приймач дискретних повідомлень

Основною задачею приймача є виділення корисного сигналу з суміші сигнал + завада. При цьому про корисний сигнал відомі деякі параметри, в невідомому параметрі закладена корисна інформація, може бути відомий динамічний діапазон параметра.

Існують такі задачі при прийомі сигналу:

1) знаходження – встановлення факту наявності сигналу, якщо в каналі зв'язку присутній адитивний шум, то задача зводиться до відповіді на запитання: чи є сигнал на вході приймача сигнал + шум чи шум?

2) розділення – відбувається передача 2 ненульових сигналів $S_1(t)$ і $S_2(t)$:

$$S_1(t) + шум,$$
$$S_2(t) + шум.$$

Задача зводиться до визначення того, яка суміш на вході приймача $S_1(t) + \omega$ чи $S_2(t) + \omega$.

3) відтворення форми – відтворення повідомлення, виникає при передачі неперервного повідомлення модульованого чи немодульованого. Задача полягає в тому, щоб отримати сигнал $y(t)$ найменш відмінний від

переданого повідомлення, причому корисне повідомлення наперед невідоме, відомо лише те, що воно належить до класу сигналів.

В цих умовах передане повідомлення можна розглядати як одну з реалізацій випадкового процесу з частково відомими статистичними характеристиками.

При відтворенні сигналу величину відхилення $y(t)$ від переданого повідомлення оцінюють, використовуючи критерії:

– критерій найбільшого відхилення:

$$\varepsilon_{\max} = \max|y(t) - m(t)|,$$

де $m(t)$ – передане повідомлення,

$y(t)$ – сигнал на вході;

– критерій середнього відхилення:

$$\varepsilon_{ср} = \frac{1}{T}\int_0^T [y(t) - m(t)]dt,$$

– критерій середньоквадратичного відхилення (по суті основний):

$$\varepsilon^2 = \frac{1}{T}\int_0^T [y(t) - m(t)]^2 dt.$$

Оцінка параметрів. Інформаційний параметр може приймати будь-яке значення з деякого інтервалу і є випадковою величиною.

Фільтрація – базова операція, яка розв'язує основні задачі.

Схема оптимального приймача буде наведена при розгляді виявлення сигналів.

Виявлення сигналів

Нехай очікуваний сигнал $s(t)$ повністю відомий, тобто відомі його форма, амплітуда, часове положення. Кінцеву функцію приймача позначимо через

$$C(u) = \begin{cases} 1 - \textit{рішення про наявність сигналу}, \\ 0 - \textit{рішення про відсутність сигналу}. \end{cases}$$

Виберемо кінцеву функцію так, щоб вагова різниця α була максимальна. Оскільки $\alpha = D - x_0 F$, а

$$D = \int_{-\infty}^{\infty} C(u) p_{сш}(u) du,$$

$$F = \int_{-\infty}^{\infty} C(u) p_{ш}(u) du,$$

де $p_ш$ і $p_{сш}$ – густини ймовірності шуму і суміші сигналу з шумом, то

$$\alpha = \int_{-\infty}^{\infty} C(u)[p_{сш}(u) - x_0 p_ш(u)]du = \int_{-\infty}^{\infty} C(u)p_ш(u)[x(u) - x_0]du,$$

де $x(u) = \dfrac{p_{сш}(u)}{p_ш(u)}$ – відношення правдоподібності, яке показує, наскільки правдоподібне припущення про прийом сигналу в порівнянні з альтернативним припущенням про його відсутність.

Для максимального значення α необхідно, щоб підінтегральний вираз був невід'ємним, що можливо за умови:

$$C(u) = 1 \quad при \quad x(u) \geq x_0,$$
$$C(u) = 0 \quad при \quad x(u) < x_0.$$

Таким чином, оптимальний алгоритм прийняття рішення приймачем зводиться до обчислення відношення правдоподібності x і порівняння його з порогом x_0.

Наприклад, для білого гауссівського шуму $n(t)$ на інтервалі $0 < t < T$ густина ймовірності дорівнює

$$p(n(t)) = ke^{-\frac{1}{N_0}\int_0^T n^2(t)dt},$$

де N_0 – спектральна густина потужності білого шуму.

Оскільки $n(t) = u(t) - \lambda s(t)$, то $p_{сш}(u(t)) = p[n(t) = u(t) - s(t)]$ і $p_ш(u(t)) = p[n(t) = u(t)]$.

Відношення правдоподібності можна спрощено записати:

$$x(u) = e^{-\frac{E}{N_0} + 2\frac{R_{сш}(0)}{N_0}}.$$

Таким чином, алгоритми виявлення повністю відомого сигналу на фоні білого гауссівського шуму такі:

$$x(u) > x_0,$$

або

$$\frac{2R_{сш}(0) - E}{N_0} \geq \ln x_0$$

чи

$$R_{сш}(0) = \int_0^T u(t)s(t)dt \geq \frac{N_0}{2}\ln x_0 - \frac{E}{2} = h_0,$$

і, як наслідок, схема оптимального приймача повинна складатись з пристрою, який обчислює взаємну кореляційну функцію (ВКП), і порогового пристрою (ПП). Ця схема наведена на рис. 6.11.

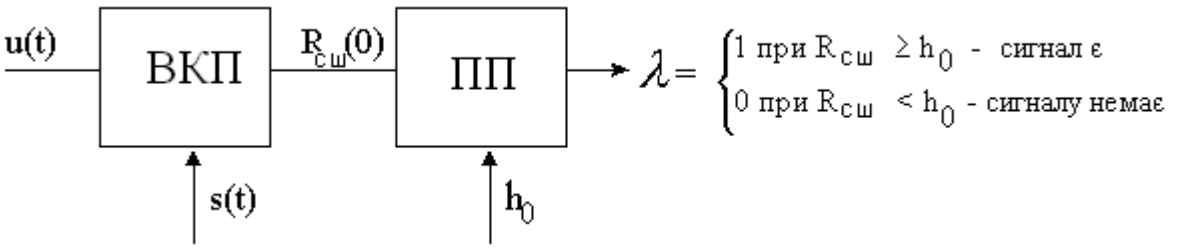


Рисунок 6.11 – Схема оптимального приймача

Узгоджена фільтрація

Як було показано вище, основною операцією при оптимальному розрізненні є обчислення функції взаємної кореляції між прийнятим коливанням $u_1(t)$ і очікуваним сигналом $s_1(t)$:

$$R(\tau) = \int_{-\infty}^{\infty} u_1(t) s_1(t-\tau) dt.$$

Ця операція виконується взаємокореляційним пристроєм (ВКП), який складається із затримуючого на час τ пристрою, помножувача і інтегратора (рис. 6.12)

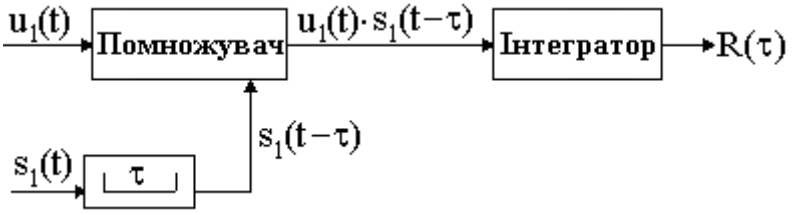


Рисунок 6.12 – Схема ВКП

Цей пристрій обчислює для заданого τ одне значення функції взаємної кореляції. Для отримання багатьох значень $R(\tau)$ необхідно або повторити процедуру для різних τ, або використати паралельне коло затримувальних пристроїв. Якщо час надходження сигналу невідомий, то достатньо обчислити лише $R(0)$, однак частіше всього час приходу є випадковою функцією, і тому доводиться аналізувати всю функцію $R(\tau)$, приймаючи рішення про наявність сигналу за її максимумом.

Можна побудувати схему ВКП у вигляді простого лінійного пристрою з постійними параметрами, на виході якого отримують неперервну функцію $R(\tau)$, а затримка часу сигналу вносить відповідну затримку у вихідний сигнал.

Такий пристрій називають лінійним фільтром. Зв'язок вхідного і вихідного сигналів в лінійному фільтрі описується інтегралом згортки:

$$u_{вих}(t) = \int_{-\infty}^{t} u_{вх}(x)h(t-x)dx,$$

де $h(t)$ – імпульсна характеристика фільтра.

Якщо підібрати лінійне коло так, щоб напруга на його виході з точністю до довільного множника k відтворювала би взаємокореляційну функцію (з деяким запізненням t_0), тобто

$$u_{вих}(t) = kR(t-t_0),$$

то його імпульсна характеристика повинна задовольняти рівність

$$h(t) = ks_1(t_0 - t).$$

Це може бути проілюстровано на такому прикладі (рис. 6.13). Імпульсна характеристика є дзеркальним відображенням сигналу.

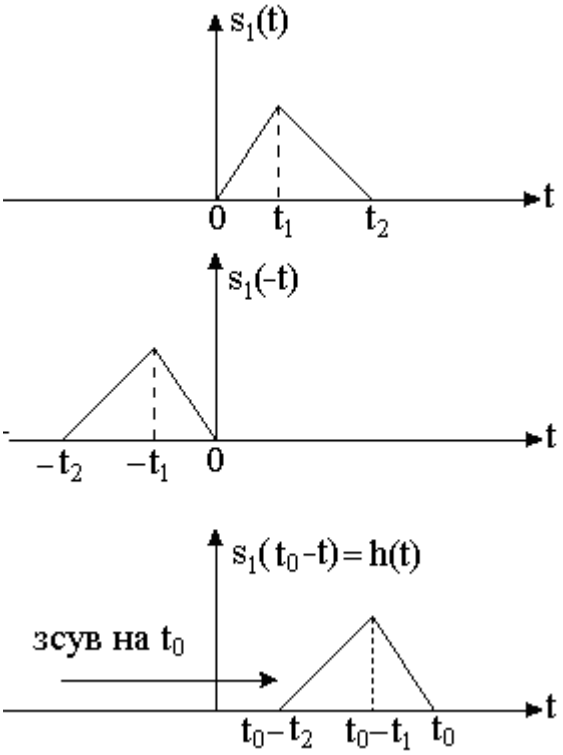


Рисунок 6.13 – Аналіз роботи оптимального фільтра

Лінійна система, яка має таку імпульсну характеристику, називається оптимальним фільтром. Вона виконує операцію обчислення взаємної кореляційної функції. Щоб сигнал на виході з'явився після приходу сигналу, повинна виконуватися умова $t_0 > t_2$. Оптимальний фільтр є найкращим за критерієм отримання на виході максимального відношення сигнал-шум при заданій формі сигналу і інтенсивності білого шуму на його виході.

При знаходженні вузькополосного радіосигналу

$$s_1(t) = S_1(t)\cos[\omega_0 t + \varphi(t)] = \mathrm{Re}\left[\dot{S}_1(t)e^{j\omega_0 t}\right],$$

де $\dot{S}_1(t) = S_1(t)e^{j\varphi(t)}$, $S_1(t)$, $\varphi(t)$ – обвідна і фаза вузькополосного сигналу, які повільно змінюються, тоді імпульсна характеристика фільтра повинна мати вигляд

$$h(t) = k\,\mathrm{Re}\left[\dot{S}_1(t - t_0)e^{-j\omega_0 t_0}\right],$$

де зірочка означає комплексно-спряжену функцію. Якщо сигнали мають дуже складну форму чи дуже велику тривалість, то побудова оптимальних фільтрів для них є дуже складною задачею.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Які бувають типи помилок при прийманні сигналу?
2. У чому сутність оптимального приймання?
3. Який зміст вкладають в поняття “розв’язати два сигнали”?
4. Наведіть коротку класифікацію завад.
5. Що являє собою завада?
6. Яка специфіка завадостійкісті сигналів з амплітудною, частотною та фазовою модуляцією?
7. Наведіть узагальнену структурну схему системи багатоканального зв’язку.
8. Які методи розподілу каналів ви знаєте? Охарактеризуйте їх.
9. За яких умов здійснюється когерентне приймання сигналів?
10. Наведіть блок-схему алгоритму розпізнавання двох сигналів?
11. Які задачі постають при прийманні сигналу?
12. Наведіть принципи виявлення сигналу.
13. Для чого використовується узгоджена фільтрація?

СЛОВНИК ТЕРМІНІВ

Автоматизована електронна система – сукупність керованого об'єкта та автоматичних пристроїв, у яких частину функцій керування виконує людина (оператор).

Автоматична електронна система – сукупність керованого об'єкта і автоматичних пристроїв, що виконують певну задачу згідно із заданою програмою без участі людини.

Аналоговий сигнал – сигнал, що може приймати будь-які значення в певних межах (наприклад, напруга може плавно змінюватися в межах від нуля до десяти вольтів).

Біт – найменша кількість інформації, двійкова система.

Вимірювальний канал – засіб вимірювання, який є складовою частиною вимірювальних систем у вигляді сукупності засобів вимірювальних операцій, засобів (каналів або ліній) зв'язку для послідовних перетворень і передачі вимірювальної інформації.

Гнучкість – це здатність системи налаштовуватися під різні задачі.

Дискретизація за часом – перетворення функції неперервного аргументу в функцію дискретного аргументу.

Дисперсія – математичне сподівання квадрата відхилення величини від математичного сподівання в певний момент часу.

Електронна підсистема – частина системи, до складу якої входять більше ніж один елемент (проміжний елемент поділу системи) і яка має певне функціональне призначення нижчого рівня, ніж система.

Електронна система – сукупність електронних компонентів, що з'єднані між собою й діють як одне ціле завдяки спеціальним сигналам управління та виконують задану функцію.

Ентропія – логарифмічна міра безладдя стану джерела повідомлень, що характеризує середній ступінь невизначеності стану цього джерела.

Ємність каналу – гранична швидкість передавання інформації цим каналом.

Завади – сигнали, що впливають на електронну систему ззовні та спотворюють корисний сигнал (наприклад, електромагнітні випромінювання від радіопередавачів або від трансформаторів).

Задача – набір функцій, виконання яких потрібно від електронної системи.

Імпульсний сигнал – сигнал, що наростає в тактовий момент, а спадає в межах даного такту.

Інтерфейс – узгодженість обміну інформацією та правил обміну інформацією, під якою розуміють електричну, логічну й конструктивну сумісність пристроїв, що беруть участь в обміні.

Інформація – сукупність відомостей про всілякі об'єкти, явища, процеси.

Канал зв'язку – сукупність передавача, приймача та лінії зв'язку, призначених для передачі інформації від відправника (джерела) до адресата.

Квантування за амплітудою – процес заміни реальних (виміряних) значень амплітуди сигналу значеннями, наближеними з деякою точністю.

Квантування за рівнем – перетворення неперервної функції в дискретну множину значень.

Кодування – подання за певними правилами дискретних повідомлень у деякі комбінації, складені з певного числа елементів-символів.

Комп'ютеризовані системи – системи, під якими розуміють сукупність технічних засобів, яка має в своєму складі керуючу ЕОМ, засоби збору, перетворення, передачі, відображення, а також спеціальне математичне та програмне забезпечення, що виконує весь комплекс оброблення інформації.

Лінія зв'язку – фізичне середовище, по якому передаються сигнали.

Мікроконтролер – обчислювально-керуючий пристрій, призначений для виконання функцій логічного контролю й керування периферійним устаткуванням, виконаний у вигляді однієї ВІС, що поєднує в собі мікропроцесорне ядро і набір вбудованих пристроїв введення-виведення

Мікропроцесор – Програмно-керований пристрій, призначений для оброблення цифрової інформації й керування процесом цього оброблення, вико-

наний у вигляді однієї (або декількох) інтегральної схеми з високим ступенем інтеграції електронних елементів.

Модель –вибраний спосіб опису об'єкта, процесу або явища, який відображає суттєві з погляду розв'язання даної задачі фактори.

Надмірність – показник ступеня відповідності можливостей системи до розв'язуваної даною системою задачі.

Простір станів – множина можливих значень випадкової величини.

Сигнал – будь-яка фізична величина (наприклад, температура, тиск повітря, інтенсивність світла, сила струму й т. д.), що змінюється в часі.

Система – група різних предметів, які об'єднані таким чином, що утворюють єдине ціле, функціонують узгоджено та підпорядковані єдиній формі управління.

Система реального часу – система, що повинна реагувати на зовнішні параметри введення і створювати нові результати виведення за обмежений час.

Технічна швидкість передачі – число елементарних сигналів (символів), переданих по каналу за одиницю часу.

Цифровий сигнал – сигнал, що може приймати тільки два (іноді – три) значення з деякими відхиленнями від цих значень.

Швидкодія – показник швидкості виконання електронною системою її функцій.

Шум – внутрішні хаотичні слабкі сигнали будь-якого електронного пристрою.

СПИСОК СКОРОЧЕНЬ

АІМ – амплітудноімпульсна модуляція.

АМ – амплітудна модуляція.

АЦП – аналогово-цифровий перетворювач.

АЧХ – амплітудно-частотна характеристика.

ВІС – велика інтегральна схема.

ВК – вимірювальний канал.

ВКП – взаємокореляційний пристрій .

ЕОМ – електронна обчислювальна машина.

ЕРС – електрорушійна сила.

ЕС – електронна система.

ІКМ – імпульсно-кодова модуляція.

КОЕС – комп’ютеризована оптико-електронна система.

КУД – кінцеве устаткування даних.

СПД – система передачі даних.

ТЧ – тональна частота.

ФІМ – фазово-імпульсна модуляція.

ФМ – фазова модуляція.

ЦАП – цифро-аналоговий перетворювач.

ЧІМ – частотно-імпульсна модуляція.

ЧМ – частотна модуляція.

ЧРК – часовий розподіл каналів.

ШІМ – широтно-імпульсна модуляція.

СПИСОК ЛІТЕРАТУРИ

1. Бабич Н. П. Основы цифровой схемотехники: учебное пособие / Н. П. Бабич, И. А. Жуков. – М. : Издательский дом «Додэка-XXI», К. : «МК-Пресс», 2007. – 480 с.
2. Баскаков С. И. Радиотехнические цепи и сигналы: учебник для вузов / Баскаков С. И. – М. : Высшая школа, 1988. – 320 с.
3. Бойко В. І. Основи технічної електроніки: У 2 кн. Кн.2 Схемотехніка: підручник / Бойко В. І., Гурій А. М., Жуйков В. Я. та ін. – К. : Вища школа, 2007. – 510 с.
4. Васюра А. С. Техніка передавання аналогової та дискретної інформації / Васюра А. С.. – Вінниця: ВДТУ, 1998. – 218 с.
5. Гольденберг Л. М. Цифровая обработка сигналов: учебное пособие для вузов / Гольденберг Л. М. – М. : Радио и связь, 1990. – 256 с.
6. Дмитриев В. И. Прикладная теория информации / Дмитриев В. И.. – М.: Высшая школа, 1989. – 320 с.
7. Кловский Д. Д. Теория передачи сигналов в задачах / Д. Д. Кловский, В. А. Шилкин. – М. : Связь, 1978. – 186 с.
8. Кричевський Р. Е. Сжатие и поиск информации / Кричевський Р. Е. – М. : Радио и связь, 1989. – 168 с.
9. Кузьмин И. В. Основы теории информации и кодирования / И. В. Кузьмин, В. А. Кедру. – К. : Вища школа, 2003. – 220 с.
10. Лукин А. Введение в цифровую обработку сигналов (Математические основы) / Лукин А. А. – М. : МГУ, Лаборатория компьютерной графики и мультимедиа, 2002. – 616 с.
11. Передача дискретних собщений: учебник для вузов //Под ред. В. П. Шувалова. – М. : Радио и связь, 1990. – 464 с.
12. Пухальский Г. И. Цифровые устройства: учебное пособие для ВТУЗов / Г. И. Пухальский, Т. Я. Новосельцев. – Спб. : Политехника, 1996. – 885 с.
13. Рябенький В. М. Цифрова схемотехніка: навч. посібник / Рябенький В. М. , Жуйков В. Я., Гулий В. Д.. – Львів : «Новий світ-2000», 2009. – 736 с.
14. Сергиенко А. Б. Цифровая обработка сигналов / Сергиенко А. Б. . – СПб. : Питер, 2003. – 608 с.
15. Системы электросвязи / Под ред. А. Н. Шувалова. – М. : Радио и связь, 1987. – 416 с.

16. Фано Р. Передача информации. Статистическая теория связи / Роберт Фано. – М. : Мир, 1965. – 438 с.
17. Хаїмзон І. Я. Техніка передачі інформації. Функціональні вузли та схеми. Частина ІІ / Хаїмзон І. Я. – Вінниця : ВДТУ, 2000. – 112 с.
18. Хемминг Р. В. Теория кодирования и теория информации / Хемминг Р. В. – М. : Радио и связь, 1983. 214 с.
19. Хуанг Т. С. и др. Быстрые алгоритмы в цифровой обработке изображений / Хуанг Т. С. и др. – М. : Радио и связь, 1984. – 224 с.
20. Цымбал В. П. Задачник по теории информации и кодированию / Цымбал В. П. – К. : Вища школа, 1976. – 312 с.
21. Чэн Ш.-К. Принципы кодирования систем визуальной информации / Чэн Ш.-К. – М. : Мир, 1994. – 408 с.
22. Шеннон К. Математическая теория связи // Работы по теории информации и кибернетике / Клод Шеннон. – М. : Иностранная литература, 1963. – С. 223-332.

Навчальне видання

Білинський Йосип Йосипович
Огородник Костянтин Володимирович
Юкиш Марина Йосипівна

ЕЛЕКТРОННІ СИСТЕМИ

Навчальний посібник

Редактор В. Дружинна
Коректор З. Поліщук

Оригінал-макет підготовлено Й. Білинським

Підписано до друку
Формат 29,7 × 42 $^{1}/_{4}$. Папір офсетний
Гарнітура Times New Roman
Друк різографічний. Ум друк.арк.
Наклад прим. Зам. №

Вінницький національний технічний університет,
навчально-методичний відділ ВНТУ.
21021, м.Вінниця, Хмельницьке шосе, 95,
ВНТУ, к. 2201.
Тел. (0432) 59-87-36.
Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009р.

Віддруковано у Вінницькому національному технічному університеті
В комп'ютерному інформаційно-видавничому центрі.
21021, м. Вінниця, Хмельницьке шосе, 95,
ВНТУ, ГНК, к.114.
Тел. (0432) 59-87-38.
Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.