

Міністерство освіти і науки України
ДВНЗ "Ужгородський національний університет"
Фізичний факультет
Кафедра твердотільної електроніки та інформаційної безпеки

Юлія Мисло

**Курс лекцій з дисципліни
"Основи завадостікості систем
захисту інформації"**

**Навчальний посібник для студентів
спеціальності "Кібербезпека та захист інформації"**

Ужгород 2024

Рецензенти:

Рубіш Василь Михайлович — доктор фізико–математичних наук, професор, академік Академії технологічних наук України, завідувач Ужгородської лабораторії матеріалів оптоелектроніки та фотоніки Інституту проблем реєстрації інформації НАН України;

Міца Олександр Володимирович — доктор технічних наук, професор, професор кафедри інформаційних упрвляючих систем та технологій

Мисло Ю.М.

Курс лекцій з дисципліни "Основи завадостійкості систем захисту інформації". Навчальний посібник для студентів спеціальності "Кібербезпека та захист інформації". Електронне видання, 2024. 79 с.

Навчальний посібник розрахований в першу чергу на студентів фізичного факультету Ужгородського національного університету, які навчаються за спеціальністю "Кібербезпека та захист інформації", а також буде корисним для студентів інших факультетів.

Затверджено та рекомендовано на засіданні кафедри твердотільної електроніки фізичного факультету ДВНЗ "Ужгородський національний університет" (протокол № 6 від 16 січня 2024 року).

Зміст

1	Загальні відомості про інформацію інформацію та системи її захисту	6
1.1	Інформація та її форми.Визначення поняття інформація. . . .	6
1.2	Види інформації	7
1.3	Кількісні міри інформації	8
1.4	Міра Хартлі	9
1.5	Міра Шеннона	10
1.6	Ентропія, як міра невизначеності вибору	11
1.6.1	Статистичні якості ентропії дискретного джерела ін- формації	11
1.6.2	Основні властивості ентропії:	12
1.6.3	Ентропія об'єднання статистично незалежних дис- кретних джерел	13
1.6.4	Властивості ентропії об'єднання:	13
1.6.5	Умовна ентропія та її властивості	13
1.6.6	Властивості умовної ентропії:	15
1.7	Ентропія безперервного джерела інформації (диференціаль- на ентропія) та її властивості	15
1.8	Безперервні розподіли з максимальною диференціальною ентропією	16
2	Дискретні повідомлення та їх джерела. Класифікація сигналів	18
2.1	Інформаційні характеристики джерел дискретних повідом- лень	18
2.2	Надмірність повідомлень	19
2.3	Продуктивність джерел дискретних повідомлень	21
3	Сигнали та їх класифікація	23
3.1	Детерміновані та випадкові сигнали.	23

3.2	Спектри дискретних сигналів	25
3.3	Спектри одиничних та періодичних імпульсних послідовностей	28
4	Способи перетворення сигналів	30
4.1	Дискретизація та модуляція аналогових сигналів. Дискретизація сигналів в часі, за рівнем та кодування безперервних сигналів.	30
4.2	Теорема Котельникова	31
4.3	Методи модуляції дискретних сигналів	32
4.4	Модуляція гармонічних коливань: амплітудна, частотна, фазова.	33
4.4.1	Амплітудна, частотна, фазова модуляція, порівняння спектрів сигналів з різними видами модуляції . . .	33
4.4.2	Імпульсні види модуляції	36
4.4.3	Амплітудно-імпульсна модуляція (АІМ)	37
4.4.4	Широтно-імпульсна модуляція (ШІМ)	38
4.4.5	Часова імпульсна модуляція	38
4.4.6	Фазо-імпульсна модуляція (ФІМ)	39
4.4.7	Частотно-імпульсна модуляція (ЧІМ)	40
5	Випадкові сигнали та їх характеристики	42
5.1	Неперервні та дискретні випадкові процеси	42
5.2	Форми сигналів	45
5.3	Спектральне подання випадкових сигналів	47
5.4	Теорема Хінчина - Вінера	50
6	Завади, їх джерела	52
6.1	Достовірність та завадостійкість	52
6.2	Завадостійке кодування	55
6.3	Зв'язок корегувальної спроможності з кодовою відстанню . .	56
6.4	Код Хемінга	58
6.5	Циклічні коди	60
6.6	Коди Боуза-Чоудхури-Хоквінгема	64
6.7	Коди Ріда-Соломона	65
7	Структура і характеристики згорткових кодів	67
7.1	Методи опису згорткових кодів	67

	5
7.2 Основні параметри й класифікація ЗК	71
8 Алгоритми декодування згорткових кодів	75
8.1 Класифікація алгоритмів декодування	75
8.2 Алгоритм Вітербі для декодування згорткових кодів	77
Бібліографія	79

Розділ 1

Загальні відомості про інформацію інформацію та системи її захисту

1.1 Інформація та її форми. Визначення поняття інформація.

Інформація – абстрактне поняття, що має різні значення залежно від контексту. Походить від латинського слова «informatio», яке має декілька значень:

- роз'яснення;
- виклад фактів, подій;
- витлумачення;
- представлення, поняття;
- ознайомлення, просвіта.

Основними **формами** подання інформації є символна, 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.

1.2 Види інформації

Інформацію можна поділити на **види** за **кількома ознаками**:

1. За способом сприйняття:

Для людини інформація поділяється на види залежно від типу рецепторів, що сприймають її.

- *Візуальна* – сприймається органами зору.
- *Аудіальна* – сприймається органами слуху.
- *Тактильна* – сприймається тактильними рецепторами.
- *Нюхова* – сприймається нюховими рецепторами.
- *Смакова* – сприймається смаковими рецепторами.

2. За формою подання: За формою подання інформація поділяється на такі види:

- *Текстова* – що передається у вигляді символів, призначених позначати лексеми мови;
- *Числова* – у вигляді цифр і знаків, що позначають математичні дії;
- *Графічна* – у вигляді зображень, подій, предметів, графіків;
- *Звукова* – усна або у вигляді запису передачі лексем мови аудіальним шляхом.

3. За призначенням:

- *Масова* – містить тривіальні відомості і оперує набором понять, зрозумілим більшій частині соціуму;

- *Спеціальна* – містить специфічний набір понять, при використанні відбувається передача відомостей, які можуть бути не зрозумілі основній масі соціуму, але необхідні і зрозумілі в рамках вузької соціальної групи, де використовується дана інформація;
- *Особиста* – набір відомостей про яку-небудь особистість, що визначає соціальний стан і типи соціальних взаємодій всередині популяції.

1.3 Кількісні міри інформації

Повідомлення, що передаються людською мовою, складаються з певного набору письмових (абетка, розділові знаки) або речових (фонетичних) знаків. Результати вимірювань позначаються системою знаків, які складаються з цифр та літер, що утворюють певний код, тобто цифри це числові значення, а літери – одиниці вимірювань. В теорії знаків, яка називається семіотика, знакові системи вивчаються на трьох основних рівнях: синтаксичному, семантичному і прагматичному.

На **синтаксичному рівні** повідомлення розглядають як символи, що абстраговані від змісту або будь-якої цінності інформації. Предметом аналізу на цьому рівні є частота появи символів, тобто знаків коду, зв'язку між ними, порядку проходження, правила побудови виразів, за допомогою яких можуть формулюватися повідомлення.

На **семантичному рівні** вивчається зміст символів повідомлення, тобто відношення повідомлення до того, що воно відображає. Синтаксично грамотно побудована фраза може бути абсолютно некоректною в семантичному відношенні.

Наприклад, речення: “Висота будинку дорівнює 1000 тон” є вірно складеним граматично, але не має ніякого змісту.

На **прагматичному рівні** символи-повідомлення розглядаються в їх відношенні до отримувача. При цьому враховуються такі характеристики повідомлення, як важливість, корисність, цінність, актуальність.

Наприклад, повідомлення “завтра очікується вітер силою 11 балів”, тобто шторм, який може викликати великі неприємності, є більш важливим за повідомлення, що “завтра очікується сильний вітер – силою 3 бала”.

В синтаксичному та семантичному відношенні ці повідомлення є однаковими. На сьогодні найбільш розроблена задача синтаксичної оцінки інформації.

При синтаксичному рівні аналіз інформації визначається як міра зменшення невизначеності знань про будь-який об'єкт (предмет) у процесі його пізнання. Для оцінки кількості інформації на синтаксичному рівні необхідно знати міру невизначеності тієї чи іншої ситуації. Інтуїтивно зрозуміло, що кількість інформації, яка буде отримана в результаті дослідження або дії, пропорційна величині різниці між невизначеністю інформації до та після дослідження.

Якщо H_1 – невизначеність знань певного питання до дослідження (апріорна), а H_2 – невизначеність, що характеризує стан знань після дослідження (апостеріорна), тоді інформація, що існує в цьому повідомленні, визначається як $I = H_1 - H_2$.

Для оцінки ступеня невизначеності знань на синтаксичному рівні розроблено велику кількість різних математичних мір, але найбільше використовуються дві міри:

- **структурна** або **логарифмічна**, запропонована *Ральфом Хартлі* в 1928 р.;
- **ймовірнісна**, запропонована *Клодом Шеноном* в 1948 р.

1.4 Міра Хартлі

При розробці логарифмічної міри Р.Хартлі запропонував наступне:

- сигнали, за допомогою яких передаються повідомлення, складаються з дискретних елементів (символів);
- елементи сигналів можуть приймати кінцеву кількість різних значень (m) (сукупність всіх m значень утворює алфавіт);
- усі символи між собою статистично незалежні і рівноймовірні;
- передача інформації відбувається за відсутності завад.

Припустимо, що символи алфавіту можуть приймати m неоднакових значень при рівній ймовірності їх появи, тобто: $p_1 = p_2 = \dots = p_i = \dots = p_m$, тобто $p_i = \frac{1}{m}$.

При послідовності з n таких символів можливо отримати N сигналів (повідомлень). При цьому $N = m^n$.

Наприклад, $m = 2$, $n = 3$, тоді $N = 2^3 = 8$. Дійсно, якщо у двійковому алфавіті всього два символи “0” чи “1” при послідовності з трьох таких

символів можемо отримати 8 слів (повідомлень), а саме: 000, 001, 010, 011, 100, 101, 110 та 111.

Кількість можливих слів повідомлень зростає в степеневій залежності від кількості символів n . Зрозуміло, що кількість інформації, що знаходиться в кожному повідомленні, пропорційна кількості символів n , і тому Р.Хартлі запропонував міру визначення інформації як логарифм можливих сигналів з однаковою кількістю символів.

$$I = \log_a N = \log_a m^n = n \log_a m \quad (1.1)$$

В залежності від основи логарифму a існують різні одиниці вимірювання інформації.

- Одиниця кількості інформації, що являє собою вибір із двох рівноймовірних подій, одержала назву двійкової одиниці, або біта (основа логарифму $a = 2$). Назва bit утворена із двох початкових і останніх літер англійського виразу binary unit, що значить двійкова одиниця.
- Якщо основа логарифма $a = 10$, то одиниця вимірювання інформації має назву – діт.
- А якщо $a = e$ (тобто основи натурального логарифму), тоді інформація вимірюється в нітах.

Логарифмічна міра для невизначеності інформації вибрана тому, що задовольняє умови **адитивності**, що дозволяє використати її для відтворення складних дослідів. Тобто, при наявності незалежних джерел інформації з кількістю N_1 і N_2 можливих повідомлень кожне із N_1 можливих значень першої величини може з'явитися одночасно із будь-яким з N_2 значень іншої величини.

Кількість усіх можливих сполучень значень двох цих величин дорівнює добутку $N = N_1 N_2$, і тоді: $I = \log N = \log_a N_1 N_2 = \log_a N_1 + \log_a N_2$. Таким чином, кількість інформації що прийдеться на одне повідомлення, дорівнює сумі кількостей інформації, яка була отримана від двох незалежних джерел, узятих окремо.

1.5 Міра Шеннона

В тому випадку, якщо ймовірність будь-якого стану або появи символу алфавіту неоднакові, тобто $p_1 \neq p_2 \neq p_m$, використовується ймовірнісна міра, яку запропонував К. Шеннон.

Міра Шеннона дозволяє визначити середню кількість інформації, яку отримують на одне повідомлення:

$$I = -n \sum_{i=1}^m p_i \log_a p_i \quad (1.2)$$

де m – кількість символів в алфавіті, n – кількість символів алфавіту в повідомленні, p_i – імовірність (відносна частота) появи символу із значенням i в повідомленні.

Виявилось, що формула, запропонована Р. Хартлі, являє собою окремий випадок більше загальної формули Шеннона.

Якщо у формулі Шеннона прийняти, що $p_1 = p_2 = \dots = p_i = \dots = p_m$, тоді $I = -n \sum_{i=1}^m p_i \log_a p_i = -nm \left(\frac{1}{m} \log_a \frac{1}{m} \right) = -n \frac{m}{m} \log_a \frac{1}{m} = -n (\log_a 1 - \log_a m) = n \log_a m$.

Знак мінус у формулі Шеннона не означає, що кількість інформації в повідомленні – від’ємна величина. Пояснюється це тим, що ймовірність p_i , відповідно до визначення, менше одиниці, але більше нуля. Відомо, що логарифм числа, меншого одиниці, тобто $\log p_i$ – величина від’ємна, то добуток ймовірності на логарифм числа буде додатним.

1.6 Ентропія, як міра невизначеності вибору

Ентропія являє собою логарифмічну міру неупорядкованості 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 $H = - \sum_{i=1}^m p_i \log p_i$. В інформаційних системах невизначеність знижується за рахунок прийнятої інформації, тому чисельно ентропія H дорівнює кількості інформації I , тобто є кількісною мірою інформації.

1.6.1 Статистичні якості ентропії дискретного джерела інформації

У загальному випадку, відповідно до теорії ймовірностей, джерело інформації однозначно й повно характеризується ансамблем станів $U = \{(u_1), (u_2), \dots, (u_n)\}$ з ймовірностями станів відповідно

$\{p(u_1), p(u_2), \dots, p(u_n)\}$ 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:
$$H = - \sum_{n=1}^N p_n \log_2 p_n.$$

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

1.6.2 Основні властивості ентропії:

1. Ентропія є величиною дійсною й невід'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 $p_n \geq 0$ 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 (ймовірності інших станів джерела дорівнюють нулю, тому що сума ймовірностей повинна бути 1).
4. Ентропія джерела із двома станами u_1 та u_2 при зміні співвідношення їхніх ймовірностей $p(u_1)$ та $p(u_2)$ визначається виразом $H(U) = -[p \log p + (1 - p) \log p]$ і досягає максимуму при рівності ймовірностей.
5. Ентропія об'єднаних статистично незалежних джерел інформації дорівнює сумі їх ентропій.
6. Ентропія максимальна при рівній ймовірності 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.

1.6.3 Ентропія об'єднання статистично незалежних дискретних джерел

Ентропія об'єднання використовується для обчислення ентропії сумісної появи статично залежних повідомлень. Наприклад, передаючи сто раз цифру 5 по каналу зв'язку з завадами, зауважили, що цифра 5 була прийнята 90 раз, цифра 6 – 8 раз і цифра 4 – 2 рази. Невизначеність виникнення комбінацій виду 5-4, 5-5, 5-6 при передачі цифри 5 може бути описана за допомогою ентропії об'єднання.

Ентропія об'єднання $H(A, B)$ – це невизначеність того, що буде послатись А, а прийматись В. Для ансамблів переданих повідомлень та прийнятих повідомлень ентропія об'єднання являє собою суму вигляду

$$H(A, B) = - \sum_i \sum_j p(a_i, b_j) \log_2 p(a_i, b_j) \quad (1.3)$$

Ентропія об'єднання та умовна ентропія пов'язані між собою наступними співвідношеннями:

$$H(A, B) = H(A) + H(B/A) = H(B) + H(A/B)$$

$$H(B/A) = H(A, B) - H(A), \quad H(A/B) = H(A, B) - H(B)$$

1.6.4 Властивості ентропії об'єднання:

1. Властивість симетрії: $H(A, B) = H(B, A)$.
2. При відсутності статичної залежності між елементами ансамблів А та В умовні імовірності перетворюються на безумовні, тоді $H(A, B) = H(A) + H(B)$.
3. При повній статичній залежності між елементами ансамблів А та В (наприклад, коли результат однієї події однозначно визначає інформацію про другу подію) умовні ентропії рівні нулю, тому $H(A, B) = H(A) = H(B)$.

1.6.5 Умовна ентропія та її властивості

На практиці часто зустрічаються взаємозалежні символи та повідомлення. При передачі змістовних повідомлень одні букви зустрічаються частіше, інші рідше, одні букви і слова часто ідуть за іншими. Наприклад,

в англійській мові найчастіше використовується буква e; у французькій після букви q завжди іде буква u, якщо q не стоїть в кінці слова. Щодо взаємодіючих систем, то переважно стан однієї з них впливає на стан іншої. В таких випадках ентропія не може визначатись лише на основі безумовних ймовірностей.

При підрахунку середньої кількості інформації на символ повідомлення взаємозалежність враховують через умовні ймовірності появи одних подій відносно інших. А отриману при цьому ентропію називають умовною ентропією.

Умовна ентропія використовується при визначенні взаємозалежності між символами первинного алфавіту, для визначення втрат при передачі інформації по каналах зв'язку, при обчисленні ентропії об'єднання.

Розглянемо текстове повідомлення. Якщо при передачі n повідомлень символ A з'явився m раз, символ B з'явився l раз, а символ A разом з символом B – k раз, то ймовірність появи символу A $p(A) = m/n$; ймовірність появи символу B $p(B) = l/n$; ймовірність сумісної появи символів A та B $p(AB) = k/n$.

Умовна ймовірність появи символу A відносно символу B $p(A/B) = \frac{p(AB)}{p(B)} = \frac{k}{l}$ і умовна ймовірність появи символу B відносно символу A $p(B/A) = \frac{p(AB)}{p(A)} = \frac{k}{m}$.

Якщо відомо умовну ймовірність, то можна легко визначити й ймовірність сумісної появи символів, використовуючи попередні формули: $p(AB) = p(B)p(A/B) = p(A)p(B/A)$.

Розглянемо формули умовної ентропії, використовуючи формулу Шеннона:

$$H(b_j/a_i) = - \sum_j p(b_j/a_i) \log_2 p(b_j/a_i) \quad (1.4)$$

$$H(a_i/b_j) = - \sum_i p(a_i/b_j) \log_2 p(a_i/b_j) \quad (1.5)$$

де індекс i вибрано для характеристики довільного стану джерела повідомлень A , індекс j вибрано для характеристики довільного стану адресату B .

Розрізняють поняття **часткової** та **загальної** умовної ентропії. Вирази (1.4) та (1.5) є частковими умовними ентропіями.

Загальна умовна ентропія повідомлення B відносно повідомлення A характеризує кількість інформації, яка міститься у будь-якому символі алфавіту, і визначається як сума ймовірностей появи символів алфавіту помножена на невизначеність, яка залишається після того як адресат прийняв

сигнал:

$$H(B/A) = - \sum_i p(a_i) H(b_j/a_i) = - \sum_i \sum_j p(a_i) p(b_j/a_i) \log_2 p(b_j/a_i) \quad (1.6)$$

Вираз (1.6) є загальним для визначення кількості інформації на один символ повідомлення для випадку нерівноймовірних та взаємозв'язаних символів.

Оскільки $p(a_i)p(b_j/a_i)$ являється імовірністю сумісної появи двох подій $p(a_i, b_j)$, то формулу (1.6) можна записати наступним чином:

$$H(B/A) = - \sum_i \sum_j p(a_i/b_j) \log_2 p(b_j/a_i) \quad (1.7)$$

1.6.6 Властивості умовної ентропії:

1. Якщо ансамблі повідомлень А та В взаємонезалежні, то умовна ентропія А відносно В рівна безумовній ентропії А і навпаки: $H(A/B) = H(A)$; $H(B/A) = H(B)$.
2. Якщо ансамблі повідомлень А та В настільки жорстко статично пов'язані, що поява одного з них означає обов'язкову появу іншого, то їх умовні ентропії рівні нулю: $H(B/A) = H(A/B) = 0$.

1.7 Ентропія безперервного джерела інформації (диференціальна ентропія) та її властивості

Для опису інформаційних властивостей безперервного джерела (сигналу) використовується поняття диференціальної ентропії. Для її отримання скористаємося формулою для ентропії дискретних повідомлень. При цьому, вираз для ентропії можна представити у вигляді $H(X) = - \sum_{i=1}^m p(x_i) \log_2 p(x_i) = - \sum_{i=1}^m f(x_i) \Delta x \log_2 [f(x_i) \Delta x] = - \sum_{i=1}^m [f(x_i) \log_2 f(x_i)] \Delta x - \log_2 \Delta x \sum_{i=1}^m f(x_i) \Delta x = H(\Delta x)$

Переходимо до границі:

$$H(X) = \lim_{\Delta x \rightarrow 0} H(\Delta X) = - \int_{-\infty}^{\infty} f(x) \log_2 f(x) dx - \lim_{\Delta x \rightarrow 0} \log_2 \Delta x.$$

Повна ентропія джерела безперервних повідомлень складається з двох доданків, один з яких визначається законом розподілу, а другий є постійною величиною, що визначає крок квантування, який впливає на точність

вимірювань. Цей член вираження визначає постійну складову і виключається з розгляду. Значення першого доданка визначається законом розподілу і характеризує диференціальну ентропію безперервного джерела (оскільки $f(x)$ – густина імовірності або диференціальний закон розподілу) $h(x) = - \int_{-\infty}^{\infty} f(x) \log_2 f(x) dx$.

Диференціальна ентропія – частина ентропії джерела безперервних повідомлень, яка залежить від густини ймовірності сигналу $x(t)$, що видається джерелом.

1.8 Безперервні розподіли з максимальною диференціальною ентропією

Різні класи фізичних явищ і процесів підпорядковуються різним законам розподілу. Безперервні сигнали повністю характеризуються законами розподілу (інтегральним або диференціальним). На будь-які реальні сигнали накладаються певні обмеження. Оскільки диференціальна ентропія залежить від густини ймовірності, визначимо, для якого закону вона максимальна. Тобто, при якому розподілі ймовірності, сигнал заданої потужності має максимальну ентропію.

Диференціальна ентропію для нормального розподілу (сигнал з обмеженою середньою потужністю) не залежить від математичного очікування і дорівнює ентропії центрованої випадкової величини. Максимальне значення для ентропії: $h(x) = \log_2 \sigma \sqrt{2\pi e}$.

Визначимо диференціальну ентропію для рівномірного розподілу, тобто сигналу з обмеженою піковою потужністю. Якщо P – пікова потужність, то $A = \sqrt{P}$ – амплітудне значення. Диференціальна ентропія для рівномірного розподілу дорівнює: $h(x) = \log_2 2A$.

Диференціальна ентропія для експоненціального розподілу (розподіл широко використовується для визначення інтенсивності відмов у радіоелектронній апаратурі): $f(x) = ke^{-kx}$ при $x > 0$.

$$h(x) = \int_0^{\infty} ke^{-kx} (\log_2 k - kx \log_2 e) dx =$$

$$= -\log_2 k \int_0^{\infty} ke^{-kx} dx + \log_2 e \int_0^{\infty} xke^{-kx} dx = -\log_2 k + \log_2 e = \log_2 \frac{e}{k}$$

Повна ентропія для експоненціального розподілу дорівнює: $H(x) = h(x) - \log_2 \Delta x = \log_2 \frac{e}{k \Delta x}$.

Розділ 2

Дискретні повідомлення та їх джерела. Класифікація сигналів

2.1 Інформаційні характеристики джерел дискретних повідомлень

Дискретне повідомлення – це буква, цифра або відлік, який з’являється після дискретизації і квантування неперервного сигналу.

Інформаційна характеристика джерела – це розподіл кількості інформації, яка припадає на кожне повідомлення алфавіту джерела. Вихідними даними для розрахунку цієї характеристики є ймовірності появи повідомлень алфавіту джерела.

Для порівняння джерел, алфавіт яких має різну кількість повідомлень, інформаційна характеристика не є зручною. Отже інформаційна характеристика виконує ілюстративну функцію. Але як і всяка інша характеристика інформаційна характеристика дає можливість визначити параметри джерела.

Середнє значення кількості інформації, яка припадає на одне повідомлення джерела.

Розглянемо методику визначення середнього значення кількості інформації, яка припадає на одне повідомлення джерела на прикладі джерела інформації, алфавіт якого складає 2 повідомлення «0» і «1». Нехай $P(1)$ і $P(0)$ ймовірності появи повідомлень «1» і «0», відповідно. Тоді кількість інформації в кожній появі повідомлення «1» (або «0») $I(1)$ (або $I(0)$) визначається так:

$$I(1) = \log_2 \frac{1}{P(1)}, \quad I(0) = \log_2 \frac{1}{P(0)}$$

Розглянемо послідовність повідомлень, наприклад, 1101111000 і позначимо кількість «1» в цій послідовності - $n(1)$ і кількість «0» - $n(0)$. Тепер середню кількість інформації, яка припадає на одне повідомлення цієї послідовності повідомлень можна порахувати за формулою:

$$\bar{I}(a) = \frac{n(1)I(1) + n(0)I(0)}{n(1) + n(0)} = P(1)I(1) + P(0)I(0).$$

На основі цієї формули сформуємо формулу для визначення середнього значення кількості інформації, яка припадає на одне повідомлення джерела при довільній кількості повідомлень джерела m :

$$\bar{I}(a) = \sum_{i=1}^m P(a_i) \log_2 \frac{1}{P(a_i)}.$$

Для практичного використання параметр середнє значення кількості інформації, яка припадає на одне повідомлення джерела має коротшу назву – **ентропія джерела**.

2.2 Надмірність повідомлень

Коли інформація від джерела інформації видається у вигляді написаного тексту і в цьому тексті є невелика кількість помилок у вигляді пропущених букв, то з такого тексту можна вилучити передавану інформацію. Отже, джерело інформації видає суттєву (існуючі букви) і несуттєву (пропущені букви) інформацію. При передачі інформації у вигляді тексту несуттєва інформація є **надлишковою**. Цей ефект називається інформаційної надлишковості джерела.

Для джерела з реальним розподілом ймовірностей появи повідомлень маємо реальне значення ентропії джерела $H(A)$ і можемо визначити максимально можливе значення ентропії джерела $H_{max}(A)$. А отже можемо визначити різницю $H_{max}(A) - H(A)$. Цей параметр називається надлишковою джерела. Але цим параметром незручно користуватися для джерел з різною кількістю повідомлень ($H_{max}(A)$ для них буде різною). Усувається ця незручність нормуванням цього показника відносно $H_{max}(A)$:

$$\rho = \frac{H_{max}(A) - H(A)}{H_{max}(A)} = 1 - \frac{H(A)}{H_{max}(A)}.$$

Параметр ρ називають коефіцієнтом надлишковості джерела. Зауважимо, що значення коефіцієнта надлишковості джерела знаходиться в межах $0 \leq \rho \leq 1$.

Коефіцієнт надлишковості характеризує питому вагу інформації джерела, яку можна не передавати і при цьому втрат інформації не буде. Другими словами, джерело має суттєву і несуттєву інформацію. Надлишковість джерела – це несуттєва інформація джерела. Проте надлишкову інформацію можна використовувати на приймальній стороні СПІ для виправлення помилок, якщо такі вносяться в інформацію в процесі її передачі через канал зв'язку.

Джерело інформації видає певну кількість інформації, яку можна представити двома способами:

1. $nH(A)$, де n – кількість повідомлень (символів) необхідна для передачі заданої кількості інформації; $H(A)$ – реальна ентропія джерела.
2. $n_0 \cdot H_{max}(A)$, де n_0 – мінімальна кількість повідомлень (символів) необхідна для передачі заданої кількості інформації; $H_{max}(A)$ – максимальна ентропія джерела.

Ці два добутки представляють одну й ту саму кількість інформації. Отже на їх основі можна скласти рівняння: $n_0 H_{max}(A) = nH(A)$. З цього рівняння отримуємо, що $\frac{H(A)}{H_{max}(A)} = \frac{n_0}{n}$. Тоді $\rho = 1 - \frac{n_0}{n} = \frac{n-n_0}{n}$. Коефіцієнт надлишковості визначає (характеризує) питому вагу зайвих повідомлень, які видає джерело інформації.

Зауваження Зайві символи вимагають додаткового часу на передачу інформації.

Використання може бути двояке:

1. Знаючи про наявність надлишковості у джерела можна розв'язувати задачу усунення надлишкової інформації, тобто **стискання інформації**.
2. Надлишкову (несуттєву) інформацію можна використовувати на приймальній стороні СПІ для виправлення помилок, якщо такі вносяться в сигнал в процесі його передачі по каналу зв'язку.

2.3 Продуктивність джерел дискретних повідомлень

Швидкість появи інформації (аналогічне поняттю швидкість руху автомобіля) представимо відношенням кількості інформації, яка міститься в послідовності повідомлень a_T до тривалості цієї послідовності T . Враховуючи, що швидкість появи інформації є величина статистична, для підвищення достовірності результату розрахунку необхідно збільшувати значення T .

Математичний вираз для швидкості появи інформації на основі сказаного буде мати такий вигляд: $R = \lim_{T \rightarrow \infty} \frac{I(a_T)}{T}$, $[\frac{\text{дв.од}}{\text{сек}}]$ або $[\frac{\text{біт}}{\text{сек}}]$.

Проведемо дослідження наведеного виразу з метою визначення умов, коли швидкість появи інформації матиме максимальне значення. Для цього представимо: $I(a_T) = nH(A)$; $T = n\bar{\tau}$, де a_T – послідовність повідомлень на виході джерела в інтервалі часу; n – кількість повідомлень в послідовності a_T ; $\bar{\tau}$ – середнє значення тривалості одного повідомлення. Тоді $R = \lim_{n \rightarrow \infty} \frac{nH(A)}{n\bar{\tau}} = \frac{H(A)}{\bar{\tau}}$.

У випадку, якщо повідомлення джерела є незалежними $\bar{\tau}$ визначається так $\bar{\tau} = \sum_{i=1}^m \tau_i P(a_i)$, де τ_i – тривалість i -го повідомлення; m – алфавіт джерела (кількість видів повідомлень, які використовуються для видачі інформації від джерела); $P(a_i)$ – ймовірність появи i -го повідомлення.

Якщо повідомлення мають фіксовану тривалість, то $\bar{\tau} = \tau$, то $R = \frac{H(A)}{\tau}$.

Тоді продуктивність джерела дискретних повідомлень можна записати в такому вигляді: $R_D = \max \frac{H(A)}{\bar{\tau}}$.

У свою чергу відношення $\frac{H(A)}{\tau}$ матиме максимальне значення якщо $H(A) = H_{\max}(A)$, а $\tau = \tau_{\min}$, тоді $R_D = \frac{H_{\max}(A)}{\tau_{\min}}$.

Зауваження Все сказане вище про параметри джерел повідомлень має місце при умові, що повідомлення джерела є незалежними.

Якщо за приклад повідомлень взяти букви українського алфавіту можна встановити, що ймовірність появи кожної букви залежить від того, яка буква була попередньою. Наприклад, ми передаємо букву «ь». Наступною буквою вже ніяк не може бути «ь» (при передачі осмисленого тексту). Тобто в загальному випадку треба розуміти що джерело передає залежні повідомлення.

Умова незалежних повідомлень виконується тоді, коли повідомлення представляють відліки неперервного випадкового сигналу, інтервал між якими Δt рівний або перевищує інтервал кореляції τ_K ($\Delta t \geq \tau_K$). У випад-

ку коли $\Delta t < \tau_K$, відліки представляють собою залежні повідомлення. Ця умова є обов'язковою при реалізації цифрових методів передачі неперервних сигналів: диференціальної імпульсно-кової модуляції, кодування передбачених значень, дельта-модуляції.

Розділ 3

Сигнали та їх класифікація

3.1 Детерміновані та випадкові сигнали.

Детерміновані сигнали – це сигнали, які можна описати явними математичними залежностями і значення яких у будь-який момент часу або в довільній точці простору (або в залежності від будь-яких інших аргументів) є апіорно відомими, або можуть бути досить точно визначені.

Сигнали, закони зміни яких неможливо описати явними математичними залежностями, оскільки вони носять випадковий характер, відносяться до випадкових. Їх сукупність оцінюється статистичними характеристиками процесу: законами розподілу ймовірностей, кореляційними функціями, спектральними густинами енергії. Реальні сигнали завжди випадкові. На практиці здебільшого мають місце квазидетерміновані сигнали, що описуються функціями з невідомими випадковими параметрами.

Як **приклад** детермінованого сигналу можна навести синусоїдальну хвилю, яка графічно описується синусоїдою. 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.

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.

Неперервні сигнали (рис. 5.2, а) зображуються функцією, безперервною в часі на відрізку спостереження $x(t)$, а дискретні (рис. 5.2, б) поступають тільки в певні моменти часу та зображаються дискретною функцією.

До елементарних детермінованих сигналів належить зокрема і одини-

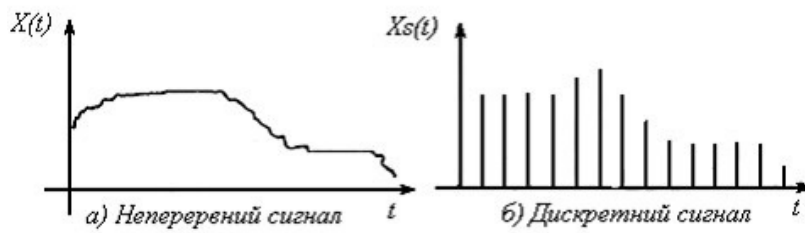
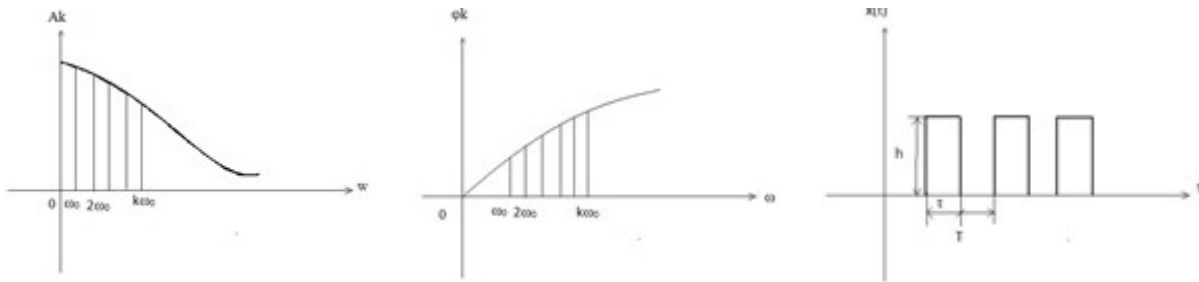


Рис. 3.1: Неперервні та дискретні сигнали

чна функція (стрибок).

Сукупність амплітуді відповідних частот гармонік прийнято називати спектром амплітуд. Сукупність початкових фаз і відповідних частот гармонік називаються спектром фаз. Спектр амплітуд і спектр фаз однозначно визначають сигнал, проте для багатьох практичних задач достатньо обмежитися розглядом тільки спектра амплітуд.



Характерною рисою спектра періодичного сигналу є його переривчастість (дискретність). Відстань між сусідніми спектральними лініями однакова і дорівнює частоті основної гармоніки.

Для опису випадкових процесів використовують методи теорії ймовірностей. У загальному випадку повною характеристикою випадкового процесу є його багатовимірна густина ймовірностей. Для стаціонарних гаусівських процесів одновимірна густина ймовірностей визначається дисперсією випадкового процесу. Для опису гаусівських процесів достатніми характеристиками є й кореляційна функція процесу. Однією з характеристик випадкового сигналу є його спектральна густина 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 густиною в обмеженій смузі частот потужність є скінченною і її можна визначити.

В більшості випадків для характеристики випадкових процесів використовують моментні функції перших двох порядків: математичне очікування, дисперсію, а також кореляційну функцію (оцінювання ступеня статичної залежності миттєвих значень процесу $U(t)$ в будь-які моменти часу t_1 та t_2).

Розглянемо спектральне подання стаціонарних **випадкових процесів**. 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:

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

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:

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

де підінтегральна функція є спектральною густиною потужності даної реалізації випадкового процесу: $W(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \left| \frac{X_T(f)}{T} \right|^2$.

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 періодичних процесів допустимо усереднювати за однією тривалою в часі реалізацією.

3.2 Спектри дискретних сигналів

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 – 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, щоб розрахувати спектр частот для сигналів змінних у часі.

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.

Зауваження Вихідний спектр отримується з вхідного простим множенням на функцію відклику системи $A(\omega)$. 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, кусково-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 , $K\omega_0$, φ_k – амплітуда, частота та початкова фаза k -тої

гармонічної складової; $\omega_0 = \frac{2\pi}{T}$ – частота основної (першої) гармоніки; T – 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 $A_k = \frac{2}{T} \int_t^{t+T} x(t) \exp\{-jK\omega_0 t\} dt$.

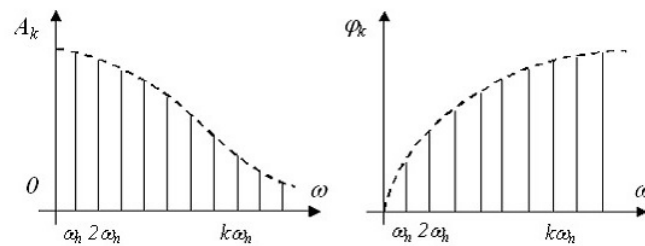


Рис. 3.2: Графічні зображення спектра амплітуд і спектра фаз періодичного сигналу

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 зменшуються і в границі при $T \rightarrow \infty$ 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 $S(j\omega) \exp(j\omega t) d\omega$ – спектральна густина сигналу; $S(\omega) = |S(j\omega)|$ – амплітудно-частотна характеристика сигналу; $\phi(\omega)$ – фазочастотна характеристика сигналу. Попередній вираз називається формулою оберненого перетворення Фур'є.

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.

3.3 Спектри одиничних та періодичних імпульсних послідовностей

Розглянемо спектри деяких найбільш важливих для радіотехніки періодичних та неперіодичних ЕРС.

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

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

$$\begin{aligned}\bar{S}(\omega) &= \bar{S}_2(\omega) + \bar{S}_1(\omega) = \bar{S}_1(\omega)(1 + e^{-j\omega\tau}) = \bar{S}_1(\omega)[1 + \cos\omega\tau - j\sin\omega\tau] = \\ &= S_1(\omega)\sqrt{(1 + \cos\omega\tau)^2 + \sin^2\omega\tau}e^{-j\psi(\omega)} = \bar{S}_1(\omega)2\cos\frac{\omega\tau}{2}e^{-j\psi(\omega)},\end{aligned}$$

де $\psi(\omega) = \arctan \frac{\sin\omega\tau}{1 + \cos\omega\tau}$.

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.

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

Розклад в ряд Фур'є періодичної послідовності прямокутних імпульсів подається у вигляді $x(t) = \frac{\tau h}{T} [1 + 2 \sum_{k=1}^{\infty} \frac{\sin \frac{K\omega_0\tau}{2}}{\frac{K\omega_0\tau}{2}} \cos K\omega_0\tau]$. Спектр амплітуд такого сигналу показаний на рис. 3.4.

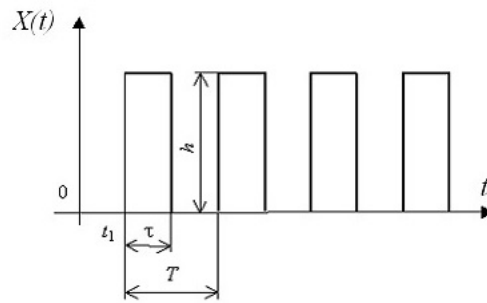


Рис. 3.3: Прямокутні імпульси

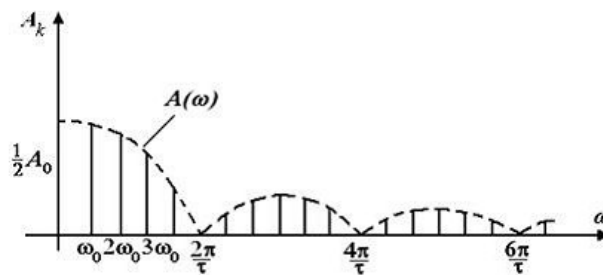


Рис. 3.4: Спектр амплітуд

Огинаюча 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.

Для періодичної послідовності імпульсів прямокутної форми тривалістю $\tau = T/2$ 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 густина такого сигналу визначається виразом $S(j\omega) = \int_{-\infty}^{\infty} x(t) \exp\{-j\omega t\} dt = \int_{-\frac{T}{2}}^{\frac{T}{2}} h \exp\{-j\omega t\} dt$. Енергія сигналу, зосереджена в смузі частот від 0 до ω_0 ,

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

Розділ 4

Способи перетворення сигналів

4.1 Дискретизація та модуляція аналогових сигналів. Дискретизація сигналів в часі, за рівнем та кодування безперервних 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 входньою 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 входного 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: $(c_1, c_2, \dots, c_n) = A[s(t)]$, де A – оператор дискретизації. Запис операції відновлення сигналу $s(t) : s'(t) = B[(c_1, c_2, \dots, c_n)]$. Вибір операторів A та B визначається необхідною точністю відновлення сигналу.

Найбільш простими є **лінійні оператори**. У загальному випадку: $c_n = \int q_n(t)s(t)dt$, де $q_n(t)$ – система вагових функцій.

Зауваження Відліки в останньому виразі пов'язані з операцією інтегрування, що забезпечує високу завадостійкість дискретизації. Однак у силу складності технічної реалізації «зваженого» інтегрування, останнє використовується досить рідко, при високих рівнях перешкод.

Більш **широке поширення** одержали методи, при яких сигнал $s(t)$ замінюється сукупністю його миттєвих значень $s_n(t)$ у моменти часу t_n . Роль вагових функцій у цьому випадку виконують решітчасті функції. Відрізок часу t між сусідніми відліками називають кроком дискретизації. Дискретизація називається рівномірною із частотою $F = \frac{1}{t}$, 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.

Оптимальними є методи дискретизації, що забезпечують мінімальний числовий ряд при заданій похибці відтворення сигналу. При неортогональних базисних функціях використовуються, в основному, статичні алгебраїчні поліноми вигляду: $s'(t) = c_n t_n$.

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 . 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).

4.2 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 відста-

ні $T_k = \frac{1}{2f_c}$. Аналітично **теорема Котельникова** записується інтерполяційним рядом $x(t) = \sum_{k=-\infty}^{\infty} x(k\Delta t) \frac{\sin[\omega_c(t-k\Delta t)]}{\omega_c(t-k\Delta t)}$, де $\Delta t = \frac{\pi}{\omega_c} = \frac{1}{2f_c}$.

Неперервна функція з обмеженим спектром може бути подана у вигляді суми нескінченно великої кількості членів, кожний з яких є множителем 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 $f = 2f_c = \frac{\omega_c}{\pi}$, 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.

4.3 Методи модуляції дискретних сигналів

Цифрова обробка сигналів (ЦОС) (digital signal processing) – це область обчислювальної техніки, що динамічно розвивається та охоплює як технічні, так і програмні засоби. Спорідненими областями для цифрової обробки сигналів є теорія інформації, зокрема, теорія оптимального прийому сигналів і теорія розпізнавання образів. При цьому в **першому** випадку основним завданням є виділення сигналу на фоні шумів і перешкод різної фізичної природи, а в **другому** – автоматичне розпізнавання, тобто класифікація та ідентифікація сигналу.

При **цифровій обробці** використовується подання сигналів у вигляді послідовностей чисел або символів. Ціль такої обробки може полягати в оцінці характерних параметрів сигналу або в перетворенні сигналу у форму, що у деякому змісті більше зручна. Формули класичного чисельного аналізу, такі, як формули для інтерполяції, інтегрування й диференціювання, безумовно є алгоритмами цифрової обробки. Цифрова обробка сигналів застосовується в таких різних областях, як біомедицина, акустика, звукова локація, радіолокація, сейсмологія, зв'язок, системи передачі даних, ядерна техніка, і багатьох інших.

Методами ЦОС є математичні співвідношення або алгоритми, відповідно до яких виконуються обчислювальні операції над цифровими си-

гнолами. До них належать алгоритми цифрової фільтрації, спектрально-кореляційного аналізу, модуляції та демодуляції сигналів, адаптивної обробки та ін.

Засобами реалізації ЦОС є жорстка логіка, програмовані логічні інтегральні схеми, мікропроцесори загального призначення, мікроконтролери, персональні комп'ютери (комп'ютерна обробка сигналів) та цифрові сигнальні процесори.

Сигнал – це інформаційна функція, що несе повідомлення про фізичні властивості, стан або поведінку будь-якої фізичної системи, об'єкта або середовища, а метою обробки сигналів у самому загальному змісті можна вважати отримання певних інформаційних відомостей, які відображені в цих сигналах (корисна або цільова інформація) і перетворення цих відомостей у форму, зручну для сприйняття і подальшого використання.

4.4 Модуляція гармонічних коливань: амплітудна, частотна, фазова.

4.4.1 Амплітудна, частотна, фазова модуляція, порівняння спектрів сигналів з різними видами модуляції

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. 4.1. 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).

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. 4.2. На діаграмі (рис. 4.2, а) показана послідовність бітів вихідної



Рис. 4.1: Амплітудно-частотна характеристика каналу тональної частоти

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. 4.2, 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 завадостійкості, але часто застосовується в поєднанні з іншим видом модуляції – фазовою модуляцією.

При **частотній модуляції** (рис. 4.2, 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.

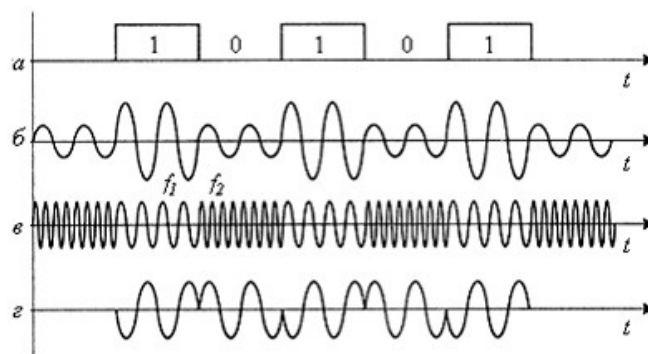


Рис. 4.2: Різні види модуляції

При **фазовій модуляції** (рис. 4.2, 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.

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. 4.3, 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 $f_0, 3f_0, 5f_0, 7f_0, \dots$, де $f_0 = \frac{N}{2}$. Амплітуди цих гармонік зменшуються досить повільно – з коефіцієнтами $1/3, 1/5, 1/7, \dots$ від амплітуди гармоніки f_0 (рис. 4.3, а). В результаті спектр потенційного коду потребує для якісної передачі широку смугу пропускання. Крім того, потрібно врахувати, що реально спектр сигналу постійно змінюється в залежності від того, які дані передаються по лінії зв'язку.

Приклад 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).

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

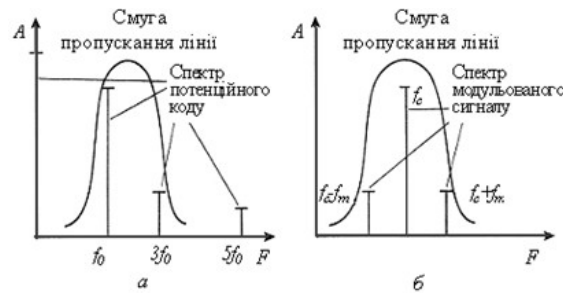


Рис. 4.3: Амплітуда модуляції неперервного процесу

інформаційного параметра синусоїди, що збігається зі швидкістю передачі даних при використанні двох рівнів амплітуди (рис. 4.3, б). Частота 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 використовуються 4 рівні амплітуди, то пропускну здатність каналу підвищується до 3100 біт/с.

При **фазовій і частотній модуляції** спектр сигналу виходить більш складним, ніж 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.

4.4.2 Імпульсні види модуляції

При імпульсній модуляції у якості несучої використовується періодична послідовність імпульсів, один з параметрів яких змінюється за законом інформаційного сигналу.

Періодична послідовність імпульсів характеризується такими параметрами: тривалістю τ ; періодом слідування T_i чи тактовою частотою $f = \frac{1}{T_i}$; амплітудою U_0 ; положенням (фазою) імпульсів відносно тактових точок t_i ; сквапністю $Q_\tau = \frac{T_i}{\tau}$, яка в імпульсних системах може перевищувати 2000.

В залежності від вибраного модулюючого параметра розрізняють наступні види імпульсної модуляції: амплітудно-імпульсну (АІМ), широтно-імпульсну (ШІМ) та часову імпульсну, різновидами якої є фазо-імпульсна

(ФІМ) і частотно-імпульсна (ЧІМ) модуляції. Можливі дискретні види ФІМ (ДФІМ) і ЧІМ (ДЧІМ). До складних видів імпульсної модуляції відносяться дельта-модуляція (ДМ) та імпульсно-кодова модуляція (ІКМ).

4.4.3 Амплітудно-імпульсна модуляція (АІМ)

При АІМ амплітуда імпульсів (рис. 4.4, а) змінюється за законом модулюючого (інформаційного) сигналу U_{Ω} (рис. 4.4, б), а інші параметри імпульсів залишаються незмінними.

Сигнал при АІМ можна записати у вигляді:

$$U_{AIM} = [U_0 + \Delta U f(t)] \sum_i F(t - t_i - iT_i) =$$

$$= U_0 [1 + m_{\alpha} f(t)] \sum_i F(t - t_i - iT_i),$$

де $f(t)$ – функція модулюючого (інформаційного) сигналу у часі; $F(t)$ – функція, що описує форму одиничного імпульсу; $m_{\alpha} = \frac{\Delta U}{U_0}$ – коефіцієнт глибини модуляції.

Розрізняють два види АІМ: неперервну, так звану АІМ 1-го роду (АІМ-1), при якій величина імпульсів змінюється відповідно до зміни модулюючого сигналу (рис. 4.4, в), та плоску – АІМ 2-го роду (АІМ-2), при якій амплітуда імпульсу визначається значеннями модулюючого сигналу в тактових точках (рис. 4.4, г). Так як тривалість імпульсів завжди набагато менше періоду модулюючого сигналу, різниця між АІМ-1 і АІМ-2 практично відсутня.

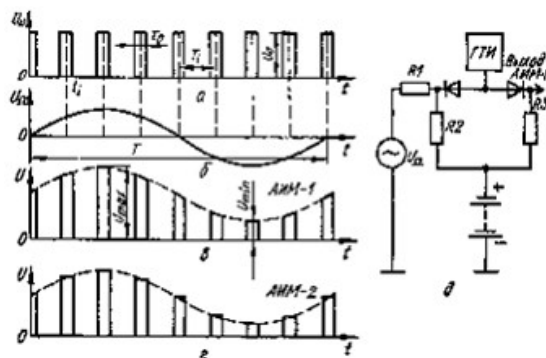


Рис. 4.4: Амплітудно-імпульсна модуляція: а – імпульси несучої частоти; б – інформаційний сигнал; в – сигнали АІМ-1; г – сигнали АІМ-2; д – схема модулятора при АІМ-1

4.4.4 Широтно-імпульсна модуляція (ШІМ)

При ШІМ тривалість імпульсів несучої частоти (рис. 4.5, а) змінюється за законом інформаційного сигналу (рис. 4.5, б) при незмінних інших параметрах імпульсів.

Розрізняють однобічну (ОШІМ) і двобічну (ДШІМ) широтно-імпульсну модуляцію. В ОШІМ зміна тривалості імпульсів виконується переміщенням тільки одного (переднього чи заднього) фронту (рис. 4.5, в), а при ДШІМ – двох фронтів, симетрично їх центрів, розміщених в тактових точках (рис. 4.5, г).

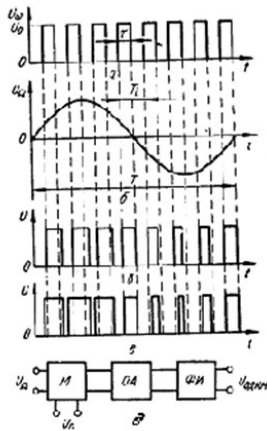


Рис. 4.5: Широтно-імпульсна модуляція: а – імпульси несучої частоти; б – інформаційний сигнал; в – ОШІМ-сигнали; г – ДШІМ-сигнали; д – схема формування сигналів ОШІМ

Схема формування сигналів ОШІМ приведена на рис. 4.5, д. ОШІМ-сигнал можна отримати з виходу модулятора M , якщо до нього підвести в якості несучої пилоподібну напругу $U_{\text{П}}$. При поданні на інший вхід M інформаційного сигналу U_{Ω} , на його виході отримуємо модульовану за амплітудою пилоподібну напругу. Обмеживши її двостороннім обмежувачем амплітуд (ОА), отримуємо імпульси, модульовані за шириною (тривалістю). Для отримання прямокутної форми імпульсів з ОШІМ на виході ОА включають формувач імпульсів (ФІ).

Завадостійкість ШІМ значно більше завадостійкості АІМ, тому ШІМ широко використовується у телевимірюваннях.

4.4.5 Часова імпульсна модуляція

При часовій імпульсній модуляції за законом модулюючої напруги змінюється положення в часі імпульсів відносно тактових точок. В залежності

від закономірностей зміни положення імпульсів відносно тактових точок розрізняють фазо-імпульсну модуляцію (ФІМ) і частотно-імпульсну модуляцію (ЧІМ).

4.4.6 Фазо-імпульсна модуляція (ФІМ)

Фазо-імпульсна модуляція (ФІМ) характеризується тим, що зсув у часі імпульсів відносно тактових точок, тобто девіація частоти слідування імпульсів (рис. 4.6, а), пропорційний амплітуді модулюючого сигналу (рис. 4.6, б) і не залежить від його частоти (рис. 4.6, в). Розрізняють ФІМ першого роду (ФІМ-1) і другого роду (ФІМ-2). При ФІМ-1 величина зсуву з часом імпульсів відносно тактових точок пропорційна значенню модулюючої напруги в моменти посилення цих імпульсів, а при ФІМ-2 – в тактових точках. Так як зсув імпульсів значно менше періоду модулюючого сигналу, то різниця між ФІМ-1 і ФІМ-2 практично відсутня.

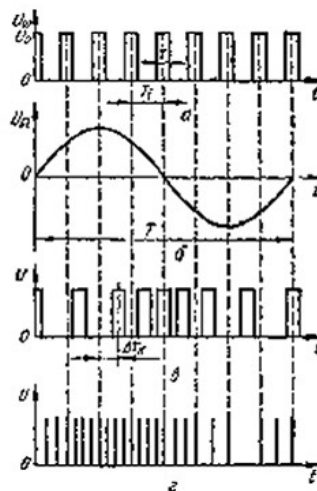


Рис. 4.6: Часово-імпульсна модуляція: а – імпульсна несуча; б – модулюючий сигнал; в – сигнал ФІМ; г – сигнал ЧІМ

Для отримання ФІМ-сигналів можна застосувати схему представлену на рис. 4.7, а. В даній схемі після отримання імпульсів, модульованих за шириною (тривалістю) на виході широтно-імпульсного модулятора ШІМ (ДІМ) (рис. 4.7, б), їх подають через диференціальну ланку **РС**, на виході якої отримують короткі гострокінцеві імпульси (рис. 4.7, в). Після обмеження їх по амплітуді обмежувачем амплітуди (ОА) отримують імпульси однієї полярності, які подаються на формуючий каскад (ФК) для отримання імпульсів, модульованих за фазою однакової тривалості і амплітуди.

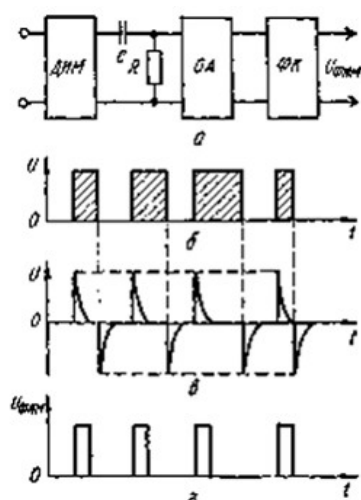


Рис. 4.7: Формування ФІМ-сигналів: а – схема формування сигналу з ФІМ; б – сигнал ДІМ; в – сигнал на виході диференціюючого ланцюга; г – сигнал ФІМ

Спектр ФІМ-сигналів подібний спектру при ОШМ лише з тією різницею, що деякі складові спектру мають дещо меншу амплітуду (бокові частоти спектру при ФІМ мають значно менші амплітуди, ніж при ОШМ, що ускладнює їх виділення при прийомі). Тому в системах з ФІМ отримано послідовність імпульсів з ФІМ, перетворюють послідовність імпульсів з АІМ чи ШІМ, з якої шляхом фільтрації виділяють модулюючий сигнал.

4.4.7 Частотно-імпульсна модуляція (ЧІМ)

ЧІМ характеризується тим, що частота слідування імпульсів несучої змінюється за законом модулюючого сигналу (рис. 4.6, г), інші параметри імпульсів при цьому залишаються незмінними. При збільшенні миттєвого значення повідомлення частота імпульсів підвищується, а при зменшенні – знижується. Таким чином здійснюється модуляція за частотою імпульсів, при якій тривалість імпульсів залишається постійною, змінюється лише інтервал між ними. Ширина смуги частот визначається тривалістю імпульсу.

Головною умовою для ЧІМ являється те, що тривалість інформаційного імпульсу (чи тривалість паузи) повинна бути більше чи рівна десяти періодам несучої послідовності.

Існує багато реалізацій ЧІМ, одна з них полягає в наступному:

- якщо в інформаційній послідовності, що складається з імпульсів, з'являється 1 (є імпульс), то частота несучої послідовності зменшу-

ється в два рази;

- якщо в інформаційній послідовності з'являється 0 (немає імпульсу), то частота несучої послідовності збільшується в два рази.

Завадостійкість ЧІМ приблизно така ж як і у ФІМ.

Розділ 5

Випадкові сигнали та їх характеристики

5.1 Неперервні та дискретні випадкові процеси

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 її кореляційна функція процесу. Однією 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.

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 $U_1 = U(t_1), \dots, U_n = U(t_n)$, взятих в різні моменти часу t_1, t_2, \dots, t_n .

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

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

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

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

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

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

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 5.1.

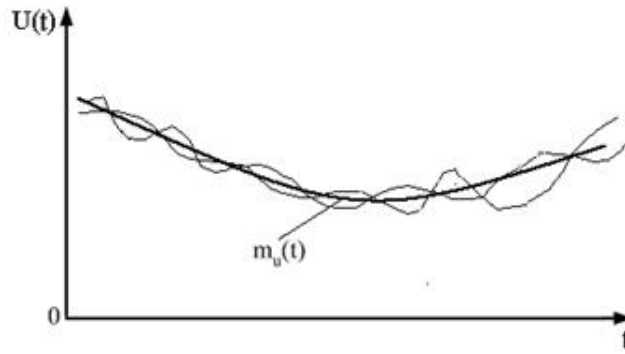


Рис. 5.1: Математичне сподівання для різних процесів

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 $U(t_2)$

$$R_U(t_1, t_2) = m_1[U(t_1), U(t_2)].$$

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 не-випадкова функція $m_x(t)$, яка при кожному даному значенні аргументу дорівнює математичному сподіванню значення випадкової функції при тому ж значенні аргументу $m_x(t) = M[X(t)]$.

Кореляційною функцією випадкової функції $X(t)$ називається не-випадкова 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

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

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

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

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

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

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

5.2 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. 5.2, а) зображуються функцією, безперервною в часі на відрізку спостереження $x(t)$, а дискретні (рис. 5.2, б) поступають тільки в певні моменти часу та зображаються дискретною функцією $x_g(t)$.

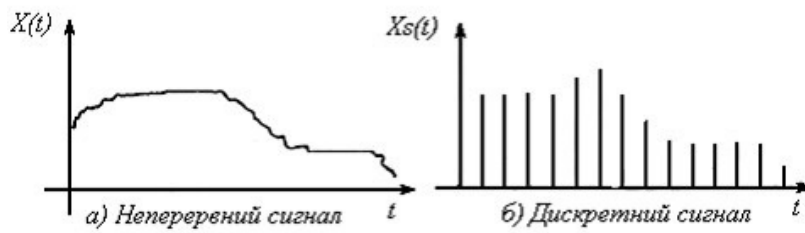
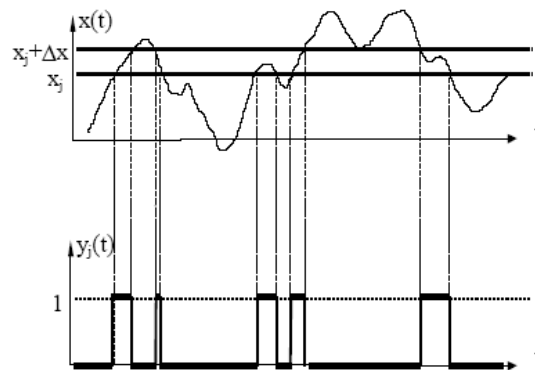


Рис. 5.2: Неперервні та дискретні сигнали

Рис. 5.3: Формування реалізацій $y_j(t)$ випадкових функцій $\eta_j(t)$ з реалізації $x(t)$ випадкового процесу $\xi(t)$

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. Сукупність детермінованих сигналів може подавати випадковий процес. Навіть при експериментальному аналізі досить доречно вводити допоміжні функції (рис. 5.3).

Існуючі аналізатори законів розподілу дозволяють визначати емпіричні функції розподілу та гистограми досліджуваних випадкових процесів. Блок-схема аналізатора наведена на рис. 5.4.

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;

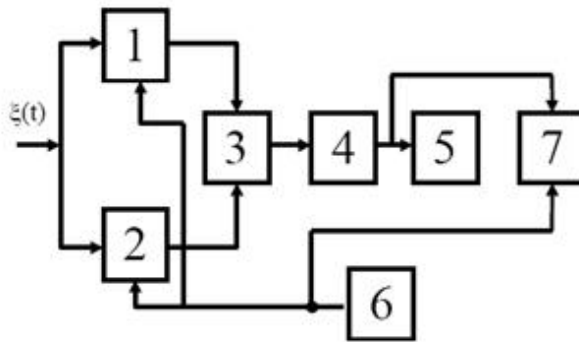


Рис. 5.4: Блок-схема аналізатора законів розподілу

- 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.

5.3 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. Відомо два поняття: стаціонарність в обмеженому розумінні і стаціонарність у широкому.

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

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

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 $t_3 - t_1 = \tau_2$

$$\begin{aligned} \omega_3(x_1, t_1; x_2, t_2; x_3, t_3) &= \omega_3(x_1, x_2, x_3; t_2 - t_1, t_3 - t_1) = \\ &= \omega_3(x_1, x_2, x_3; \tau). \end{aligned}$$

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, кореляційна функція стаціонарного процесу також залежить тільки від різниці часу $t_2 - t_1 = \tau_1$.

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

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

де $\tau = t_2 - t_1$.

Дисперсія стаціонарної випадкової функції постійна

$$D_x(t) = K_x(t_1, t_2) = k_x(0) = \text{const.}$$

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 $Y_k, Z_k, (k = 0, 1, 2, \dots, n)$ – 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:

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

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:

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

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

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

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

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

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

5.4 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 отримуємо 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:

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

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

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

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

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:

$$R(\tau) = 2 \int_0^{\infty} W(f)\cos(2\pi f\tau)df,$$

$$W(f) = 2 \int_0^{\infty} R(\tau)\cos(2\pi f\tau)d\tau.$$

Розділ 6

Завади, їх джерела

6.1 Достовірність та завадостійкість

Достовірність – ступінь відповідності прийнятої інформації тій, що передана. Оцінка достовірності – відношення кількості правильно прийнятих сигналів ($k_{\text{пп}}$) до загальної кількості переданих ($k_{\text{пер}}$) сигналів за певний проміжок часу (при достатній кількості переданих повідомлень): $O_{\text{д}} = \frac{k_{\text{пп}}}{k_{\text{пер}}}$.

Невідповідність отриманої інформації переданій зумовлена спотвореннями, які бувають лінійними, нелінійними, випадковими. Усі випадкові спотворення викликаються завадами в каналі та апаратурі зв'язку.

Завади можуть або подавляти корисний сигнал або утворювати неправдивий сигнал.

Вплив завади на сигнал може бути двояким:

1. Якщо завада $\zeta(t)$ складається з корисним сигналом $S(t)$, і на вхід приймача діє їх сума, то така завада називається адитивною: $X(t) = S(t) + \zeta(t)$.
2. Якщо ж результуючий сигнал утворюється за рахунок добутку завади й сигналу, то така завада називається мультиплікативною: $X(t) = S(t) \cdot \zeta(t)$.

Мультиплікативні завади виникають при використанні радіозв'язку. Більшість систем передачі інформації використовують провідний зв'язок, де мають місце тільки адитивні завади. Вони поділяються за типом і за джерелом завад.

Типи завад – імпульсні, флуктуаційні, гармонічні.

Джерела завад – зовнішні і внутрішні. До внутрішніх відносять теплові та фон наведення. До зовнішніх – промислові, атмосферні, космічні.

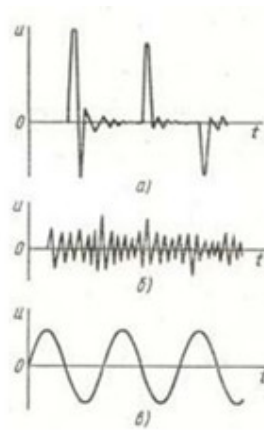


Рис. 6.1: Форма адитивних завад: а) – імпульсні; б) – флуктаційні; в) – гармонічні

За формою адитивні завади поділяються на імпульсні, флуктаційні та гармонічні (рис. 6.1).

Імпульсні завади – послідовні імпульси з випадковою амплітудою, тривалістю й моментом появи окремих імпульсів.

Флуктуаційні завади або шуми мають вигляд безперервних випадково змінних коливань. Їх найважливіші характеристики: потужність сигналу завади, закон розподілу амплітуд, вигляд енергетичного спектра або функції кореляції. Найчастіше зустрічаються флуктуаційні завади, амплітуди яких підпорядковуються закону нормального розподілу (для таких завад імовірність того, що амплітуда викиду перевищить потрібне значення ефективної напруги дуже мала). Флуктуаційні завади не мають постійної складової. При рівномірному за частотою спектру завад в лінії зв'язку ефективна напруга флуктуаційних завад на виході приймача пропорційна квадратному кореню із смуги пропускання каналу зв'язку: $U_{\text{зав}}^{\text{ef}} = \sqrt{\Delta F}$, а потужність пропорційна смузі: $P_{\text{зав}} = \Delta F$.

Для імпульсних завад потужність та амплітуда завади пропорційна смузі пропускання: $U_{\text{зав}} = k_u \Delta F$ та $P_{\text{зав}} = k_p \Delta F$.

На амплітуду й потужність завад суттєвий вплив має смуга пропускання каналу. Вона також впливає і на тип завади (для однакових завад у лінії зв'язку на вході вузькосмугового приймача завади можуть мати флуктуаційний характер, а для широкосмугового – імпульсний). Енергетичний спектр завади характеризує її розподіл за потужністю в діапазоні частот.

Завадостійкість – це здатність правильно сприймати інформацію, незважаючи на вплив завад.

Розглянемо завадостійкість елементарного сигналу для випадку впливу

флуктуаційних і імпульсних завад. Елементарний сигнал може приймати два значення: $A_1(t)$, $A_2(t)$.

Трансформування сигналу – це невиявлена зміна повідомлення, що виникає під час передачі під впливом завад і спричиняє прийом спотвореного сигналу. Імовірність заглушення каналу позначається через $P_{\text{заглуш}}$ або P_{10} (перетворює 1 у 0). Імовірність спотвореної неправдивої завади $P_{\text{н}}$ або P_{01} (перетворює 0 у 1). Таке порівняння полягає в визначенні різниці між прийнятим сигналом $x(t)$ із кожним зразковим $A_1(t)$ і $A_2(t)$ за певний проміжок часу t , де t – тривалість сигналу. Тоді знаходять:

$$I_1 = \int_0^{\tau} [x(t) - A_1]^2 dt; \quad I_2 = \int_0^{\tau} [x(t) - A_2]^2 dt.$$

Після цього констатують, що було передано сигнал, для якого I_i є мінімальним, тобто якщо $I_2 - I_1 > 0$, то вважають прийнятим сигнал $A_1(t)$, і навпаки – якщо $I_2 - I_1 < 0$ – тоді сигнал $A_2(t)$.

Розглянемо на прикладі принцип функціонування **ідеального приймача Котельникова**. По лінії зв'язку передають два елементарних сигналу $A_1(t) = 1$, наприклад – 5В та $A_2(t) = 0$, відповідно – 0,5В.

1. Спочатку було передано сигнал 1 (5В), який в лінії зв'язку було спотворено, і на вхід інтеграторів надійшов послаблений сигнал – 3В, тоді, для $\tau = 1$ маємо на інтеграторах:

$$I_1 = [3 - 5]^2 = 4\text{Вт};$$

$$I_2 = [3 - 0,5]^2 = 6,25\text{Вт}.$$

Після суматора $I_2 - I_1 = 6,25 - 4 = 2,25 > 0$, і тому робимо висновок, що по лінії зв'язку було передано сигнал $A_1(t) = 1$.

2. У другому випадку було передано сигнал $A_2(t) = 0$ (0,5В), середнє амплітудне значення якого за рахунок завад в лінії зв'язку збільшилося до 1В. Тоді для маємо на інтеграторах:

$$I_1 = [1 - 5]^2 = 16\text{Вт};$$

$$I_2 = [1 - 0,5]^2 = 0,025\text{Вт}.$$

Після суматора $I_2 - I_1 = 0,025 - 16 = -15,975 < 0$, і тому робимо висновок, що по лінії зв'язку було передано сигнал $A_2(t) = 0$.

Для оцінки завадостійкості передачі існує поняття питомої завади: $\sigma_0 = \frac{U_{\text{сер кв}}^{\text{зав}}}{\sqrt{\Delta F}}$, де $U_{\text{сер кв}}^{\text{зав}}$ – середнє квадратичне значення амплітуди завади.

При цьому величина, що характеризує потенційну завадостійкість дорівнює відношенню енергії сигналу до значення питомої завади: $a_0 = \frac{\int_0^T [A_1(t) - A_2(t)]^2 dt}{\sigma_0}$. Методи боротьби із завадами є одночасно методами підвищення завадостійкості сигналу і поділяються на категорії:

- що зменшують енергію завад (визначенні джерел завад, місця їх виникнення і зменшенні потужності їх виливу за рахунок: віддалення джерел завад від каналів зв'язку (не дозволяється прокладати разом силові та інформаційні кабелі); використання схем заглушення завад (використання фільтрів, іскрогасящих ланок та ін.));
- що базуються на збільшенні завадостійкості переданих сигналів (базується або на підвищенні енергії сигналу, що збільшує коефіцієнт завадостійкості, або в забезпеченні завадостійкості передачі шляхом: завадостійкого кодування; передачі інформації з повторенням; використання завадостійких методів модуляції; використання зворотного зв'язку.).

Метод передачі інформації із її повторенням використовують при відсутності зворотного зв'язку. Суть цього методу полягає у передаванні одного і того ж повідомлення декілька разів, запам'ятовуванні прийнятих повідомлень, порівнянні їх поелементно і складанні остаточного повідомлення, в якому включені елементи що вибрані «по більшості». Метод передачі із використанням зворотнього зв'язку полягає у тому, що порівнюються сигнали передачі по прямому і зворотньому каналах, і випадку їх співпадіння, повідомлення приймається.

6.2 Завадостійке кодування

Під завадостійкими кодами розуміють коди, що дозволяють виявляти або 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** (рекурентні або ланцюгові).

У загальному випадку кожна з N дозволених комбінацій може трансформуватися в будь-яку з N_0 можливих комбінацій, тобто усього є $N < N_0$ можливих варіантів передачі, із них:

- N варіантів безпомилкової передачі;
- $N(N - 1)$ варіантів переходу в інші дозволені комбінації;
- $N(N_0 - N)$ варіантів переходу в заборонені комбінації.

Таким чином, не всі перекручування можуть бути виявлені. Частка помилкових комбінацій, що виявляються, складає $\frac{N(N_0 - N)}{NN_0} = 1 - \frac{1}{N}$.

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 $(N_0 - N)$ випадках, рівних кількості заборонених комбінацій. Частка помилкових комбінацій, що виправляються, від загального числа помилкових комбінацій, що виявляються, складає $\frac{N_0 - N}{N(N_0 - N)} = \frac{1}{N}$.

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

Нехай необхідно побудувати код, що виявляє всі помилки кратністю t і нижче. Побудувати такий код – це означає із множини N_0 можливих вибрати N дозволених комбінацій так, щоб будь-яка з них у сумі за модулем два з будь-яким вектором помилок із вагою $W_t \leq t$ не дала б у результаті ніякої іншої дозволеної комбінації. Для цього необхідно, щоб найменша кодова відстань задовольняла умову $d_{min} \geq t + 1$.

У загальному випадку для усунення помилок кратності σ кодова відстань повинна задовольняти умову $d_{min} \geq 2\sigma + 1$.

Для виправлення всіх помилок кратності не більше σ і одночасно виявлення всіх помилок кратності не більше t і (при $t \geq \sigma$) кодова відстань повинна задовольняти умову $d_{min} \geq t + \sigma + 1$.

6.3 Зв'язок корегувальної спроможності з кодовою відстанню

Підвищення коригуючої здатності коду досягається при зберіганні n за рахунок зменшення множини N дозволених комбінацій (або зменшення кількості k інформаційних 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 $E2^k = EN$. Тому, що кількість помилкових комбінацій дорівнює $N_0 - N$, то код забезпечує виправлення не більше $N_0 - N$ комбінацій. Отже, необхідну умову для можливості виправлення помилок можна записати у вигляді:

$$NE \leq N_0 - N,$$

і отримаємо $N_0 \geq (1 + E)N$, або $N \leq \frac{2^n}{1+E}$. Ця формула 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 $E = C_n^1 = n$. У цьому випадку залежність набуде вигляду: $N = 2 \leq \frac{2^n}{1+n}$.

При побудові коду доцільно користуватися табл. 6.1. Потрібно при цьому мати на увазі, що код повинен також задовольняти умову $d_{min} > 3$.

Табл. 6.1: 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 $E_1 = C_n^1$;
- кількість можливих дворазових помилок $E_2 = C_n^2$;
- кількість можливих l -кратних помилок $E_l = C_n^l$.

Загальна кількість помилок $E = \sum_{i=1}^l C_n^i$. При цьому залежність набуде вигляду: $N \leq \frac{2^n}{1 + \sum_{i=1}^l C_n^i}$.

Ця умова є нижньою оцінкою для довжини коригуючого коду, тобто вона визначає необхідну мінімальну довжину коду n , що забезпечує ви-

правлення 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.

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.

6.4 Код Хемінга

Код Хемінга відноситься до систематичних кодів, в яких з n символів, які утворюють комбінацію, n_0 символів є інформаційними, а останні $k = n - n_0$ є 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 кодї Хемінга знаходять на основі таких міркувань. При передачі комбінації по каналу з шумами може бути спотворений довільний з n 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 2^k подій. Для цього повинна бути використана умова: $2^k \geq n + 1 = n_0 + k + 1$.

В таблиці 6.2 подана залежність між k і n_0 , яка отримана з цієї невірності, де k – число контрольних символів в кодї Хемінга, n_0 – число інформаційних символів.

Табл. 6.2: 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

$$\begin{array}{ccccccc} k_1 & k_2 & a_{04} & k_3 & a_{03} & a_{02} & a_{01} \\ |a_1| & |a_2| & a_3 & |a_4| & a_5 & a_6 & a_7 \end{array} .$$

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 складають 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 кожному перевірному 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.

6.5 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 $a^0 a^1 a^2 \dots a^{n-1}$ є дозволеною комбінацією циклічного коду, то комбінація $a^{n-1} a^0 a^1 a^2 \dots a^{n-2}$ також належить цьому коду.

Циклічні коди зручно розглядати, подаючи комбінацію двійкового коду не у вигляді послідовностей нулів і одиниць, а у вигляді полінома від фіктивної змінної x : $G(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_1x + a_0$, 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 полінома 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 x ; **множення** та **ділення** здійснюється за звичайними правилами множення та ділення логічних функцій, отримані при цьому коефіцієнти при рівних степенях змінної x додаються за модулем два.

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:

$$\begin{aligned} U_1(x) &= P(x); \\ U_2(x) &= P(x)x - C_2(x^n + 1); \\ U_3(x) &= P(x)x^2 - C_3(x^n + 1); \\ &\dots\dots\dots \\ U_n(x) &= P(x)x^{n-1} - C_n(x^n + 1), \end{aligned}$$

де C_2, C_3, \dots, C_n – коефіцієнти, що приймають значення 1 при $P(x)x^i > (x^n - 1)$ і значення 0 при $P(x)x^i < (x^n - 1)$.

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 (якщо під впливом завад дозволена кодова 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, A) -код може бути отриманий шляхом множення простого A -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 одно-

член 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:

$$\frac{x^{n-k}G(x)}{P(x)} = Q(x) + \frac{R(x)}{P(x)},$$

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.

На практиці зазвичай використовується перший спосіб одержання циклічного коду.

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 $P(x)$. Степінь утворюючого полінома повинен дорівнювати числу перевірних розрядів ρ .

Оскільки в циклічному коді розпізнавачами 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.

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 $2^\rho - 1$ ненульових залишків, то необхідною умовою виправлення будь-якої одиничної помилки є виконання нерівності: $2^\rho - 1 \geq n$.

Звідси знаходять степінь утворюючого полінома: $\rho = n - k \geq \log_2(n + 1)$ і загальну довжину n кодової комбінації.

Оскільки утворюючий многочлен $P(x)$ повинен входити як співмножник в розкладання двочлена $(x^n + 1) = (x^{2z-1} + 1)$ можна вибрати утворюючий поліном.

Однак не всякий многочлен 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.

6.6 Коди Боуза-Чоудхури-Хоквингема

Коди Боуза-Чоудхури-Хоквингема (БЧХ) являють собою великий клас кодів, здатних виправляти кілька помилок і займають помітне місце в теорії і практиці кодування. Інтерес до кодів БЧХ визначається щонайменше наступними чотирма обставинами:

1. серед кодів БЧХ при невеликих довжинах існують гарні (але, як правило, не кращі з відомих) коди;
2. відомі відносно прості й конструктивні методи їх кодування і декодування (хоча якщо єдиним критерієм є простота, то перевага варто віддати іншим кодам);
3. коди Ріда-Соломона, що є широко відомим підкласом недвійкових кодів, мають певні оптимальні властивості і прозору вагову структуру;
4. повне розуміння кодів БЧХ, як видно, є найкращою відправною крапкою для вивчення багатьох інших класів кодів.

Одним із класів циклічних кодів, здатних виправляти багатократні помилки, є коди БЧХ.

Примітивним кодом БЧХ, що виправляє t_u помилок, називається код довжиною $n = qt - 1$ над $GF(q)$, для якого елементи $\alpha^1, \alpha^2, \dots, \alpha^{2t_u}$ є коренями породжуючого багаточлена. Тут примітивний елемент $GF(qt)$. Породжуючий багаточлен визначається з виразу $g(x) = \text{НОК}[f_1(x), f_2(x), \dots$

$\dots, f_{2t_u}(x)]$, де $f_1(x), f_2(x), \dots$ - мінімальні багаточлени корінь $g(x)$. Число перевірочних елементів коду БЧХ задовольняє співвідношенню $r = n - k \leq mt_u$.

На практиці при визначенні значень породжуючого багаточлена користуються спеціальною таблицею мінімальних багаточленів, і виразом для породжуючого багаточлена $g(x) = f_1(x)f_3(x)\dots f_j(x)$. При цьому робота здійснюється в наступній послідовності.

По заданій довжині коду n і кратності помилок t_u , що виправляють, визначають:

- з виразу $n = m - 1$ значення параметра m , що є максимальним ступенем співмножників $g(x)$;
- з виразу $j = 2t_u - 1$ максимальний порядок мінімального багаточлена, що входить у число співмножників $g(x)$;
- користуючись таблицею мінімальних багаточленів, визначається вираз для $g(x)$ залежно від m і j . Для цього з колонки, що відповідає параметру m , вибираються багаточлени з порядками від 1 до j , які в результаті перемножування дають значення $g(x)$.

Примітка. У виразі для $g(x)$ утримуються мінімальні багаточлени тільки для непарних ступенів α , тому що звичайно відповідні їм мінімальні багаточлени парних ступенів α мають аналогічні вирази. Наприклад, мінімальні багаточлени елементів відповідають мінімальному багаточлену елемента α^1 , мінімальні багаточлени елементів відповідають мінімальному багаточлену і т.і.

6.7 Коди Ріда-Соломона

Коди Ріда-Соломона (РС) (Reed-Solomon) – підклас недвійкових кодів БЧХ із параметрами: символи коду вибираються з поля $GF(q)$, $q = 2^m$, m – ціле, довжина блоку (число q -ічних символів) $N = (q - 1)$, кількість інформаційних символів $k = (N - 2q_{\text{випр}})$, кодова відстань $d_{\text{min}} = (2q_{\text{исп}})$. Можливо також розширення коду до $N = q$ або до $N = (q + 1)$. Алгебраїчна структура коду РС описується з використанням арифметики поля Галуа $GF(2m)$. При фіксованій швидкості коду й довжині блоку коди РС забезпечують найбільшу мінімальну віддаль між комбінаціями. Їх зручно

використовувати для виправлення пакетів помилок, а також у каскадних системах кодування як зовнішні коди.

Ефективне використання циклічних властивостей дозволених комбінацій циклічних кодів дозволяє реалізувати досить прості алгоритми декодування ЦК. Уважається, що складність реалізації алгоритмів декодування циклічних кодів описується степінною функцією $C_D = n^k$, де k – мале число, величина якого залежить від конкретної реалізації алгоритму.

Розділ 7

Структура і характеристики згорткових кодів

7.1 Методи опису згорткових кодів

Згорткові коди утворюють підклас неперервних кодів. Найменування згортковий код походить від того, що результат кодування на виході кодера утворюється як згортка інформаційної послідовності, що кодується, з імпульсною реакцією кодера.

Кодер ЗК містить один або кілька регістрів елементів затримки й перетворювач інформаційних послідовностей у кодові послідовності. Процес кодування виконується неперервно.

Схема простого кодера показана на рис. 7.1

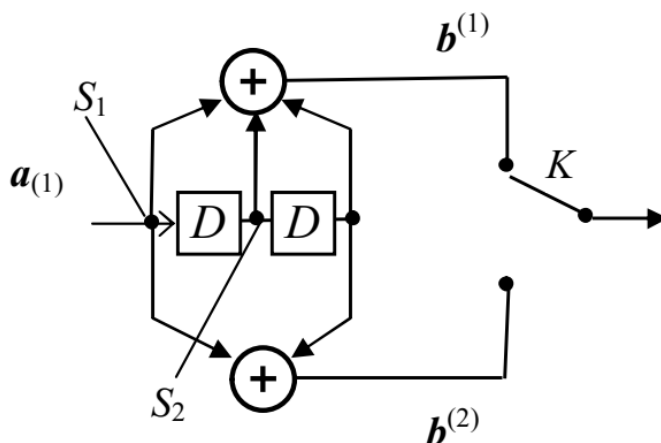


Рис. 7.1: Кодер ЗК

Інформаційні двійкові символи a надходять на вхід регістра з K елементами затримки D . На виходах суматорів за модулем 2 утворюються кодові

символи $\mathbf{b}(1)$ і $\mathbf{b}(2)$. Входи суматорів з'єднані з певними входами (виходами) елементів регістра кодера. Комутатор \mathbf{K} на виході кодера встановлює черговість послідовності кодів символів у канал. За час одного інформаційного символу на виході утвориться два кодів символи.

Швидкість коду $R_{\text{код}} = k/n$, де k – число інформаційних символів, що одночасно надходять на входи кодера, а n – число відповідних їм кодів символів на виходах кодера. Швидкість коду в цьому прикладі дорівнює $R_{\text{код}} = 1/2$. Можливе кодування й з іншими швидкостями.

Згортковий кодер, як автомат з кінцевим числом станів, може бути описаний діаграмою станів. Станом прийнято вважати набір символів на входах елементів затримки регістра. Наприклад, символами (S_1, S_2) позначені стани кодера на рис.

Діаграма станів являє собою спрямований граф, що описує всі можливі переходи кодера з одного стану в інший, а також містить символи виходів кодера, які супроводжують ці переходи.

Приклад діаграми станів кодера показано на рис. 7.2

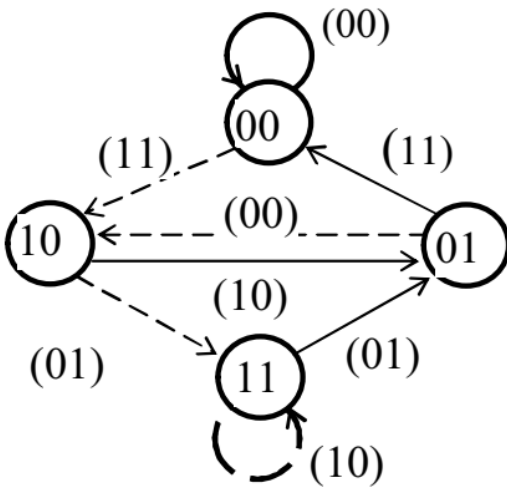


Рис. 7.2: Діаграма станів ЗК

У кружках зазначені чотири можливих стани кодера $(S_1S_2) = 00, 10, 11, 01$, стрілками – можливі переходи. Символи біля стрілок позначають символи на виході кодера $(\mathbf{b}(1)\mathbf{b}(2))$, що відповідають даному переходу. Суцільними лініями відзначені переходи, що мають місце при надходженні на вхід кодера інформаційного символу 0, і пунктирними – при надходженні символу 1. Спочатку кодер перебуває в стані 00, і надходження на його вхід інформаційного символу $\mathbf{a} = \mathbf{0}$ переводить його також у стан 00. При цьому на виході кодера будуть символи $\mathbf{b}(1)\mathbf{b}(2) = \mathbf{00}$. На діагра-

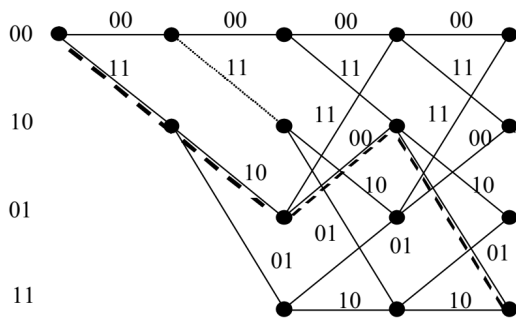


Рис. 7.3: Решітчаста діаграма ЗК (7,5)

мі цей перехід позначається петлею "00 що виходить зі стану 00 і знову повертається в цей стан.

Далі, при надходженні символу $a = 1$ кодер переходить у стан 10, при цьому на виході будуть символи $b(1)b(2) = 11$. Цей перехід позначається пунктирною лінією зі стану 00 у стан 10. Далі, можливе надходження на вхід кодера інформаційних символів 0 або 1. При цьому кодер переходить у стан 01 або 11, а символи на виході будуть 10 або 01, відповідно. Процес побудови діаграми закінчується, коли будуть переглянуті всі можливі переходи з кожного стану в усі інші.

Решітчаста діаграма (решітка) є розгорненням діаграми станів у часі. На решітці стани показані вузлами, а переходи – з'єднуючими їхніми лініями (вітками). Після кожного переходу з одного стану в інший відбувається зсув на один крок вправо.

Приклад решітчастої діаграми показано на рис.

Решітчаста діаграма дає наочне подання всіх дозволених шляхів, які є аналогами дозволених кодових комбінацій блокових кодів. По них може просуватися стан кодера при кодуванні.

Кожної інформаційної послідовності на вході кодера відповідає єдиний шлях по решітці.

Побудову решітчастої діаграми зручно робити з використанням діаграми станів. Вихідним є нульовий стан $S_1S_2 = 00$. Далі, з надходженням чергового інформаційного символу $a = 0$ або 1 можливі переходи в стан **00** або **10**, позначувані вітками (00) і (11) відповідно. Процес слід продовжити, причому через 3 кроки, черговий фрагмент решітки буде повторюватися.

Пунктиром показаний шлях по решітці **11100001...**, що відповідає надходженню навхід кодера інформаційної послідовності **1011.....**

Для опису роботи кодера послідовності символів на вході й виході зручно представляти з використанням оператора затримки D у вигляді не-

скінченних рядів:

$$a_{(i)}(D) = a_{(i)0}D^0 + a_{(i)1}D^1 + a_{(i)2}D^2 + \dots$$

$$b^{(j)}(D) = b_0^{(j)}D^0 + b_2^{(j)}D^2 + b_3^{(j)}D^3 + \dots$$

Тут індекси в дужках позначають:

i – номер входу кодера, $1 \leq i \leq k$;

j – номер виходу кодера, $1 \leq j \leq n$. Індекси без дужок (0, 1, 2, ...) позначають дискретні моменти часу.

Для опису згорткового кодування використовують поняття багаточлена, що породжує. Згортковий код буде повністю задано, якщо відомо схему кодера:

- кількість входів кодера **k**;
- кількість виходів кодера **n**;
- довжина кожного з регістрів K_i ;
- зазначені зв'язки суматорів з комірками регістрів.

Для кодів зі швидкістю $R_{\text{код}} = 1/n$ зв'язки j -го суматора ($1 \leq j \leq n$) з комірками регістра зсуву описуються шляхом подання багаточлена, що породжує:

$$g^{(j)}(D) = g_0^{(j)} + g_1^{(j)}D + g_2^{(j)}D^2 + \dots + g_\nu^{(j)}D^\nu \quad (7.1)$$

причому, $g_k^{(j)} = 1$, якщо зв'язок j -го суматора з k -ою коміркою регістра існує, і $g_k^{(j)} = 0$, якщо такого зв'язку немає.

Процес кодування може бути представлений як множення багаточлена вхідної інформаційної послідовності $a_{(i)}D$ на багаточлен, що породжує, $g_{(i)}^{(j)}(D)$, що описує зв'язки комірок i -го регістра кодера з входами j -го суматора модулем 2:

$$b^{(j)} = a_{(i)}g_{(i)}^{(j)}(D). \quad (7.2)$$

Наприклад, кодер рис.7.1 характеризується багаточленами, що породжують, $g^{(1)}(D) = 1 + D + D^2$ і $g^{(2)}(D) = 1 + D^2$ або, записуючи послідовність коефіцієнтів g^k у вигляді двійкових комбінацій, одержуємо $g^{(1)} = (111)$ і $g^{(2)} = (101)$.

Для довгих кодів часто використовують восьмеричну форму запису. У цьому випадку багаточлени, що породжують, будуть представлені так: $g^{(1)} = (7)$ і $g^{(2)} = (5)$, або $G = (g^{(1)}, g^{(2)}) = (7, 5)$.

7.2 Основні параметри й класифікація ЗК

Швидкість коду визначається як

$$R_{\text{код}} = k/n, \quad (7.3)$$

де k – кількість інформаційних символів, що одночасно надходять на k входів кодера, n – кількість відповідних їм кодових символів на n виходах кодера.

Використовують кілька характеристик для визначення довжини пам'яті при кодуванні. Довжина регістра, що кодує (ДКР) K дорівнює кількості елементів затримки, що містяться в схемі кодера. ДКР часто застосовують для визначення пам'яті при кодуванні зі швидкістю $R_{\text{код}} = 1/n$, коли кодер містить один регістр. Кодер, зображений на рис. 7.1, має ДКР $K = 3$. Якщо кодер містить кілька входів ($k > 1$), то довжини регістрів, підключених до кожного входу, можуть бути різні. У цьому випадку визначають довжину кодового обмеження.

Довжина кодового обмеження (ДКО) по кожному вході визначається старшим степенем відповідних багаточленів, що породжують

$$\nu_i = \max[\text{deg}_{(i)}^{(j)}(D)].$$

Результуюча довжина кодового обмеження кодера визначається сумою:

$$\nu = \sum_{i=1}^k \nu_i. \quad (7.4)$$

Для кодів з одним регістром пам'яті ($k = 1$) величини ДКО й ДКР зв'язані простим співвідношенням:

$$\nu = K \quad (7.5)$$

Для порівняння складності алгоритмів декодування ЗК використовують характеристику складності. Оскільки, розвиток решітчастої діаграми складається в повторенні того самого кроку (див. рис. 7.3), складність діаграми прийнято визначати кількістю гілок на кроці решітчастої діаграми. Число станів решітки визначаються числом змінних $K = \nu$ на входах елементів регістра. При використанні коду з основою m всі можливі комбінації цих змінних утворюють набір станів кодера. Загальне число станів дорівнює $S = m^K$. З кожного стану виходять (i в кожний стан також входять)

m^k гілок. У підсумку складність одного кроку решітки можна визначити кількістю гілок на цьому кроці

$$W = m^{\nu+k} \quad (7.6)$$

Завадостійкість декодування залежить від дистанційних властивостей кодових послідовностей на виході кодера. При цьому для двійкових кодів найчастіше віддаль між послідовностями оцінюють у метриці Хеммінга.

Вільна віддаль згорткового коду d_f – мінімальна віддаль між двома довільними напівнескінченими послідовностями на виході кодера, що відрізняються в першій гілці.

Для коротких кодів вільну віддаль можна визначити по діаграмі станів. Якщо діаграма двійкового коду задана, то вільна віддаль коду дорівнює мінімальній вазі Хеммінга шляху по діаграмі зі стану 00 у цей же стан (крім петлі в цьому стані). На діаграмі рис. 7.2 видно, що вільна віддаль $d_f = 5$.

По величині вільної віддалі ЗК судять про коректувальні властивості згорткових кодів. Зокрема, якщо два шляхи на виході кодера ЗК, що виходять із одного стану решітчастої діаграми, розрізняються в метриці Хеммінга на величину d_f , то при декодуванні по мінімуму віддалі за аналогією з випадком декодування блокових кодів, кратність помилок, що виправляються, визначається виразом

$$q_{\text{вип}} \leq \frac{d_f - 1}{2}, \quad (d_f - \text{непарне}) \quad (7.7)$$

Вільна віддаль використовується для оцінки завадостійкості декодування згорткових кодів із застосуванням алгоритмів максимальної правдоподібності або близьких до них (алгоритм Вітербі й ін.).

Деякі ЗК мають властивість катастрофічності. Катастрофічним ЗК називається такий код, у якого вхідна інформаційна послідовність нескінченної ваги дає на виході кодера послідовність кінцевої ваги. При використанні катастрофічного коду кінцеве число помилок у каналі викликає нескінченне число помилок при декодуванні. Із цієї причини використовувати на практиці катастрофічні коди не рекомендується, а таблиці ЗК відомостей про катастрофічні коди не містять, оскільки при пошуку багаточленів, що породжують, кращих кодів катастрофічні коди відкидаються.

У систематичному коді на \mathbf{k} (з \mathbf{n} можливих) виходах кодера присутні інформаційні послідовності переданих символів, а на інші $(\mathbf{n} - \mathbf{k})$ виходах

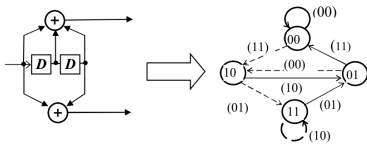


Рис. 7.4: Кодер ЗК(7,5) і діаграма станів

– послідовності додаткових символів, сформованих як лінійні комбінації інформаційних символів. При швидкості $R_{\text{код}} = 1/2$ багаточлени, що породжують, систематичний код мають вигляд

$$g^{(1)} = 1 \text{ і } g^{(2)}(D) = g_0^{(2)} + g_1^{(2)}D + g_2^{(2)}D^2 + \dots + g_\nu^{(2)}D^\nu.$$

Систематичні коди дозволяють одержати на прийомній стороні оцінку інформаційних символів, не роблячи декодування або яку-небудь іншу обробку прийнятих символів. Несистематичним кодам така властивість не властива.

Як і у випадку блокових кодів, застосування згорткового кодування зі швидкістю $R_{\text{код}} = k/n$ приводить до розширення спектра сигналу в каналі. При цьому коефіцієнт розширення смуги визначається виразом:

$$K_F = n/k. \quad (7.8)$$

При малих швидкостях кодів значне розширення смуги стає неприйнятним, тому намагаються застосовувати коди з високою швидкістю. Практично, вибір параметрів згорткових кодів роблять на основі компромісу, виходячи з необхідного енергетичного виграшу кодування й припустимого коефіцієнта розширення спектра сигналу в каналі.

Вправа 7.1.

Аналіз взаємозв'язків параметрів ЗК.

Використовуючи послідовну модифікацію структури вихідного кодера ЗК (7, 5) (рис. 7.1) і відповідних йому діаграм станів і решітки (7.2 і 7.3), установимо взаємозв'язок параметрів кодера k , n , $R_{\text{код}}$, S , і багаточленів, що породжують, із вільною віддалю коду d_f .

Розглянемо кілька варіантів ЗК:

1. Вихідний згортковий код (7, 5) – див. рис. 7.4.

Параметри коду: $k = 1$; $n = 2$; $K = 2$; $R_{\text{код}} = 1/2$; для двійкового коду з $m = 2 \rightarrow S = m = 2^2 = 4$; вільна віддаль $d_f = 5$; код несистематичний.

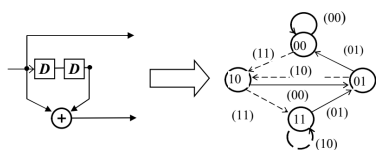


Рис. 7.5: Кодер ЗК(7,5) і діаграма станів

2. Формування систематичного коду (1, 5) – див. рис. 7.5. Модифікуємо перший багаточлен вихідного коду, залишивши один зв'язок. Діаграма станів частково зміниться. Число станів залишиться колишнім, оскільки склад регістра кодера не змінився. Змінюються ненульові гілки: відповідно до зміни першого багаточлена, що породжує, на місці першого символу гілки варто проставити перший символ стану, у яке спрямована ця гілка. Швидкість коду також не змінилася.

Параметри коду:

$$k = 1;$$

$$n = 2;$$

$$K = 2;$$

$$R_{\text{код}} = 1/2;$$

для двійкового коду $m = 2 \rightarrow S = 4$;

вільна віддаль зменшилася: $d_f = 3$;

код систематичний.

Розділ 8

Алгоритми декодування згорткових кодів

8.1 Класифікація алгоритмів декодування

При оптимальній обробці з метою винесення рішення прийняту з каналу послідовність символів необхідно зіставити з усіма можливими послідовностями, що передаються. Оскільки число можливих послідовностей довжини N двійкового коду дорівнює 2^N , то при більших довжинах послідовностей декодер стає неприпустимо складним, а оптимальне декодування практично важко реалізованим. Однак, саме при великих N можливе значне підвищення надійності передачі, тому що дія шуму усереднюється на довгій послідовності. Тому важливою є проблема зниження складності алгоритмів декодування ЗК.

Відомі дві групи методів декодування згорткових кодів:

1. Алгебраїчні методи декодування засновані на використанні алгебраїчних властивостей кодових послідовностей. У ряді випадків ці методи приводять до простих реалізацій кодека. Такі алгоритми є неоптимальними, тому що використовувані алгебраїчні процедури декодування призначені для виправлення конкретних (і не всіх) конфігурацій помилок у каналі. Алгебраїчні методи ототожнюють із поелементним прийманням послідовностей, що для кодів з надмірністю, як відомо, дає гірші результати, ніж «**приймання у цілому**». Найбільш простим з алгебраїчних алгоритмів є **алгоритм порогового декодування** згорткових кодів. Цей алгоритм далекий від оптимального й тому рідко використовується, а використовується, у першу чергу, у системах з високою швидкістю передачі інформації.

2. Імовірнісні методи декодування значно ближчі до оптимального «приймання в цілому», тому що в цьому випадку декодер оперує з величинами, пропорційними апостеріорним імовірностям, оцінює й порівнює ймовірності різних гіпотез і на цій основі виносить рішення про передані символи.

Алгебраїчні алгоритми оперують із кінцевим алфавітом вхідних даних, для одержання яких на виході неперервного каналу необхідно виконати квантування прийнятого сигналу із шумом. Процеси обробки сигналів у пристрої, що вирішує, на виході демодулятора протилежних сигналів показані на рис. 8.1, де представлені:

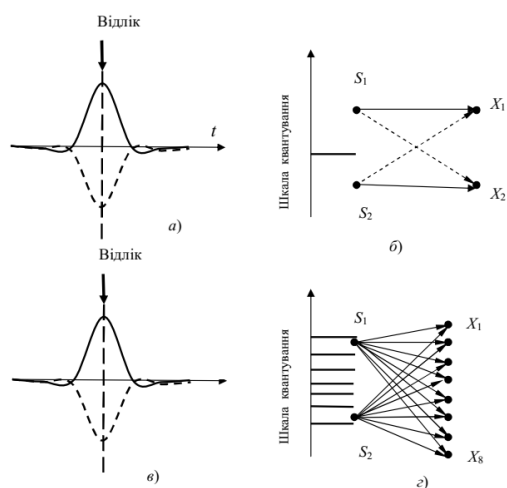


Рис. 8.1: Процеси обробки сигналів

а, в – форми протилежних сигналів у момент відліку на вході пристрою демодулятора, що вирішує.

б – шкала квантування й граф переходів при жорсткому рішенні.

г – шкала квантування й граф переходів при м'якому рішенні.

У найпростішому випадку роблять **квантування** кожного канального символу у відліковий момент часу **на два рівні** (іменоване в літературі як «жорстке рішення» на виході демодулятора). При **жорсткому** рішенні число рівнів квантування $L = 2$. При цьому жорстке рішення представлене одним двійковим символом. Це показано на рис. 8.1, б.

При «**м'якому рішенні**» число рівнів квантування $L > 2$ (рис. 8.1, г). При м'якому рішенні вихід квантувача більш точно описує величину відліку сигналу з завадою, що підвищує завадостійкість декодування. Відомі два основні ймовірнісні алгоритми декодування згорткових кодів, а також їхні різні модифікації.

1. Алгоритм послідовного декодування забезпечує доволіно малу ймовірність помилки при ненульовій швидкості передачі повідомлень по каналу. При послідовному декодуванні виконується пошук шляху на кодових решітках, що відповідають переданій інформаційній послідовності. Послідовне декодування використовується для декодування довгих згорткових кодів.

Іншим різновидом імовірнісних алгоритмів є алгоритм, заснований на принципі динамічного програмування, і відомий як алгоритм Вітербі. Принцип динамічного програмування був сформульований в 1940 р. Р.Беллманом. З тих пір він знайшов широке використання у теорії керування й теорії кіл. В 1970 р. динамічне програмування, у формі алгоритму декодування ЗК (алгоритму Вітербі), було застосовано А. Вітербі до проблем телекомунікації.

2. Алгоритм Вітербі знаходить широке застосування й реалізує **пошук максимально правдоподібного шляху** на кодових решітках з відкиданням частини найменш правдоподібних варіантів шляхів на кожному кроці декодування. Алгоритм Вітербі характеризується сталістю обчислювальної роботи, однак складність декодера Вітербі росте, як при всіх переборних алгоритмах, за експонентним законом від довжини кодового обмеження згорткового коду. Тому алгоритм Вітербі використовується для декодування коротких згорткових кодів.

8.2 Алгоритм Вітербі для декодування згорткових кодів

Розглянемо алгоритм Вітербі на прикладі коду зі швидкістю $R_{\text{код}} = 1/n$.

Нехай, починаючи з моменту часу $\mathbf{t} = \mathbf{0}$, на вхід кодера подається інформаційна послідовність довжиною в L символів $a = (a_0, a_1, \dots, a_{L-1})$... На виході кодера буде послідовність символів $b_L = (b_0, b_1, \dots, b_{L-1})$. Стан кодера в момент \mathbf{t} визначають як набір з ν інформаційних символів $w_t = (a_0, a_1, \dots, a_{t-L+1})$. Решітчаста діаграма коду однозначно зв'язує інформаційну послідовність \mathbf{a} , послідовність станів кодера \mathbf{w} і послідовність символів на його виході \mathbf{b} .

Кожній гілці b_t у каналі відповідає сигнал, що може бути представлений набором координат $S_t = (S_t^1, S_t^2, \dots, S_t^N)$, де N – розмірність простору сигналів.

У каналі діє адитивна завада. Тоді послідовність прийнятого сигналу на

входідекодера буде дорівнювати:

$$X_L = S_L + n,$$

де $S_L = (S_0, S_1, \dots, S_{L-1})$; $n = (n_0, n_1, \dots, n_{L-1})$; $n_t = (n_t^{(1)}, n_t^{(2)}, \dots, n_t^{(N)})$ – N -вимірний вектор завади.

Бібліографія

- [1] Білинський Й.Й., Огородник К.В., Юркиш М.Й. Електронні системи. Навчальний посібник. Вінниця: ВНТУ, 2011. 208 с.
- [2] Первунінський С.М., Чепинога А.В. Потенціальна завадостійкість цифрового автокореляційного демодулятора Ланге–Мюллера. Інфокомунікаційні та комп'ютерні технології. № 1 (01), 2021. с. 90-99.
- [3] Іщенко М.О., Хорольський І.В. Завадостійкість системи бездротового зв'язку в каналах з простово-часовим кодуванням. Наукові парці. Комп'ютерні технології. Випуск 161. Том 173. 2011. с. 111-117
- [4] С'янов О.М., Марченко С.В. Конспект лекцій з дисципліни "Спеціальні розділи з теорії передавання інформації". Кам'янське, ДДТУ, 2018. 77с.
- [5] Лаптев О.А., Собчук В.В., Савченко В.А. Метод підвищення завадостійкості системи виявлення, розпізнавання і локалізації цифрових сигналів в інформаційних системах. Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка, (66), с. 90–104. <https://doi.org/10.17721/2519-481X/2020/66-09>