

ODESA STATE ACADEMY OF
TECHNICAL REGULATION AND QUALITY



ВІД ЯКОСТІ ОСВІТИ ДО ЯКОСТІ ЖИТТЯ!



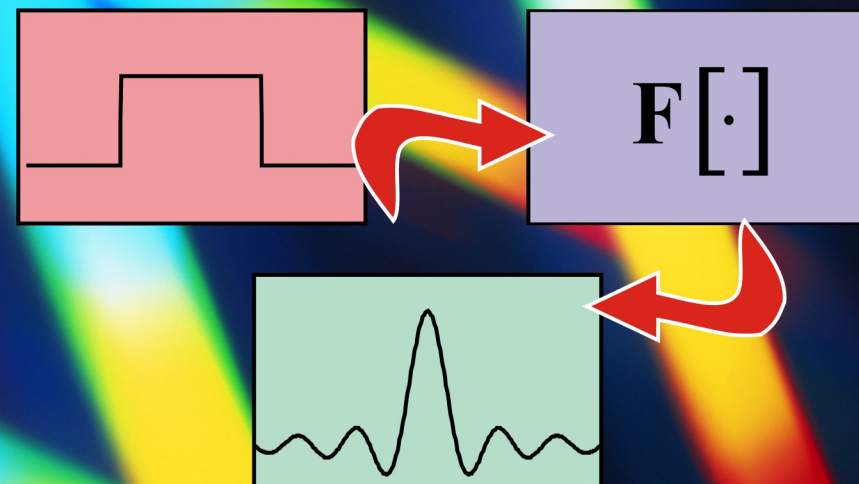
65020, м. Одеса, вул. Ковальська, 15
тел.: (048) 726-68-92 факс. (048) 726-76-95

e-mail: odatry@gmail.com
www.kachestvo.od.ua

МЕТОДИ ТА ЗАСОБИ ОБРОБКИ СИГНАЛІВ
2014

Братченко Г. Д., Перелігін Б. В., Казакова Н. Ф.,
Банзак О. В., Григор'єв Д. В.

МЕТОДИ ТА ЗАСОБИ ОБРОБКИ СИГНАЛІВ



**Міністерство освіти і науки України
Одеська державна академія технічного регулювання та
якості**

**Братченко Г. Д., Перелигін Б. В., Банзак О. В.,
Казакова Н. Ф., Григор'єв Д. В.**

МЕТОДИ ТА ЗАСОБИ ОБРОБКИ СИГНАЛІВ

Навчальний посібник

За загальною редакцією доктора технічних наук
Г. Д. Братченка

Одеса
2014

ББК 32.81
Б 87
УДК 621.391

Братченко Г. Д., Перелигін Б. В., Банзак О. В., Казакова Н. Ф., Григор'єв Д. В. Методи та засоби обробки сигналів. Навчальний посібник. – Одеса: Типографія-видавництво «Плутон», 2014. – 452 с.

ISBN 978-966-342-972-4

Навчальний посібник призначений для підготовки студентів галузі знань «Метрологія, вимірювальна техніка та інформаційно-вимірювальні технології». Предметом вивчення навчальної дисципліни є методи та засоби обробки сигналів в аналогових і цифрових приладах і системах. В навчальному посібнику викладені теоретичні основи обробки сигналів в лінійних системах та окремі поняття щодо нелінійних перетворень сигналів, які є важливими для фахівців в галузі метрології, вимірювальної техніки та інформаційно-вимірювальних технологій для розуміння складних процесів, що відбуваються при обробці сигналів в засобах вимірювальної техніки, медичних приладах і системах.

*Рекомендовано Вченою радою Одеської державної академії
технічного регулювання та якості (протокол № 5 від 27.11.14)*

Рецензенти:

В.А. Мокрицький – д-р техн. наук, професор
(Одеський національний політехнічний університет, м. Одеса)
Б.О. Дем'янчук – д-р техн. наук, доцент
(Військова академія, м. Одеса)

ISBN 978-966-342-972-4

© Братченко Г. Д., Перелигін Б. В., Банзак О. В.,
Казакова Н. Ф., Григор'єв Д. В., 2014

ПЕРЕЛІК ПРИЙНЯТИХ СКОРОЧЕНЬ

АІМ	– амплітудно-імпульсна модуляція
АКФ	– автокореляційна функція сигналу
АМ	– амплітудна модуляція
АМ _н	– амплітудна маніпуляція
АП	– аналоговий перетворювач
АР	– авторегресійний
АЦП	– аналого-цифровий перетворювач
АЧХ	– амплітудно-частотна характеристика системи
БМ	– балансний модулятора
ВКФ	– взаємна кореляційна функція
ВС	– вимірювальний сигнал
ВЧ	– високочастотний
ДПФ	– дискретне перетворення Фур'є
ДР	– диференційне рівняння
ЕОМ	– електронно-обчислювальна машина
ЗВ	– засіб вимірювання
ЗВТ	– засіб вимірювальної техніки
ІВС	– інформаційна вимірювальна система
КАМ	– квадратурна модуляція
КАМ _н	– квадратурна маніпуляція
КІХ	– кінцева імпульсна характеристика
КМ	– кутова модуляція
ЛЧМ	– лінійно-частотно-модульований
МНК	– метод найменших квадратів
МП	– мікропроцесор
НЧ	– низькочастотний
ОДВП	– обернене дискретне вейвлет-перетворення
ОНВП	– обернене неперервне вейвлет-перетворення
ПЦОС	– процесор цифрової обробки сигналів
ПЦС	– процесор цифрових сигналів
ПДВП	– пряме дискретне вейвлет-перетворення
ПНВП	– пряме неперервне вейвлет-перетворення
РЛС	– радіолокаційна станція
РНК	– рекурсивний алгоритм найменших квадратів

СКВ	– середнє квадратичне відхилення
СКП	– середньоквадратична похибка
СМП	– сигнальний мікропроцесор
СЧАН	– спектрально-часовий аналіз
УФ	– узгоджений фільтр
ФВ	– фізична величина
ФІМ	– фазово-імпульсна модуляція
ФМ	– фазова модуляція
ФМн	– фазова маніпуляція
ФНЧ	– фільтр низьких частот
ФЧХ	– фазочастотна характеристика системи
ЦАП	– цифро-аналоговий перетворювач
ЦОС	– цифрова обробка сигналів
ЦПОС	– цифровий процесор обробки сигналів
ЦСП	– цифровий сигнальний процесор
ЧІМ	– часово-імпульсна модуляція
ЧМн	– частотна маніпуляція
ШІМ	– широтно-імпульсна модуляція
ШПФ	– швидке перетворення Фур'є
МАС	– операція «множення з накопиченням»

ВСТУП.....	8
1. ВИМІРЮВАЛЬНІ СИГНАЛИ ТА ЇХ ХАРАКТЕРИСТИКА.....	11
1.1. Поняття сигналу. Класифікація вимірювальних сигналів.....	11
1.2. Математичні моделі сигналів. Спектральне подання сигналів.....	25
1.3. Тестові сигнали. Дельта функція. Функція одиночного стрибка.....	30
1.4. Енергія і потужність сигналу.....	36
1.5. Перетворення типу сигналів: аналогові і цифрові сигнали.....	37
2. СПЕКТРАЛЬНИЙ АНАЛІЗ ПЕРІОДИЧНИХ ТА НЕПЕРІОДИЧНИХ СИГНАЛІВ.....	58
2.1. Розкладання в ряд Фур'є періодичних сигналів.....	58
2.2. Приклади розкладання сигналів в ряд Фур'є.....	68
2.3. Інтегральне перетворення Фур'є.....	76
3. КОРЕЛЯЦІЙНИЙ АНАЛІЗ СИГНАЛІВ.....	97
3.1. Автокореляційна функція.....	97
3.2. Взаємна кореляційна функція.....	100
3.3. Зв'язок між кореляційними функціями і спектрами сигналів.....	103
4. АНАЛІЗ ВИПАДКОВИХ СИГНАЛІВ.....	107
4.1. Характеристики випадкових сигналів.....	107
4.2. Аналіз випадкових сигналів. Кореляційні функції випадкових сигналів.....	117
4.3. Стаціонарні та ергодичні випадкові процеси.....	122
4.4. Нестационарні випадкові процеси.....	128
4.5. Спектральні характеристики випадкових процесів.....	130
4.6. Приклади стаціонарних випадкових процесів.....	136
5. ДИСКРЕТНІ СИГНАЛИ ТА МЕТОДИ ЇХ АНАЛІЗУ... ..	143
5.1. Моделі дискретних сигналів. Дискретизація періодичних сигналів. Дискретне перетворення Фур'є.....	143
5.2. Швидке перетворення Фур'є.....	151
5.3. Пряме та обернене z-перетворення, його властивості.	159

6. ЛІНІЙНІ СИСТЕМИ.....	172
6.1. Аналогові лінійні системи.....	169
6.2. Дискретні лінійні системи.....	186
6.3. Проектування цифрових лінійних систем.....	200
7. МОДУЛЬОВАНІ СИГНАЛИ І НЕЛІНІЙНІ СИСТЕМИ	222
7.1. Нелінійні системи. Безінерційні нелінійні перетворювачі.....	222
7.1.1. Властивості нелінійних систем.....	222
7.1.2. Вплив стаціонарних випадкових сигналів на безінерційні нелінійні системи	237
7.2. Модульовані сигнали	251
7.2.1. Амплітудно-модульовані сигнали	253
7.2.2. Сигнали з кутовою модуляцією.....	266
7.2.3. Модульовані сигнали з кількома змінними параметрами.....	277
7.3. Демодуляція сигналів.....	285
8. ВЕЙВЛЕТ-АНАЛІЗ СИГНАЛІВ.....	308
8.1. Поняття про спектрально-часовий аналіз.....	308
8.2. Загальні відомості про вейвлет-перетворення.....	311
8.3. Властивості вейвлет-перетворення.....	330
8.4. Класифікація та приклади вейвлетів.....	337
8.5. Можливості вейвлет-аналізу.....	342
8.6. Вейвлет-обробка сигналів.....	347
8.7. Приклади вейвлет-обробки.....	362
9. ОПТИМАЛЬНІ І АДАПТИВНІ ЛІНІЙНІ СИСТЕМИ... ..	372
9.1. Оптимальна лінійна обробка сигналів	372
9.2. Оптимальна лінійна обробка цифрових сигналів	385
9.3. Адаптивна обробка сигналів.....	392
9.3.1. Оптимальне рекурсивне калманівське оцінювання... ..	397
9.3.2. Векторний фільтр Калмана.....	402
9.4. Адаптивні алгоритми для фільтрів з кінцевою імпульсною характеристикою.....	407
9.4.1. Рекурсивний алгоритм найменших квадратів.....	408
9.4.2. Алгоритм РНК з експоненціальним зважуванням....	412
9.4.3. Обчислювальна складність.....	415

9.4.4. Стохастична інтерпретація.....	416
9.4.5. Адаптивний алгоритм методу найменших квадратів	422
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ	435
Додаток А.....	439
Додаток Б.....	442

ВСТУП

Обробка сигналів має велике фундаментальне і прикладне значення в сучасній метрології й інформаційно-вимірювальній техніці. Її методи широко застосовуються для розробки і дослідження засобів вимірювальної техніки різного призначення, а засоби – для їх аналогової або апаратно-програмної реалізації.

Навчальна дисципліна «Методи та засоби обробки сигналів» охоплює широке коло теоретичних питань, основи яких викладаються на лекційних заняттях і більш детально досліджуються на практичних і лабораторних заняттях та в процесі самостійної роботи студентів. На лекційних заняттях через обмеженість часу розглядаються базові, ключові питання дисципліни, розуміння яких дозволить самостійно вивчити повний курс дисципліни, представлений в даному посібнику.

Основними завданнями дисципліни «Методи та засоби обробки сигналів» є вивчення теоретичних основ обробки сигналів в лінійних системах та окремих понять щодо нелінійних перетворень сигналів, серед яких: класифікація сигналів та їх математичні моделі; спектральний аналіз періодичних та неперіодичних детермінованих сигналів; кореляційний та спектральний аналіз детермінованих та випадкових сигналів; дискретні і цифрові сигнали та методи їх аналізу; принципи цифрової обробки сигналів; випадкові сигнали та їх статистичні характеристики; лінійна фільтрація детермінованих та випадкових сигналів; основи модуляції і нелінійного перетворення сигналів; спектральний аналіз сигналів з амплітудною та з кутовою модуляцією; вейвлет-аналіз сигналів; основи оптимальної і адаптивної обробки сигналів в лінійних системах.

Метою викладання навчальної дисципліни є навчити студентів правильно орієнтуватись в складних процесах, що відбуваються при обробці сигналів у приладах і системах.

Успішному освоєнню дисципліни сприятиме знання основ математичного аналізу і теорії ймовірностей, положень базової навчальної дисципліни «Радіотехнічні ланцюги і сигнали».

Матеріал курсу розділений на 9 глав. Зміст їх викладається на лекційних заняттях. Можливість самостійного закріплення матеріалу за допомогою навчального посібника дозволяє студентам під час лекцій більшою мірою зосередитись на розумінні досліджуваних питань. Матеріали глав 1-6 та додатки підготовлені д.т.н, с.н.с. Братченко Г.Д. та Григор'євим Д.В., главу 7 підготувала к.т.н., доцент Банзак О.В., главу 8 – к.т.н., доцент Перелигін Б.В. та главу 9 підготували Братченко Г.Д., к.т.н., доцент Казакова Н.Ф. та Григор'єв Д.В. В главі 9 також використані ілюстративні матеріали надані д.т.н., професором Скачковим В.В.

За результатами вивчення дисципліни студенти повинні засвоїти: основні поняття теорії сигналів і систем; типи сигналів та їх класифікацію; основи спектрального і кореляційного аналізу детермінованих і випадкових сигналів; теоретичні основи обробки аналогових і цифрових сигналів в лінійних системах; основні види модульованих сигналів, способи модуляції, які використовуються при передачі виміральної інформації; основні поняття про нелінійні системи та проходження випадкових сигналів через безінерційні нелінійні кола, основи вейвлет-аналізу сигналів та оптимальної і адаптивної обробки сигналів в лінійних системах. Це дозволить їм вміти: класифікувати сигнали; складати та застосовувати їх математичні моделі у часовій і частотній областях; розкласти в ряд Фур'є періодичні сигнали; обчислювати прямі і обернені перетворення Фур'є детермінованих неперіодичних аналогових і цифрових сигналів; обчислювати прямі і обернені z-перетворення цифрових сигналів; розраховувати основні кореляційні та числові характеристики детермінованих і випадкових сигналів; будувати оптимальні лінійні системи.

Навчальний посібник рекомендується бакалаврам, магістрам і спеціалістам галузі знань «Метрологія, вимірювальна техніка та інформаційно-вимірювальні технології», а також аспірантам спеціальностей 05.01.02 стандартизація сертифікація та метрологічне забезпечення та 05.11.01 – прилади і методи вимірювання механічних величин.

Глава 1

ВИМІРЮВАЛЬНІ СИГНАЛИ ТА ЇХ ХАРАКТЕРИСТИКА

1.1. Поняття сигналу. Класифікація вимірювальних сигналів

Термін «сигнал» (від *лат. signum* – знак) застосовується в багатьох галузях науки і техніки, однак його значення значно ширше, ніж дослівний переклад з латині. Це, передусім, фізичний процес або явище, яке повідомляє про певну подію, стан об'єкта, передає команду керування, оповіщення, а в переносному значенні – попередження, повідомлення про щось небезпечне. В системах зв'язку сигнал переносить повідомлення певного змісту (воно складається з різного роду знаків або символів). Його, зазвичай, потрібно передати на деяку відстань користувачу. У випадку передачі телеметричних даних повідомлення є функцією у часі, наприклад, вимірних значень температури, тиску повітря тощо [1]. Тобто в цьому випадку в сигналі передається повідомлення про результати вимірювань певних параметрів оточуючого середовища. Саме такі повідомлення є найбільш характерними для метрології, де прийняті такі визначення [2]:

сигнал – це фізичний процес, властивості якого визначаються взаємодією між матеріальним об'єктом та засобом його дослідження;

параметр сигналу – одна із властивостей сигналу, яка є фізичною величиною (ФВ);

інформативний параметр сигналу – це параметр сигналу, який функціонально пов'язаний з досліджуваною або вимірюваною величиною чи той, що має задане значення;

вимірювальний сигнал (ВС) – сигнал, один чи кілька параметрів якого є інформативними;

вхідний (вимірювальний) сигнал – ВС сигнал, що виникає під час взаємодії об'єкта вимірювання та засобу вимірювальної техніки (ЗВТ);

зразковий сигнал – ВС один чи кілька параметрів якого мають задане значення.

Засобами дослідження об'єктів можуть бути засоби вимірювань (тобто ЗВТ, які реалізують процедуру вимірювання), до яких відносять [2]: кодові ЗВ (засіб вимірювання); вимірювальні прилади; вимірювальні канали та вимірювальні системи.

На виході ЗВ отримується сигнал вимірювальної інформації. При цьому вхідний сигнал в ЗВ певним чином обробляється і перетворюється у вихідний сигнал у зручній для спостерігача формі (рис. 1.1). Для цього за необхідності ВС може перетворюватись до зручного для його передавання по каналах зв'язку вигляду, кодуватись та моделювати деякий несучий сигнал, який передається каналом передавання (дротовим або просторовим). На приймальному боці (боці користувача) сигнал демодулюється, декодується і перетворюється до зручного для спостерігача вигляду.

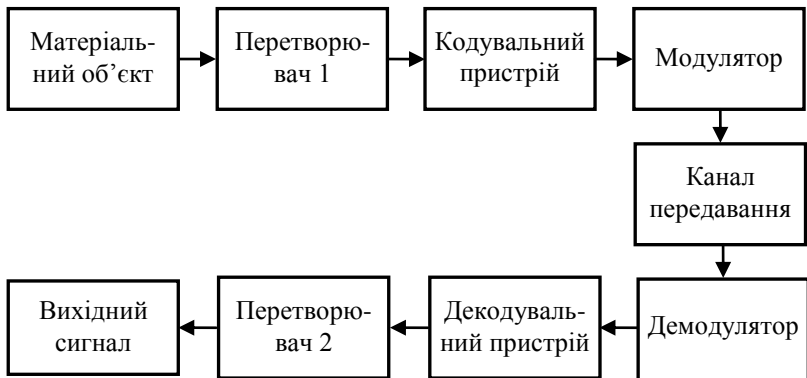


Рисунок 1.1 – Узагальнена схема обробки сигналу в інформаційній вимірювальній системі

В процесі вивчення дисципліни «Методи та засоби обробки сигналів» не розглядається будь-який зв'язок між фізичним виглядом сигналу та смисловим змістом тієї інформації, яка в ньому може бути закладена, та її цінності. Це виходить за її рамки. Подібні задачі вирішуються, наприклад, в межах теорії інформації, теорії розпізнавання образів.

У загальному розумінні сигнал є залежністю однієї величини від іншої, тобто з математичної точки зору є *функцією*. Зазвичай розглядаються часові та частотні залежності. Однак, можливі, наприклад, і просторові залежності (широко застосовуються при геометричних вимірюваннях при оптичній обробці сигналів, в радіолокації, в гідролокації). Фізична природа сигналу також може бути різною. Частіш за все це напруга, рідше – струм, але можливі й інші фізичні величини.

В загальному випадку сигнал може бути багатомірним, тобто представляти собою упорядковану сукупність одновимірних сигналів, деяким чином пов'язаних між собою. Наприклад, вони можуть одночасно залежати від кількох часових і просторових параметрів, які потребують вимірювання. Такі сигнали отримують в різних діапазонах електромагнітних і акустичних хвиль.

Вимірювальні сигнали надзвичайно різноманітні. Їхню класифікацію за різними ознаками наведено на рис. 1.2 [3-5].

За характером вимірювань інформативного і часового параметрів вимірювальні сигнали поділяють на *аналогові, дискретні та цифрові*.

Аналогові сигнали – описуються неперервною або кусково-неперервною функцією $u_a(t)$, причому ця функція та її аргумент t можуть набувати будь-яких значень на заданих інтервалах $u \in (u_{min}, u_{max})$ і $t \in (t_{min}, t_{max})$. Термін «*аналоговий сигнал*» підкреслює, що такий сигнал є «аналогічним», тобто повністю подібним фізичному процесу, який його породжує. Одномірний аналоговий сигнал наочно представляється графіком (осцилограмою), що може бути неперервним або з

точками розриву (рис. 1.3, *a* – сигнал неперервний у часі та за станами) [3].

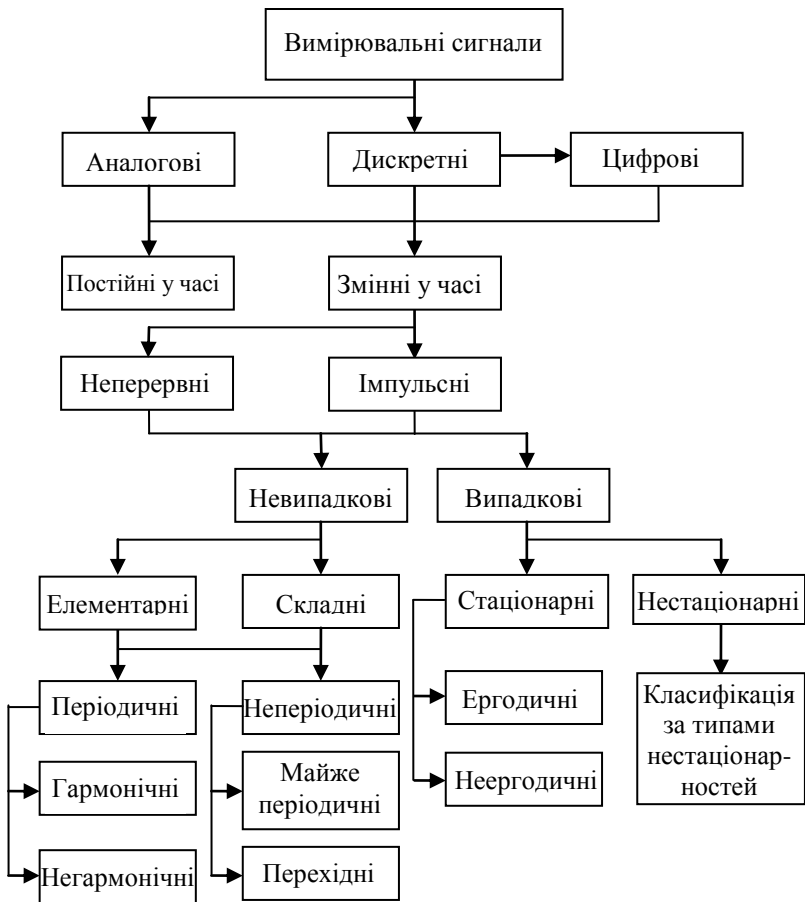


Рисунок 1.2 – Класифікація вимірювальних сигналів

Дискретні сигнали – змінюються дискретно у часі або за рівнем. У першому випадку він може набувати у дискретні моменти часу nT , де число $T = \text{const}$ – інтервал (період) дискретизації; $n = 0, 1, 2, \dots$ – ціле число, будь-яких значень $u_n \in (u_{\min}, u_{\max})$, які називають *вибірками* або *відліками* (рис. 1.3,

б). У другому випадку значення сигналу $u_a(t)$ існують у будь-який момент $t \in (t_{min}, t_{max})$, але вони можуть набувати обмежений ряд станів $h_i = iq$, кратних кванту q , $i = 0, 1, 2, \dots, I$ (рис. 1.3, в - сигнал неперервний у часі та квантований за рівнем або за розміром).

Цифрові сигнали – квантовані за рівнем і дискретні за часом сигнали $Y(nT)$, які описуються квантованими решітчастими функціями (квантованими послідовностями), які набувають у дискретні моменти часу nT лише скінченний ряд дискретних значень – рівнів квантування h_1, h_2, \dots, h_I , (рис. 1.3, г – сигнал дискретизований у часі та квантований за станом). Характерним для таких сигналів є представлення їх відліків за рівнем у формі чисел, зазвичай двійкової розрядності.

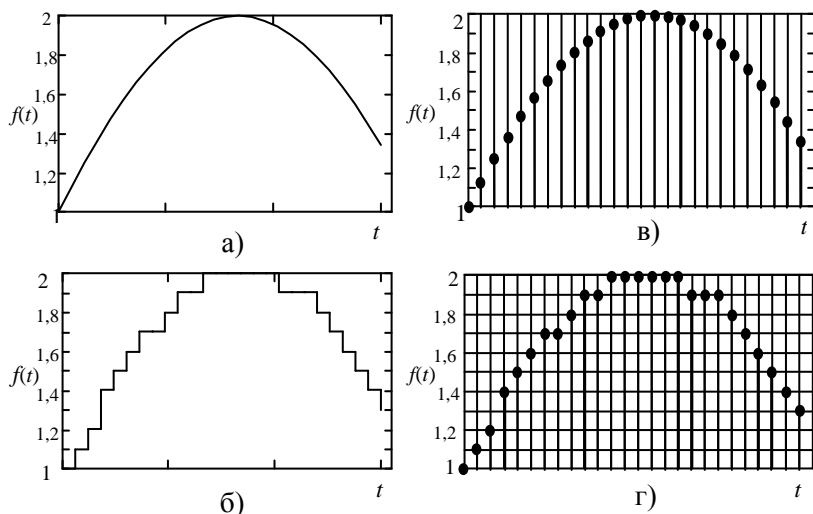


Рисунок 1.3 – Вимірювальні сигнали: а) аналоговий; б) дискретний за часом; в) квантований за рівнем; г) цифровий

За характером зміни у часі сигнали поділяють на *постійні*, з незмінними у часі значеннями, і *змінні*, значення яких змінюються з часом. Постійні сигнали є найпростішим видом ВС. Змінні сигнали можуть бути неперервними за часом та імпульсними.

Неперервним називають сигнал, який існує на незчисленній множині точок часової осі і триває нескінченно довго ($-\infty < t < \infty$), параметри такого сигналу змінюються неперервно; *імпульсні сигнали* (застосовується також термін *фінитні сигнали* – сигнали обмеженої тривалості) – це сигнали кінцевої енергії, істотно відмінні від нуля впродовж обмеженого інтервалу часу сумірного з часом завершення перехідного процесу в системі, для впливу на яку цей сигнал призначений. Розрізняють також *відеоімпульси* і *радіоімпульси*. Якщо $u_B(t)$ відеоімпульс (ВІ), то йому відповідатиме радіоімпульс (РІ) $u_P(t) = u_B(t) \cos(\omega_0 t + \phi_0)$ (частота ω_0 і початкова фаза ϕ_0 є довільними). При цьому $u_B(t)$ – обвідна РІ, $\cos(\omega_0 t + \phi_0)$ – його заповнення (в загальному випадку частота і початкова фаза також можуть бути змінними у часі).

За ступенем наявності *апріорної інформації* змінні ВС поділяють на *невипадкові* (детерміновані, квазидетерміновані) та *випадкові*.

Детермінований сигнал - це сигнал, закон зміни якого відомий, а модель не містить невідомих параметрів. Миттєві значення такого сигналу відомі у будь-який момент часу. Детермінованими (з відомим ступенем точності) є сигнали на виході мір.

Приклад: вихідний сигнал генератора низькочастотного синусоїдального сигналу характеризується значеннями амплітуди і частоти, які встановлені на його пристроях керування (похибки встановлення цих параметрів визначаються метрологічними характеристиками генератора).

Квазидетерміновані сигнали – це сигнали з частково відомим характером зміни за часом, тобто з одним або кількома невідомими параметрами. Вони найцікавіші з точки зору метрології. Переважна більшість ВС є квазидетермінованими.

Детерміновані та квазидетерміновані сигнали поділяють на елементарні (прості), що описуються простими математичними формулами, і складні. До *елементарних* належать постійні та гармонічні сигнали, а також сигнали, що описуються одиничною і δ -функцією. До *складних* сигналів належать імпульсні і модульовані сигнали. В радіотехніці також широко застосовується ділення таких сигналів на складні і прості. Для складних імпульсних модульованих сигналів (шумоподібних, широкосмугових) база $B = \Delta f \tau_i \gg 1$, де Δf ширина смуги частот, де зосереджена основна енергія сигналу, τ_i – тривалість імпульсу. Для простих сигналів виконується умова $B \approx 1$.

Сигнали можуть бути періодичними і неперіодичними.

Періодичними називають сигнали, миттєві значення яких повторюються через постійний інтервал часу $u(t) = u(t + nT)$, n – довільне ціле число. Період T сигналу – це параметр, який дорівнює найменшому такому інтервалу часу; частота періодичного сигналу – є оберненою періоду: $f = 1/T$.

Приклад: Положення проекції точки P , яка рівномірно обертається по колу з кутовою частотою ω , на вісь ординат (рис. 1.4) [6].

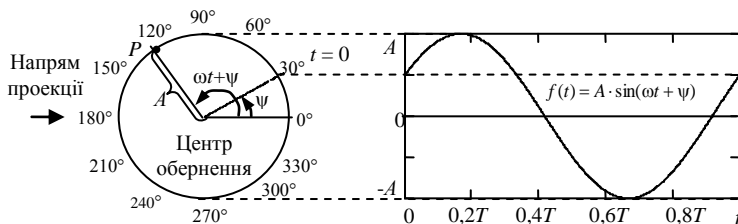


Рисунок 1.4 – Положення проекції точки, яка обертається, на вертикальну вісь є гармонічною функцією $f(t)$

Періодичні сигнали бувають гармонічними (рис. 1.4), тобто такими, які мають тільки одну гармоніку, і негармонічними (полігармонічними), спектр яких складається

з множини гармонічних складових. До гармонічних сигналів належать сигнали, що описуються функцією синуса або косинуса. Решта сигналів є полігармонічними. В главі 2 буде доведено, що періодичні сигнали загального вигляду розкладаються в ряд Фур'є, тобто мають дискретні спектри з сукупності гармонік. *Гармоніка* – це гармонічний сигнал з амплітудою і початковою фазою, яка дорівнює відповідним значенням амплітудного і фазового спектра періодичного сигналу при деякому значенні частоти, яка є кратною f . Періодичний сигнал можна достатньо точно апроксимувати сумою ряду таких гармонік.

В метрології широко застосовуються, наприклад, такі періодичні сигнали: прямокутний, пилкоподібний, трикутний (рис. 1.5).

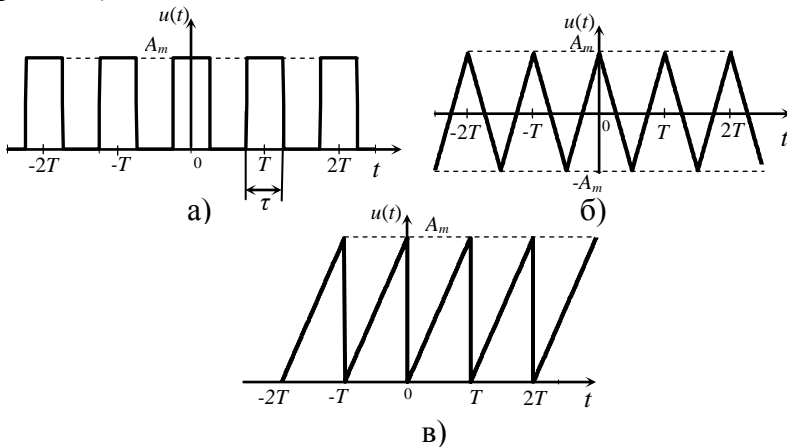


Рисунок 1.5 – Приклади періодичних сигналів:
 а) прямокутний; б) трикутний; в) пилкоподібний

Наявність вищих гармонік у спектрі періодичного сигналу кількісно описується коефіцієнтом гармонік, що характеризує відмінність форми певного періодичного сигналу від гармонічної. Він дорівнює відношенню середнього квадратичного значення сигналу суми вищих гармонік до середнього квадратичного значення першої гармоніки:

$$K_r = \sqrt{\sum_{i=2}^n U_i^2} / U_1,$$

де U_i – амплітуда i -тої гармоніки періодичного сигналу, $i = \overline{1, n}$.

Неперіодичні сигнали поділяють на *майже періодичні* і *перехідні*.

Майже періодичними називають сигнали, значення яких приблизно повторюються при додаванні до часового аргументу належним чином вибраного числа – майже періоду. Прикладом такого сигналу є сума періодичних функцій з некрatними періодами, наприклад $u(t) = \cos(\omega_0 t) + \cos(\sqrt{3}\omega_0 t)$. Майже періодичні функції, як і періодичні, також мають дискретний спектр, але тільки амплітудний з некрatними частотами.

Перехідні сигнали – описують перехідні процеси у фізичних системах. Перехідні неперіодичні сигнали мають неперервні спектри. Наприклад для прямокутного імпульсу виду $\sin(x)/x$.

Безумовно, на практиці може зустрітись будь-яке сполучення цих типів сигналів (неперіодичних і періодичних).

Випадковий сигнал – це змінна у часі ФВ, миттєве значення якої є випадковою величиною. Їх класифікація, яка подана на рис. 1.2 [5], буде детальніше розглянута в главі 4. Нижче наведено класифікацію *перешкод (завад)*, які є різновидами випадкових сигналів [3, 4].

Під *перешкодою (завадою)* розуміють сигнал однорідний з вимірювальним, що діє одночасно з ним. Наявність перешкод приводить до появи додаткової похибки (чи невизначеності) вимірювання. Між сигналами і перешкодами немає принципової різниці. Можлива ситуація, коли корисний сигнал для однієї вимірювальної системи є задовим для іншої (наприклад, в системах стільникового зв'язку при повторному використанні частот в ретрансляційних пунктах; в

радіолокаційних системах, які працюють на однакових частотах, випромінювані однією радіолокаційною станцією (РЛС) сигнали утворюють несинхронну імпульсну заваду для інших РЛС). Тобто, наявність у вимірювальній системі каналу передавання робить принципово можливим утворення для неї перешкод сигналами подібними до тих, що передаються цими каналами.

За місцем виникнення перешкоди поділяють на зовнішні та внутрішні. Причиною виникнення *зовнішніх перешкод* є природні процеси і робота різних технічних пристроїв. Останні створюють так звані *індустріальні перешкоди*.

Природними *джерелами зовнішніх перешкод (шумів)* є: випромінювання позаземних джерел або шум космічного простору; шум, зумовлений флуктуаційним характером поглинання радіохвиль в атмосфері Землі; шум, зумовлений тепловим випромінюванням Землі; перешкоди від грозових розрядів в атмосфері.

Основними джерелами індустріальних перешкод є: працюючі електричні апарати (високочастотна апаратура для дугового зварювання, газорозрядні пристрої, динамо-машини, генератори електричного струму, індукційна й комутаційна апаратура); системи запалювання двигунів внутрішнього згоряння різного призначення; лінії електропередач. Спектральна густина і спектральний склад таких перешкод суттєво різняться навіть у поблизу розташованих районах. Найбільш високою є їх густина у великих містах і наближених до них районах. Крім цих джерел, що утворюють нерегулярні заводські сигнали, є також доволі широка група технічних пристроїв, які є джерелами відносно регулярних високочастотних (ВЧ) коливань не призначених для випромінювання: системи розгортки електронно-променевих трубок; різного роду промислові, медичні ВЧ установки і т.п.

Внутрішні перешкоди зумовлені процесами, що виникають під час роботи самого ЗВТ.

Залежно від виду вмикання джерел перешкод і ВС в еквівалентних схемах ЗВТ розрізняють перешкоди загального

виду (синфазні) і перешкоди нормального (послідовного) виду. Джерело перешкод *загального виду* ввімкнено між загальними точками (корпусами) схем об'єкта вимірювань і ЗВ, тому вони є мультиплікативними. Джерело перешкод *нормального виду* ввімкнено послідовно у вхідний ланцюг ЗВТ, такі перешкоди по відношенню до ВС є адитивними.

Джерелом внутрішніх перешкод можуть бути, наприклад, електронно-обчислювальні машини (ЕОМ), які є складовими інформаційних вимірювальних систем (ІВС). Існує два основні види перешкод, генерованих ЕОМ [7]:

1. Перешкоди, генеровані швидкісними дискретними пристроями, – системної плати. За характером, це високочастотні перешкоди з частотами кратними тактовим частотам вузлів ЕОМ (центральный процесор, відеопроцесор, мікросхеми північного моста і південного моста, модулі пам'яті) і її складових частин (ядро, кеш тощо), а спектральний розподіл визначається ще й алгоритмами роботи програмного забезпечення і процесора.

2. Перешкоди генеровані джерелами живлення ЕОМ: широкосмугові – комутаційні (їх генерують: блок живлення комп'ютера; регульовані джерела живлення процесорів і пам'яті (інвертори));

низькочастотні: генерує блок живлення із-за нелінійності навантаження.

За *видом частотного спектру* перешкоди поділяють на білий і рожевий шуми. Спектральні складові *білого шуму* рівномірно розподілені за всім частотним діапазоном; у *рожевого шуму* усталеною є спектральна потужність, що припадає на декаду частоти.

За *основними властивостями* перешкоди можна поділити на три види: флуктуаційні, зосереджені та імпульсні. *Флуктуаційні* перешкоди відображають хаотичну, випадкову зміну з часом сигналу, однорідного з вимірюваним, у будь-якому місці вимірювального пристрою. Такі перешкоди часто називають *шумом*.

Приклад: внутрішні шуми вимірювальних електронних підсилювачів.

Розрізняють такі основні види шумів [3]:

- *тепловий* (шум Джонсона) – генерується будь-яким резистором, що розміщений у вимірювальному ланцюзі, який устанавлює нижню границю напруги шумів будь-якого вимірювального перетворювача, що має вихідний опір (за власними властивостями близький до білого шуму);

- *дробовий*, зумовлений рухом електронів – дискретних носіїв електричного струму в електровакуумних і напівпровідникових приладах (має рівномірний спектр, тобто є білим);

- *флікер-шуми* – виникають у напівпровідникових матеріалах, гранульованих резисторах, напилених конденсаторах. Причиною їх виникнення є нестационарні процеси, які пов'язані в основному із захопленням носіїв ловушками (домішковими атомами, енергетичні рівні яких поблизу рівня Фермі). Ці шуми зумовлюють низькочастотні (НЧ) флуктуації провідності середовища, крізь яке протікає струм. Спектральна потужність на декаду частоти приблизно усталена: шум усталеного резистора, пропорційний силі струму, що проходить через нього; шум струму бази транзистора (рожеві шуми).

Вплив флуктуаційної перешкоди може бути зменшений для періодичних ВС при підсумовуванні (накопиченні) сигналу у кількох періодах (рис. 1.6).

Накопичення сигналів $f_k(t) = s(t) + n_k(t)$, де $n_k(t)$ – реалізація шумового процесу у k -му періоді (рис. 1.6, а), у кількох періодах ілюструється на рис. 1.6, б.

Завдяки випадковому характеру перешкоди, відліки якої розподілені за нормальним законом з математичним сподіванням рівним нулю та середньоквадратичним відхиленням $0,3A$, при підсумовуванні достатньо великої кількості періодів сигналу середнє значення перешкоди наближається до нуля, тому [6]

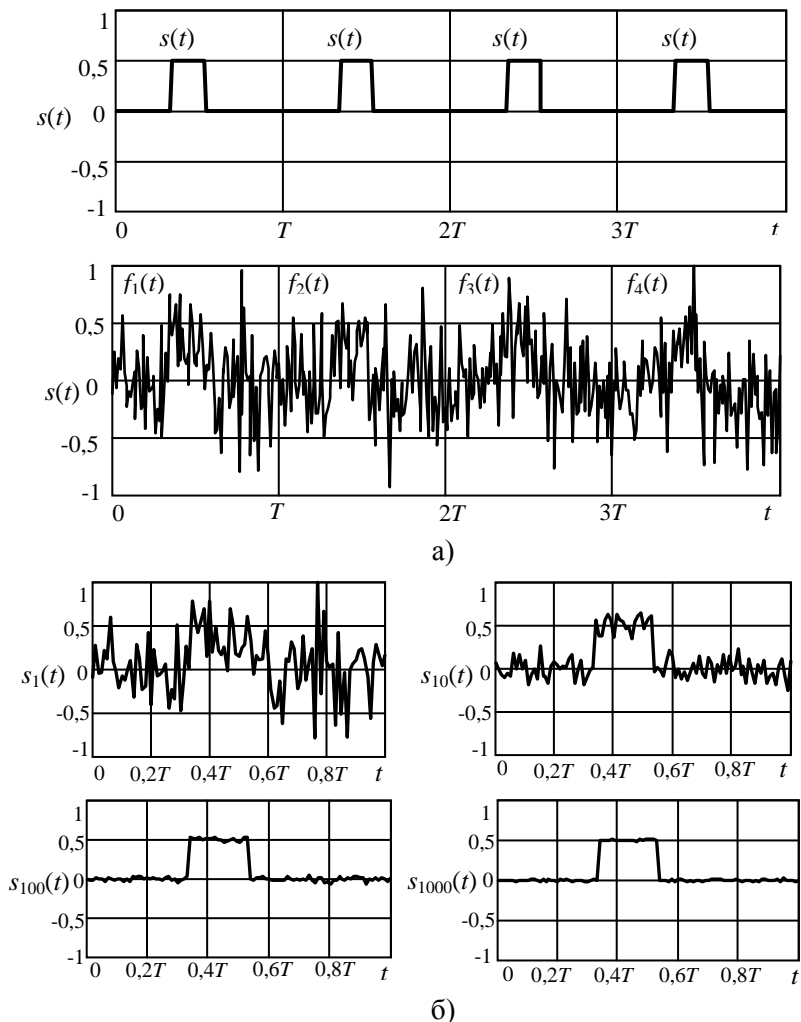


Рисунок 1.6 – Періодична послідовність імпульсів:
 а) без впливу та за наявності шумів; б) покращення
 виділення періодичного сигналу шляхом міжперіодного
 накопичення за 10, 100 та 1000 періодів

$$s_N(t) = \frac{1}{N} \sum_{k=1}^N f_k(t) \xrightarrow{N \rightarrow \infty} s(t).$$

Зосередженими називають перешкоди, основна частина потужності яких зосереджена на окремих ділянках діапазону частот, які менші, ніж смуга пропускання ЗВ. Перешкоди ЗВ від промислової силової мережі частотою 50 Гц є зосередженими. Ефективність їх заглушення великою мірою визначається наявністю апріорних даних про частотний спектр.

Імпульсними перешкодами називають регулярну або хаотичну послідовність імпульсних сигналів, які однорідні з вимірювальним сигналом. Джерелами таких перешкод є цифрові комутаційні елементи ЗВТ або працюючого поряд з ними пристрою. Характерний приклад імпульсних перешкод – перешкоди від пристроїв запалювання двигунів внутрішнього згоряння (групи імпульсів з випадковими амплітудами і часовими інтервалами тривалістю в межах від кількох мікросекунд до кількох мілісекунд, тривалість окремих імпульсів 1...6 нс [8]).

Перешкоди від ліній електропередач також мають випадковий характер і форму непрямокутних імпульсів, тривалість яких суттєво перевищує тривалість імпульсів перешкод від пристроїв запалювання двигунів внутрішнього згоряння. Вони виникають, завдяки перехідним процесам викликаним електричними розрядами, які хаотично виникають на поверхнях провідників і ізоляторів лінії.

Імпульсні струми, що виникають при таких розрядах, розповсюджуються уздовж лінії, яка може мати вигляд коаксіального хвилеводу або одиночної лінії над земною поверхнею. Імпульсні та зосереджені перешкоди часто називають *наводками*.

Оскільки основним наслідком дії перешкоди є поява додаткової похибки вимірювань, то прагнуть усунути, або принаймні зменшити, їхній вплив на ЗВТ. Для цього доцільно, якщо це можливо, усунути причини їхнього виникнення. Способи боротьби з перешкодами значною мірою залежать від їх спектрального складу, виду ВС і перешкоди.

Більшість електричних перешкод можна усунути шляхом екранування, заземлення ЗВ, застосування спеціальних фільтрів. Теплові шуми можна значно зменшити шляхом охолодження їхнього джерела. Однак, у цілому, боротьба з перешкодами є надзвичайно складною і потребує глибокого аналізу їхніх можливих джерел. В деяких випадках застосовують монолітні кам'яні столи для усунення сторонньої вібрації, розміщення ЗВ або їхніх частин у термостатах, електричне й електромагнітне екранування помешкань для усунення електромагнітних наводок. Ураховуючи наявність перешкод, які не можуть бути повністю усунені, обробка сигналів потребує такої оптимізації, щоб за цих умов забезпечувати найкращу точність вимірювання інформативних параметрів ВС. Принципи такої оптимізації будуть розглянуті в главі 9.

1.2. Математичні моделі сигналів. Спектральне подання сигналів

Для описання ВС в метрології застосовуються математичні моделі такого вигляду [3, 4]

$$y = f(x, a, b, c, \dots),$$

де: y – основний інформативний параметр сигналу; x – незалежний аргумент сигналу; a, b, c – кількісні параметри сигналу, один чи кілька з яких є інформативними.

В теорії сигналів, як і в метрології, найбільш поширені часові ($x = t$ – час) та частотні ($x = \omega$ – кутова частота) математичні моделі. Вигляд моделі, яка обирається, визначається зручністю її застосування для розв'язування конкретної задачі. Тому в загальному випадку можуть застосовуватись часово-частотні, просторово-частотні та інші моделі сигналів.

Математичні моделі у часовій області $y = f(t, a, b, c, \dots)$ є функціями, які достатньо точно описують зміни у часі

напруги (струму тощо) y , а один або кілька з параметрів – a, b, c, \dots залежить від вимірюваних фізичних величин. Часова форма зображення дозволяє вимірювати такі параметри як тривалість, період, амплітуда, параметри форми та інші характеристики періодичних та неперіодичних сигналів, а також визначити такі важливі їх характеристики як потужність і енергія сигналу. В загальному випадку багатомірного сигналу його модель можна подати у вигляді m -мірного вектора

$$\mathbf{Y}(t) = [y_1(t), y_2(t), \dots, y_m(t)]^T,$$

де T – символ транспонування вектора, m – відповідає кількості одновимірних сигналів, які можуть бути зареєстровані одночасно в кількох вимірювальних каналах або послідовно, як це, наприклад, показано на рис. 1.6, б. Наприклад, сигнал $y_i(t)$ реєструється в i -му каналі багатоканального ЗВ або ж в i -му періоді спостереження періодичного сигналу.

Прикладами математичних моделей елементарних сигналів є постійний сигнал $y(t) = A$, де A – єдиний параметр постійного сигналу; розглянутий вище гармонічний сигнал (синусоїда або косинусоїда). В розділі 1.3 розглядаються моделі й інших видів елементарних сигналів, які застосовуються у вимірювальній техніці.

Математичні моделі складних детермінованих і випадкових сигналів будуть вводиться поступово в процесі вивчення навчальної дисципліни.

Оскільки на практиці корисний ВС $y(t)$ зазвичай спостерігається на тлі перешкод $n(t)$, тому математична модель ВС в загальному випадку може бути представлена

$$y = f(t, n, a, b, c, \dots).$$

В теорії сигналів часто застосовують моделі [1]

адитивної завади

$$y = f(t, a, b, c, \dots) + n(t),$$

мультиплікативної завади

$$y = kf(t, a, b, c, \dots),$$

де $k(t)$ – випадковий процес, та їх сполучення

$$y = kf(t, a, b, c, \dots) + n(t).$$

Разом з часовим широко використовують *спектральне (частотне)* представлення сигналів. Воно є дуже зручним для визначення параметрів апаратури передавання і оброблення сигналів. Так, для періодичного сигналу $y(t)$ є справедливим його розкладання у тригонометричний ряд

$$y(t) = A_0 + \sum_{n=1}^{\infty} A_n \cos(n\omega_1 t + \psi_n),$$

де: A_0 – стала складова; $\omega_1 = 2\pi/T$, T – період сигналу; A_n , ψ_n – амплітуда і фаза n -ї гармоніки.

Кожна з гармонічних складових сигналу, яка має назву гармоніки, може бути представлена графічно у вигляді двох залежностей: амплітуди і фази від частоти (рис. 1.7, а). Множини значень $A_n(\omega)$ і $\psi_n(\omega)$ утворюють відповідно *амплітудний* і *фазовий спектри*, які характеризують властивості періодичного сигналу $y(t)$ у частотній області (рис. 1.7, б). Такий спектр називають *лінійчатим* або *дискретним*. Ця та інші форми представлення спектра періодичного сигналу, а також методика їх обчислення докладніше розглянуті у підрозділі 2.1.

Окрім дійних, в теорії сигналів широко застосовуються комплексні сигнали. За теоремою Ейлера $e^{j\pm x} = \cos x \pm j\sin x$, де $j = \sqrt{-1}$ – уявна одиниця, тому гармонічний сигнал $y(t) = A\cos(\omega t + \psi)$ можна подати у вигляді [9]

$$y(t) = \operatorname{Re}[\dot{y}(t)] \text{ або } y(t) = \frac{1}{2} \dot{y}(t) + \frac{1}{2} \dot{y}^*(t), \quad -\infty > t > \infty,$$

де $\dot{y}(t) = Ae^{j(\omega t + \psi)}$ – комплексний гармонічний сигнал,
 $\dot{y}^*(t) = Ae^{-j(\omega t + \psi)}$ – комплексно-спряжений з ним сигнал. Таке
 представлення полегшує вирішення багатьох задач
 перетворення сигналів і дає вірний кінцевий результат у
 вигляді дійсної частини комплексного рішення.

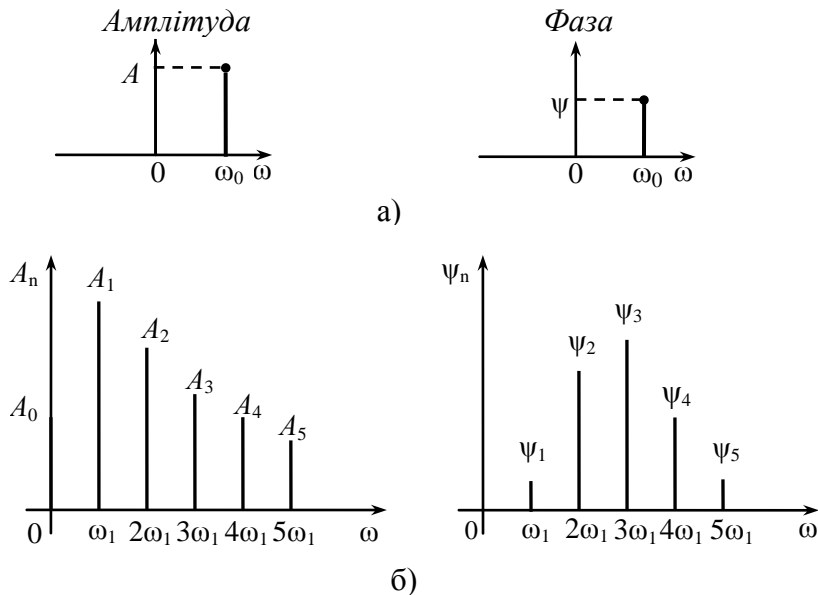


Рисунок 1.7 – Амплітудний і фазовий спектри гармоніки (а) та типового періодичного сигналу (б)

Узагальненим символічним поданням сигналу будь-якої складності є аналітичний сигнал [1, 5]

$$\dot{y}(t) = y(t) + j\tilde{y}(t) = a(t)e^{j\theta(t)},$$

$$\text{де } y(t) = \operatorname{Re}[\dot{y}(t)] = a(t) \cos\theta(t),$$

$$\tilde{y}(t) = \operatorname{Im}[\dot{y}(t)] = a(t) \sin\theta(t),$$

$a(t)$ – обвідна сигналу $y(t)$, $\theta(t)$ – його миттєва фаза.

Сигнал $y(t)$ та спряжений з ним сигнал $\tilde{y}(t)$ пов'язані між собою прямим і оберненим перетвореннями Гільберта [10]:

$$\tilde{y}(t) = y(t) * \frac{1}{\pi t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{y(\tau) d\tau}{t - \tau} = \mathbf{H}[y(t)];$$

$$y(t) = -\tilde{y}(t) * \frac{1}{\pi t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tilde{y}(\tau) d\tau}{\tau - t} = \mathbf{H}^{-1}[\tilde{y}(t)],$$

де «*» – символ згортки.

Для спектрального описання неперіодичних сигналів $y(t)$ застосовують комплексну спектральну густину $\dot{S}(\omega)$, її модуль $|\dot{S}(\omega)|$ і аргумент $\arg \dot{S}(\omega)$.

Спектральна густина визначається за допомогою інтеграла Фур'є

$$\dot{S}(\omega) = \int_{-\infty}^{\infty} y(t) e^{-j\omega t} dt = |\dot{S}(\omega)| e^{j \arg \dot{S}(\omega)} = \operatorname{Re}[\dot{S}(\omega)] - j \operatorname{Im}[\dot{S}(\omega)]$$

де $\operatorname{Re}[\dot{S}(\omega)]$ і $\operatorname{Im}[\dot{S}(\omega)]$ – дійсна й уявна частини спектральної густини, які визначаються за виразами:

$$\operatorname{Re}[\dot{S}(\omega)] = \int_{-\infty}^{\infty} y(t) \cdot \cos \omega t dt = A(\omega);$$

$$\operatorname{Im}[\dot{S}(\omega)] = \int_{-\infty}^{\infty} y(t) \cdot \sin \omega t dt = B(\omega).$$

Модуль і аргумент спектральної густини визначаються відповідно за формулами:

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

Спектральне зображення сигналу дає змогу оцінити його частотний діапазон, тобто граничні частоти, між якими

розміщені всі або основні, що мають найбільші амплітуди, гармонічні складові сигналу. Частотний діапазон є важливою характеристикою сигналу, що визначає необхідну смугу пропускання ЗВ для передавання сигналів з потрібною точністю. Докладніше питання щодо спектрального описання неперіодичних сигналів розглянуті у підрозділі 2.2.

1.3. Тестові сигнали. Дельта функція. Функція одиничного стрибка

Найбільш поширеними в метрології елементарними тестовими сигналами, окрім гармонічного сигналу, є одинична функція (функція включення) або функція Хевісайда і функція Дірка або δ -функція. Такі сигнали застосовуються для динамічного представлення сигналів і обчислення реакції фізичної системи на відомий вхідний сигнал.

Функція Хевісайда описується рівнянням (рис. 1.8)

$$\sigma(t-t_0) = \begin{cases} 0, & t < t_0, \\ 1/2, & t = t_0, \\ 1, & t > t_0. \end{cases}$$

Вона має один параметр – момент часу t_0 . Функція може бути отримана, наприклад, в результаті обчислення границі при $\xi \rightarrow 0$ функції вигляду [10]

$$v(t-t_0) = \begin{cases} 0, & t-t_0 < -\xi, \\ 1/2 \left[\frac{(t-t_0)}{\xi} + 1 \right], & -\xi \leq t-t_0 \leq \xi, \\ 1, & t-t_0 > \xi. \end{cases}$$

або іншого вигляду функцій, що змінюються в границях від 0 до 1 і мають параметр, при спрямуванні якого до певного граничного значення забезпечується практично миттєвий перехід від 0 до 1.

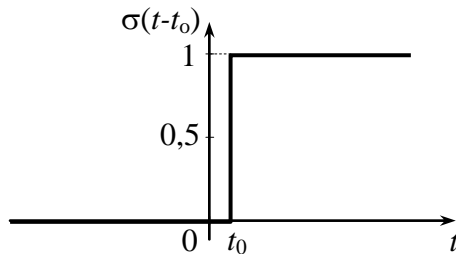


Рисунок 1.8 – Функція Хевісайда

Функція Хевісайда може бути застосована для наближеного представлення довільного сигналу $s(t)$ ступінчастою функцією (подібно до рис. 1.3, б)

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

$$= s_0 \sigma(t) + \sum_{k=1}^{\infty} (s_k - s_{k-1}) \sigma(t - k\Delta t).$$

Якщо крок дискретизації Δt спрямувати до нуля, тоді $k\Delta t$ буде наближатись до неперервної змінної t' , а малі прирости сигналу $(s_k - s_{k-1})$ представлятимуть диференціали $ds = \frac{ds}{dt'} dt'$. Підсумовування при цьому перетвориться в інтегрування в границях від нуля до нескінченності. В результаті отримаємо динамічне представлення довільного сигналу у вигляді [10]

$$s(t) = s_0 \sigma(t) + \int_0^{\infty} \frac{ds}{dt'} \sigma(t - t') dt'.$$

Функція Дірака або δ -функція описується рівнянням

$$\delta(t - t_0) = \begin{cases} 0, & t \neq t_0, \\ \infty, & t = t_0. \end{cases}$$

Приклади зображення δ -функції подані на рис. 1.9.

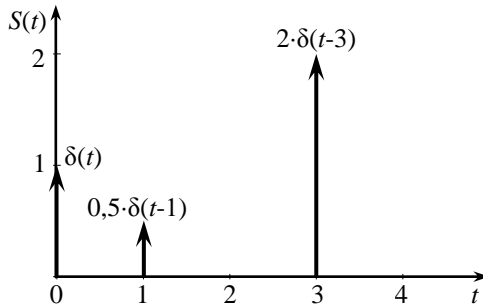


Рисунок 1.9 – Графік сигналу $s(t) = \delta(t) + 0,5\delta(t-1) + 2\delta(t-3)$

Функцію Дірака можна розглядати як граничну функцію однопараметричного сімейства неперервних функцій густин ймовірностей, наприклад нормального розподілу з нескінченно малим середньоквадратичним

$$\delta(t-t_0) = \lim_{\gamma \rightarrow 0} \frac{\exp\left[-(t-t_0)^2 / (2\gamma^2)\right]}{\gamma\sqrt{2\pi}}$$

або рівномірного розподілу, який можна представити з використанням функцій Хевісайда у вигляді

$$\delta(t-t_0) = \lim_{\gamma \rightarrow 0} \frac{1}{\gamma} \left[\sigma\left(t-t_0 + \frac{\gamma}{2}\right) - \sigma\left(t-t_0 - \frac{\gamma}{2}\right) \right].$$

Фізична розмірність δ -функції, як і частоти, $1/c$.

Функція Дірака має спектр нескінченної ширини, один параметр – момент часу t_0 і такі властивості:

- $\int_{-\infty}^{\infty} \delta(t-t_0) dt = \int_{t_0-\varepsilon}^{t_0+\varepsilon} \delta(t-t_0) dt = 1,$

де ε - будь-яке скільки завгодно мале число.

- Одинична і δ -функції пов'язані між собою співвідношеннями:

$$\sigma(t-t_0) = \int_0^t \delta(t-t_0) dt; \delta(t-t_0) = \frac{d\sigma(t-t_0)}{dt}.$$

Ця властивість, як і перша, є наслідком властивостей спектральних густин, з яких граничним переходом може бути отримана δ -функція.

- Фільтрувальна дія, яка описується рівнянням

$$\int_{-\infty}^{\infty} s(t)\delta(t-t_0) dt = s(t_0).$$

Остання властивість використовується для динамічного представлення сигналу $s(t) = \int_{-\infty}^{\infty} s(\tau)\delta(t-\tau) d\tau$, яке відповідає апроксимації сигналу зімкненою послідовністю прямокутних імпульсів при спрямуванні їх тривалості до нуля, а також дискретизованої у часі функції з кроком дискретизації Δt

$$s_d(t) = \sum_{n=-\infty}^{\infty} s(n\Delta t)\delta(t-n\Delta t).$$

Гармонічний сигнал описується наведеним вище (див. рис. 1.4) рівнянням синусоїди $f(t) = A \cdot \sin(\omega t + \psi)$ або косинусоїди $f(t) = A \cdot \cos(\omega t + \psi)$. Параметрами такого сигналу є: амплітуда A , колова частота $\omega = 2\pi/T$ (період T або частота $f = 1/T$) і початкова фаза ψ .

У ЗВ використовують ВС різноманітної форми. Розглянемо деякі найбільш поширені.

Прямокутні імпульси. Поодинокий ідеальний прямокутний імпульс тривалістю τ з амплітудою A_m , який з'являється в момент часу t_0 , описується рівнянням (рис. 1.10)

$$s(t) = A_m [\sigma(t-t_0) - \sigma(t-t_0 - \tau)],$$

тобто він формується як різниця двох одиничних функцій, зсунутих за часом на тривалість імпульсу τ .

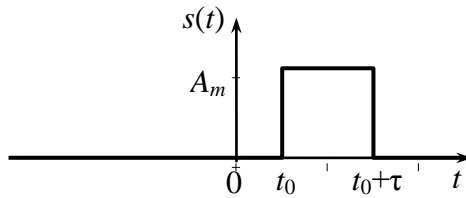


Рисунок 1.10 – Прямокутний імпульс

Послідовність прямокутних імпульсів є сумою поодиноких імпульсів з періодом слідування T (рис. 1.5, а)

$$s(t) = \sum_{k=-\infty}^{\infty} A_m [\sigma(t - kT + \tau/2) - \sigma(t - kT - \tau/2)].$$

Для її описання потрібно знати три параметри: амплітуду A_m , тривалість τ і період T . Відношення періоду до тривалості прямокутного імпульсу називають *шпаруватістю* q , а обернену до нею величину – *коефіцієнтом заповнення*. Послідовність імпульсів з шпаруватістю $q = 2$ називають *меандром*.

Ідеальних прямокутних імпульсів у природі не існує, а у реальних імпульсів час зміни сигналу від нульових до амплітудних значень (і навпаки) завжди має скінченну тривалість. Отже, у реальних імпульсів форма наближається до трапецеїдальної.

Трапецеїдальний імпульс також є ідеалізацією реальних імпульсів, які відрізняються від трапеції складом вершини імпульсу, викидами на вершині і у паузі та іншими особливостями, врахованими у системі параметрів реального прямокутного імпульсу. Для ВС, який близький за формою до трапеції, прийнято визначати його амплітуду (висоту) A , а з часових параметрів *тривалість імпульсу* τ_i , *тривалість фронту* τ_ϕ та *тривалість зрізу* τ_z (рис. 1.11).

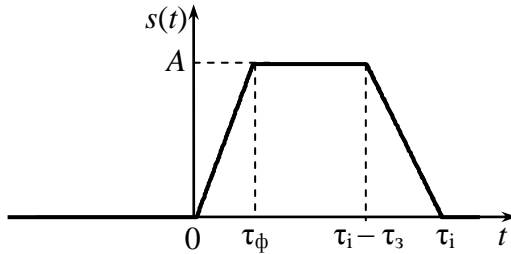


Рисунок 1.11 – Трапецеїдальний імпульс

Сигнали з лінійними ділянками. Під час побудови ЗВ широко застосовують періодичні сигнали з лінійними ділянками. Це передусім лінійний знакозмінний і однополярний лінійно-змінний (пилкоподібний) сигнали (рис. 1.5).

Періодичний лінійний знакозмінний сигнал (рис. 1.5, б) описується рівнянням

$$u(t) = \sum_{n=-\infty}^{\infty} s(t+nT),$$

де T – період сигналу, який при $t \in [-T/2; T/2]$ має вид

$$s(t) = \frac{4A_m}{T} t \left[\sigma\left(t + \frac{T}{2}\right) - 2\sigma(t) + \sigma\left(t - \frac{T}{2}\right) \right] + A_m \left[\sigma\left(t + \frac{T}{2}\right) - \sigma\left(t - \frac{T}{2}\right) \right],$$

а пилкоподібний сигнал (рис. 1.5, в) рівнянням

$$u(t) = \frac{A_m}{T} \sum_{n=-\infty}^{\infty} (t-nT) \left[\sigma(t-nT) - \sigma(t-(n+1)T) \right].$$

1.4. Енергія і потужність сигналу

З курсу фізики відоме співвідношення для визначення потужності, що розсіюється резистором, $P = \frac{U^2}{R}$, де U – постійна напруга; R – опір резистора. При цьому за час T в ньому виділяється теплова енергія $E = \frac{U^2 T}{R}$.

Якщо напруга є змінною у часі, тобто є сигналом $s(t)$, тоді потужність, що розсіюється на цьому резисторі, також залежатиме від часу, тобто є *миттєвою потужністю* (англ. instantaneous power) [11]

$$p(t) = \frac{s^2(t)}{R}.$$

Для обчислення енергії, яка виділяється за час T , потрібно виконати інтегрування миттєвої потужності

$$E = \int_0^T p(t) dt = \frac{1}{R} \int_0^T s^2(t) dt.$$

Вводять також поняття *середньої потужності* (англ. average power) за заданий проміжок часу

$$P_{\text{сеп}} = \frac{E}{T} = \frac{1}{RT} \int_0^T s^2(t) dt.$$

В усі ці вирази входить R – опір навантаження. В теорії сигналів для порівняння різних сигналів, коли невідоме і є неважливим конкретне значення R , приймають його величину рівною 1, виключаючи таким чином з наведених вище формул. В результаті отримуємо визначення *миттєвої потужності*, *енергії* та *середньої потужності*, які прийняті в теорії сигналів [11]:

$$p(t) = s^2(t); E = \int_0^T s^2(t) dt; P_{\text{сеп}} = \frac{1}{T} \int_0^T s^2(t) dt.$$

«Потужність» при такому записі має розмірність V^2 , енергія – V^2/c . Таким чином, в даному випадку маємо чисельні значення миттєвої потужності, енергії та середньої потужності, які розсіюються на резисторі з опором 1 Ом. За потреби нескладно перерахувати ці значення для довільного значення опору R .

Якщо енергія сигналу нескінченна, застосовуючи граничний перехід, визначають його середню потужність

$$P_{\text{сеп}} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s^2(t) dt.$$

Для періодичних сигналів такий перехід не необхідним, достатньо обмежитись тривалістю, яка дорівнює періоду сигналу.

Середньоквадратичне значення (*англ.* root mean square) сигналу є коренем квадратним з середньої потужності

$$\sigma_s = \sqrt{P_{\text{сеп}}} = \sqrt{\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s^2(t) dt}.$$

З точки зору класифікації сигналів в теорії сигналів важливим є клас сигналів з *інтегрованим квадратом*. Інакше їх називають сигналами з *обмеженою енергією*. Для таких сигналів є справедливим співвідношення

$$\int_{-\infty}^{\infty} s^2(t) dt < \infty.$$

Багато важливих співвідношень в теорії сигналів отримуються за умови кінцевої енергії сигналів, що аналізуються. Якщо ж ця умова не виконується змінюють підхід до рішення задачі (наприклад, визначення кореляційної функції для сигналів з кінцевою і нескінченною енергією розглядаються в главі 3) або застосовують апарат узагальнених функцій (Фур'є-аналіз не інтегрованих сигналів).

1.5. Перетворення типу сигналів: аналогові і цифрові сигнали

В розділі 1.1 розглядалась класифікація та математичні моделі деяких сигналів. За цією класифікацією сигнали можуть бути аналоговими (неперервними або кусково-неперервними), дискретними й цифровими. В цьому розділі будуть розглянуті теоретичні питання щодо перетворення типів сигналів для отримання з аналогового – цифрового і навпаки. Такі перетворення застосовуються для реалізації цифрової обробки сигналів (ЦОС) із застосуванням спеціалізованих або універсальних цифрових засобів обчислювальної техніки (рис. 1.12 [11]).

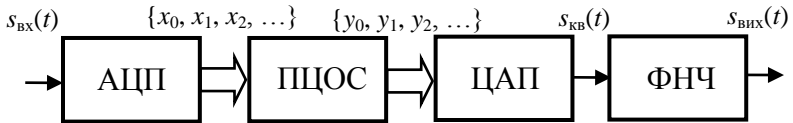


Рисунок 1.12 – Узагальнена структурна схема системи цифрової обробки сигналів

Часова дискретизація вхідного аналогового сигналу $s_{вх}(t)$ (див. рис. 1.3, а) та його квантування за рівнем виконуються в аналого-цифровому перетворювачі (АЦП). Аналоговий фільтр низьких частот (ФНЧ) (*англ.* low-pass filter) для попередньої обробки сигналу (на рис. 1.12 не показаний) і АЦП разом утворюють пристрій, що має назву «кодер». На виході АЦП кодера сигнал є послідовністю чисел $\{x_0, x_1, x_2, \dots\}$ певної розрядності (див. рис. 1.3, г), яка поступає в процесор цифрової обробки сигналів (ПЦОС), де виконується необхідна його обробка. В процесорі реалізуються потрібні математичні операції над вхідними відліками; раніше отримані відліки і проміжні результати можуть зберігатись в пам'яті процесора для використання в наступних обчисленнях. Цифрові відліки вихідного сигналу ПЦОС $\{y_0, y_1, y_2, \dots\}$ за потреби можуть бути використані для отримання аналогового вихідного сигналу за допомогою цифро-аналогового перетворювача ЦАП $s_{кв}(t)$ (див. рис. 1.3, в) і (за необхідності)

згладжувального ФНЧ. Пристрій, що об'єднує ЦАП і згладжувальний ФНЧ, називається декодером. Декодер за отриманим цифровим сигналом формує аналоговий сигнал $s_{\text{вих}}(t)$, тобто виконує обернене порівняно з кодером перетворення. Сигнал $s_{\text{вих}}(t)$ максимально наближається до вигляду аналогового сигналу, що міг би бути отриманий у разі обробки сигналу $s_{\text{вих}}(t)$ аналоговими пристроями, подібні до яких були в чисельному вигляді реалізовані в ЦОС.

За допомогою ЦОС можуть бути здійснені усі відомі операції над сигналами. Наприклад, на основі цифрової обробки сигналів в метрології, медицині, радіотехніці вирішуються наступні завдання: генерація, модуляція, формування і корекція сигналів, фільтрація і оцінювання, аналіз спектрів, виявлення і розпізнавання сигналів, кодування і декодування, обробка зображень тощо. Сучасні цифрові ЗВ (осцилографи, аналізатори спектра, вимірювальні генератори тощо) дозволяють вимірювати (відтворювати) характеристики сигналів в діапазоні надвисоких частот з шириною смуги частот до кількох гігагерць.

Перевагами ЦОС порівняно з аналоговою є:

- багатофункціональність обробки, тобто можливість використання одного і того ж процесора для виконання різних операцій обробки шляхом зміни програми;
- мультиплексування, тобто можливість обробки сигналів від декількох датчиків в проміжках між суміжними тактами, що особливо актуально при реалізації багатоканальних інформаційно-вимірювальних систем;
- відсутність принципових обмежень на складність обробки (принципово можливо забезпечити необхідну точність обробки при будь-якій складності алгоритмів);
- зняття обмежень на види обробки, які властиві аналоговій техніці, тобто можливість реалізації будь-яких формально описуваних перетворень.

Реалізація ЦОС в сучасних ЗВ виконується із застосуванням цифрових сигнальних процесорів (ЦСП) (*англ.*

digital signal processor, DSP; сигнальний мікропроцесор – СМП; процесор цифрових сигналів – ПЦС; цифровий процесор обробки сигналів – ЦПОС) – спеціалізованих мікропроцесорів, призначених для цифрової обробки сигналів (зазвичай в реальному масштабі часу) [12]. На рис. 1.13 процесор ЦОС – це обчислювач, який саме й реалізується на ЦСП (DSP).

Особливості архітектури сигнальних процесорів порівняно з мікропроцесорами (МП) загального застосування пов'язані з прагненням максимально прискорити виконання типових задач ЦОС: цифрова фільтрація, перетворення Фур'є, пошук сигналів і т.п. Математично ці завдання зводяться до поелементного перемножування багатоелементних векторів дійсних чисел і подальшого підсумовування добутків (наприклад, при цифровій фільтрації вихідний сигнал фільтра з кінцевою імпульсною характеристикою дорівнює сумі добутків коефіцієнтів фільтра на вектор відліків сигналу, аналогічні обчислення проводяться при пошуку максимумів кореляційних і автокореляційних функцій цифрових сигналів). Тому ЦСП оптимізовані за швидкодією для виконання саме таких операцій [12]:

1. Операція «множення з накопиченням» (*англ.* Multiply-Accumulate, MAC) ($Y = Y + A \times B$), де Y , A , B – елементи масивів дійсних чисел, з автоматичним розрахунком адрес елементів масивів, зазвичай реалізована апаратно і виповнюється за один машинний цикл (рівнозначні 2-5 звичайним командам МП і мікроконтролерів).

2. Апаратна реалізація багаторазового повторення заданого набору команд, тобто циклів заздалегідь призначеної довжини без використання лічильників циклу і команд перевірки обнуління лічильника циклу – ознаки виходу з циклу;

3. Можливість одночасної в одному машинному такті вибірки команди і двох операндів для максимально швидкого виконання команди MAC. Для цього ЦСП має кілька портів

звернення до пам'яті (незалежних областей пам'яті, кожна зі своїм комплектом шин адреси і даних);

4. Підтримка векторно-конвеєрної обробки за допомогою генераторів адресних послідовностей.

Оцифровані відліки вхідного сигналу записуються в апаратно організовані кільцеві буфери. Кільцевий буфер схожий на стек, але на відміну від стека він ніколи не переповнюється.

Зміна покажчиків в кільцевому буфері на запис в буфер (стирання старих записів, додавання нових) і читання виконується апаратно. Елементи кільцевого буфера можуть бути аргументами MAC команди, причому перехід до наступного елементу кільцевого буфера відбувається автоматично за покажчиками при виконанні поточної MAC команди. При роботі з даними зовнішньої пам'яті застосовується прямий доступ до пам'яті (ПДП), для чого використовується контролер ПДП або контролер кільцевого буфера з ПДП. Для асинхронного запису і читання в буфер часто використовується двохпортова пам'ять.

Для роботи з синфазними каналами в деяких ЦСП існує два ядра й два набори буферів і контролерів керування кільцевими буферами.

Теорія ЦОС вивчає принципи і методи побудови алгоритмів, програм і технічних засобів цифрової обробки сигналів. Залежно від засобів реалізації (апаратні, програмні або змішані) розглядаються завдання розробки і реалізації алгоритмів ЦОС, методи синтезу фізично можливих операторів для різноманітних прикладних завдань, процедури для досягнення потрібної точності виконання алгоритмів.

Як програмний, так і апаратний способи реалізації ЦОС базуються або на часовому, або на спектральному підході. Розглянемо спочатку принципи часового підходу. Цей підхід потребує, в першу чергу, коректного перетворення аналогового сигналу в дискретний цифровий, тобто представлення його сукупністю відліків у часі, яка є адекватною вхідному сигналу.

Постановка завдання дискретизації

Представлення неперервного сигналу $s(t)$ на інтервалі $[T_1, T_2]$ вектором відліків (координат) $(c_1, c_2, \dots, c_N)^T$ (« T » – операція транспонування) може бути записане у вигляді

$$(c_1, c_2, \dots, c_N)^T = D[s(t)], \quad (1.1)$$

де D - оператор дискретного представлення сигналу, що реалізується пристроєм, званим дискретизатором. Аналогічно можна записати операцію відновлення функції $u(t)$, так звану відтворюючою функцією, за вектором відліків $(c_1, c_2, \dots, c_N)^T$

$$u(t) = D^{-1}[\{c_1, c_2, \dots, c_N\}], \quad (1.2)$$

де D^{-1} – оператор відновлення, обернений до оператора D , який реалізується пристроєм відновлення сигналу. Очевидно, відтворююча функція $u(t)$ відображає початковий сигнал $s(t)$ з деякою поточною похибкою наближення (відновлення): $\varepsilon(t) = s(t) - u(t)$. Завдання дискретизації в математичному плані зводиться до спільного вибору пари операторів D і D^{-1} , що забезпечують задану точність відновлення сигналу.

У загальному випадку лінійної відтворюючої функції вона може бути представлена апроксимуючим поліномом

$$u(t) = \sum_{i=1}^N c_i \varphi_i(t), \quad (1.3)$$

де $\{\varphi_i(t)\}_{i=1}^N$ – система базисних функцій. Частіш за все обирають ортогональні базисні функції, для яких скалярний

добуток $(\varphi_i, \varphi_j) = \int_{-\infty}^{\infty} \varphi_i(t) \varphi_j(t) dt = \begin{cases} 0, & \text{якщо } i \neq j; \\ c, & \text{якщо } i = j, \end{cases}$ де

$c = \text{const}$. Якщо $c = 1$, базисні функції до того ж є ортонормованими. Такі базисні функції забезпечують

збіжність ряду $\sum_{i=1}^N c_i \varphi_i(t)$ до функції $s(t)$ при $N \rightarrow \infty$. У разі

відновлення аналогового сигналу за дискретними відліками з'являється можливість обмежити їх кількість відповідно до заданої похибки відновлення.

Основні завдання, що вирішуються при дискретизації сигналу:

- вибір кроку дискретизації при заданій похибці дискретизації (наближення, відновлення);
- оцінка істотності відліків, що забезпечують представлення сигналу $s(t)$ із заданою точністю, і усунення можливої надмірності.

Алгоритм рішення задачі дискретизації:

- по наявних відліках, використовуючи базисні ортогональні функції і методи інтерполяції або екстраполяції, будують з тією або іншою мірою наближення початковий сигнал, тобто отримують відтворюючу функцію $u(t)$;
- оцінюють ефективність дискретизації при вибраному методі апроксимації, визначаючи похибку дискретизації (відновлення) із заданим кроком дискретизації або крок дискретизації при заданій похибці.

Вибір системи базисних функцій у складі апроксимуючого полінома (1.3) визначається вимогою забезпечення простоти технічної реалізації апаратних (програмних) засобів дискретизації і відновлення сигналу. При неортогональних представленнях сигналу часто застосовують степеневі алгебраїчні поліноми виду

$$u(t) = \sum_{i=0}^N a_i t^n,$$

де a_i – дійсні коефіцієнти.

Як ортогональні базисні функції використовують функції відліків, поліноми Лежандра, функції Уолша й інші. Якщо базисні функції вибрані так, що значення апроксимуючого полінома співпадають зі значеннями вибірки початкового сигналу в моменти їх відліку, то такий поліном називають інтерполяційним. З точки зору скорочення числа відліків інтерполяційні методи відновлення сигналу більш прийнятні, проте для їх реалізації потрібна затримка сигналу на інтервал інтерполяції, що у ряді випадків неприпустимо. Наприклад, у системах управління, що працюють в реальному часі, використовують екстраполяційні методи, які не вимагають затримки сигналу при проведенні операцій знаходження значень вибірок і відновлення сигналу (*екстраполяція* – (від лат. *extra* – ззовні, по за, крім і лат. *polire* – прикладжую, виправляю, змінюю) – особливий тип апроксимації, коли функція апроксимується за межами інтервалу, а не між заданими значеннями. Інакше, екстраполяція – наближене визначення значень функції в точках, що не належать відрізу, по значенням в точках у його середині [13]).

При обраному способі побудови відтворюючої функції повинна забезпечуватись задана точність відновлення сигналу. Залежно від цільового призначення отримуваної інформації використовують різні критерії точності наближення $u(t)$ до $s(t)$, що визначають чисельні значення похибки відновлення і, отже, дискретизації.

Для оцінки точності відновлення (похибки дискретизації) можуть застосовуватись наступні критерії:

1. *Критерій найбільшого відхилення* – встановлює максимальне значення допустимої похибки відновлення ε_m

$$\varepsilon_m = \max | \varepsilon(t) |_{t \in \Delta_i} = \max | s(t) - u(t) |_{t \in \Delta_i},$$

де Δ_i – ділянка апроксимації.

2. *Критерій середньоквадратичного наближення* – задає значення допустимої середньоквадратичної похибки відновлення σ_δ

$$\sigma_\delta \geq \sigma = \sqrt{\frac{1}{\Delta_i} \int_{\Delta_i} \varepsilon^2(t) dt},$$

де σ – середньоквадратична похибка наближення. Зазвичай цей критерій використовується для функцій, інтегрованих в квадраті.

3. *Імовірнісний критерій* – визначає допустиму ймовірність P_0 того, що поточна похибка наближення $\varepsilon(t)$ не перевищить деякого певного значення ε_0

$$P[\varepsilon(t) \leq \varepsilon_0] = P_0.$$

Важливою ознакою процесу дискретизації є регулярність відліків. Якщо крок дискретизації постійний в усьому діапазоні перетворення, дискретизація вважається рівномірною. Методи рівномірної дискретизації застосовуються найширше. Проте невідповідність кроку дискретизації характеру поведінки конкретної реалізації досліджуваного процесу на окремих ділянках часто призводить до значної надмірності відліків.

Якщо крок дискретизації змінюється, наприклад, залежно від швидкості зміни сигналу або за заданою програмою, дискретизацію називають нерівномірною. Розрізняють адаптивні методи нерівномірної дискретизації, при яких крок Δt_i змінюється залежно від поточної зміни параметрів сигналу, і програмовані методи, в яких зміна кроку дискретизації виробляється оператором на основі аналізу інформації, що поступає, або відповідно до заздалегідь встановленої програми. Адаптивні методи нерівномірної дискретизації

забезпечують ефективно усунення надмірності у відліках сигналів. Інтервал дискретизації в цьому випадку тісно пов'язаний з поточними значеннями параметрів реалізації сигналу.

Вибір частоти дискретизації виконується за допомогою теореми Котельникова.

Теорема Котельникова (доведення див. в додатку А) [1].

Довільний сигнал $s(t)$, неперервний спектр якого не містить частот вище $f_b = \omega_b/2\pi$, може бути точно відновлений (інтерпольований) за його відліками, які узяті через рівні інтервали часу $t = k\Delta t = k\pi/\omega_b$. Здійснюється це відновлення за допомогою ряду

$$s(t) = \sum_{k=-\infty}^{\infty} s(k\Delta t) \frac{\sin \omega_b(t - k\Delta t)}{\omega_b(t - k\Delta t)}, \quad (1.4)$$

який має назву ряду Котельникова. В ньому роль коефіцієнтів виконують відліки функції $s(k\Delta t)$, а базисними є функції виду

$$\varphi_k(t) = \frac{\sin \omega_b(t - k\Delta t)}{\omega_b(t - k\Delta t)}. \quad (1.5)$$

Функції (1.5) утворюють ортогональний базис в просторі як детермінованих сигналів з обмеженим (фінітним) спектром Фур'є, тобто для яких $\dot{S}(\omega) = 0$ при $|\omega| > \omega_b$, так і для випадкових сигналів з фінітним спектром середньої потужності [1].

Теорема Котельникова, таким чином, встановлює принципову можливість повного відновлення детермінованого сигналу з обмеженим спектром і вказує граничне значення кроку (інтервалу) дискретизації, при якому таке відновлення можливе.

Отже, неперервна функція з обмеженим спектром може бути представлена безліччю своїх значень (відліків), узятих в моменти часу $t = k\Delta t = k\pi/\omega_b$.

Базисні функції $\varphi_k(t)$ називають функціями відліків. Вони мають наступні властивості:

1. Оскільки при будь-яких цілих числах n і k справедливі співвідношення $\omega_B(n\Delta t - k\Delta t) = (n - k)\omega_B\Delta t = (n - k)\pi$, тому очевидно

$$\frac{\sin \omega_B(t - k\Delta t)}{\omega_B(t - k\Delta t)} = \begin{cases} 1, & \text{якщо } t = k\Delta t, \\ 0, & \text{якщо } t = n\Delta t, \quad n \neq k. \end{cases}$$

відліків і формування сигналу на виході відновлювального фільтру приведені на рис 1.13.

Завдяки цій властивості сигнали $s(t)$ з обмеженим спектром можуть бути точно представлені своїми дискретними відліками без втрати інформації.

Кожна з функцій $\varphi_k(t)$ має необмежену протяжність в часі і досягає свого найбільшого значення, рівного 1, в моменти часу $t = k\Delta t$. Відносно цього моменту часу функція $\varphi_k(t)$ симетрична. У будь-які інші моменти часу, що кратні Δt , функція $\varphi_k(t)$ дорівнює нулю.

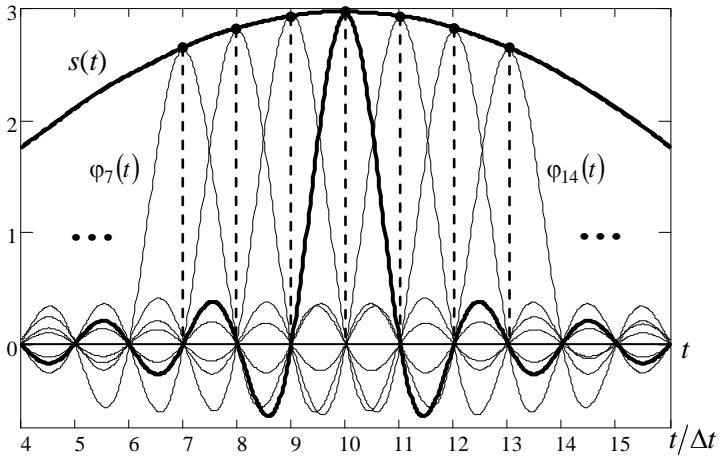


Рисунок 1.13 – Формування фрагменту сигналу на виході відновлювального фільтру

2. Функції відліків ортогональні з вагою 1 на нескінченно великому інтервалі часу

$$\int_{-\infty}^{\infty} \Phi_k(t) \Phi_l(t) dt = \begin{cases} \pi/\omega_B, & k=l, \\ 0, & k \neq l. \end{cases}$$

Кожну функцію відліку можна розглядати як реакцію (відгук) ідеального фільтру нижніх частот з частотою зрізу $f_3 = \omega_B/2\pi$ на дельта-імпульс, що приходить у момент часу $t = k\Delta t$ і має площу, яка дорівнює $s(k\Delta t)$. Такий фільтр має імпульсну характеристику $h(t) = \frac{\sin \omega_B t}{\omega_B t}$, яка збігається з функцією відліку при $k = 0$.

Таким чином, для отримання вихідного неперервного сигналу $s(t)$ з дискретизованого сигналу

$$s_d(t) = \sum_{n=-\infty}^{\infty} s(n\Delta t) \delta(t - n\Delta t) \quad (\text{див. розд. 1.2})$$

потрібно цю послідовність пропустити крізь ідеальний фільтр нижніх

частот з граничною частотою ω_b . На виході фільтру, в якому виконується операція згортки сигналу з імпульсною характеристикою фільтру, відновлений сигнал $u(t)$ матиме вигляд (1.4), тобто $u(t) = s(t)$ (див. рис. 1.13). При обчисленні згортки застосовується фільтрувальна властивість δ -функції.

Практичні аспекти використання теореми Котельникова.

Важлива особливість теореми Котельникова полягає в її конструктивному характері: вона не лише вказує на можливість розкладання сигналу у відповідний ряд, але й визначає спосіб відновлення безперервного сигналу, заданого своїми дискретними значеннями (відліками). З її допомогою може бути вибраний оптимальний крок дискретизації реального сигналу і оцінена похибка дискретизації, що виникає при цьому. Проте використання теореми як точного твердження по відношенню до реальних сигналів натрапляє на ряд принципових труднощів. *По-перше*, реальний сигнал має кінцеву тривалість і, отже, має необмежений спектр. Проте через реальні властивості джерел сигналів і обмеженість смуги пропускання реальних приладів і систем спектр сигналу з тією або іншою мірою точності можна вважати обмеженим деякою граничною частотою. Найчастіше граничне значення частоти $f_{гр} = \omega_{гр}/2\pi$ визначають на основі енергетичного критерію, за яким практичну ширину спектру сигналу вибирають так, щоб в ній була зосереджена значна частина енергії сигналу. Для цього використовують рівність Парсеваля, що дозволяє визначити енергію сигналу E_s або через функцію $s(t)$, що описує реальний сигнал тривалістю T , або через модуль її спектральної густини $|\dot{S}(\omega)|$

$$E_s = \int_{-\infty}^{\infty} |s(t)|^2 dt = \int_0^T |s(t)|^2 dt = \frac{1}{\pi} \int_0^{\infty} |\dot{S}(\omega)|^2 d\omega.$$

Практична ширина спектру сигналу, яка зосереджена в діапазоні частот від 0 до деякого значення $\omega_{\text{гр}}$, визначається з співвідношення

$$\frac{1}{\pi} \int_0^{\omega_{\text{гр}}} |\dot{S}(\omega)|^2 d\omega = \frac{\eta}{\pi} \int_0^{\infty} |\dot{S}(\omega)|^2 d\omega, \quad (1.6)$$

де $\omega_{\text{гр}}$ – гранична частота, що визначає верхнє значення спектру сигналу; η – коефіцієнт досить близький до 1 (на практиці його значення залежно від вимог до якості відтворення сигналу вибирають в інтервалі від 0,9 до 0,998). Значення $\eta = 0,95$ означає, що в смузі частот від 0 до $\omega_{\text{гр}}$ міститься 95 % енергії сигналу. Значення граничної частоти знаходять, вирішуючи трансцендентне рівняння (1.6).

Обмеження спектру реального сигналу, природно, призводить до спотворення сигналу. Таким чином, відновлення обмеженого в часі сигналу по відліках відповідно до теореми Котельникова за умови примусового обмеження спектру сигналу можливо тільки приблизно. Точність такого наближення може бути оцінена як абсолютним значенням похибки, званою енергією похибки

$$\Delta E = \frac{1}{\pi} \int_{\omega_{\text{гр}}}^{\infty} |\dot{S}(\omega)|^2 d\omega,$$

так і відносною похибкою: $\gamma = \frac{\Delta E}{E_s}$, де $E_s = \frac{1}{\pi} \int_0^{\infty} |\dot{S}(\omega)|^2 d\omega$.

Похибка дискретизації виникає не лише за рахунок примусового обмеження спектру, але і за рахунок кінцевого числа відліків на інтервалі тривалості сигналу $T = T_2 - T_1$, яке відповідно до теореми Котельникова складе $2f_{\text{гр}} T + 1 \approx 2f_{\text{гр}} T$ (оскільки зазвичай $f_{\text{гр}} T \gg 1$). Ця складова є наслідком нехтування вкладом нескінченного числа функцій відліків,

відповідних вибіркам за межами інтервалу (T_1, T_2) . Для реальних сигналів теорему Котельникова є наближеною (для спрощення T_1 і T_2 вважаються кратними Δt):

$$s(t) \approx \sum_{k=T_1/\Delta t}^{T_2/\Delta t} s(k\Delta t) \frac{\sin \omega_{\text{гр}}(t - k\Delta t)}{\omega_{\text{гр}}(t - k\Delta t)}. \quad (1.7)$$

Не дивлячись на вище перелічені труднощі, теорема Котельникова (у зарубіжних джерелах – теорема Шенона-Найквіста) широко використовується в процесі перетворення аналогових сигналів в цифрову форму, а величину $B = 2f_{\text{гр}} T$ називають *базою* сигналу. Цей параметр відіграє важливе значення в теорії й техніці передавання дискретних повідомлень [1].

На практиці ця теорема має величезне значення. Наприклад, відомо, що діапазон звукових сигналів, що сприймаються людиною, не перевищує 20 кГц. Отже, при дискретизації записаних звукових сигналів з частотою не менш 40 кГц можна точно відновити вихідний аналоговий сигнал за його цифровими відліками, що й виконується в програвачах компакт-дисків для відновлення звуку. Частота дискретизації звукового сигналу при запису на компакт-диск становить 44 100 Гц.

В той же час, важливо щоб при дискретизації сигналу з частотою $2f_{\text{гр}}$ у ньому не виявилось складових з частотами вищими за $f_{\text{гр}}$ (*частота Найквіста*, яка дорівнює половині частоти дискретизації), оскільки це може викликати появу у спектрі відновленого після дискретизації сигналу складових, яких не було у вихідному сигналі. Це є наслідком ефекту, що має назву *аліасинг* (англ. *aliasing*) – накладання або неможливість розрізнення різних неперервних сигналів при їхній дискретизації [14] (рис. 1.14).

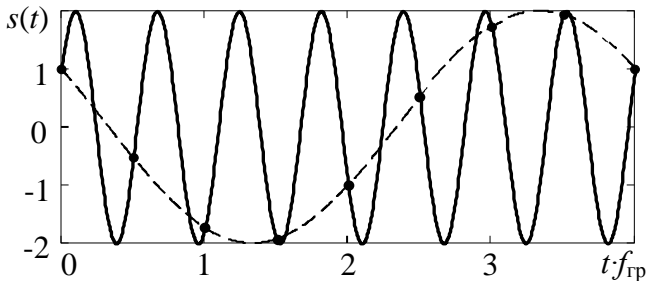


Рисунок 1.14 – Два різні синусоїдальні сигнали в результаті дискретизації не відрізняються: високочастотний з частотою $1,75f_{гр}$ (суцільна лінія) і низькочастотний $0,25f_{гр}$ (пунктир)

Припустимо, що при записі акустичного сигналу оркестру в приміщенні було присутнє джерело ультразвукового сигналу з частотою $f_1 = 32,05$ кГц. При запису з типовою частотою дискретизації сигналу на виході мікрофона $2f_{гр} = 44,1$ кГц в динаміках при прослуховуванні такого запису з використанням ЦАП буде чути завадовий сигнал на частоті

$$2f_{гр} - f_1 = f_{гр} - (f_1 - f_{гр}) = 44,1 - 32,05 = 12,05 \text{ (кГц)}.$$

Як видно, в результаті дискретизації сигнали з частотами, що лежать вище половини частоти дискретизації, сприйматимуться з частотами «дзеркальними» відносно $f_{гр}$ в нижній частині спектру і складаються з присутніми там гармоніками. Аналогічно в головний частотний діапазон дискретних сигналів «відображаються» й усі високочастотні шуми, присутні у вихідному аналоговому сигналі. Таким чином, високочастотні складові накладаються на низькочастотні, в результаті цього відновлений аналоговий сигнал буде спотворюватись.

Для запобігання аліасингу слід підвищувати частоту дискретизації для вірного відображення усіх частотних

складових сигналу або перед оцифровуванням обмежувати спектр сигналу фільтрами низьких частот, які пропускають без зміни всі частоти, нижче заданої, і пригнічують в сигналі частоти, вище заданої. Ця гранична частота називається частотою зрізу (*англ.* cut-off frequency) фільтра. Частота зрізу антиаліасингових фільтрів встановлюється рівною половині частоти дискретизації. Такі ФНЧ майже завжди вбудовуються у реальні АЦП.

Окрім забезпечення потрібного періоду дискретизації, для вірного відтворення сигналу є необхідним також узгодження (приблизна рівність) діапазону зміни сигналу та робочого діапазону АЦП $\bar{V}_{АЦП}$ при квантуванні сигналу. Для ілюстрації цього положення моделювались два синусоїдальні сигнали виду $s_i(t) = A_i \sin(2\pi ft)$ з частотою $f = 0,05$ Гц і амплітудами $A_1=2,5$ В, $A_2=0,4$ В. Для АЦП були вибрані параметри: $\bar{V}_{АЦП} = 5$ В; число розрядів цифрового коду з виходу АЦП (не враховуючи знаку) $L_A = 4$. При цьому розмір кванту $q = \bar{V}_{АЦП} / (2^{L_A} - 1) = 0,33$ В. В результаті квантування спостерігається значне спотворення форми сигналу $s_2(t)$, амплітуда якого значно менша діапазону $\bar{V}_{АЦП}$, – перетворений синусоїдальний сигнал $s_{2кв}(t)$ є послідовністю прямокутних імпульсів змінної полярності (рис. 1.15) [15].

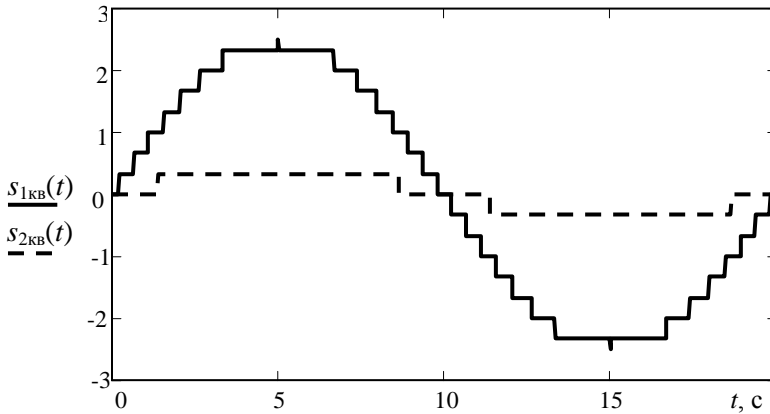


Рисунок 1.15 – Результати квантування синусоїдальних сигналів $s_1(t) = 2,5 \sin(2\pi ft)$ (суцільна лінія) і $s_2(t) = 0,4 \sin(2\pi ft)$ (пунктир)

Таким чином, для систем ЦОС є суттєвими такі параметри АЦП [15]: 1) $t_{\text{АЦП}}$ – час перетворення аналогової напруги $s(t)$ в цифровий код (повинна виконуватись нерівність $t_{\text{АЦП}} < \Delta t$); 2) L_A – число розрядів цифрового коду з виходу АЦП (не враховуючи знаку); 3) $\bar{V}_{\text{АЦП}}$ – робочий діапазон АЦП по вхідній напрузі, цей параметр встановлюється стандартним рядом значень – частіш за все $\bar{V}_{\text{АЦП}} = 1$ В або 5 В.

Якщо діапазон зміни сигналу $[X_{\min}, X_{\max}]$ узгоджений з діапазоном квантувача АЦП $[Y_{\min}, Y_{\max}]$, тоді функція перетворення ідеального квантувача (рис. 1.16, а) може бути представлена математичною моделлю [16]

$$Y = q \cdot \text{int}[X/q + 0,5 \text{sign } X],$$

де $\text{int}[*]$ – функція виділення цілої частини числа, а $\text{sign } X$ – дорівнює 1, якщо $X \geq 0$, і мінус 1 для від'ємних X , повинна відтворювати залежність $Y = X$ за умови $q \rightarrow 0$. На

рис. 1.16, а X_{li} і X_{ri} – ліва і права границі i -ї підмножини $[X_{li}, X_{ri})$.

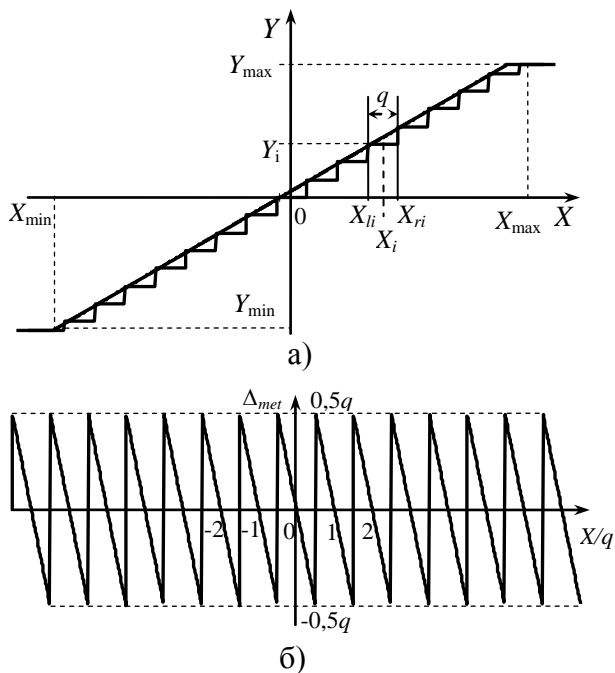


Рисунок 1.16 – Функція квантування ідеального квантувача (а) та його методична похибка (б)

Методична похибка АЦП $\Delta_{met} = Y - X$ (похибка квантування) змінюється в інтервалі $[-q/2, q/2]$ (рис. 1.16, б)

$$\begin{aligned} \Delta_{met} &= q \cdot \text{int}[X/q + 0,5 \text{sign } X] - X = \\ &= 0,5q \text{sign } X - q \text{Fr}[X/q + 0,5 \text{sign } X], \end{aligned}$$

де $\text{Fr}[X]$ – функція виділення дробової частки числа X .

Для цифрового ЗВ слід враховувати не ідеальність процесу квантування, яка характеризується додатковою інструментальною похибкою. Джерелом цієї похибки можна вважати додатковий аналоговий перетворювач (АП)

(рис. 1.17). Математична модель процесу квантування в загальному вигляді враховує відмінність коефіцієнту перетворення АП при різних рівнях сигналу X і має вигляд [16]

$$\Delta_{ins} = X\beta(X) + \alpha(X),$$

де $\beta(X) \ll 1$ – деяка, в загальному випадку нелінійна, функція вхідної величини X ; $\alpha(X)$ – зсув нульового рівня.

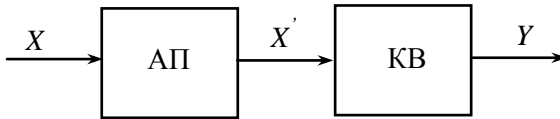


Рисунок 1.17 – Еквівалентна схема цифрового засобу вимірювання

Функцію перетворення цифрового ЗВ можна представити співвідношенням

$$Y = q \cdot \text{Int} \left[\frac{X + \Delta_{ins}(X)}{q} + 0,5 \text{sign}(X + \Delta_{ins}(X)) \right],$$

яке враховує дві складові похибки цифрового ЗВ: інструментальну пов'язану з властивостями елементів схеми ЗВ і можливою диференціальною нелінійністю АЦП (хоча цією складовою зазвичай можна нехтувати); методичну – за рахунок принципу аналого-цифрового перетворення.

Специфіка цифрових ЗВ, які застосовуються для вимірювання різного роду фізичних величин є предметом вивчення інших дисциплін.

Контрольні запитання

1. Дайте визначення поняття «сигнал» та «вимірювальний сигнал».

2. Назвіть основні елементи узагальненої схеми обробки сигналу в інформаційній вимірювальній системі та їх призначення.

3. За якими ознаками класифікують сигнали?

4. Як поділяються сигнали за характером вимірювань інформативного і часового параметрів? Охарактеризуйте їх відмінності.

5. Вкажіть відмінності детермінованого, квазидетермінованого і випадкового сигналів.

6. Охарактеризуйте відмінності зосереджених та імпульсних перешкод.

7. Чим вимірювальний сигнал відрізняється від сигналу? Приведіть приклади вимірювальних сигналів, які використовуються в різних розділах науки та техніки.

8. Поясніть відмінності аналогового, дискретного і цифрового сигналів.

9. Наведіть різновиди та джерела виникнення перешкод. Зовнішні та внутрішні перешкоди.

10. Які бувають типи математичних моделей вимірювальних сигналів? Як обрати математичну модель для вирішення конкретної задачі?

11. Який вигляд має математична модель багатомірного сигналу?

12. Скільки та яких параметрів необхідно знати для описання кожного з елементарних вимірювальних сигналів?

13. Як обчислюються енергія, миттєва і середня потужність сигналу кінцевої тривалості?

14. Дайте визначення операції квантування. Де і яким чином вона використовується в метрології? Що таке похибка квантування?

15. Дайте визначення дискретизації. Як проводиться дискретизація вимірювальних сигналів. Що стверджує теорема Котельникова?

16. Які критерії можуть застосовуватись для оцінки точності відновлення сигналу при його дискретизації?

17. В чому полягає сутність явища аліасингу? Як зменшити його вплив?

18. Запишіть математичні вирази функцій Хевісайда і Дірака та динамічне представлення довільного сигналу з їх застосуванням.

19. Побудуйте графік залежності амплітуди сигналу від часу для одиночного прямокутного імпульсу тривалістю $\tau_1 = 3$ мс, амплітудою $A = 2$ В, зсувом переднього фронту імпульсу відносно початку відліку часу $\tau_3 = 4$ мс. Записати його математичну модель з використанням функції Хевісайда.

20. Побудуйте графік залежності напруги від часу періодичного сигналу $s(t) = A \cdot \sin(\omega_0 t + \varphi)$. Подайте його комплексне представлення та перехід до дійсного у двох формах та відповідні спектральні представлення цього сигналу.

Глава 2

СПЕКТРАЛЬНИЙ АНАЛІЗ ПЕРІОДИЧНИХ ТА НЕПЕРІОДИЧНИХ СИГНАЛІВ

2.1. Розкладання в ряд Фур'є періодичних сигналів

В розділі 1.5 показано, що для відновлення аналогового сигналу з дискретного може бути застосований ряд Котельникова (1.4). Цей ряд є розкладанням довільного сигналу $s(t)$ по базисним функціям відліків

$$\varphi_k(t) = \frac{\sin \omega_B(t - k\Delta t)}{\omega_B(t - k\Delta t)},$$

які є ортогональними на нескінченному часовому інтервалі.

В практиці обробки сигналів широко застосовуються розкладання сигналів по базисним функціям різного вигляду. Частіш за все їх обирають ортогональними або ортонормованими. При цьому умова ортогональності може виконуватись не тільки на необмеженій, але й на обмеженій ділянці часу.

Якщо система базисних функцій визначена на обмеженому відрізку часу $[t_1, t_2]$ і є ортонормованою, тоді скалярний добуток базисних функцій [10, 17, 18]

$$(\varphi_i, \varphi_j) = \int_{t_1}^{t_2} \varphi_i(t) \varphi_j(t) dt = \begin{cases} 1, & \text{якщо } i = j; \\ 0, & \text{якщо } i \neq j. \end{cases} \quad (2.1)$$

В цьому випадку розглядається простір сигналів з обмеженою енергією, який є гільбертовим простором H . Розкладання довільного сигналу $s(t) \in H$ в ряд

$$s(t) = \sum_{i=0}^{\infty} c_i \varphi_i(t) \quad (2.2)$$

має назву узагальненого ряду Фур'є. Коефіцієнти цього ряду можна знайти множенням обох частин рівняння на базисну функцію, наприклад з індексом k , і подальшим інтегруванням рівняння за часом на інтервалі $[t_1, t_2]$

$$\int_{t_1}^{t_2} s(t) \varphi_k(t) dt = \sum_{i=0}^{\infty} c_i \int_{t_1}^{t_2} \varphi_i(t) \varphi_k(t) dt.$$

З умови ортонормованості базисних функцій (2.1) у правій частині рівняння залишиться лише один член суми з номером k

$$c_k = \int_{t_1}^{t_2} s(t) \varphi_k(t) dt = (s, \varphi_k).$$

Сукупність коефіцієнтів c_k узагальненого ряду Фур'є називається спектром сигналу $s(t)$ за ортонормованою системою базисних функцій $\{\varphi_k(t)\}$. Спектр повністю визначає сигнал $s(t)$. Для ортогональної системи

$c_k = \frac{1}{\|\varphi_k\|^2} \int_{t_1}^{t_2} s(t) \varphi_k(t) dt$, де $\|\varphi_n\| = \sqrt{\int_{t_1}^{t_2} \varphi_n(t) \varphi_n^*(t) dt}$ – норма функції $\varphi_n(t)$.

Якщо система функцій $\varphi_n(t)$ приймає комплексні значення, то умова ортогональності функцій має вигляд

$$(\varphi_i, \varphi_j) = \int_{t_1}^{t_2} \varphi_i(t) \varphi_j^*(t) dt = \begin{cases} c, & \text{якщо } i = j; \\ 0, & \text{якщо } i \neq j, \end{cases}$$

де функція $\varphi^*(t)$ – комплексно-спряжена до функції $\varphi(t)$.

При застосуванні повної ортонормованої послідовності (базису) функцій енергія сигналу представленого співвідношенням (2.2) згідно *тотожності Парсеваля* є

$$E = \sum_{n=0}^{\infty} |c_n|^2, \quad (2.3)$$

у разі ж тільки їх ортогональності має місце *рівність Парсеваля*

$$E = \sum_{n=0}^{\infty} |c_n|^2 \|\varphi_n\|^2. \quad (2.4)$$

Потужність сигналу при цьому

$$P = \frac{E}{t_2 - t_1} = \frac{1}{t_2 - t_1} \sum_{n=0}^{\infty} |c_n|^2 \|\varphi_n\|^2.$$

Співвідношення (2.3) і (2.4) для обчислення E відомі як *теорема Парсеваля*.

Як вже йшлося в розділі 1.5, вибір системи базисних функцій у складі апроксимуючого поліному визначається вимогами забезпечення простоти технічної реалізації апаратних (програмних) засобів дискретизації і відновлення сигналу. В [9, 10] доведено, що узагальнений ряд Фур'є при заданій системі функцій $\varphi_n(t)$ і фіксованій кількості доданків ряду забезпечує найкращу апроксимацію (у розумінні мінімуму середньоквадратичної похибки) даної функції $s(t)$, а саме

$$M = \int_{t_1}^{t_2} \left[s(t) - \sum_{n=1}^N a_n \varphi_n(t) \right]^2 dt$$

досягає мінімуму, коли $a_n = c_n$. При цьому є вірною *нерівність Бесселя*

$$\|s(t)\|^2 \geq \sum_{n=1}^N |c_n|^2 \|\varphi_n\|^2,$$

тобто енергія (потужність) сигналу, який апроксимується лінійною комбінацією ортогональних базисних функцій, завжди є не меншою за енергію (потужність) апроксимуючої функції.

Дуже поширеним на практиці як при аналоговій, так і при цифровій обробці сигналів є застосування ортогонального базису, який утворений гармонічними функціями з кратними частотами. Такий базис забезпечує точне розкладання сигналу і, крім цього, гармонічне коливання (функція часу) – єдине, що зберігає свою форму при проходженні через лінійні кола з постійними параметрами.

Розкладанню в ряд Фур'є піддаються періодичні сигнали. Вони представляються сумою гармонічних функцій або комплексних експонент з частотами, що утворюють арифметичну прогресію. Для того, щоб таке розкладання існувало, фрагмент сигналу $s(t)$ тривалістю в один період $[t_1, t_2]$ повинен задовольняти умовам Діріхле [11]:

- інтервал $[t_1, t_2]$ можна розділити на кінцеву кількість інтервалів, на яких функція $s(t)$ неперервна і монотонна, тобто число розривів першого роду (стрибків) має бути кінцевим;
- якщо t_0 є точкою розриву, то повинні існувати $s(t+0)$ і $s(t-0)$.

Ряд Фур'є може бути застосований для представлення не лише періодичних сигналів, але й сигналів кінцевої тривалості. При цьому обумовлюється часовий інтервал, для якого будується ряд Фур'є, а в інші моменти часу сигнал вважається рівним нулю. Для розрахунку коефіцієнтів ряду такий підхід фактично означає періодичне продовження сигналу за межами даного інтервалу.

Залежно від конкретної форми базисних функцій розглянемо три форми запису ряду Фур'є.

Перший випадок – тригонометрична синус-косинусна форма.

Системою ортонормованих базисних функцій на інтервалі періоду сигналу $[-T/2, T/2]$ у цьому випадку є $\varphi_0 = 1/\sqrt{T}$,

$\varphi_1 = \sqrt{2/T} \sin \omega_1 t$, $\varphi_2 = \sqrt{2/T} \cos \omega_1 t$, $\varphi_3 = \sqrt{2/T} \sin 2\omega_1 t$,
 $\varphi_4 = \sqrt{2/T} \cos 2\omega_1 t$, $\varphi_5 = \sqrt{2/T} \sin 3\omega_1 t$, $\varphi_6 = \sqrt{2/T} \cos 3\omega_1 t$, ..., де
 $\omega_1 = 2\pi/T$. Ряд Фур'є має вигляд

$$s(t) = \sum_{i=0}^{\infty} c_i \varphi_i(t) = \frac{a_0}{2} + \sum_{i=1}^{\infty} (a_i \cos i\omega_1 t + b_i \sin i\omega_1 t). \quad (2.5)$$

Гармонічне коливання з частотою $\omega_i = i\omega_1$ є i -ю гармонікою сигналу. Коефіцієнти ряду a_i і b_i розраховуються по формулах

$$a_0 = \frac{2}{T} \int_{-T/2}^{T/2} s(t) dt, \quad a_i = \frac{2}{T} \int_{-T/2}^{T/2} s(t) \cos i\omega_1 t dt,$$

$$b_i = \frac{2}{T} \int_{-T/2}^{T/2} s(t) \sin i\omega_1 t dt. \quad (2.6)$$

Таким чином, у загальному випадку періодичний сигнал є сумою постійної складової і нескінченного ряду гармонік з кратними частотами $\omega_i = i\omega_1$ ($i = 1, 2, \dots$).

Константа a_0 розраховується по загальній формулі для a_i . Заради цієї спільності введена дещо дивна, на перший погляд, форма запису постійного доданку (з діленням на два). Цей доданок є середнім значенням сигналу на періоді:

$$\frac{a_0}{2} = \frac{1}{T} \int_{-T/2}^{T/2} s(t) dt.$$

Зауважимо, що границі інтегрування не обов'язково мають бути $[-T/2, T/2]$, як в формулах (2.6). Важливо, щоб інтервал інтегрування мав тривалість T – результат від цього принципово не зміниться. В окремих випадках може виявитись зручніше виконувати інтегрування, наприклад, від 0 до T або від $-T$ до 0.

Оскільки інтегрування ведеться в симетричних границях, якщо $s(t)$ парна функція, тоді усі коефіцієнти b_i дорівнюватимуть нулю, й у формулі (2.5) будуть присутні тільки косинусні доданки. Якщо $s(t)$ непарна функція, дорівнювати нулю будуть, навпаки, косинусні коефіцієнти a_i , і в формулі залишаться лише синусні доданки.

Другий випадок – тригонометрична косинусна (дійсна) форма.

Деяка незручність синус-косинусної форми ряду Фур'є полягає в тому, що для кожного значення індексу підсумовування i (тобто для кожної гармоніки з частотою $\omega_i = i\omega_1$) у формулі фігурують два доданки – синусний і косинусний. Використавши тригонометричне співвідношення $\cos(\alpha - \beta) = \cos\alpha \cos\beta + \sin\alpha \sin\beta$, можна i -у гармоніку представити у вигляді

$$\begin{aligned} & a_i \cos i\omega_1 t + b_i \sin i\omega_1 t = \\ & = \sqrt{a_i^2 + b_i^2} \left(\frac{a_i}{\sqrt{a_i^2 + b_i^2}} \cos i\omega_1 t + \frac{b_i}{\sqrt{a_i^2 + b_i^2}} \sin i\omega_1 t \right) \\ & = A_i \cos(i\omega_1 t + \psi_i), \end{aligned}$$

де

$A_i = \sqrt{a_i^2 + b_i^2}$ – амплітуда і $\operatorname{tg} \psi_i = -\frac{b_i}{a_i}$. При застосуванні

формули $\psi_i = -\operatorname{arctg} \left(\frac{b_i}{a_i} \right)$ для знаходження початкової фази

гармоніки слід враховувати знаки коефіцієнтів a_i, b_i :

$$\psi_i = \begin{cases} -\operatorname{arctg}(b_i/a_i), & \text{якщо } a_i > 0; \\ \pi + \operatorname{arctg}(b_i/a_i), & \text{якщо } b_i > 0 \wedge a_i < 0; \\ -\pi + \operatorname{arctg}(b_i/a_i), & \text{якщо } b_i < 0 \wedge a_i < 0. \end{cases}$$

Це дозволяє трансформувати ряд (2.5) в іншу еквівалентну форму

$$s(t) = \frac{a_0}{2} + \sum_{i=1}^{\infty} A_i \cos(i\omega_1 t + \psi_i). \quad (2.7)$$

Ця форма застосовується для спектрального представлення періодичного сигналу у вигляді двох залежностей амплітуд A_i і фаз ψ_i гармонік від частоти $\omega_i = i\omega_1$.

Якщо $s(t)$ є парною функцією, фази ψ_i можуть набувати тільки значень 0 і π (оскільки $b_i = 0$), а якщо $s(t)$ – функція непарна, то можливі значення для фази – $\pm\pi/2$ ($a_i = 0$).

Третій випадок – комплексно-експоненційна форма.

Ця форма представлення ряду Фур'є виходить з представлення косинуса у вигляді напівсуми комплексних експонент $\cos x = (e^{jx} + e^{-jx})/2$, $\sin x = (e^{jx} - e^{-jx})/2j$ (таке представлення витікає з формули Ейлера $e^{jx} = \cos x + j\sin x$).

Застосувавши це перетворення до косинусної форми ряду Фур'є, отримаємо суми комплексних експонент з додатними і від'ємними показниками степені

$$s(t) = \frac{a_0}{2} + \sum_{i=1}^{\infty} \frac{A_i}{2} [\exp(ji\omega_1 t + j\psi_i) + \exp(-ji\omega_1 t - j\psi_i)].$$

Далі будемо розглядати експоненти зі знаком «мінус» в показнику як члени ряду з від'ємними номерами. Постійний доданок $a_0/2$ при цьому стане членом ряду з нульовим номером. В результаті матимемо комплексну форму запису ряду Фур'є

$$s(t) = \sum_{i=-\infty}^{\infty} \dot{C}_i \exp(ji\omega_1 t). \quad (2.8)$$

Комплексні коефіцієнти ряду (2.8) пов'язані з амплітудами A_i і фазами ψ_i , які є в дійсній формі запису ряду Фур'є, наступними співвідношеннями:

$$\dot{C}_i = \frac{1}{2} A_i e^{j\psi_i}; \quad A_i = 2|\dot{C}_i|; \quad \psi_i = \arg(\dot{C}_i).$$

Нескладно виглядають і формули зв'язку \dot{C}_i з коефіцієнтами a_i і b_i синус-косинусної форми ряду Фур'є, оскільки $\psi_i = -\arctg\left(\frac{b_i}{a_i}\right)$ та $A_i = \sqrt{a_i^2 + b_i^2}$:

$$\dot{C}_i = \frac{1}{2}(a_i - jb_i); \quad a_i = 2\operatorname{Re}(\dot{C}_i); \quad b_i = 2\operatorname{Im}(\dot{C}_i).$$

Звідси отримаємо формулу для безпосереднього розрахунку коефіцієнтів \dot{C}_i ряду Фур'є в комплексній формі

$$\dot{C}_i = \frac{2}{T} \int_{-T/2}^{T/2} s(t) \frac{(\cos i\omega_1 t - j \sin i\omega_1 t)}{2} dt = \frac{1}{T} \int_{-T/2}^{T/2} s(t) e^{-ji\omega_1 t} dt, \quad (2.9)$$

оскільки $e^{-jx} = \cos x - j \sin x$.

Система ортонормованих базисних функцій в цьому випадку має вигляд $\{\varphi_i\} = \exp(ji\omega_1 t) / \sqrt{T}$, $i = 0, \pm 1, \pm 2, \dots$, що можна безпосередньо перевірити

$$(\varphi_i, \varphi_k) = \frac{1}{T} \int_{-T/2}^{T/2} e^{ji\omega_1 t} e^{-jk\omega_1 t} dt = \begin{cases} 1, & \text{якщо } i = k, \\ 0, & \text{якщо } i \neq k. \end{cases}$$

Якщо $s(t)$ є парною функцією, коефіцієнти ряду \dot{C}_i будуть дійсними, а якщо $s(t)$ – функція непарна, коефіцієнти ряду виявляться чисто уявними.

Зазначимо також, що на кожному відкритому проміжку, на якому функція $s(t)$ має розрив в момент часу t_0 , ряд Фур'є сходиться до напівсуми правої і лівої границь

$$s'(t_0) = \frac{1}{2} \left(\lim_{t \rightarrow t_0 - 0} s(t) + \lim_{t \rightarrow t_0 + 0} s(t) \right),$$

де $s(t)$ – початковий сигнал, $s'(t)$ – сума ряду Фур'є для нього.

Сукупність амплітуд гармонік ряду Фур'є часто називають амплітудним спектром, а сукупність їх фаз – фазовим спектром. Ці поняття не слід плутати з амплітудною і фазочастотною характеристиками, які відносяться не до сигналів, а до ланцюгів.

Якщо аналізований сигнал $s(t)$ є дійсним, то його амплітудний і фазовий спектри мають симетрію:

$$|\dot{C}_{-i}| = |\dot{C}_i^*|; \quad \Psi_{-i} = -\Psi_i; \quad \dot{C}_{-i} = \dot{C}_i^*.$$

Оскільки середня потужність періодичного сигналу

$$P_{\text{ср}} = \frac{1}{T} \int_{-T/2}^{T/2} |s(t)|^2 dt,$$

отримаємо варіанти тотожності Парсеваля для трьох форм ряду

$$\begin{aligned} P_{\text{ср}} &= \frac{1}{T} \int_{-T/2}^{T/2} s(t) s^*(t) dt = \frac{1}{T} \int_{-T/2}^{T/2} s(t) \sum_{i=-\infty}^{\infty} \dot{C}_i^* e^{-ji\omega_1 t} dt = \\ &= \sum_{i=-\infty}^{\infty} \dot{C}_i^* \frac{1}{T} \int_{-T/2}^{T/2} s(t) e^{-ji\omega_1 t} dt = \sum_{i=-\infty}^{\infty} |\dot{C}_i|^2 = \\ &= \frac{1}{4} a_0^2 + \frac{1}{2} \sum_{i=1}^{\infty} A_i^2 = \frac{1}{4} a_0^2 + \frac{1}{2} \sum_{i=1}^{\infty} (a_i^2 + b_i^2). \end{aligned} \quad (2.10)$$

Таким чином, середня потужність періодичного сигналу може бути обчислена як сума потужностей його сталої та гармонічних складових.

Узагальнимо властивості розкладання в ряд Фур'є періодичних сигналів на прикладі комплексно-експоненційого ряду Фур'є [9]:

1. Спектр періодичного сигналу повністю визначають дві характеристики – амплітудна і фазова, тобто, модулі та аргументи комплексних коефіцієнтів ряду Фур'є.

2. Спектр періодичного сигналу *дискретний або лінійчастий*. Він складається з окремих ліній (точок) на відповідних частотах $\dots, -2\omega_1, -\omega_1, 0, \omega_1, 2\omega_1, \dots$.

3. Спектри періодичних сигналів *еквідистантні*, тобто сусідні спектральні лінії знаходяться одна від одної на однаковій відстані $\omega_1 = 2\pi/T$. Частота ω_1 називається частотою першої гармоніки.

4. Коефіцієнти ряду Фур'є, а отже й спектри *симетричні* відносно частоти $\omega = 0$. Оскільки

$$\begin{aligned} \dot{C}_i &= \frac{1}{T} \int_{-T/2}^{T/2} s(t) e^{-ji\omega_1 t} dt = \frac{1}{T} \int_{-T/2}^{T/2} s(t) \cos i\omega_1 t dt - \\ &- j \frac{1}{T} \int_{-T/2}^{T/2} s(t) \sin i\omega_1 t dt, \end{aligned}$$

тому, якщо сигнал $s(t)$ є дійсною функцією часу, перша складова правої частини коефіцієнта \dot{C}_i є дійсною, а друга складова – уявною, і тоді можна записати

$$\dot{C}_i = \begin{cases} (a_i - jb_i)/2, & i > 0, \\ (a_i + jb_i)/2, & i < 0. \end{cases}$$

З цього виразу виходить, що $\dot{C}_i = \dot{C}_{-i}^*$. Оскільки $\dot{C}_i = \frac{1}{2} A_i e^{j\psi_i}$ отримаємо, що $|\dot{C}_i| = |\dot{C}_{-i}^*|$ та $A_{-i} = A_i, \psi_{-i} = -\psi_i$.

Таким чином, якщо сигнал $s(t)$ є дійсною функцією часу, тоді модулі коефіцієнтів ряду Фур'є мають властивість парної симетрії відносно i , тобто частоти, а аргументи – властивість непарної симетрії відносно i (частоти).

5. Якщо $s(t)$ є дійсною й парною функцією часу, то уявна частина коефіцієнтів ряду Фур'є \dot{C}_i дорівнює нулю, тобто коефіцієнти синус-косинусного ряду Фур'є $b_i = 0$.

З умови $\operatorname{tg} \psi_i = -b_i/a_i$ виходить, що ψ_i може набувати лише двох значень: 0 або $\pm\pi$.

Якщо $s(t)$ є дійсною і непарною функцією часу, то дійсна частина коефіцієнтів ряду Фур'є \dot{C}_i дорівнює нулю, тобто коефіцієнти синус-косинусного ряду Фур'є $a_i = 0$. У цьому випадку ψ_i може набувати лише двох значень: $\pi/2$ або $-\pi/2$.

6. *Теорема Парсеваля для періодичних сигналів.* Середня потужність періодичного сигналу є сумою потужностей сталої складової і потужностей гармонічних складових сигналу.

$$P_{\text{ср}} = \sum_{i=-\infty}^{\infty} |\dot{C}_i|^2 = \frac{1}{4} a_0^2 + \frac{1}{2} \sum_{i=1}^{\infty} A_i^2 = \frac{1}{4} a_0^2 + \frac{1}{2} \sum_{i=1}^{\infty} (a_i^2 + b_i^2). \quad (2.11)$$

2.2. Приклади розкладання сигналів в ряд Фур'є

Застосуємо ряд Фур'є для аналізу конкретних сигналів.

Послідовність прямокутних імпульсів.

Розглянемо послідовність прямокутних імпульсів з амплітудою A , тривалістю τ і періодом слідування (повторення) T . Початок відліку часу приймемо розташованим в середині імпульсу (рис. 2.1)

$$s(t) = \sum_{k=-\infty}^{\infty} A [\sigma(t + \tau/2 - kT) - \sigma(t - \tau/2 - kT)].$$

Цей сигнал є парною функцією, тому для його представлення зручніше використовувати синус-косинусну форму ряду Фур'є – в ній будуть присутні тільки косинусні доданки

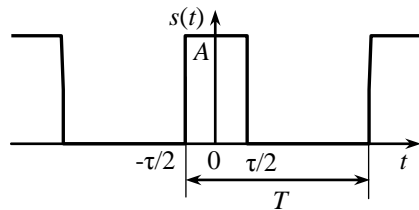


Рисунок 2.1 – Періодична послідовність прямокутних імпульсів

$$a_i = \frac{2}{T} \int_{-\tau/2}^{\tau/2} A \cos \frac{2\pi i}{T} t dt = \frac{2A}{\pi i} \sin \left(\frac{\pi i \tau}{T} \right).$$

В отриманій формулі можна помітити, що тривалість імпульсів і їх період слідування входять в неї не відокремлено, а виключно у вигляді відношення. Цей параметр – відношення періоду до тривалості імпульсів – називають шпаруватістю (щільністю) послідовності імпульсів і означають буквою q : $q = T/\tau$. Введемо цей параметр в отриману формулу для коефіцієнтів ряду Фур'є, а потім приведемо формулу до виду $\sin(x)/x$

$$a_i = \frac{2A}{\pi i} \sin \left(\frac{\pi i \tau}{T} \right) = \frac{2A}{q} \frac{\sin(\pi i/q)}{\pi i/q}.$$

При такій формі запису добре помітно чому дорівнює значення постійної складової ряду: оскільки при $x \rightarrow 0$ $\frac{\sin(x)}{x} \rightarrow 1$, тому $\frac{a_0}{2} = \frac{A}{q} = \frac{A\tau}{T}$.

У зарубіжній літературі замість шпаруватості використовується обернена до неї величина τ/T – коефіцієнт заповнення (англ. duty cycle).

Тепер можна записати і саме представлення періодичної послідовності прямокутних імпульсів у вигляді ряду Фур'є

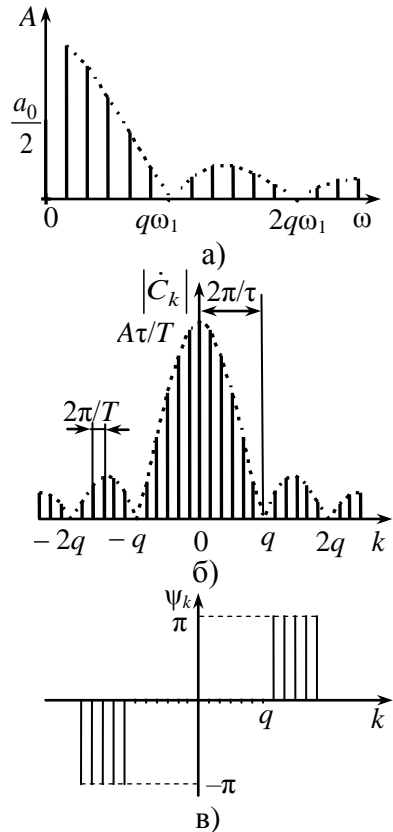


Рисунок 2.2 – Спектр періодичної послідовності прямокутних імпульсів: амплітудний а) дійсна форма і б) комплексна форма; в) фазовий спектр

$$s(t) = \frac{A}{q} + \sum_{i=1}^{\infty} \frac{2A}{q} \frac{\sin\left(\frac{\pi i}{q}\right)}{\frac{\pi i}{q}} \cos \frac{2\pi i}{T} t = \frac{A}{q} + \sum_{i=1}^{\infty} \frac{2A}{\pi i} \sin\left(\frac{\pi i}{q}\right) \cos \frac{2\pi i}{T} t.$$

Амплітуди гармонічних доданків ряду залежать від номера гармоніки по закону $\sin(x)/x$ (рис. 2.2). Графік функції $\sin(x)/x$ має пелюстковий характер. Говорячи про ширину цих пелюсток, слід підкреслити, що для графіків дискретних спектрів періодичних сигналів можливі два варіанти градування горизонтальної осі – в номерах гармонік і в частотах. На рис. 2.2, а градування осі абсцис виконане за частотою, а на рис. 2.2, б (комплексне представлення) – відповідає номерам гармонік, частотні ж параметри спектру нанесені на графік за допомогою розмірних ліній.

Отже, ширина пелюсток, виміряна у кількості гармонік, дорівнює шпаруватості послідовності. При $i = q$

маємо $\sin\left(\frac{\pi i}{q}\right) = 0$. Звідси витікає важлива властивість спектру

послідовності прямокутних імпульсів – в ньому відсутні (мають нульові амплітуди) гармоніки з номерами, кратними шпаруватості.

Відстань по частоті між сусідніми гармоніками дорівнює частоті слідування (повторення) імпульсів – $2\pi/T$. Ширина пелюсток спектру, виміряна в одиницях частоти, складає $2\pi/\tau$, тобто обернено пропорційна тривалості імпульсів. Це, як ми побачимо далі, прояв загального закону (для вузькосмугових, немодульованих сигналів) – чим коротше сигнал, тим ширше його спектр.

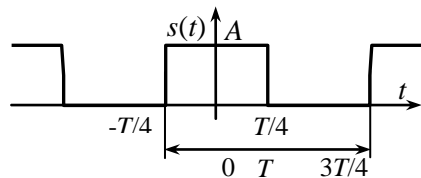


Рисунок 2.3 – Меандр

Меандр є поширеним тестовим сигналом – це послідовність прямокутних імпульсів зі шпаруватістю, що дорівнює двом, коли тривалості імпульсів і проміжків між ними є рівними (рис. 2.3).

У цьому випадку $q = 2$ і коефіцієнти ряду

$$a_i = \frac{2A}{\pi i} \sin\left(\frac{\pi i}{2}\right) = A \frac{\sin\left(\frac{\pi i}{2}\right)}{\frac{\pi i}{2}} = \begin{cases} A, & i = 0, \\ 0, & i = 2m, m \neq 0, \\ \frac{2A}{\pi i}, & i = 4m + 1, \\ -\frac{2A}{\pi i}, & i = 4m - 1. \end{cases}$$

Тут m – довільне ціле число.

Таким чином, в спектрі меандру присутні тільки непарні гармоніки. Крім цього, для представленого на рис. 2.3 випадку функція є парною. Тому згідно з розглянутою вище властивістю 5 в розкладанні такого меандру присутні лише косинусні складові. Представлення меандру у вигляді ряду Фур'є з урахуванням вказаного має вид

$$s(t) = \frac{A}{2} + \frac{2A}{\pi} \left[\cos\left(\frac{2\pi}{T}t\right) - \frac{1}{3} \cos\left(3\frac{2\pi}{T}t\right) + \frac{1}{5} \cos\left(5\frac{2\pi}{T}t\right) - \dots \right].$$

Гармоніки, з яких складається меандр, мають амплітуди, обернено пропорційні до номерів гармонік, і знаки, що чергуються.

Якщо меандр має різнополярні півперіоди, представлення періодичної послідовності прямокутних імпульсів типу меандр у вигляді ряду Фур'є можна виконати як для непарної функції (рис. 2.4, а), так і для парної (рис. 2.4, б) [18].

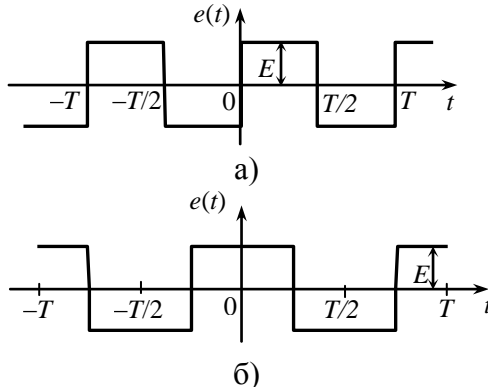


Рисунок 2.4 – Меандр з різнополярними імпульсами

У випадку непарної функції (рис. 2.4, а)

$$e(t) = \sum_{n=1,3,5,\dots}^{\infty} b_n \cos\left(n \frac{2\pi}{T} t - \frac{\pi}{2}\right) =$$

$$= \frac{4E}{\pi} \left[\sin\left(\frac{2\pi}{T} t\right) + \frac{1}{3} \sin\left(3 \frac{2\pi}{T} t\right) + \frac{1}{5} \sin\left(5 \frac{2\pi}{T} t\right) + \dots \right].$$

У разі парної функції (рис. 2.4, б)

$$e(t) = \frac{4E}{\pi} \left[\cos\left(\frac{2\pi}{T} t\right) - \frac{1}{3} \cos\left(3 \frac{2\pi}{T} t\right) + \frac{1}{5} \cos\left(5 \frac{2\pi}{T} t\right) - \dots \right].$$

На відміну від попереднього випадку відсутня постійна складова, що пояснюється змінністю полярності сигналу на половині періоду, а цей доданок, як вже зазначалося, є

середнім значенням сигналу на періоді: $\frac{a_0}{2} = \frac{1}{T} \int_{-T/2}^{T/2} e(t) dt$. Як і

раніше, сума ряду включає лише непарні гармоніки (рис. 2.5). Амплітуди цих гармонік обернено пропорційні їх номеру.

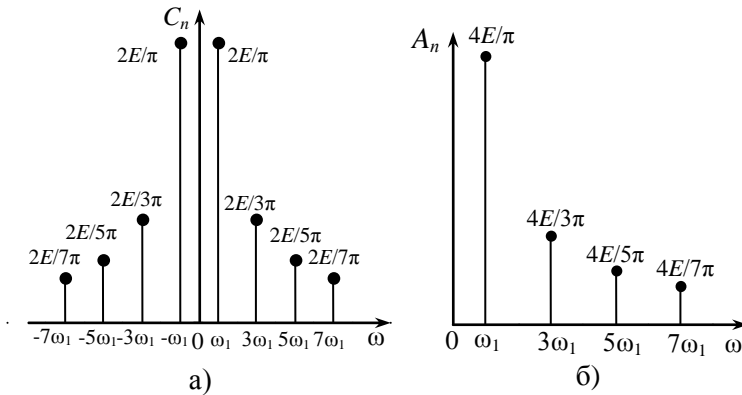


Рисунок 2.5 – Коефіцієнти ряду Фур'є меандру з різнополярними імпульсами (рис. 2.4): а) комплексного; б) дійсного

Взагалі, послідовність прямокутних імпульсів погано підходить для представлення рядом Фур'є – вона містить стрибки, а сума будь-якого числа гармонік з будь-якими амплітудами завжди є неперервною функцією. Тому поведінка ряду Фур'є в околицях розривів представляє особливий інтерес.

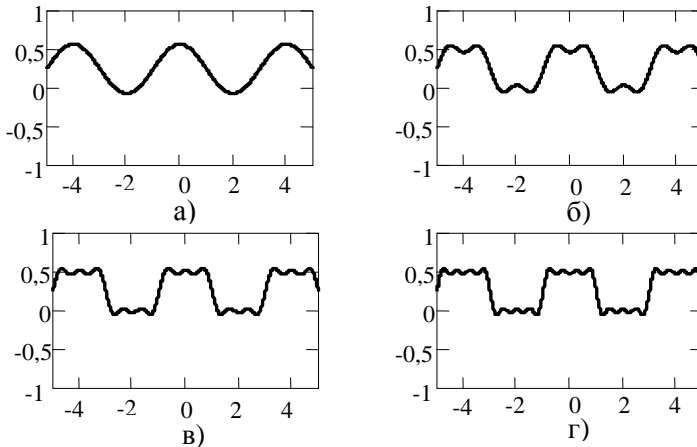


Рисунок 2.6 – Проміжні стадії підсумовування членів ряду Фур'є меандру: а) одного; б) трьох; в) п'яти; г) семи

На графіках рис. 2.6 добре видно, що в околиці точки розриву підсумовування ряду Фур'є дає похилу ділянку, причому крутизна нахилу зростає із зростанням кількості підсумовуваних гармонік. У самій точці розриву ряд Фур'є сходиться до напівсуми правої і лівої границь.

На ділянках, що примикають до розриву, сума ряду Фур'є дає помітні пульсації, причому на графіках рис. 2.6 помітно, що амплітуда цих пульсацій не зменшується із зростанням числа підсумовуваних гармонік – пульсації лише стискаються по горизонталі, наближаючись до точки розриву. Це явище, яке властиве рядам Фур'є для будь-яких сигналів з розривами першого роду (стрибками), має назву ефекту Гіббса. Амплітуда першого (найбільшого) викиду складає приблизно 9 % від величини стрибка [11,18].

Розглянемо деякі інші приклади розкладання періодичних сигналів в ряд Фур'є [11].

Пилкоподібний, який в межах періоду T описується лінійною функцією (рис. 2.7),

$$s(t) = \sum_{i=-\infty}^{\infty} \frac{2A}{T} (t-iT) \left[\sigma\left(t-iT + \frac{T}{2}\right) - \sigma\left(t-iT - \frac{T}{2}\right) \right],$$

де A – амплітуда сигналу,

$\sigma(t)$ – функція Хевісайда.

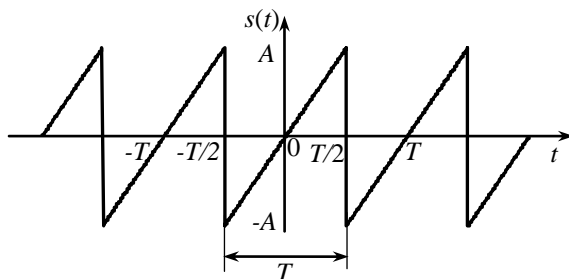


Рисунок 2.7 – Двополярний пилкоподібний сигнал

Ряд Фур'є цього сигналу виглядає наступним чином:

$$s(t) = \frac{4A}{\pi} \left[\begin{array}{l} \sin\left(\frac{2\pi}{T}t\right) - \frac{1}{2}\sin\left(2\frac{2\pi}{T}t\right) + \frac{1}{3}\sin\left(3\frac{2\pi}{T}t\right) - \\ \frac{1}{4}\sin\left(4\frac{2\pi}{T}t\right) + \frac{1}{5}\sin\left(5\frac{2\pi}{T}t\right) + \dots \end{array} \right],$$

де коефіцієнти ряду $b_i = \frac{2}{T} \int_{-T/2}^{T/2} \frac{2A}{T} t \sin i\omega_1 t dt = -\frac{2A}{\pi i} (-1)^i$.

У розглянутих вище спектрів прямокутного і пилкоподібного періодичних сигналів є одна загальна риса – амплітуди гармонік із зростанням їх номерів убувають пропорційно до i . У наступного сигналу швидкість загасання спектру буде іншою, а чому, ми обговоримо після розрахунку коефіцієнтів ряду Фур'є для нього.

Послідовність трикутних імпульсів (рис. 2.8).

Отримаємо розкладання в ряд Фур'є. Строго кажучи, імпульси в попередньому сигналі теж були трикутними, але в даному випадку вони матимуть не пилкоподібну, а симетричну форму (рис. 2.8):

$$s(t) = A \left(1 - 4 \frac{|t - iT|}{T} \right), \left(i - \frac{1}{2} \right) T < t \leq \left(i + \frac{1}{2} \right) T.$$

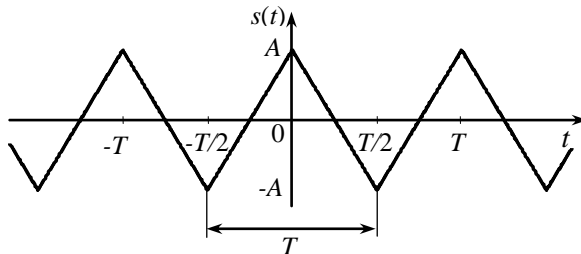


Рисунок 2.8 – Послідовність трикутних імпульсів

Обчислимо коефіцієнти ряду Фур'є (сигнал є парною функцією, тому в синус-косинусній формі ряду Фур'є (1.6) будуть присутні тільки косинусні доданки):

$$a_k = \frac{2}{T} \int_{-T/2}^{T/2} A \left(1 - 4 \frac{|t|}{T} \right) \cos \left(\frac{2\pi k}{T} t \right) dt = \frac{4A}{(\pi k)^2} [1 - (-1)^k] =$$

$$= \begin{cases} 0, & k = 2m, \\ \frac{8A}{(\pi k)^2}, & k = 2m + 1. \end{cases}$$

Як і у разі меандру, присутні тільки парні гармоніки. Ряд Фур'є має вигляд

$$s(t) = \frac{8A}{\pi^2} \left[\cos \left(\frac{2\pi}{T} t \right) + \frac{1}{3^2} \cos \left(3 \frac{2\pi}{T} t \right) + \frac{1}{5^2} \sin \left(5 \frac{2\pi}{T} t \right) + \dots \right].$$

Таким чином, для трикутного періодичного сигналу амплітуди гармонік убувають пропорційно квадрату номерів гармонік k , на відміну від послідовностей прямокутних і пілкоподібних імпульсів, де діяла залежність $1/k$. Це прояв загального правила: швидкість убування спектру залежить від міри гладкості сигналу [11]. Граничним випадком є гармонічний сигнал, який без втрати неперервності диференціюється нескінченну кількість разів. Швидкість убування спектру цього сигналу нескінченна, оскільки ряд Фур'є для гармонічного сигналу містить тільки одну гармоніку.

2.3. Інтегральне перетворення Фур'є

Розглянутий в розділі 2.1 метод рядів Фур'є може бути узагальнений і на випадок неперіодичних сигналів. Для отримання спектрального представлення неперіодичного сигналу застосуємо комплексну форму ряду Фур'є (2.8).

Нехай аналізований сигнал $s(t)$ є імпульсним сигналом обмеженої тривалості. Якщо уявно доповнити цей сигнал

послідовністю таких самих сигналів з періодом T , тоді такий сигнал може бути представлений рядом Фур'є

$$s_{\text{пер}}(t) = \sum_{i=-\infty}^{\infty} \dot{C}_i \exp(ji\omega_1 t) \quad (2.12)$$

з коефіцієнтами (2.9): $\dot{C}_i = \frac{1}{T} \int_{-T/2}^{T/2} s(t) e^{-ji\omega_1 t} dt$.

Для повернення до одиночного імпульсу слід спрямувати період T цієї послідовності до нескінченності при цьому матимемо:

1. Частоти сусідніх гармонік $i\omega_1$ і $(i-1)\omega_1$ будуть скільки завгодно близькими, тому у формулах (2.9) і (2.12) дискретну змінну $i\omega_1$ можна змінити на неперервну ω – поточну частоту, а значення частоти $\omega_1 \rightarrow d\omega$.

2. Амплітудні коефіцієнти \dot{C}_i стануть нескінченно малими, оскільки у знаменнику (2.9) є нескінченно велике значення періоду T .

Запишемо ряд (2.12) для цього випадку

$$\begin{aligned} s(t) &= \lim_{T \rightarrow \infty} s_{\text{пер}}(t) = \lim_{T \rightarrow \infty} \sum_{i=-\infty}^{\infty} \left[\int_{-T/2}^{T/2} s(t) e^{-ji\omega_1 t} dt \right] \exp(ji\omega_1 t) \frac{\omega_1}{2\pi} = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(j\omega t) \left[\int_{-\infty}^{\infty} s(t) e^{-j\omega t} dt \right] d\omega, \end{aligned}$$

де ураховано, що $\frac{1}{T} = \frac{\omega_1}{2\pi}$.

Таким чином ми отримали подвійний інтеграл. Внутрішній інтеграл за визначенням є спектральною густиною сигналу $s(t)$, яка обчислюється за допомогою перетворення Фур'є:

$$\dot{S}(\omega) = \int_{-\infty}^{\infty} s(t) e^{-j\omega t} dt. \quad (2.13)$$

Також з цього граничного представлення сигналу $s(t)$ маємо й *обернене перетворення Фур'є* для відновлення цього сигналу за його спектральною густиною

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}(\omega) e^{j\omega t} d\omega. \quad (2.14)$$

Якщо використовувати не кругову частоту ω , а циклічну частоту f , формули прямого і оберненого перетворення Фур'є стають більш симетричними, відрізняючись лише знаком у показнику експоненти:

$$\dot{S}(f) = \int_{-\infty}^{\infty} s(t) e^{-j2\pi ft} dt; \quad s(t) = \int_{-\infty}^{\infty} \dot{S}(f) e^{j2\pi ft} df.$$

Як вже вказувалось в розділі 1.1, *спектральну густину* можна представити через її модуль і аргумент або сумою дійсної й уявної функцій

$$\dot{S}(\omega) = |\dot{S}(\omega)| e^{j \arg \dot{S}(\omega)} = \operatorname{Re}[\dot{S}(\omega)] + j \operatorname{Im}[\dot{S}(\omega)] = A(\omega) - jB(\omega).$$

Модуль і аргумент спектральної густини визначають відповідно за формулами:

$$|\dot{S}(\omega)| = \sqrt{A^2(\omega) + B^2(\omega)}; \quad \theta(\omega) = \arg \dot{S}(\omega) = -\arctg \frac{B(\omega)}{A(\omega)},$$

$$\text{де } A(\omega) = \int_{-\infty}^{\infty} s(t) \cos \omega t dt \text{ та } B(\omega) = \int_{-\infty}^{\infty} s(t) \sin \omega t dt.$$

Умови існування спектральної густини сигналу (2.13).

Для існування спектральної густини сигналу $s(t)$, $t \in (-\infty, \infty)$, як це доведено в математиці, необхідне виконання двох умов:

1) на будь-якому кінцевому інтервалі функція $s(t)$ повинна задовольняти умовам Діріхле;

2) функція $s(t)$ абсолютно інтегрована на усій числовій осі, тобто

$$\int_{-\infty}^{\infty} |s(t)| dt = Q < \infty.$$

Умови Діріхле є достатніми, хоча й не є необхідними. Цим умовам, наприклад, задовольняють всі сигнали з обмеженою енергією. Друга умова додатково звужує клас допустимих сигналів. Наприклад, гармонічний сигнал, який існує на нескінченній осі часу не є абсолютно інтегрованим. Однак, завдяки застосуванню узагальнених функцій (Хевісайда, Дірака), також вдається обчислювати спектральні густини неінтегрованих, в тому числі періодичних, сигналів, що буде застосовуватись в подальшому.

Якщо аналізований сигнал $s(t)$ – дійсна функція, то відповідна йому спектральна густина (функція) $\dot{S}(\omega)$ є "спряжено-симетричною" відносно нульової частоти. Це означає, що значення спектральної густини (функції) на частотах ω і $-\omega$ є комплексно-спряженими по відношенню одне до одного $\dot{S}(-\omega) = \dot{S}^*(\omega)$.

Якщо $s(t)$ – парна функція, тоді, як і у разі ряду Фур'є, спектр буде чисто дійсним (і, отже, буде парною функцією). Якщо, навпаки, $s(t)$ – функція непарна, тоді спектральна функція $\dot{S}(\omega)$ буде чисто уявною (і непарною).

Модуль спектральної функції часто називають амплітудним спектром, а її аргумент – фазовим спектром. Легко показати, що для дійсного сигналу амплітудний спектр є парною функцією частоти, а фазовий – непарною:

$$|\dot{S}(-\omega)| = |\dot{S}(\omega)|, \theta(-\omega) = -\theta(\omega).$$

Отже, перетворення Фур'є (2.11) ставить у відповідність сигналу, заданому в часі, його спектральну функцію (густину). При цьому здійснюється перехід з часової області в частотну. Перетворення Фур'є є взаємно-однозначним, тому представлення сигналу в частотній області (спектральна функція) містить рівно стільки ж інформації, скільки й вихідний сигнал заданий в часовій області.

Розглянемо приклади розрахунку перетворення Фур'є для деяких сигналів, що часто зустрічаються при рішенні різних задач.

Прямокутний імпульс

Почнемо з прямокутного імпульсу центрованого відносно початку відліку часу (рис. 2.9):

$$s(t) = A[\sigma(t + \tau/2) - \sigma(t - \tau/2)].$$

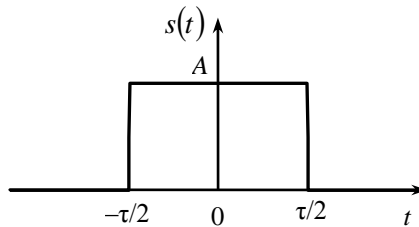


Рисунок 2.9 – Прямокутний імпульс

Обчислимо його спектральну функцію

$$\begin{aligned} \dot{S}(\omega) &= \int_{-\infty}^{\infty} s(t)e^{-j\omega t} dt = A \int_{-\tau/2}^{\tau/2} e^{-j\omega t} dt = A \int_{-\tau/2}^{\tau/2} (\cos\omega t - j\sin\omega t) dt = \\ &= 2A \int_0^{\tau/2} \cos\omega t dt = \frac{2A}{\omega} \sin \frac{\omega\tau}{2}. \end{aligned}$$

Отриманий результат можна представити у вигляді

$$\dot{S}(\omega) = A\tau \frac{\sin(\omega\tau/2)}{\omega\tau/2}. \quad (2.15)$$

Як видно, спектр сигналу є функцією виду $\sin(x)/x$ (рис. 2.10).

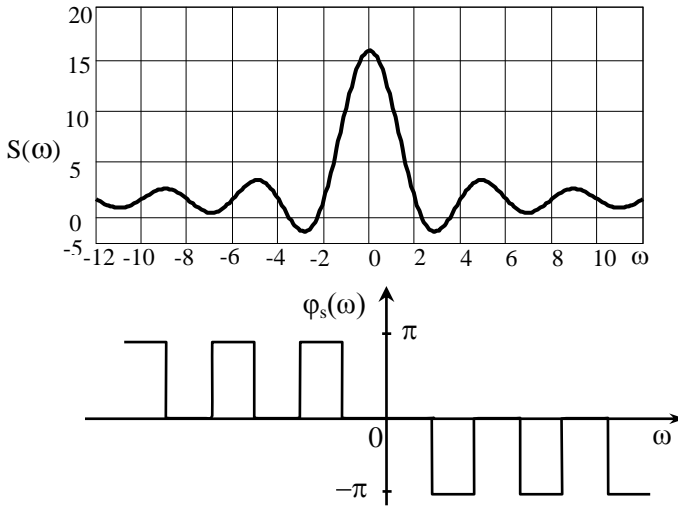


Рисунок 2.10 – Спектральна густина і фазовий спектр прямокутного відеоімпульсу з параметрами: $\tau = \pi$ с; $A = 5$ В

Амплітудний спектр має пелюстковий характер, ширина пелюсток складає $2\pi/\tau$, тобто обернено пропорційна тривалості імпульсу. Значення спектральної густини на нульовій частоті дорівнює площі імпульсу – $\dot{S}(0) = A\tau$ (на графіку $A\tau \approx 15,7$). Спектральна функція є дійсною, тому фазовий спектр набуває лише два значення – 0 або π залежно від знаку функції $\sin(x)/x$. Значення фази π і $-\pi$ однакові, різні знаки для фазового спектру при $\omega > 0$ і $\omega < 0$ використані з метою представлення у вигляді непарної функції згідно з його загальними властивостями.

Розглянемо випадок, коли імпульс (рис. 2.1) зсувається в часі. Нехай імпульс починається в нульовий момент часу (рис. 2.11).

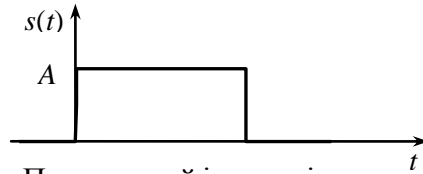


Рисунок 2.11 – Прямокутний імпульс із затримкою в часі на $\tau/2$

Обчислимо перетворення Фур'є і побудуємо графіки амплітудного і фазового спектрів цього сигналу, урахувавши, що

$$\sin x = \frac{1}{2j} \left(e^{jx} - e^{-jx} \right),$$

$$\begin{aligned} \dot{S}(\omega) &= A \int_0^{\tau} e^{-j\omega t} dt = \frac{A}{j\omega} \int_{-j\omega\tau}^0 e^x dx = \frac{A}{j\omega} \left(e^0 - e^{-j\omega\tau} \right) = \\ &= \frac{A\tau}{j\omega\tau} \left(e^{j\frac{\omega\tau}{2}} - e^{-j\frac{\omega\tau}{2}} \right) e^{-j\frac{\omega\tau}{2}} = A\tau \frac{\sin(\omega\tau/2)}{\omega\tau/2} e^{-j\frac{\omega\tau}{2}}. \end{aligned} \quad (2.16)$$

З формул (2.15) і (2.16) і графіків (рис. 2.11 і 2.12) видно, що після затримки імпульсу в часі його амплітудний спектр залишився тим самим, а фазовий отримав лінійно залежний від частоти зсув.

Цей приклад демонструє прояв властивості перетворення Фур'є, що стосується зміни спектру при зсуві сигналу в часі. Ця властивість буде сформульована в загальному вигляді при розгляді властивостей перетворення Фур'є.

Строго кажучи, спектр цього сигналу тягнеться до нескінченності, лише поступово згасаючи. Тому вводять поняття ефективної ширини спектру. Як видно з графіків, спектр має пелюстковий характер і ширина головної пелюстки дорівнює $2\pi/\tau$. Саме її приймають за ефективну ширину спектру, яка є обернено пропорційною тривалості імпульсу. Це загальне співвідношення: чим коротше сигнал, тим ширше його спектр. Добуток же ефективних значень тривалості сигналу і ширини його спектру B (база сигналу) залишається рівним деякій константі, залежній від конкретного способу визначення цих параметрів. У розглянутому прикладі B , очевидно, дорівнює 2π . Взагалі, для сигналів простої

форми (що не мають складної внутрішньої імпульсної структури) величина бази незалежно від способу визначення ефективних значень тривалості і ширини спектру складає декілька одиниць.

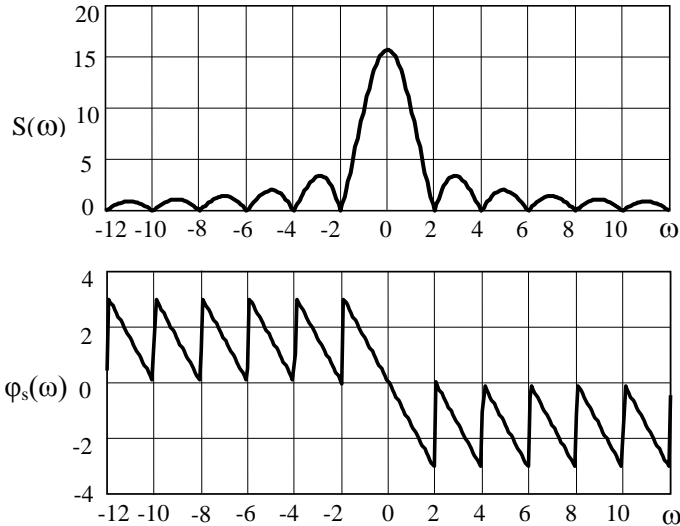


Рисунок 2.12 – Амплітудний і фазовий спектри прямокутного імпульсу із затримкою

Тривалість сигналу і ширина його спектру підкоряються співвідношенню невизначеності, тобто добуток цих параметрів (база сигналу) не може бути меншим за одиницю. Обмежень максимального значення бази сигналу не існує. Звідси витікає, що можна сформувати сигнал великої тривалості, що одночасно має і широкий спектр (такі сигнали називають широкосмуговими, або складними, або сигналами з великою базою). А ось короткий сигнал з вузьким спектром, згідно із співвідношенням невизначеності, існувати не може.

Властивості перетворення Фур'є.

Під властивостями перетворення Фур'є розуміється взаємна відповідність трансформацій сигналів і їх спектрів. Добре знання властивостей перетворення Фур'є дозволяє

передбачати приблизний (а іноді і точний) вид спектру аналізованого сигналу і таким чином контролювати правдоподібність результату при обчисленнях на комп'ютері.

У цьому розділі розглядатимемо два абстрактні сигнали, $f(t)$ і $g(t)$, і вважати, що їх спектральні функції рівні $\dot{F}(\omega)$ і $\dot{G}(\omega)$.

1. Лінійність

Перетворення Фур'є є лінійним інтегральним перетворенням. Сенса властивості лінійності можна сформулювати так: спектр суми дорівнює сумі спектрів. Говорячи математичною мовою, лінійна комбінація сигналів має спектр у вигляді такої ж (з тими ж коефіцієнтами) лінійної комбінації їх спектральних функцій (густин):

$$\text{якщо } s(t) = \alpha f(t) + \beta g(t), \text{ тоді } \dot{S}(\omega) = \alpha \dot{F}(\omega) + \beta \dot{G}(\omega).$$

2. Затримка

Розглянемо, як впливає на спектральну функцію затримка сигналу в часі. Нехай τ – час затримки, тоді сигнал $s(t) = f(t - \tau)$. Цього спектральна функція зміниться таким чином:

$$\begin{aligned} \dot{S}(\omega) &= \int_{-\infty}^{\infty} f(t - \tau) e^{-j\omega t} dt = \\ &= e^{-j\omega\tau} \int_{-\infty}^{\infty} f(t - \tau) e^{-j\omega(t - \tau)} d(t - \tau) = \dot{F}(\omega) e^{-j\omega\tau}. \end{aligned}$$

Результат показує, що спектр незатриманого сигналу $f(t)$ виявився помноженим на комплексну експоненту виду $e^{-j\omega\tau}$. Таким чином, амплітудний спектр сигналу не змінюється (оскільки модуль такої комплексної експоненти дорівнює 1; до того ж здоровий глузд підказує, що співвідношення між амплітудами спектральних складових із за зсуву сигналу в часі змінитися не повинно). Фазовий спектр

набуває додаткового доданку $-\omega t$, який лінійно залежить від частоти.

Зазначимо, що таке перетворення сигналу може бути виконане лінійною системою з постійними параметрами. Такі системи будуть розглядатись в главі 6.

3. Зміна масштабу осі часу

Розглядаючи конкретні приклади, ми вже вказали загальне на практиці правило: чим коротше сигнал, тим ширше його спектр. Тепер поглянемо та це правило з боку строгих теоретичних позицій. Якщо змінити тривалість сигналу $f(t)$, зберігаючи його форму, то новий сигнал $s(t)$ слід записати як $s(t) = f(at)$.

При $|a| > 1$ сигнал стискується, при $|a| < 1$ – розтягується, якщо $a < 0$, додатково відбувається дзеркальне відображення сигналу відносно вертикальної осі. Подивимось, як таке перетворення позначається на спектрі:

$$\dot{S}(\omega) = \int_{-\infty}^{\infty} f(at) e^{-j\omega t} dt = \frac{1}{a} \int_{-\infty}^{\infty} f(x) e^{-j\frac{\omega}{a}x} dx = \frac{1}{a} \dot{F}\left(\frac{\omega}{a}\right).$$

Отже, зміна тривалості сигналу веде до зміни ширини спектру в протилежний бік (аргумент t на a множиться, а ω – ділиться) у сполученні із збільшенням (при розтягуванні, $a < 1$) або зменшенням (при стискуванні, $a > 1$) рівня спектральних складових.

Отримана формула справедлива для $a > 0$. При $a < 0$ використана заміна змінної $at = x$ викличе перестановку границь інтегрування і, як наслідок, зміну знаку у результаті:

$$\dot{S}(\omega) = -\frac{1}{a} \dot{F}\left(\frac{\omega}{a}\right), \quad a < 0.$$

Об'єднуючи обидва випадки, отримаємо

$$\dot{S}(\omega) = \frac{1}{|a|} \dot{F}\left(\frac{\omega}{a}\right), \quad a \neq 0.$$

У окремому випадку $a = -1$ формула має вигляд

$$\dot{S}(\omega) = \dot{F}(-\omega) = \dot{F}^*(\omega).$$

Отже, дзеркальне відображення сигналу відносно початку відліку часу призводить до дзеркального відображення спектру відносно нульової частоти. Для дійсного сигналу це відповідає комплексному сполученню спектру.

В даному випадку результат не отримується множенням вихідного спектру на деяку функцію. Це означає, що зміна тривалості сигналу не може бути здійснена лінійною системою з постійними параметрами.

4. Диференціювання сигналу

Розглянемо вплив диференціювання сигналу в часовій області на його спектр. Для цього скористуємось визначенням поняття похідної

$$s(t) = \frac{df}{dt} = \lim_{\varepsilon \rightarrow 0} \frac{f(t + \varepsilon) - f(t)}{\varepsilon}.$$

Застосуємо до цього виразу перетворення Фур'є:

$$\begin{aligned} \dot{S}(\omega) &= \int_{-\infty}^{\infty} \lim_{\varepsilon \rightarrow 0} \frac{f(t + \varepsilon) - f(t)}{\varepsilon} e^{-j\omega t} dt = \lim_{\varepsilon \rightarrow 0} \frac{\dot{F}(\omega) e^{j\omega\varepsilon} - \dot{F}(\omega)}{\varepsilon} = \\ &= \dot{F}(\omega) \lim_{\varepsilon \rightarrow 0} \frac{e^{j\omega\varepsilon} - 1}{\varepsilon} = j\omega \dot{F}(\omega). \end{aligned}$$

Спектр похідної обчислюється множенням спектру вихідного сигналу на $j\omega$. Таким чином, при диференціюванні низькі частоти ослабляються, а високі підсилюються. Фазовий спектр сигналу зсувається на 90° для додатних частот і на -90° для від'ємних. Множник $j\omega$ називають оператором диференціювання сигналу в частотній області.

Теорема дуальності вказує на можливість взаємної заміни змінних частоти і часу, тобто якщо $s(t) \Leftrightarrow \dot{S}(f)$, тоді

$\dot{S}(t) \Leftrightarrow s(-f)$. З її допомогою можна довести теорему диференціювання в частотній області

$$\frac{d}{dt} s(t) \Leftrightarrow j2\pi f \dot{S}(f), \text{ тоді } \frac{d}{dt} \dot{S}(f) \Leftrightarrow -j2\pi t s(t).$$

Звідки випливає теорема множення сигналу на t

$$t \cdot s(t) \Leftrightarrow -\frac{1}{j2\pi} \frac{d}{df} \dot{S}(f).$$

5. Інтегрування сигналу

Інтегрування, як відомо, є операцією, оберненою до диференціювання. Тому, виходячи з результатів, отриманих в попередньому розділі, здавалося б, можна очікувати наступний результат

$$\dot{S}(\omega) = \frac{\dot{F}(\omega)}{j\omega}.$$

Проте ця формула справедлива лише для сигналів, що не містять постійної складової, у яких

$$\dot{F}(0) = \int_{-\infty}^{\infty} f(t) dt = 0.$$

В загальному ж випадку результат повинен містити також доданок у вигляді дельта-функції на нульовій частоті. Множник перед дельта-функцією пропорційний постійній складовій сигналу

$$\dot{S}(\omega) = \frac{\dot{F}(\omega)}{j\omega} + \pi \dot{F}(0) \delta(\omega).$$

Отже, при інтегруванні вихідного сигналу високі частоти ослабляються, а низькі посилюються. Фазовий спектр сигналу зміщується на -90° для додатних частот і на 90° для від'ємних. Множник $1/(j\omega)$ називають оператором інтегрування в частотній області.

6. Спектр згортки сигналів

Згортка сигналів є дуже часто використовуваною в вимірювальній техніці інтегральною операцією. Вона описує, зокрема, проходження сигналу через лінійну систему з постійними параметрами (детальніше це обговорюється в главі 6)

$$s(t) = \int_{-\infty}^{\infty} f(\tau) g(t-\tau) d\tau.$$

Піддамо таку конструкцію перетворенню Фур'є:

$$\begin{aligned} \dot{S}(\omega) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\tau) g(t-\tau) d\tau e^{-j\omega t} dt = \\ &= \int_{-\infty}^{\infty} f(\tau) e^{-j\omega\tau} d\tau \int_{-\infty}^{\infty} g(t-\tau) e^{-j\omega(t-\tau)} d(t-\tau) = \dot{F}(\omega) \dot{G}(\omega). \end{aligned}$$

Отриманий результат дуже важливий, він часто використовується на практиці: *спектр згортки дорівнює добутку спектрів*.

У випадку, коли сигнал є добутком двох інших $s(t) = f(t)g(t)$ можна кожну з цих функцій представити їх оберненими перетвореннями Фур'є від відповідних спектральних густин кожного з сигналів

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{F}(\omega) \exp(j\omega t) d\omega, \\ g(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{G}(\omega) \exp(j\omega t) d\omega \end{aligned}$$

і отримати наступне співвідношення

$$\begin{aligned} \dot{S}(\omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) \left[\int_{-\infty}^{\infty} \dot{G}(x) \exp(jxt) dx \right] e^{-j\omega t} dt = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{G}(x) \left[\int_{-\infty}^{\infty} f(t) e^{-j(\omega-x)t} dt \right] dx. \end{aligned}$$

Інтеграл у квадратних дужках по змінній t є спектральною густиною функції $f(t)$ при частоті $(\omega - x)$, тобто $\dot{F}(\omega - x)$, звідки маємо

$$\dot{S}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{G}(x) \dot{F}(\omega - x) dx.$$

Таким чином, спектр добутку двох функцій часу $f(t)g(t)$ дорівнює (з коефіцієнтом $1/2\pi$) згортці їх спектрів $\dot{F}(\omega)$ і $\dot{G}(\omega)$, де

$$\int_{-\infty}^{\infty} \dot{G}(x) \dot{F}(\omega - x) dx = \dot{G}(\omega) * \dot{F}(\omega)$$

або інакше

$$f(t)g(t) \Leftrightarrow \frac{1}{2\pi} \dot{G}(\omega) * \dot{F}(\omega).$$

Неважко переконатися, що операція згортки комутативна

$$\dot{G}(\omega) * \dot{F}(\omega) = \dot{F}(\omega) * \dot{G}(\omega).$$

Якщо сигнали $f(t)$ і $g(t)$ будуть однаковими, тоді з попереднього рівняння отримаємо рівність

$$\int_{-\infty}^{\infty} f(t)g(t)dt = \int_{-\infty}^{\infty} f^2(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{F}(x) \dot{F}(\omega - x) dx,$$

яка у випадку, коли $\omega = 0$, прийме вигляд

$$\int_{-\infty}^{\infty} f^2(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{F}(x) \dot{F}(-x) dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{F}(\omega) \dot{F}^*(\omega) d\omega.$$

Остаточно, можна записати рівність Парсеваля для неперіодичного сигналу у вигляді

$$E = \int_{-\infty}^{\infty} f^2(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\dot{F}(\omega)|^2 d\omega. \quad (2.17)$$

Ця рівність встановлює зв'язок між енергією сигналу та його спектральною густиною. Залежність $|\dot{F}(\omega)|^2$ при цьому має розмірність енергії, що припадає на 1 Гц, і може трактуватись як спектральна густина енергії сигналу.

7. Множення сигналу на гармонічну функцію

Помножимо вихідний сигнал $f(t)$, спектр якого нам відомий, на гармонічну функцію: $s(t) = f(t)\cos(\omega_0 t + \varphi_0)$ і поглянемо, що станеться із спектром сигналу:

$$\begin{aligned} s(t) &= \int_{-\infty}^{\infty} f(t)\cos(\omega_0 t + \varphi_0)e^{-j\omega t} dt = \\ &= \int_{-\infty}^{\infty} f(t)\frac{e^{j(\omega_0 t + \varphi_0)} + e^{-j(\omega_0 t + \varphi_0)}}{2} e^{-j\omega t} dt = \\ &= \frac{1}{2} \int_{-\infty}^{\infty} f(t)e^{j\varphi_0} e^{-j(\omega - \omega_0)t} dt + \frac{1}{2} \int_{-\infty}^{\infty} f(t)e^{-j\varphi_0} e^{-j(\omega + \omega_0)t} dt = \\ &= \frac{1}{2} e^{j\varphi_0} F(\omega - \omega_0) + \frac{1}{2} e^{-j\varphi_0} F(\omega + \omega_0). \end{aligned}$$

Видно, що спектр розділився на два доданки, які мають у два рази менший рівень (множник 1/2), зміщених на ω_0 праворуч ($\omega - \omega_0$) і ліворуч ($\omega + \omega_0$) по осі частот. Крім того, при кожному доданку є множник, що враховує початкову фазу гармонічного коливання. Практичне застосування цієї властивості розглядатиметься при обговоренні властивостей сигналів з амплітудною модуляцією в главі 7.

8. Зв'язок перетворення Фур'є і коефіцієнтів ряду Фур'є

Нехай $s(t)$ – сигнал кінцевої тривалості, а $\dot{S}(\omega)$ – його спектральна густина. Отримаємо на основі $s(t)$ періодичний

сигнал, узявши період повторення T не менше тривалості сигналу: $s_T(t) = \sum_{k=-\infty}^{\infty} s(t-kT) [\sigma(t+\tau/2-kT) - \sigma(t-\tau/2-kT)]$.

З порівняння формул для розрахунку перетворення Фур'є сигналу $s(t)$

$$\dot{S}(\omega) = \int_{-\infty}^{\infty} s(t)e^{-j\omega t} dt \quad \text{і} \quad \dot{C}_k = \frac{1}{T} \int_{-T/2}^{T/2} s(t)e^{-jk\omega_1 t} dt,$$

для розрахунку коефіцієнтів ряду Фур'є сигналу $s_T(t)$, можна помітити, що ці формули передбачають обчислення одного і того ж інтегралу. Відмінність полягає в тому, що для розрахунку коефіцієнтів ряду Фур'є в підінтегральний вираз підставляються не довільні, а дискретні значення частоти $\omega_k = 2\pi k/T$ і, крім того, результат інтегрування ділиться на період сигналу T . Таким чином, між спектральною функцією $\dot{S}(\omega)$ поодинокого імпульсу і коефіцієнтами \dot{C}_k ряду Фур'є для періодичної послідовності таких імпульсів існує простий зв'язок:

$$\dot{C}_k = \frac{1}{T} \dot{S}\left(\frac{2\pi k}{T}\right).$$

Ця формула справедлива і у тому випадку, якщо період слідування імпульсів менше їх тривалості (тобто якщо сусідні імпульси періодичної послідовності перекриваються).

Далі розглянемо *приклад* Фур'є-аналізу неінтегрованих сигналів.

При введенні поняття перетворення Фур'є були вказані умови його застосовності: виконання умов Діріхле і абсолютна інтегрованість сигналу. Проте у ряді випадків можна застосувати перетворення Фур'є і до сигналів, що цим умовам не задовольняють, і отримати при цьому цілком осмислений і практично корисний результат.

Отже, зараз ми скористаємося перетворенням Фур'є для спектрального аналізу таких сигналів, до яких воно формально непридатне.

Дельта-функція.

Передусім обчислимо перетворення Фур'є для сигналу у вигляді дельта-функції (про її властивості йшла мова в розділі 1.2, і її фільтрувальна властивість (1.1) зараз застосовується):

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

Спектр дельта-функції є константою, тобто є рівномірним в нескінченній смузі частот.

Це цілком узгоджується із загальним співвідношенням між тривалістю сигналу і шириною його спектру: дельта-імпульс має нескінченно малу тривалість, а його спектр нескінченно широкий. З отриманого результату виходить, що дельта-функцію можна записати у вигляді оберненого перетворення Фур'є таким чином

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

Це корисне співвідношення використаємо при аналізі наступного сигналу.

Постійний за часом сигнал (константа).

Оскільки ми вже знаємо, що спектром дельта-функції є константа, завдяки дуальності перетворення Фур'є, можна відразу ж сказати, що спектром постійного сигналу $s(t) = A$ буде δ -функція частоти. Перевіримо це, скориставшись отриманим вище співвідношенням:

$$\dot{S}(\omega) = \int_{-\infty}^{\infty} A e^{-j\omega t} dt = 2\pi A \delta(\omega).$$

Припущення повністю підтвердились. Знову добре простежується обернена пропорційність між тривалістю

сигналу і шириною його спектру: нескінченно протяжний сигнал має нескінченно вузький спектр.

Функція *одиночного стрибка* (див. розділ 1.2) є інтегралом від δ -функції, тому, відповідно до властивостей перетворення Фур'є, отримаємо

$$\dot{S}(\omega) = \int_{-\infty}^{\infty} \sigma(t) e^{-j\omega t} dt = \pi \delta(\omega) + \frac{1}{j\omega}.$$

Оскільки функція $\sigma(t)$ має ненульову (рівну 1) постійну складову, то в повній відповідності з формулою

$$\dot{S}(\omega) = \frac{\dot{F}(\omega)}{j\omega} + \pi \dot{F}(0) \delta(\omega),$$

яка приведена в п. 5 «Властивостей перетворення Фур'є», в спектрі з'являється додатковий доданок у вигляді дельта-функції на нульовій частоті.

Гармонічний сигнал.

Розрахуємо спектр гармонічного сигналу загального вигляду: $s(t) = A \cos(\omega_0 t + \varphi_0)$.

Для розрахунку спектральної функції представимо косинус у вигляді напівсуми комплексних експонент і

скористаємось формулою $\delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} d\omega$:

$$\begin{aligned} \dot{S}(\omega) &= \int_{-\infty}^{\infty} A \cos(\omega_0 t + \varphi_0) e^{-j\omega t} dt = \\ &= \frac{1}{2} \int_{-\infty}^{\infty} A e^{j\varphi_0} e^{-j(\omega - \omega_0)t} dt + \frac{1}{2} \int_{-\infty}^{\infty} A e^{-j\varphi_0} e^{-j(\omega + \omega_0)t} dt = \\ &= A \pi e^{j\varphi_0} \delta(\omega - \omega_0) + A \pi e^{-j\varphi_0} \delta(\omega + \omega_0). \end{aligned}$$

Результат, як бачимо, є парою дельта-функцій, розташованих на частотах $\pm\omega_0$. Множники при них

відбивають амплітуду і початкову фазу (тобто комплексну амплітуду) гармонічного сигналу.

Той же результат можна було б отримати, застосувавши до спектру постійного за часом сигналу властивість перетворення Фур'є 7 щодо множення сигналу на гармонічну функцію.

Комплексна експонента.

Цей сигнал, на відміну від попередніх, не є дійсним:

$$s(t) = A \exp(j\omega_0 t).$$

Результат обчислення його спектру легко передбачити. Розглянутий раніше гармонічний сигнал мав спектральну функцію у вигляді двох дельта-функцій, оскільки косинус за допомогою формули Ейлера представлявся напівсумою двох комплексних експонент. Тому, спектром комплексної експоненти повинна бути поодинокі дельта-функція

$$\dot{S}(\omega) = \int_{-\infty}^{\infty} A e^{j\omega_0 t} e^{-j\omega t} dt = 2A\pi\delta(\omega - \omega_0).$$

Результат, як видно, відповідає очікуванню. Потрібно мати на увазі те, що оскільки сигнал не є дійсним, його спектр втрачає властивість симетрії.

Довільний періодичний сигнал.

Як відомо, періодичний сигнал з періодом T може бути представлений у вигляді ряду Фур'є

$$s(t) = \sum_{i=-\infty}^{\infty} \dot{C}_i \exp\left(j \frac{2\pi i}{T} t\right).$$

Після обчислення спектрів гармонічного сигналу і комплексної експоненти стає ясно, що спектральна функція такого сигналу є набором дельта-функцій, розташованих на частотах гармонік ряду Фур'є

$$\dot{S}(\omega) = \sum_{i=-\infty}^{\infty} 2\pi \dot{C}_i \delta\left(\omega - \frac{2\pi i}{T}\right).$$

Множники при дельта-функціях дорівнюють відповідним коефіцієнтам ряду Фур'є \dot{C}_i , помноженим на 2π .

Контрольні запитання

1. Який вигляд має розкладання сигналу в узагальнений ряд Фур'є?
2. Отримайте співвідношення для обчислення коефіцієнтів узагальненого ряду Фур'є, якщо базисні функції є ортонормованими на інтервалі $[t_1, t_2]$.
3. Сформулюйте й доведіть теорему Парсеваля для узагальненого ряду Фур'є.
4. Які три форми запису тригонометричного ряду Фур'є вам відомі?
5. Приведіть вираз ряду Фур'є тригонометрично-косинусної форми.
6. Опишіть властивості розкладання в ряд Фур'є періодичних сигналів.
7. Наведіть співвідношення теореми Парсеваля для трьох форм розкладання періодичного сигналу в ряд Фур'є.
8. Яким умовам повинен задовольняти фрагмент сигналу для того щоб існувало розкладання в ряд Фур'є?
9. Виведіть співвідношення для обчислення коефіцієнтів розкладання в ряд Фур'є послідовності прямокутних імпульсів.
10. Отримайте представлення меандру у вигляді ряду Фур'є.
11. Застосуйте ряд Фур'є для аналізу прямокутного, пилкоподібного та трикутного періодичних сигналів.

12. Побудуйте амплітудний і фазовий спектри послідовності прямокутних імпульсів типу меандр.

13. Запишіть формули прямого та оберненого перетворення Фур'є.

14. Назвіть дві умови існування спектральної густини сигналу.

15. Порівняйте властивості розкладання в ряд Фур'є періодичних сигналів та перетворення Фур'є.

16. Що розуміють під властивостями перетворення Фур'є?

17. В чому полягає властивість лінійності перетворення Фур'є?

18. Доведіть властивість затримки сигналу

19. Знайдіть спектральну густину трикутного імпульсу поданого на рис. 3.3.

20. Приведіть спектри неінтегрованих сигналів: функції Дірака; функції Хевісайда; гармонічного сигналу; періодичного сигналу.

Глава 3

КОРЕЛЯЦІЙНИЙ АНАЛІЗ СИГНАЛІВ

Кореляційний аналіз разом із спектральним відіграє велику роль в теорії сигналів. Перш за все, його сенс полягає в кількісному вимірюванні міри подібності різних сигналів. Для цього застосовуються кореляційні функції, з розгляду яких ми й почнемо цей розділ. Крім цього, кореляційні властивості є важливою складовою при виборі різновиду сигналу для вирішення тих чи інших завдань. Вони дають уяву про зміну рівня сигналу на виході пристрою його обробки при відмінності параметрів від очікуваних, що важливо для задач виявлення сигналів та вимірювання їх параметрів на тлі перешкод.

3.1. Автокореляційна функція

Завдання порівняння сигналу з його зсунутою в часі копією, є типовим для локаційних методів вимірювання дальності до об'єкта (радіолокація, гідролокація). В цьому випадку затриманий сигнал, що відбився від опромінюваного об'єкта, має практично таку ж форму, що й зондувальний. В пристрої вимірювання часу затримки відбитого сигналу виконується порівняння цього сигналу з копіями зразкового (зондувального) сигналу зсунутими на фіксовані відрізки часу. В результаті порівняння в кожному з каналів двох сигналів обирається канал, на виході якого схожість сигналів є найкращою. За номером цього каналу отримується значення часу затримки відбитого сигналу і відповідно дальності до об'єкта.

Якість роботи такого пристрою буде залежати від того, якою мірою будуть відрізнятись між собою сигнали при порівняно невеликих відносних зсувах. Для кількісної оцінки

такої відмінності застосовується автокореляційна функція сигналу (АКФ).

Автокореляційна функція (англійський термін – correlation function, CF) детермінованого сигналу $s(t)$ з кінцевою енергією є функцією часу τ відносного зсуву двох копій сигналу, яка є скалярним добутком

$$B_s(\tau) = \int_{-\infty}^{\infty} s(t)s(t-\tau)dt. \quad (3.1)$$

Автокореляційна функція є мірою схожості між сигналом і його зсунутою копією – чим більше значення АКФ, тим ця схожість сильніша. Крім того, АКФ має наступні властивості [11]:

1. Значення АКФ при $\tau = 0$ дорівнює енергії сигналу, тобто інтегралу від його квадрата

$$B_s(0) = \int_{-\infty}^{\infty} s(t)s(t)dt = \int_{-\infty}^{\infty} s^2(t)dt = E.$$

2. АКФ є парною функцією свого аргументу τ

$$B_s(\tau) = B_s(-\tau).$$

3. Значення АКФ при $\tau = 0$ є максимально можливим значенням: $B_s(\tau) \leq B_s(0)$.

4. Із зростанням абсолютного значення τ АКФ сигналу з кінцевою енергією затухає $\lim_{|\tau| \rightarrow \infty} B_s(\tau) = 0$.

5. Якщо сигнал $s(t)$ не містить особливостей у вигляді δ -функцій, його АКФ не може мати розривів (тобто обов'язково буде неперервною функцією).

6. Якщо сигнал – напруга, то розмірність його АКФ дорівнює $V^2 \cdot c$.

Розглянемо обчислення АКФ на прикладі прямокутного імпульсу тривалістю T , який розпочинається в нульовий момент часу (рис. 3.1):

- при $0 \leq \tau \leq T$ $B_s(\tau) = \int_{\tau}^T A^2 dt = A^2(T - \tau)$;
- при $-T \leq \tau \leq 0$ $B_s(\tau) = \int_0^{T+\tau} A^2 dt = A^2(T + \tau)$;

- при $|\tau| > T$ $B_s(0) = 0$.

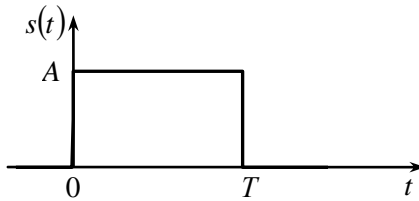


Рисунок 3.1 – Прямокутний імпульс

В результаті отримаємо

$$B_s(\tau) = \begin{cases} A^2(T - |\tau|), & \text{якщо } |\tau| \leq T; \\ 0, & \text{якщо } |\tau| > T. \end{cases}$$

Графік АКФ прямокутного імпульсу показаний на рис. 3.2.

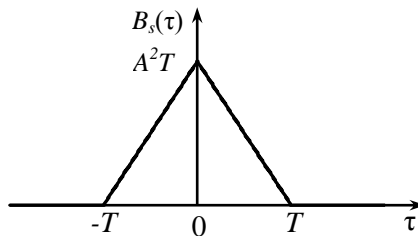


Рисунок 3.2 – Кореляційна функція прямокутного імпульсу

У разі періодичного сигналу (і взагалі будь-якого сигналу з нескінченною енергією) скористатися наведеним вище визначенням АКФ не вдасться. Тому АКФ періодичного сигналу з періодом T обчислюють, усереднюючи добуток зсунутих копій в межах одного періоду:

$$B_s(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} s(t)s(t - \tau)dt. \quad (3.2)$$

Набір властивостей АКФ (3.2) дещо змінюється порівняно з (3.1):

1. При $\tau = 0$ її значення дорівнює не енергії, а середній потужності аналізованого сигналу

$$B_s(0) = \frac{1}{T} \int_{-T/2}^{T/2} s^2(t) dt = P_{cp}.$$

2. Властивість парності зберігається $B_s(\tau) = B_s(-\tau)$.

3. Значення АКФ при $\tau = 0$, як і раніше, є максимально можливим $B_s(\tau) \leq B_s(0)$.

4. АКФ періодичного сигналу є періодичною функцією з тим же періодом, що й сам сигнал, $B_s(\tau + T) = B_s(\tau)$.

5. Якщо сигнал не містить дельта-функцій, його АКФ буде неперервною функцією.

6. Розмірність АКФ періодичного сигналу – квадрат розмірності сигналу (B^2 , якщо сигнал – напруга).

Як приклад, обчислимо АКФ гармонічного сигналу

$$s(t) = A \cos(\omega_0 t + \varphi_0).$$

Період такого сигналу дорівнює $2\pi/\omega_0$. Тому кореляційний інтеграл

$$\begin{aligned} B_s(\tau) &= \frac{\omega_0}{2\pi} \int_{-\pi/\omega_0}^{\pi/\omega_0} A \cos(\omega_0 t + \varphi_0) A \cos(\omega_0(t - \tau) + \varphi_0) dt = \\ &= \frac{A^2}{2} \cos(\omega_0 \tau). \end{aligned}$$

Як видно, АКФ гармонічного сигналу також є гармонічною функцією. Важливо також, що отриманий результат не залежить від початкової фази (параметр φ_0 до отриманого виразу не увійшов). Це прояв загальної властивості всіх кореляційних функцій, про яку піде мова далі в підрозділі 3.3.

3.2. Взаємна кореляційна функція

Якщо АКФ показує міру схожості між зсунутими копіями одного й того ж сигналу, то взаємна кореляційна функція (ВКФ; англійський термін – cross-correlation function, CCF) дозволяє

обчислювати аналогічну величину для зсунутих екземплярів двох різних сигналів.

Загальний вигляд формули кореляційної функції зберігається, але під інтегралом стоїть добуток двох різних сигналів, один з яких затриманий на якийсь час τ

$$B_{12}(\tau) = \int_{-\infty}^{\infty} s_1(t) s_2(t - \tau) dt. \quad (3.3)$$

Очевидно, що АКФ є окремим випадком ВКФ, коли обидва сигнали однакові: $s_1(t) = s_2(t) = s(t)$.

Як приклад, обчислимо ВКФ прямокутного і трикутного імпульсів (див. рис. 3.1 і 3.3):

$$s_1(t) = \begin{cases} A, & 0 \leq t \leq T, \\ 0, & t < 0, t > T; \end{cases} \quad s_1(t) = \begin{cases} A \frac{t}{T}, & 0 \leq t \leq T, \\ 0, & t < 0, t > T. \end{cases}$$

- при $0 \leq \tau \leq T$

$$B_{12}(\tau) = \int_{\tau}^T A^2 \frac{t - \tau}{T} dt = \frac{A^2}{2T} (T - \tau)^2;$$

- при $-T \leq \tau \leq 0$

$$B_{12}(\tau) = \int_0^{T+\tau} A^2 \frac{t - \tau}{T} dt = \frac{A^2}{2T} (T^2 - \tau^2);$$

- при $|\tau| > T$

$$B_{12}(\tau) = 0.$$

Об'єднуючи результати, можна записати

$$B_{12}(\tau) = \begin{cases} \frac{A^2}{2T} (T - \tau)^2, & 0 \leq \tau \leq T, \\ \frac{A^2}{2T} (T^2 - \tau^2), & -T \leq \tau \leq 0, \\ 0, & |\tau| > T. \end{cases}$$

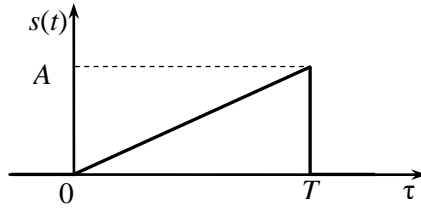


Рисунок 3.3 – Трикутний імпульс

Графік отриманої ВКФ представлений на рис. 3.4.

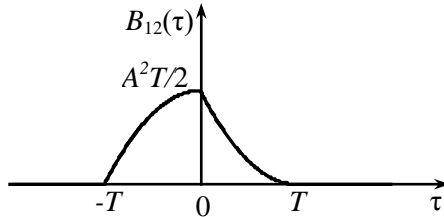


Рисунок. 3.4 – ВКФ прямокутного і трикутного імпульсів

Властивості ВКФ дещо відрізняються від властивостей АКФ:

1. $|B_{12}(\tau)| \leq \sqrt{E_1 E_2}$, де E_1 і E_2 – енергії сигналів.
 2. $B_{12}(-\tau) = B_{21}(\tau)$, тобто зміна знаку τ рівносильна взаємній перестановці сигналів.
 3. Значення ВКФ при $\tau = 0$ не є особливим; її максимум може мати місце в будь-якій точці на осі τ .
 4. Із зростанням абсолютного значення τ ВКФ сигналів з кінцевою енергією згасає: $\lim_{|\tau| \rightarrow \infty} B_{12}(\tau) = 0$.
 5. Якщо сигнали $s_1(t)$ і $s_2(t)$ не містять особливостей у вигляді дельта-функцій, їх ВКФ не матиме розривів (тобто обов'язково буде неперервною функцією).
 6. Якщо сигнали – напруга, то розмірність їх ВКФ є $V^2 \cdot c$.
- Для періодичних сигналів поняття ВКФ зазвичай не застосовується, хоча воно може бути введено у разі, якщо сигнали $s_1(t)$ і $s_2(t)$ мають однаковий період.

3.3. Зв'язок між кореляційними функціями і спектрами сигналів

Оскільки кореляційні функції, як і спектри є інтегральними перетвореннями аналізованих сигналів, логічно припустити, що ці характеристики якимось пов'язані одна з одною. Для виявлення цього зв'язку піддамо взаємну кореляційну функцію перетворенню Фур'є, вважаючи, що сигнали $s_1(t)$ і $s_2(t)$ мають спектральні функції $\dot{S}_1(\omega)$ і $\dot{S}_2(\omega)$:

$$\begin{aligned} \int_{-\infty}^{\infty} B_{12}(\tau) e^{-j\omega\tau} d\tau &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_1(t) s_2(t-\tau) e^{-j\omega\tau} dt d\tau = \\ &= \int_{-\infty}^{\infty} s_1(t) e^{-j\omega t} \int_{-\infty}^{\infty} s_2[-(\tau-t)] e^{-j\omega(\tau-t)} d(\tau-t) dt = \dot{S}_1(\omega) \dot{S}_2^*(\omega). \end{aligned} \quad (3.4)$$

При обчисленні другого інтегралу враховано властивість 3 перетворення Фур'є у випадку $a = -1$ (див. підрозділ 2.2).

Таким чином, ВКФ пов'язана перетворенням Фур'є з так званим взаємним спектром сигналів, який для сигналів $s_1(t)$ і $s_2(t)$ є добутком їх спектральних функцій, одна з яких піддана комплексному спряженню

$$\dot{S}_{12}(\omega) = \dot{S}_1(\omega) \dot{S}_2^*(\omega). \quad (3.5)$$

Звідси можна зробити дуже важливий висновок: якщо спектри сигналів не перекриваються, то їх взаємний спектр дорівнює нулю на усіх частотах, а це означає, що дорівнює нулю і їх ВКФ при будь-яких часових зсувах τ . Таким чином, сигнали із спектрами, що не перекриваються, є некорельованими. Приймавши $s_1(t) = s_2(t) = s(t)$, отримуємо аналогічний результат для АКФ

$$\int_{-\infty}^{\infty} B_s(\tau) e^{-j\omega\tau} d\tau = \dot{S}(\omega) \dot{S}^*(\omega) = |\dot{S}(\omega)|^2. \quad (3.6)$$

Отже, АКФ сигналу пов'язана перетворенням Фур'є з квадратом модуля спектральної функції, або з *енергетичним спектром сигналу*.

Таким чином, обернене перетворення Фур'є енергетичного спектру сигналу дозволяє обчислити його АКФ

$$B_s(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\dot{S}(\omega)|^2 e^{j\omega\tau} d\omega. \quad (3.7)$$

Розглянемо *приклад* обчислення АКФ сигналу, який має рівномірний і обмежений за частотою енергетичний спектр

$$W_u(\omega) = \begin{cases} W_0, & -\omega_B \leq \omega \leq \omega_B, \\ 0, & |\omega| > \omega_B, \end{cases}$$

де ω_B – верхня гранична частота спектру.

За формулою (3.7) маємо АКФ

$$B_u(\tau) = \frac{W_0}{2\pi} \int_{-\omega_B}^{\omega_B} e^{j\omega\tau} d\omega = \frac{W_0}{\pi} \int_0^{\omega_B} \cos \omega\tau d\omega = \frac{W_0\omega_B}{\pi} \frac{\sin \omega_B \tau}{\omega_B \tau}.$$

Отримана АКФ сигналу з прямокутним спектром подібно до спектру прямокутного імпульсу має пелюстковий характер. Ширину головної пелюстки часто приймають за інтервал кореляції, який отримуємо з співвідношення $\omega_B \tau_K = \pi$, тобто $\tau_K = \pi/\omega_B = 1/(2f_B)$.

Кореляційні властивості сигналу, таким чином, можуть бути оцінені за його енергетичним спектром. Чим ширше спектр сигналу, тим вужче головна пелюстка його АКФ і тим він краще з точки зору вимірювання, наприклад, моменту появи відбитого сигналу в процесі виявлення цілі при застосуванні локаційних методів. Окрім цієї властивості, важливою кореляційною властивістю сигналу є й рівень бічних пелюсток сигналу. Для виявлення слабких сигналів поруч з більш потужними потрібно обирати такі сигнали, які забезпечують порівняно низький рівень бічних пелюсток відносно головної пелюстки.

З (3.6) також видно, що *АКФ сигналу не залежить від його фазового спектру*. Отже, сигнали, амплітудні спектри яких однакові, а фазові різняться, матимуть однакову АКФ. Тобто маємо ще один наслідок – *за АКФ неможливо відновити вихідний сигнал* (із-за втрати інформації про фазу).

Взаємний спектр сигналів $s_1(t)$ і $s_2(t)$ (3.5) є прямим перетворенням Фур'є їх взаємної кореляційної функції (3.4). Відповідно, якщо відомий взаємний спектр сигналів, їх ВКФ може бути обчислена за формулою оберненого перетворення Фур'є

$$B_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}_{12}(\omega) e^{j\omega\tau} d\omega. \quad (3.8)$$

Підстановка значення $\tau = 0$ в (3.5) з розкриттям виразів для ВКФ і взаємного спектру дає рівність

$$\int_{-\infty}^{\infty} s_1(t) s_2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}_1(\omega) \dot{S}_2^*(\omega) d\omega. \quad (3.9)$$

Вираз (3.9) може бути поданий у більш загальному вигляді

$$(s_1, s_2) = \frac{1}{2\pi} (\dot{S}_1, \dot{S}_2). \quad (3.10)$$

Рівність (3.10) має назву *узагальненої формули Релея*. Вона є справедливою й випадку, коли сигнал комплексний. Узагальнена формула Релея може трактуватись наступним чином: *скалярний добуток двох сигналів з точністю до коефіцієнта дорівнює скалярному добутку їх спектральних густин*.

Якщо сигнали однакові ($s_1(t) = s_2(t) = s(t)$) з (3.9) отримаємо співвідношення для обчислення енергії сигналу як в часовій, так і в частотній області, яке є різновидом *рівності Парсеваля* для фінітних сигналів (див. також (2.4))

$$E = \int_{-\infty}^{\infty} s^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\dot{S}(\omega)|^2 d\omega. \quad (3.11)$$

Контрольні запитання

1. Дайте визначення автокореляційної функції .
2. Перелічіть властивості автокореляційної функції.
3. Вкажіть особливості визначення автокореляційної функції та її властивостей для періодичного сигналу.

4. Знайдіть вираз для автокореляційної функції прямокутного радіоімпульсу $s(t) = A[\sigma(t) - \sigma(t - \tau)] \cos(\omega_0 t + \varphi_n)$.

5. Знайдіть автокореляційну функцію періодичного сигналу, представленого рядом Фур'є $s(t) = \sum_{n=0}^{\infty} A_n \cos(n\Omega t + \varphi_n)$.

6. Дайте визначення взаємної кореляційної функції

7. Чим відрізняються властивості взаємної кореляційної функції порівняно з АКФ.

8. Знайдіть вираз для взаємної кореляційної функції прямокутних імпульсів $s_1(t) = A[\sigma(t) - \sigma(t - \tau)]$ і $s_2(t) = A[\sigma(t) - \sigma(t - 2\tau)]$.

9. Виявіть зв'язок між кореляційними функціями і спектрами сигналів.

10. Чим є взаємний спектр сигналів?

11. Який можна зробити висновок, якщо спектри сигналів не перекриваються?

12. Що дозволяє обчислити обернене перетворення Фур'є енергетичного спектру сигналу?

13. Обчисліть АКФ автокореляційну функцію, який має рівномірний і обмежений за частотою енергетичний спектр.

14. Якими кореляційними властивостями характеризується сигнал?

15. Запишіть співвідношення, яке пов'язує взаємну кореляційну функцію і взаємний спектр двох сигналів.

16. Чи залежить автокореляційна функція від фазового спектру сигналу?

17. Чи можливо відновити вихідний сигнал за допомогою його автокореляційної функції?

18. Яку назву має рівняння $(s_1, s_2) = \frac{1}{2\pi} (\dot{S}_1, \dot{S}_2)$?

19. Як може трактуватись узагальнена формула Релея?

20. Запишіть рівність Парсеваля для фінітних сигналів. В чому полягає його фізична сутність?

Глава 4

АНАЛІЗ ВИПАДКОВИХ СИГНАЛІВ

4.1. Характеристики випадкових сигналів

На відміну від детермінованих сигналів, форма яких відома, миттєві значення випадкових сигналів заздалегідь не відомі і можуть бути передбачені лише з деякою ймовірністю, яка менша за одиницю. Характеристики таких сигналів є ймовірнісними. В розділі 1.1 розглянуті приклади сигналів, що потребують ймовірнісного описання, серед яких різні види шумових сигналів, а також адитивні, мультиплікативні та комбіновані суміші шумів і сигналів, що несуть вимірвальну інформацію.

У технічній літературі терміни «випадковий сигнал» і «випадковий процес» зазвичай використовуються як синоніми. Застосовується також термін «стохастичний процес».

Випадковий процес (сигнал) $X(t)$ – це випадкова функція, миттєве значення якої в будь-який момент часу $t = t_0$ є випадковою величиною $X(t_0)$ [11, 17]. Або інакше – це сукупність (ансамбль) $\{x_k(t)\}$, $k = 1, 2, \dots$, всіх можливих вибірових функцій (реалізацій), які породжуються певним випадковим явищем.

Одна з цих реалізацій, наприклад $x_k(t)$, стає повністю відомою після реєстрації (прийому) сигналу. Вона вже не випадкова, а детермінована функція часу. На рис. 4.1 наведено приклад сукупності реалізацій випадкового сигналу.

Випадкові сигнали в багатьох випадках можуть бути доволі простими сигналами, в яких випадковим є один або кілька параметрів цього сигналу. Такі випадкові сигнали називають квазідетермінованими (див. підрозділ 1.1).

Для аналізу властивостей і характеристик випадкового сигналу, його перетворень в системах обробки сигналів потрібно задати його математичну модель. Така модель може описувати можливі реалізації випадкового сигналу разом із вказівкою відносної частоти їх появи.

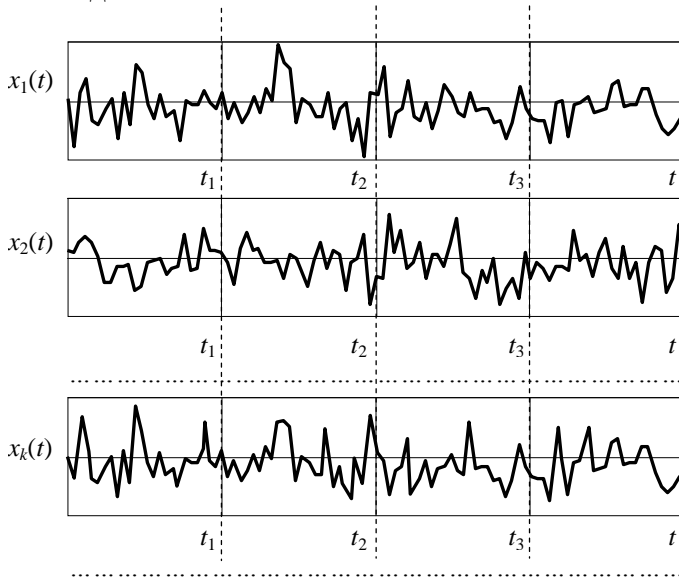


Рисунок 4.1 – Реалізації випадкового сигналу

Випадковий процес може залежати не лише від одного аргументу – часу, але й від кількох, наприклад часу і просторових координат (приклад – швидкість вітру). Такі випадкові процеси називають *випадковими полями* [19]. Випадкові поля можуть бути скалярними і векторними.

Тому додатково до введеної в підрозділі 1.1 класифікації сигналів випадкові процеси можуть поділятися на [19]:

скалярний $X(t)$ – випадковий процес, область значень якого є множиною в просторі дійсних чисел;

векторний $\mathbf{X}(t)$ – випадковий процес, область значень якого є множиною у відповідному координатному просторі;

скалярне випадкове поле $X(\mathbf{r}, t)$ – випадковий процес, область значень якого є множиною з дійсних чисел у відповідному координатному просторі;

векторне випадкове поле $\mathbf{X}(\mathbf{r}, t)$ – випадковий процес, що представляється у вигляді компонент, які є скалярними полями;

Розглянемо як приклад модель гармонічного сигналу з випадковою початковою фазою.

Така модель скалярного випадкового сигналу використовується у багатьох практичних задачах. Реалізації сигналу є гармонічними коливаннями з відомою амплітудою і частотою, але з випадковою початковою фазою. Кожна з них може бути записана як

$$s(t) = A \cos(\omega_0 t + \varphi),$$

де A – амплітуда (детермінована), ω_0 – частота (детермінована) і φ – випадкова початкова фаза, яка в більшості практично цікавих випадків може вважатися рівномірно розподіленою на інтервалі $0 \dots 2\pi$, тобто ймовірності, що має наступну густину:

$$p_\varphi(\varphi) = \begin{cases} 1/2\pi, & 0 \leq \varphi < 2\pi, \\ 0, & \text{в інших випадках.} \end{cases}$$

Графіки декількох реалізацій цього випадкового сигналу косинусоїди, що зміщені одна відносно одної за часом (рис. 4.2).

Конкретний вигляд реалізації на рис. 4.2 визначається значенням лише однієї випадкової величини – початкової фази. Тобто маємо приклад квазидетермінованого випадкового сигналу.

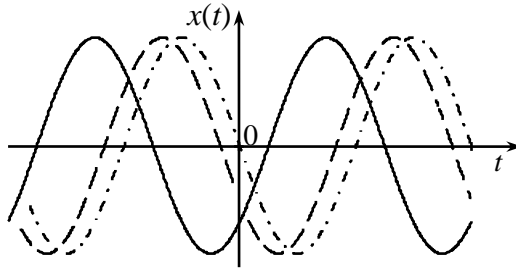


Рисунок 4.2 – Реалізації гармонічного сигналу з випадковою початковою фазою

Ймовірнісні характеристики випадкових сигналів

Нехай випадковий процес $X(t)$ заданий ансамблем реалізацій $\{x_1(t), x_2(t), \dots, x_k(t), \dots\}$. Вибравши довільний момент часу t_1 , зафіксуємо значення, що приймаються усіма реалізаціями: $\{x_1(t_1), x_2(t_1), \dots, x_k(t_1), \dots\}$ (див. рис. 4.1). Сукупність цих значень утворює одновимірний переріз випадкового процесу і є випадковою величиною $X(t_1)$.

Нагадаємо коротко основні характеристики випадкових величин. При цьому для одновимірних перерізів випадкових процесів вони в загальному випадку залежать від обраного моменту часу t_1 (на рис. 4.1 також показані перерізи в моменти часу t_2 і t_3).

Функціональні характеристики

Функція розподілу ймовірності (англ. cumulative distribution function, CDF) дорівнює ймовірності того, що у момент часу t_1 значення випадкового процесу не перевищує x

$$F(x, t_1) = P(X(t_1) \leq x). \quad (4.1)$$

Функція $F(x, t_1)$ є неубутною. Її значення лежать в діапазоні $0 \leq F(x, t_1) \leq 1$. Для граничних значень x виконуються співвідношення: $F(-\infty, t_1) = 0$ і $F(\infty, t_1) = 1$.

Імовірність потрапляння значення випадкового процесу в інтервал $[a, b]$ дорівнює різниці значень функції розподілу на кінцях цього інтервалу

$$P(a \leq X(t_1) \leq b) = F(b, t_1) - F(a, t_1).$$

Одновимірна густина ймовірності (англ. probability density function, PDF) є похідною від функції розподілу

$$p(x, t_1) = dF(x, t_1) / dx. \quad (4.2)$$

Відповідно добуток

$$p(x, t_1)dx = P(x \leq X(t_1) \leq x + dx),$$

дорівнює ймовірності потрапляння значення випадкового процесу $X(t_1)$ в нескінченно малий інтервал $[x, x + dx]$. Звідси витікає, що густина ймовірності є невід'ємною функцією $p(x, t_1) \geq 0$. Для розрахунку ймовірності потрапляння значення $X(t_1)$ у довільний інтервал $[a, b]$, необхідно обчислити інтеграл

$$P(a \leq X(t_1) \leq b) = \int_a^b p(x, t_1) dx.$$

Оскільки випадкова величина обов'язково набуває якого-небудь значення, повинна виконуватись умова нормування

$$\int_{-\infty}^{\infty} p(x, t_1) dx = P(-\infty < X(t_1) < \infty) = 1.$$

Знаючи густину ймовірності, можна розрахувати й функцію розподілу

$$F(x, t_1) = \int_{-\infty}^x p(x, t_1) dx. \quad (4.3)$$

Якщо випадкова величина є дискретною і приймає фіксовані значення $\{x_1, x_2, \dots, x_k, \dots\}$ з відповідними

ймовірностями $\{P_1, P_2, \dots, P_k, \dots\}$, тоді її густину ймовірності можна подати як суму дельта-функцій

$$p(x, t_1) = \sum_i P_i \delta(x - x_i),$$

для якої також виконується умова невід'ємності $p(x, t_1) \geq 0$.

Числові характеристики

Знання одновимірної густини ймовірності $p(x, t_1)$ дозволяє здійснити статистичне усереднення як самої величини $X(t_1)$, так і будь-якої функції від неї. Під статистичним усереднюванням (*англ.* ensemble averaging) мається на увазі усереднення по множині (по ансамблю реалізацій) в якому-небудь перерізі процесу, тобто у фіксований момент часу.

Найбільш поширені в практичних застосуваннях числові характеристики для перерізу випадкового сигналу (процесу) у певний момент часу є математичне сподівання, дисперсія і середнє квадратичне відхилення.

Математичне сподівання (*англ.* mean value) – це теоретична оцінка середнього зваженого значення випадкового процесу у момент часу t

$$m_x(t) = M \{X(t)\} = \int_{-\infty}^{\infty} xp(x, t)dx. \quad (4.4)$$

Математичне сподівання деякої функції f від випадкової величини x , що має густину ймовірності $p_x(x)$ обчислюється за формулою

$$M \{f(x)\} = \int_{-\infty}^{\infty} f(x)p_x(x)dx. \quad (4.5)$$

Формула (4.4) для математичного сподівання є окремим випадком формули (4.5) при $f(x) = x$.

Дисперсія (*англ.* variance) характеризує середню потужність відхилень випадкового процесу від його

середнього значення $m_x(t)$, званих *флуктуаціями* (англ. fluctuation)

$$D_x(t) = M \left\{ [X(t) - m_x(t)]^2 \right\} = M \left\{ X^2(t) \right\} - m_x^2(t) = \int_{-\infty}^{\infty} x^2 p(x,t) dx - m_x^2(t). \quad (4.6)$$

Середнє квадратичне відхилення (СКВ) (англ. standard deviation) є квадратним коренем з дисперсії і слугує амплітудною мірою розкиду значень випадкового процесу в момент часу t відносно математичного сподівання

$$\sigma_x(t) = \sqrt{D_x(t)} = \sqrt{M \left\{ [X(t) - m_x(t)]^2 \right\}} = \sqrt{M \left\{ X^2(t) \right\} - m_x^2(t)}. \quad (4.7)$$

Дисперсія випадкової величини X часто позначається як σ_x^2 .

Розглянемо ймовірнісні характеристики найбільш поширених в метрологічній практиці *рівномірного і нормального законів розподілу* випадкових величин.

Для *рівномірного розподілу* (англ. uniform distribution) густина ймовірності є константою на деякому інтервалі $[a, b]$ (рис. 4.3, а). Ця константа згідно з умовою нормування дорівнює $1/(b-a)$.

Функція розподілу згідно (4.2) на інтервалі $[a, b]$ лінійно зростає від 0 до 1 (див. рис. 4.3, б):

$$F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases}$$

Математичне сподівання цілком очікувано відповідає середині інтервалу можливих значень випадкової величини

$$m_x = \int_a^b x \frac{1}{b-a} dx = \frac{a+b}{2}.$$

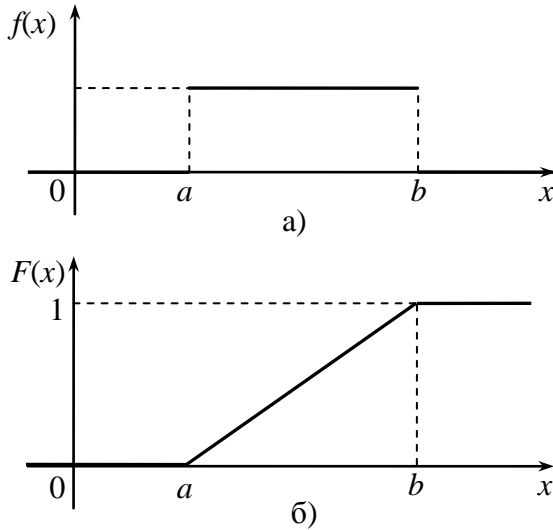


Рисунок 4.3 – Рівномірний закон розподілу:
а) густина ймовірності; б) функція розподілу

Загалом, якщо функція густини ймовірності є симетричною, тоді значення математичного сподівання завжди збігається з центром симетрії.

Для розрахунку дисперсії спочатку визначимо середній квадрат випадкової величини

$$M\{X^2\} = \int_a^b x^2 \frac{1}{b-a} dx = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}.$$

Тепер за формулою (4.6) розрахуємо дисперсію

$$D_x = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

Отже дисперсія рівномірно розподіленої випадкової величини дорівнює одній дванадцятій квадрата ширини інтервалу розподілу $[a, b]$. Середнє квадратичне відхилення відповідно до (4.7) є пропорційним цій ширині

$$\sigma_x = \sqrt{D_x} = \frac{b-a}{2\sqrt{3}}.$$

Нормальний (гаусівський) закон розподілу випадкових величин (англ. normal distribution, Gaussian distribution) дуже зручний для аналізу і часто зустрічається на практиці. Він характерний для заводів в інформаційно-вимірювальних каналах радіотехнічних засобів (зв'язку, радіолокації) та засобів вимірювань радіотехнічних величин. Одновимірна густина ймовірності випадкової величини з нормальним розподілом визначається виразом

$$p(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{(x-m_x)^2}{2\sigma_x^2}\right),$$

де m_x і σ_x^2 – відповідно математичне сподівання і дисперсія випадкової величини. На рис. 4.4 приведений графік побудований згідно цього виразу при $m_x = 0$ і $\sigma_x = 1$.

Нормальний закон розподілу є дуже поширеним в природі. Це пояснюється тим, що при підсумовуванні досить великого числа рівнопотужних статистично незалежних або слабко залежних випадкових величин, що мають довільні густини розподілу ймовірності, густина розподілу суми прямує до нормальної. Це твердження має назву центральної граничної теореми, яка сформульована Ляпуновим А. М. у 1901 р. Дуже корисною для математичного аналізу властивістю нормального розподілу є те, що з некорельованості випадкових гаусівських величин виходить їх статистична незалежність (різниця між цими поняттями пояснюється далі).

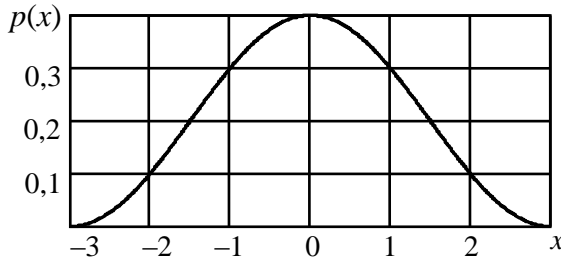


Рисунок 4.4 – Густина ймовірності випадкової величини з нормальним законом розподілу

Функція розподілу для нормального закону не виражається через елементарні функції. У вітчизняній літературі прийнято виражати її через інтеграл ймовірності, який має вигляд

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt.$$

При цьому параметри розподілу $m_x = 0$ і $\sigma_x = 1$. Для нормального закону з довільним математичним сподіванням m_x і СКВ σ_x функція розподілу пов'язана з інтегралом ймовірності наступним чином

$$F(x) = \Phi\left(\frac{x - m_x}{\sigma_x}\right).$$

В математичних пакетах, наприклад Matlab, Mathcad, є вбудовані функції для обчислення поширених законів розподілу: нормального, рівномірного та інших. Також мають датчики випадкових чисел, які дозволяють формувати послідовності випадкових чисел з поширеними на практиці законами розподілів.

Одновимірні густини ймовірності і пов'язані з нею числові характеристики несуть важливу інформацію про властивості випадкового процесу $X(t)$. Однак, вони не дають

вичерпну уяву про його внутрішню структуру і швидкість зміни в часі. Для описання часових характеристик використовують багатовимірні густини ймовірностей, а у випадку обмеження двовимірною густиною широко застосовується кореляційна функція випадкового процесу. Способи описання багатовимірних густин, кореляційних, а також спектральних характеристик випадкових процесів будуть розглянуті далі.

4.2. Аналіз випадкових сигналів. Кореляційні функції випадкових сигналів

Як вже зазначалося, для описання поведінки випадкового процесу в часі недостатньо одновимірної густини ймовірності. Значно більше відомостей можна отримати, маючи в розпорядженні два перерізи випадкового процесу в довільні моменти часу t_1 і t_2 (див. рис. 4.1). Два перерізи утворюють двовимірну випадкову величину $\{X(t_1), X(t_2)\}$, яка описується двовимірною густиною ймовірності $p(x_1, x_2, t_1, t_2)$. Добуток $p(x_1, x_2, t_1, t_2)dx_1dx_2$ є ймовірністю того, що реалізація випадкового процесу $X(t)$ у момент часу t_1 потрапляє в нескінченно малий інтервал шириною dx_1 в околиці x_1 , а в момент часу t_2 – в нескінченно малий інтервал шириною dx_2 в околиці x_2 .

Природним узагальненням є n -мірний переріз випадкового процесу, що призводить до n -мірної густини ймовірності $p(x_1, x_2, \dots, x_n, t_1, t_2, \dots, t_n)$. При $n \rightarrow \infty$ така функція є вичерпною ймовірнісною характеристикою випадкового процесу.

Описання властивостей випадкових процесів за допомогою багатовимірної густини ймовірності високої розмірності може бути дуже детальним, але натикається на

серйозні математичні труднощі. В той же час, багато задач, пов'язаних з описанням випадкових сигналів, вирішується на основі двовимірної густини ймовірності.

Зокрема, двовимірна густина ймовірності $p(x_1, x_2, t_1, t_2)$ дозволяє визначити важливу характеристику випадкового процесу – його *коваріаційну функцію*

$$K_x(t_1, t_2) = M \{x(t_1)x(t_2)\},$$

яка є статистично усередненим добутком значень випадкової функції $X(t)$ в моменти часу t_1 і t_2 .

Для кожної реалізації випадкового процесу добуток $x(t_1)x(t_2)$ є деяким числом. Сукупність реалізацій утворює безліч випадкових чисел, розподіл яких характеризується двовимірною густиною ймовірності $p(x_1, x_2, t_1, t_2)$. Якщо ця густина ймовірності відома, операція усереднення по множині задається формулою

$$K_x(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2, t_1, t_2) dx_1 dx_2. \quad (4.8)$$

Часто при аналізі випадкових процесів основний інтерес представляє їх флуктуаційна складова. У таких випадках застосовується *кореляційна функція*, що є статистично усередненим добутком значень центрованої випадкової функції $X(t) - m_x(t)$ в моменти часу t_1 і t_2

$$\begin{aligned} R_x(t_1, t_2) &= M \{ [x(t_1) - m_x(t_1)] [x(t_2) - m_x(t_2)] \} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x(t_1) - m_x(t_1)] [x(t_2) - m_x(t_2)] p(x_1, x_2, t_1, t_2) dx_1 dx_2 = \\ &= K_x(t_1, t_2) - m_x(t_1) m_x(t_2). \end{aligned} \quad (4.9)$$

Кореляційна функція характеризує міру статистичного зв'язку тих значень випадкового процесу, які спостерігаються при $t = t_1$ і $t = t_2$. При $t_1 = t_2 = t$ вираз (4.9) відповідає

визначенню дисперсії випадкового процесу X (див. формулу (4.6)). Отже, при поєднанні перерізів функція кореляції дорівнює дисперсії

$$R_x(t, t) = \int_{-\infty}^{\infty} x^2 p(x, t) dx - m_x^2(t) = D_x(t). \quad (4.10)$$

При користуванні іноземною літературою потрібно знати, що в ній використовується зворотна термінологія – $K_x(t_1, t_2)$ називається кореляційною (*англ.* correlation), а $R_x(t_1, t_2)$ – ковариаційною функцією (*англ.* covariance). Втім, при аналізі центрованих (що мають нульове математичне сподівання) випадкових процесів вирази (4.9) і (4.10) збігаються й кореляційна і ковариаційна функції співпадають.

Як приклад розрахуємо кореляційну функцію вже згаданого гармонічного сигналу з випадковою рівномірно розподіленою початковою фазою. Легко переконатися, що цей випадковий процес є центрованим, тобто його математичне сподівання не залежить від часу і дорівнює нулю

$$m_x(x) = \int_{-\infty}^{\infty} x(t) p_{\varphi}(\varphi) d\varphi = \int_0^{2\pi} A \cos(\omega_0 t + \varphi) \frac{1}{2\pi} d\varphi = 0.$$

Тому ковариаційна і кореляційна функції цього процесу співпадають. Оскільки реалізації цього випадкового процесу є функціями, які залежать від однієї випадкової величини φ , для усереднювання добутку не потрібно застосовувати двовимірну густину ймовірності – достатньо скористатися формулою (4.5) для усереднення довільної функції від випадкової величини

$$\begin{aligned} R_x(t_1, t_2) &= K_x(t_1, t_2) = \int_{-\infty}^{\infty} x(t_1) x(t_2) p_{\varphi}(\varphi) d\varphi = \\ &= \int_0^{2\pi} A \cos(\omega_0 t_1 + \varphi) A \cos(\omega_0 t_2 + \varphi) \frac{1}{2\pi} d\varphi = \end{aligned}$$

$$= \frac{A^2}{2\pi} \left[\int_0^{2\pi} \frac{1}{2} \cos[\omega_0(t_1 + t_2) + 2\varphi] d\varphi + \int_0^{2\pi} \frac{1}{2} \cos[\omega_0(t_1 - t_2)] d\varphi \right].$$

В першому доданку інтегрування виконується по двох періодах функції косинус, тому цей інтеграл дорівнює нулю. У другому доданку підінтегральна функція не залежить від змінної інтегрування φ , результат її інтегрування дорівнює добутку підінтегрального виразу і довжини проміжку інтегрування 2π . Остаточно отримуємо

$$R_x(t_1, t_2) = K_x(t_1, t_2) = \frac{A^2}{2} \cos[\omega_0(t_1 - t_2)].$$

Видно, що кореляційна функція цього випадкового процесу гармонічно залежить від інтервалу між аналізованими моментами часу. При збігові моментів часу $t_1 = t_2$ отримуємо величину дисперсії випадкового процесу

$$D_x(t) = R_x(t, t) = \frac{A^2}{2}.$$

Отримане значення дисперсії цього випадкового сигналу дорівнює середній потужності детермінованого гармонічного сигналу.

Некорельованість і статистична незалежність

Якщо спільно розглядати дві випадкові величини X_1 і X_2 , між ними може існувати або не існувати статистичний зв'язок. Відсутність такого зв'язку означає, що густина ймовірності однієї випадкової величини не залежить від того, якого значення набуває інша випадкова величина. Двовимірна густина ймовірності при цьому є добутком одновимірних густин

$$p(x_1, x_2) = p_1(x_1)p_2(x_2).$$

Ця умова називається *умовою статистичної незалежності*. За наявності статистичного зв'язку між випадковими величинами статистичні властивості кожної з них залежать від значення, що приймається іншою

випадковою величиною. Цей зв'язок може бути сильним або слабким, лінійним або нелінійним. Мірою лінійного статистичного зв'язку між випадковими величинами є коефіцієнт кореляції

$$r_{12} = \frac{M\{X_1 X_2\} - M\{X_1\}M\{X_2\}}{\sqrt{D\{X_1\}D\{X_2\}}}. \quad (4.11)$$

Можна показати, що $|r_{12}| \leq 1$. Граничні значення ± 1 досягаються, якщо реалізації випадкових величин жорстко пов'язані лінійним співвідношенням $x_2 = ax_1 + b$, де a і b – деякі константи. Знак коефіцієнта кореляції при цьому співпадає зі знаком множника a .

Рівність коефіцієнта кореляції нулю свідчить про відсутність лінійного статистичного зв'язку між випадковими величинами (при цьому говорять про їх некорельованість). Як видно з виразу (4.11), при цьому математичне сподівання добутку випадкових величин дорівнює добутку їх математичних сподівань $M\{X_1 X_2\} = M\{X_1\}M\{X_2\}$.

Легко показати, що із статистичної незалежності виходить некорельованість випадкових величин. Зворотне твердження в загальному випадку невірне – некорельовані випадкові величини можуть бути залежними.

Класичним прикладом останнього твердження є пара випадкових величин $x_1 = \cos \varphi$ і $x_2 = \sin \varphi$, де φ – випадкова величина, рівномірно розподілена на інтервалі $0..2\pi$. Очевидно, що x_1 і x_2 залежать одна від одної; проте їх коефіцієнт кореляції виявляється рівним нулю.

Для випадкових процесів можуть бути введені умовні густини ймовірності. Наприклад, коли значення процесу $X(t_1)$ в момент часу t_1 розглядається при відомому його значенні в момент t_2 маємо

$$p(x_1, t_1 | x_2, t_2) = p_2(x_1, x_2, t_1, t_2) / p_1(x_2, t_2),$$

де

$$p_1(x_2, t_2) = \int_{-\infty}^{\infty} p_2(x_1, x_2, t_1, t_2) dx_1.$$

Умовна густина ймовірності $p(x_1, t_1 | x_2, t_2)$ вміщує більше (загалом не менше) відомостей, ніж безумовна – $p_1(x_1, t_1)$. У разі, коли додаткова інформація про $X(t_2) = x_2$ не додає інформації про $X(t_1)$ виконується рівність $p(x_1, t_1 | x_2, t_2) = p(x_1, t_1)$ і маємо вираз

$$p_2(x_1, x_2, t_1, t_2) = p(x_1, t_1) p_1(x_2, t_2),$$

який є необхідною і достатньою умовою незалежності значень випадкового процесу в моменти часу t_1 і t_2 . Ця умова виконується для багатьох фізично існуючих процесів при $|t_2 - t_1| \rightarrow \infty$.

4.3. Стаціонарні та ергодичні випадкові процеси

В загальному випадку, як вже говорилося, ймовірнісні і кореляційні характеристики випадкових процесів залежать від одного або декількох моментів часу, в які ці характеристики визначаються. Проте існує клас випадкових процесів, у яких залежність характеристик від часу відсутня. Крім того, для деяких випадкових процесів не обов'язково робити усереднювання по ансамблю реалізацій – можна обмежитися розглядом однієї реалізації і її усереднюванням в часі.

Розглянемо такі випадкові процеси.

Стаціонарні випадкові процеси – це випадкові процеси, статистичні характеристики яких однакові в усіх часових перерізах.

Випадковий процес є *строго стаціонарним* (чи *стаціонарним у вузькому сенсі*), якщо його багатовимірна густина ймовірності $p(x_1, x_2, \dots, x_n, t_1, t_2, \dots, t_n)$ довільної розмірності n не змінюється при одночасному зсувові усіх

часових перерізів t_1, t_2, \dots, t_n уздовж осі часу на будь-яку однакову величину τ

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

Якщо ж обмежитись вимогами, щоб від часового зсуву не залежали лише одновимірна і двовимірна густини ймовірності, то такий випадковий процес буде *стаціонарний в широкому сенсі*.

Зрозуміло, що із стаціонарності у вузькому сенсі витікає стаціонарність в широкому сенсі, але не навпаки. Для стаціонарного випадкового процесу математичне сподівання і дисперсія не залежать від часу, а кореляційна функція залежить не від самих моментів часу, а тільки від інтервалу між ними $\tau = t_2 - t_1$: $R_x(t_1, t_2) = R_x(t_2 - t_1) = R_x(\tau)$.

З цієї причини при записі статистичних параметрів стаціонарного випадкового процесу можна опускати позначення фіксованих моментів часу: $m_x, D_x, K_x(\tau), R_x(\tau)$.

Як і для детермінованих сигналів, кореляційна функція стаціонарного випадкового процесу є парною

$$R_x(-\tau) = R_x(\tau).$$

Крім того, абсолютні значення цієї функції при будь-яких τ не перевищують її значення при $\tau = 0$ (нагадаємо, що це значення дорівнює дисперсії випадкового процесу):

$$|R_x(\tau)| \leq R_x(0) = D_x.$$

Часто зручно використовувати *коефіцієнт кореляції* (його також називають *нормованою кореляційною функцією*)

$$r_x(\tau) = R_x(\tau) / D_x.$$

Для коефіцієнта кореляції виконуються співвідношення $r_x(0) = 1, |r_x(\tau)| \leq 1, r_x(-\tau) = r_x(\tau)$.

Функції $R_x(\tau)$ і $r_x(\tau)$ характеризують зв'язок (кореляцію) між значеннями $X(t)$, розділеними проміжком τ . Чим повільніше убувають ці функції із зростанням абсолютного значення τ , тим більше проміжок, протягом якого

спостерігається статистичний зв'язок між миттєвими значеннями випадкового процесу, і тим повільніше, плавніше змінюються в часі його реалізації. Легко побачити, що розглянутий вище гармонічний випадковий процес з випадковою початковою фазою є стаціонарним в широкому сенсі. Дійсно, залежні від одновимірної густини ймовірності

математичне сподівання $m_x(x) = \int_0^{2\pi} A \cos(\omega_0 t + \varphi) \frac{1}{2\pi} d\varphi = 0$ і

дисперсія $D_x(t) = R_x(t, t) = \frac{A^2}{2}$ не залежать від часу, а

кореляційна функція $R_x(t_1, t_2) = K_x(t_1, t_2) = \frac{A^2}{2} \cos(\omega_0(t_1 - t_2))$,

звичай залежна від двовимірної густини ймовірності, залежить лише від інтервалу між даними моментами часу

$$R_x(\tau) = K_x(\tau) = \frac{A^2}{2} \cos(\omega_0 \tau).$$

Коефіцієнт кореляції такого випадкового процесу дорівнює

$$r_x(\tau) = \frac{R_x(\tau)}{D_x} = \cos(\omega_0 \tau).$$

Тут слід зазначити, що стаціонарним буде будь-який випадковий процес, реалізації якого є періодичними функціями, ідентичними за формою і лише з різними «початковими фазами», тобто положенням початку відліку часу в межах періоду. При цьому принциповою є рівномірність розподілу початкової фази в межах періоду. Дійсно, нехай у гармонічного процесу початкова фаза рівномірно розподілена в межах половини кола – на інтервалі $0 \dots \pi$. Математичне сподівання процесу в цьому випадку буде

$$m_x(t) = \int_{-\infty}^{\infty} x(t) p_{\varphi}(\varphi) d\varphi = \int_0^{\pi} A \cos(\omega_0 t + \varphi) \frac{1}{\pi} d\varphi = -\frac{2A \sin(\omega_0 t)}{\pi}.$$

Результат обчислень показує, що математичне сподівання процесу залежить від часу, отже, він не є стаціонарним.

Важливим прикладом стаціонарного багатовимірного випадкового сигналу є гаусівський випадковий процес. Він широко застосовується для описання випадкових сигналів при радіотехнічних вимірюваннях. За визначенням n -вимірна густина ймовірності такого процесу залежить від $n-1$ часових аргументів $\tau_i = t_i - t_1$, $i = 2, 3, \dots, n$ [10]:

$$p(x_1, \dots, x_n, \tau_1, \dots, \tau_{n-1}) = \frac{1}{\sigma^n (2\pi)^{n/2} |\mathbf{r}|^{1/2}} \times \exp \left[-\frac{1}{2|\mathbf{r}|\sigma^2} \sum_{i,j=1}^n A_{ij} (x_i - m)(x_j - m) \right], \quad (4.12)$$

де $|\mathbf{r}|$ – визначник матриці \mathbf{r}

$$\mathbf{r} = \begin{pmatrix} 1 & r_{12} & \dots & r_{1n} \\ r_{21} & 1 & \dots & r_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ r_{n1} & r_{n2} & \dots & 1 \end{pmatrix},$$

елементи якої визначаються нормованою функцією кореляції $r_{ij} = r(\tau_i - \tau_j) = R(\tau_i - \tau_j) / \sigma^2$, де

$$R(\tau_i - \tau_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - m)(x_j - m) p(x_1, \dots, x_n, t_1 + \tau_i, t_1 + \tau_j) dx_1 \dots dx_n,$$

тут також прийняті позначення $x_i = x(t_1 + \tau_i)$, $i = 2, 3, \dots, n$,

t_1 – довільний початковий момент часу,

A_{ij} – алгебраїчне доповнення елемента r_{ij} визначника $|\mathbf{r}|$.

Двовимірна густина для гаусівського процесу матиме вигляд

$$p(x_1, x_2, \tau) = \frac{1}{2\pi\sigma^2\sqrt{1-r^2(\tau)}} \times \\ \times \exp\left\{-\frac{(x_1-m)^2 - 2r(\tau)(x_1-m)(x_2-m) + (x_2-m)^2}{2\sigma^2[1-r^2(\tau)]}\right\}.$$

Подальше спрощення аналізу випадкових процесів досягається при виконанні умови ергодичності процесу.

Стационарний випадковий процес називається *ергодичним* (англ. ergodic), якщо при визначенні будь-яких його моментних функцій статистичне усереднення по ансамблю реалізацій еквівалентне усереднюванню за часом (англ. time averaging) однієї, теоретично нескінченно довгої, реалізації $x(t)$.

Позначивши усереднення кутовими дужками, запишемо вирази для обчислення важливих статистичних характеристик ергодичного процесу за його єдиною реалізацією $x(t)$ (ергодичний процес обов'язково є стаціонарним, але не навпаки)

$$m_x(x) = \langle x(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt.$$

$$D_x = \langle [x(t) - m_x]^2 \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt - m_x^2.$$

$$R_x(\tau) = \langle [x(t) - m_x][x(t-\tau) - m_x] \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t-\tau) dt - m_x^2$$

Математичне сподівання ергодичного випадкового процесу дорівнює постійній складовій будь-якої його

реалізації, а дисперсія має наочний фізичний зміст потужності флуктуаційної складової.

Достатньою умовою ергодичності випадкового процесу, стаціонарного в широкому сенсі, є наближення до нуля його кореляційної функції із зростанням часового зсуву τ :

$$\lim_{\tau \rightarrow \infty} R_x(\tau) = 0.$$

При експериментальному дослідженні випадкових процесів доступне, як правило, спостереження однієї реалізації сигналу, а не усього ансамблю. Якщо процес, що вивчається, є ергодичним, то реалізація достатньої довжини є «типovим» представником статистичного ансамблю. Згідно з приведеними вище формулами по цій єдиній реалізації можна визначити математичне сподівання, дисперсію і кореляційну функцію ергодичного випадкового процесу. На практиці інтегрування виконується, природно, не в нескінченних границях, а на кінцевому інтервалі, довжина якого має бути тим більше, чим вище вимоги до точності результатів вимірювання.

Як приклад перевіримо ергодичність вже розглянутого стаціонарного гармонічного процесу з випадковою початковою фазою. Його кореляційна функція

$$R_x(\tau) = K_x(\tau) = \frac{A^2}{2} \cos(\omega_0 \tau)$$

із зростанням τ не наближається до нуля, тобто умова $\lim_{\tau \rightarrow \infty} R_x(\tau) = 0$ не виконується. Проте це лише достатня, але не необхідна умова, тому її невиконання ще не означає неергодичності процесу. Перевіримо ергодичність згідно з визначенням, обчисливши усереднені за часом параметри:

$$\langle x(t) \rangle = \frac{\omega_0}{2\pi} \int_{-\pi/\omega_0}^{\pi/\omega_0} A \cos(\omega_0 \tau + \varphi) dt = 0 = m_x,$$

$$\begin{aligned} \langle [x(t) - m_x]^2 \rangle &= \frac{\omega_0}{2\pi} \int_{-\pi/\omega_0}^{\pi/\omega_0} A^2 \cos^2(\omega_0 t + \varphi) dt = \frac{A^2}{2} = D_x, \\ \langle [x(t) - m_x][x(t - \tau) - m_x] \rangle &= \\ &= \frac{\omega_0}{2\pi} \int_{-\pi/\omega_0}^{\pi/\omega_0} A^2 \cos(\omega_0 t + \varphi) \cos(\omega_0(t - \tau) + \varphi) dt = \\ &= \frac{A^2}{2} \cos(\omega_0 \tau) = R_x(\tau). \end{aligned}$$

Той факт, що реалізації даного процесу є періодичними функціями, дозволяє спростити обчислення, замінивши усереднювання по нескінченному проміжку часу усереднюванням по одному періоду, який в даному випадку дорівнює $2\pi/\omega_0$.

Отже, параметри, обчислені усереднюванням за часом, співпали з параметрами, отриманими раніше шляхом статистичного усереднювання. Тому, гармонічний випадковий процес з випадковою початковою фазою є ергодичним.

Слід також зазначити, що будь-який випадковий процес, реалізації якого є періодичними функціями, ідентичними за формою і такими, що відрізняються лише рівномірно розподіленою в межах періоду «початковою фазою», буде не тільки стаціонарним, але й ергодичним.

4.4. Нестационарні випадкові процеси

Нестационарними є всі випадкові процеси, для яких не виконуються умови стаціонарності. Загалом, це більшість процесів, які зустрічаються на практиці. В деяких порівняно простих випадках існує можливість перетворення нестационарного випадкового процесу в стаціонарний. Серед таких процесів виділимо:

1. *Адитивні нестационарні процеси* – це процеси виду $\eta(t) = \xi(t) + \varphi(t)$, де $\xi(t)$ – стаціонарний випадковий процес; $\varphi(t)$ – деяка детермінована функція за умови $\varphi(t) \neq \text{const}$.

2. *Мультиплікативні нестационарні процеси* описуються одним з наступних співвідношень: $\eta(t) = \xi(t)\varphi(t)$, де $\xi(t)$ – деякий стаціонарний випадковий процес; $\varphi(t)$ – довільна детермінована функція, $\varphi(t) \neq \text{const}$.

3. *Комбіновані нестационарні процеси* – це процеси виду $\eta(t) = \varphi_1(t)\xi(t) + \varphi_2(t)$, де $\xi(t)$ – стаціонарний випадковий процес; $\varphi_1(t), \varphi_2(t)$ – деякі детерміновані функції, для яких $\varphi_1(t) \neq \text{const}$ і $\varphi_2(t) \neq \text{const}$.

Процеси 1-3 – це такі процеси, в яких принципово можливо виділити, так званий, детермінований тренд і звести процес до стаціонарного.

4. *Нестационарні процеси із змінною в часі частотною структурою* $\eta(t) = \xi(t^n)$, де $\xi(t)$ – деякий стаціонарний випадковий процес, n – показник степені. Такі процеси можуть бути представлені як стаціонарні, наприклад у разі нерівномірного взяття відліків процесу у відповідності з залежністю t^n .

Комбінуванням елементарних моделей 1, 2, 4 можна досягти кращої відповідності для моделей представлення реальних нестационарних процесів.

Аналіз деяких нестационарних процесів можна спростити, якщо на обмежених часових ділянках вони наближаються до стаціонарних процесів. Це, так звані *квазістаціонарні* (майже стаціонарні) *процеси*, для яких справедливі співвідношення

$$M\{\xi(t)\} = \begin{cases} m_1, & \text{якщо } t_0 - T/2 \leq t \leq t_0 + T/2, \\ 0, & \text{при інших } t. \end{cases}$$

$$R_{\xi}(t_1, t_2) = \begin{cases} R_{\xi}(t_2 - t_1), & \text{якщо } t_0 - T/2 \leq (t_1, t_2) \leq t_0 + T/2, \\ 0, & \text{при інших } (t_1, t_2). \end{cases}$$

Тут T – інтервал фактичного існування випадкової функції $\xi(t)$; t_0 – деякий довільний момент часу.

4.5. Спектральні характеристики випадкових процесів

Кожна окрема реалізація випадкового процесу є детермінованою функцією, до якої можна застосувати перетворення Фур'є. При цьому різні реалізації будуть, природно, мати різні спектри. Нас же цікавлять статистично усереднені характеристики випадкових процесів. Спробуємо знайти середнє значення спектральної густини випадкового процесу (горизонтальною рискою тут і далі позначається операція статистичного усереднювання по ансамблю реалізацій)

$$\overline{S_x(\omega)} = \overline{\int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt} = \int_{-\infty}^{\infty} \overline{x(t)}e^{-j\omega t} dt = \int_{-\infty}^{\infty} m_x(t)e^{-j\omega t} dt$$

Як видно, усереднена спектральна густина випадкового процесу є спектром його детермінованої складової (математичного сподівання). Для центрованих процесів $m_x(t) = 0$ і $\overline{S_x(\omega)} = 0$. Таким чином, усереднене значення спектральної густини не несе ніякої інформації про флуктуаційну, тобто власне випадкову складову процесу. Це свідчить про випадковість і незалежність фаз спектральних складових в різних реалізаціях випадкового процесу.

Можна, проте, розглянути спектральну густина потужності випадкового процесу, оскільки потужність не залежить від співвідношення фаз спектральних складових.

Розглянемо центрований випадковий процес і виділимо з його ансамблю яку-небудь реалізацію $x(t)$ на обмеженому інтервалі часу $[-T/2, T/2]$. Застосувавши до цієї реалізації

перетворення Фур'є, знайдемо її спектральну густину $\dot{X}_T(\omega)$. Енергію E_T даного відрізка реалізації згідно з рівністю Парсеваля (1.24) можна обчислити як

$$E_T = \int_{-T/2}^{T/2} x^2(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\dot{X}_T(\omega)|^2 d\omega.$$

Розділивши цю енергію на T , отримаємо *середню потужність* P_T реалізації на цьому часовому інтервалі

$$P_T = \frac{E_T}{T} = \left\langle x^2(t) \right\rangle_{|t| \leq T/2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|\dot{X}_T(\omega)|^2}{T} d\omega.$$

Усереднивши цю потужність по ансамблю реалізацій при одночасному збільшенні тривалості проміжку часу ($T \rightarrow \infty$) отримаємо спектральну густину середньої потужності

$$M \left\{ \left\langle x^2(t) \right\rangle \right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} M \left\{ \lim_{T \rightarrow \infty} \frac{|\dot{X}_T(\omega)|^2}{T} \right\} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} W(\omega) d\omega,$$

де функція $W(\omega) = M \left\{ \lim_{T \rightarrow \infty} \frac{|\dot{X}_T(\omega)|^2}{T} \right\}$ є *спектральною*

густиною середньої потужності розглянутої реалізації (*англ. термін* - power spectral density, PSD), застосовуються також терміни «*спектральна густина потужності*» або «*спектр потужності*».

У випадку ергодичного процесу функція $W(\omega)$ може бути знайдена без усереднювання по ансамблю реалізацій шляхом усереднювання за часом однієї реалізації, яка в цьому випадку характеризуватиме весь процес в цілому. Середня потужність будь-якої його реалізації дорівнює дисперсії процесу

$$D_x = \frac{1}{2\pi} \int_{-\infty}^{\infty} W(\omega) d\omega.$$

Функція $W(\omega)$ – дійсна, вона не містить інформації про фази спектральних складових і не дозволяє відновити окремі реалізації випадкового процесу. Крім того, з визначення

спектральної густини очевидно, що $W(\omega)$ є невід'ємною і парною функцією частоти.

Тут не наводяться приклади розрахунку спектру випадкового процесу згідно з його визначенням, оскільки такого роду розрахунок рідко потрібний на практиці. Як правило, обчислення спектру випадкового процесу виконується на основі його кореляційної функції за допомогою теорему Вінера-Хинчина, мова про яку піде далі.

Як розподіл спектральної густини потужності, так і вид кореляційної функції пов'язані зі швидкістю зміни випадкового процесу в часі. Знайдемо зв'язок між цими двома характеристиками.

Як відомо, кореляційна функція детермінованого сигналу пов'язана перетворенням Фур'є з його енергетичним спектром. Застосуємо цю властивість до відрізка реалізації випадкового процесу тривалістю T

$$\int_{-T/2}^{T/2} x(t)x(t-\tau)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\dot{X}_T(\omega)|^2 e^{j\omega\tau} d\omega.$$

Розділимо обидві частини цієї рівності на T і спрямуємо T до нескінченності

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t-\tau)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{|\dot{X}_T(\omega)|^2}{T} e^{j\omega\tau} d\omega. \quad (4.12)$$

Якщо вважати даний процес ергодичним, то в лівій частині останньої рівності стоїть кореляційна функція процесу, отримана шляхом усереднювання за часом. У правій частині під інтегралом міститься вираз $W(\omega)$ для спектральної густини потужності випадкового процесу. З урахуванням цього

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W(\omega) e^{j\omega\tau} d\omega. \quad (4.13)$$

У разі неергодичного процесу до обох частин рівності (4.12) необхідно додатково застосувати усереднювання по

ансамблю реалізацій, що приведе до того ж самому результату.

Таким чином, кореляційна функція центрованого стаціонарного випадкового процесу і його спектральна густина потужності пов'язані перетвореннями Фур'є. Тобто при відомій кореляційній функції спектральна густина потужності може бути визначена як

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W(\omega) e^{j\omega\tau} d\omega. \quad (4.14)$$

Співвідношення (4.13) і (4.14) складають зміст *теорема Вінера-Хинчина*, доведеної в тридцять років минулого століття незалежно один від одного американським і радянським математиками Н. Вінером і О. Я. Хинчиним.

Оскільки як $W(\omega)$, так і $R(\tau)$ є парними дійсними функціями, можна відмовитись від комплексної форми запису перетворення Фур'є і перейти до напівнескінчених границь інтегрування:

$$R(\tau) = \frac{1}{\pi} \int_0^{\infty} W(\omega) \cos(\omega\tau) d\omega, \quad W(\omega) = 2 \int_0^{\infty} R(\tau) \cos(\omega\tau) d\tau. \quad (4.15)$$

При $\tau = 0$ маємо $R(0) = \frac{1}{\pi} \int_0^{\infty} W(\omega) d\omega = \sigma^2$. Дисперсія стаціонарного випадкового процесу, таким чином, є сумою внесків від усіх ділянок частотної вісі. На відміну від детермінованих імпульсних сигналів $W(\omega)$ характеризує питому міру потужності, а не частку енергії, яка припадає на одиничну смугу частот.

Теорема *Вінера-Хинчина* у вигляді (4.15) стосується випадкових процесів з нульовим середнім. У загальному ж випадку існує зв'язок

$$W_x(\omega) = \int_{-\infty}^{\infty} K_x(\tau) \exp(-\omega\tau) d\tau, \quad K_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W_x(\omega) e^{j\omega\tau} d\omega. \quad (4.16)$$

Для зручності розрахунків також вводять односторонній спектр потужності $F(\omega)$ випадкового процесу

$$F(\omega) = \begin{cases} 0, & \omega < 0, \\ W(\omega)/\pi, & \omega \geq 0. \end{cases} \quad (4.17)$$

Представлення $F(\omega)$ за (4.17) дозволяє знаходити дисперсію стаціонарного випадкового процесу інтегруванням тільки по фізично існуючим додатним частотам

$$\sigma^2 = R(0) = \int_0^{\infty} F(\omega) d\omega. \quad (4.18)$$

В техніці також широко застосовується односторонній спектр потужності $N(f)$, який характеризує середню потужність випадкового процесу на кожному з інтервалів частот в 1 Гц

$$N(f) = \begin{cases} 0, & f < 0, \\ 2W(2\pi f), & f \geq 0. \end{cases} \quad (4.19)$$

В цьому випадку

$$\sigma^2 = \int_0^{\infty} N(f) df. \quad (4.20)$$

Випадкові процеси, які розглядаються в теорії сигналів, часто мають функцію кореляції, що наближається до нуля із збільшенням часового зсуву τ

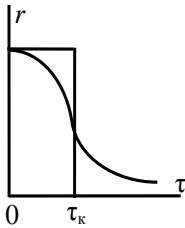


Рисунок 4.5 – Площі обох фігур однакові

(нагадаємо, що це є достатньою умовою ергодичності процесу) (рис. 4.5). Чим швидше убуває функція $R(\tau)$, тим слабкішим є статистичний зв'язок між миттєвими значеннями випадкового сигналу в два не співпадаючі моменти часу. Числовою характеристикою для оцінки «швидкості зміни» реалізацій

випадкового процесу слугує *інтервал кореляції* τ_k , який визначається за формулою

$$\tau_k = \frac{1}{R(0)} \int_0^{\infty} |R(\tau)| d\tau = \int_0^{\infty} |r(\tau)| d\tau.$$

Якщо відома інформація про поведінку якої-небудь реалізації випадкового процесу «у минулому», то можливий ймовірнісний прогноз випадкового процесу на час порядку τ_k . Прогнозування на інтервал часу, який суттєво перевищує τ_k виявляється безрезультатним, оскільки миттєві значення стають практично некорельованими.

З іншого боку, досліджуваний випадковий процес характеризується спектром густини потужності $W(\omega)$, що має максимальне значення W_{\max} . Подібно до методики знаходження інтервалу кореляції замінимо подумки цей випадковий процес іншим, у якого спектральна густина потужності постійна і дорівнює W_{\max} в межах деякої смуги частот, яка обрана з умови рівності дисперсій (тобто середніх потужностей) обох процесів. Ширина цієї смуги частот називається *ефективною шириною спектру* випадкового процесу і має вигляд:

$$W_{\max} \Delta\omega_{\text{еф}} = \int_{-\infty}^{\infty} W(\omega) d\omega, \quad \Delta\omega_{\text{еф}} = \frac{1}{W_{\max}} \int_{-\infty}^{\infty} W(\omega) d\omega.$$

Ефективну ширину спектру випадкового процесу можна визначити й іншими способами, наприклад, виходячи з умови зменшення значень спектру потужності на межі цього частотного інтервалу до рівня $0,1 \cdot W_{\max}$. У будь-якому випадку величини τ_k і $\Delta\omega_{\text{еф}}$ пов'язані відомим з властивостей перетворення Фур'є співвідношенням невизначеності $\Delta\omega_{\text{еф}} \tau_k \sim 2\pi$. Для ілюстрації цього співвідношення на рис. 4.6 наведені приклади реалізацій двох випадкових процесів (рис. 4.6, а) та відповідні їм кореляційні функції (рис. 4.6, б) і спектри густини середньої потужності (рис. 4.6, в).

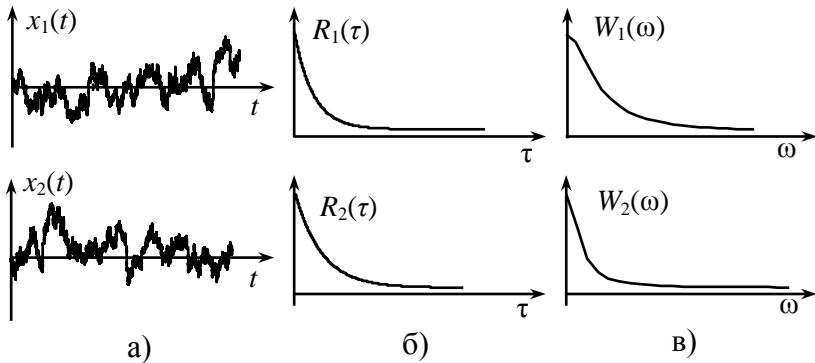


Рисунок 4.6 – Взаємзв’язок між видом реалізацій випадкових процесів (а), їх кореляційними функціями (б) і спектрами (в)

4.6. Приклади стаціонарних випадкових процесів

Білий шум

В радіовимірюваннях та радіотехніці «білим шумом» (*англ.* white noise), подібно до білого світла (воно має постійну інтенсивність у межах видимого діапазону), називають стаціонарний випадковий процес, спектральна густина потужності якого постійна на всіх частотах

$$W(\omega) = W_0 = \text{const}.$$

Згідно з теоремою Вінера-Хинчина кореляційна функція білого шуму є дельта-функцією

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

тобто дорівнює нулю усюди, окрім точки $\tau=0$. Дисперсія білого шуму нескінченно велика.

Значення білого шуму у різні моменти часу некорельовані, оскільки, яким би не був малим інтервал τ , сигнал за цей час може змінитися на будь-яку величину.

Білий шум є абстрактною математичною моделлю і фізично існувати не може. Це пояснюється передусім

нескінченністю його дисперсії (тобто середньої потужності). Проте в тих випадках, коли смуга пропускання досліджуваної системи істотно вужча за ефективну ширину спектру шуму, який на неї впливає, можна для спрощення аналізу приблизно замінити реальний випадковий процес білим шумом.

Вузькосмуговий випадковий процес – особливий клас випадкових процесів, спектр яких зосереджений у відносно вузькій смузі поблизу деякої частоти ω_0 (рис. 4.7). Він відіграє важливу роль в радіовимірюваннях. Розглянемо статистичні властивості таких процесів.

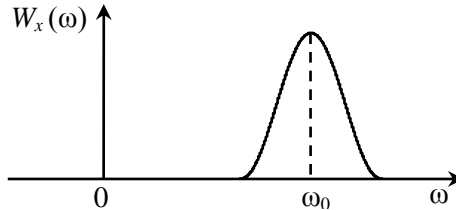


Рисунок 4.7 – Спектральна густина потужності вузькосмугового випадкового процесу

Отже, нехай $x(t)$ – стаціонарний, нормальний і центрований вузькосмуговий випадковий процес, спектр густини потужності якого $W_x(\omega)$ (рис. 4.7).

Згідно з теоремою Вінера-Хинчина кореляційна функція і спектральна густина потужності випадкового процесу пов'язані між собою перетворенням Фур'є. Вузькосмуговий характер спектру $W_x(\omega)$ вказує на те, що кореляційна функція $R_x(\tau)$ має вигляд вузькосмугового радіосигналу

$$R_x(\tau) = R_0(\tau) \cos[\omega_0 \tau + \varphi_0(\tau)],$$

де $R_0(\tau)$ і $\varphi_0(\tau)$ – змінюються повільно (в порівнянні з функцією $\cos(\omega_0 \tau)$). Вузькосмуговий спектр і осцилюючий характер кореляційної функції означають, що окремі

реалізації вузькосмугового випадкового процесу є квазігармонічними коливаннями

$$x(t) = A(t) \cos[\omega_0 t + \varphi(t)],$$

у яких обвідна $A(t)$, як і початкова фаза $\varphi(t)$ є випадковими функціями, що змінюються в часі повільно в порівнянні з $\cos(\omega_0 t)$. Щоб визначити статистичні параметри обвідної і початкової фази, розглянемо комплексний аналітичний сигнал $\dot{z}(t)$, який відповідає дійсному випадковому процесу $x(t)$

$$\dot{z}(t) = x(t) + jx_{\perp}(t),$$

де $x_{\perp}(t)$ – спряжений випадковий процес, реалізації якого пов'язані з реалізаціями процесу $x(t)$ перетворенням Гільберта

$$x_{\perp}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(t')}{t-t'} dt'.$$

За допомогою спряженого сигналу можна визначити миттєві значення обвідної і повної фази вузькосмугового сигналу:

$$A(t) = |\dot{z}(t)| = \sqrt{x^2(t) + \hat{x}^2(t)},$$

$$j(t) = \arg \dot{z}(t) = \begin{cases} \arctg[\hat{x}(t)/x(t)], & x(t) \geq 0, \\ \arctg[\hat{x}(t)/x(t)] + \pi, & x(t) < 0. \end{cases}$$

Розглянемо статистичні властивості спряженого процесу. По-перше, визначимо його математичне сподівання, застосувавши усереднювання до формули перетворення Гільберта і потім помінявши усереднювання й інтегрування місцями:

$$M\{x_{\perp}(t)\} = M\left\{\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(t')}{t-t'} dt'\right\} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{M\{x(t')\}}{t-t'} dt' = 0.$$

Результат дорівнює нулю, оскільки процес $x(t)$ є центрованим.

Далі, оскільки процес $x(t)$ нормальний, а перетворення Гільберта є лінійним інтегральним перетворенням, то нормальним буде і спряжений процес $x_{\perp}(t)$.

З властивостей перетворення Гільберта випливає, що спектри конкретних реалізацій процесів $x(t)$ й $x_{\perp}(t)$ пов'язані наступним чином:

$$\dot{S}_{x_{\perp}}(\omega) = \begin{cases} -j\dot{S}_x(\omega), & \omega > 0, \\ 0, & \omega = 0, \\ j\dot{S}_x(\omega), & \omega < 0, \end{cases}$$

звідки видно, що енергетичні спектри реалізацій процесів $x(t)$ й $x_{\perp}(t)$ збігаються, а отже, збігаються і спектральні густини потужностей цих процесів: $W_{x_{\perp}}(\omega) = W_x(\omega)$. Кореляційні функції пов'язані зі спектрами густини потужності оберненим перетворенням Фур'є, тому вони теж рівні: $R_{x_{\perp}}(\tau) = R_x(\tau)$.

Нам залишилося з'ясувати, чи є статистичний зв'язок між процесами $x(t)$ і $x_{\perp}(t)$. Обмежимося при цьому розрахунком кореляції між ними в співпадаючі моменти часу, тобто обчислимо $R_{xx_{\perp}}(0)$:

$$\begin{aligned} R_{xx_{\perp}}(0) &= M\{x(t)x_{\perp}(t)\} = M\left\{x(t)\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x(t')}{t-t'}dt'\right\} = \\ &= M\left\{\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x(t)x(t')}{t-t'}dt'\right\}. \end{aligned}$$

Далі, як і раніше, внесемо операцію статистичного усереднення під знак інтегралу, а потім використовуємо заміну змінної $\tau = t - t'$:

$$R_{xx_{\perp}}(0) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{M\{x(t)x(t')\}}{t-t'} dt' = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{R_x(t-t')}{t-t'} dt' =$$

$$= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{R_x(\tau)}{\tau} dt' = 0.$$

Результат інтегрування дорівнює нулю, оскільки $R_x(\tau)$ є парною функцією, а весь підінтегральний вираз, отже, - непарної. Таким чином, процеси $x(t)$ і $x_{\perp}(t)$ в співпадаючі моменти часу некорельовані. Оскільки вони, крім того, є нормальними, то з некорельованості витікає статистична незалежність.

Відзначимо ще раз, що дисперсія випадкового процесу може бути обчислена двома способами:

1. На основі відомої одновимірної густини ймовірності

$$D_x(t) = M\{[X(t) - m_x(t)]^2\} = \int_{-\infty}^{\infty} x^2 p(x,t) dx - m_x^2(t).$$

2. За відомою кореляційною функцією

$$R_x(t, t) = K_x(t, t) - m_x^2(t) = D_x(t).$$

При цьому випадковий процес, наприклад, гаусівський може мати різного виду спектральну густину потужності, а білий шум – довільну функцію розподілу.

Контрольні запитання

1. Дайте визначення поняттю «випадковий процес».
2. Що необхідно знайти для аналізу властивостей і характеристик випадкового сигналу?
3. Які випадкові процеси називають випадковими полями?
4. Наведіть класифікацію випадкових процесів.
5. Перелічіть імовірнісні характеристики, які можуть бути отримані в одному перерізі випадкового процесу (сигналу) та наведіть формули для їх обчислення.

6. Які існують функціональні характеристики випадкових сигналів?

7. Дайте визначення функції розподілу ймовірності.

8. Яка функція є похідною від функції розподілу ймовірності?

9. Які ви знаєте найбільш поширені в практичних застосуваннях числові характеристики для перерізу випадкового сигналу?

10. Дайте визначення математичного сподівання випадкового сигналу.

11. Приведіть ймовірнісні характеристики найбільш поширених в метрологічній практиці рівномірного і нормального законів розподілу випадкових величин.

12. Яка характеристика випадкового процесу є вичерпною для її описання.

13. Дайте визначення ковариаційної та кореляційної функцій. В чому їх відмінність?

14. Виведіть співвідношення для кореляційної функції періодичного випадкового сигналу, для якого рівноймовірним є часовий зсув в межах його періоду (скористатись розкладанням сигналу в ряд Фур'є та співвідношенням для кореляційної функції гармонічного сигналу).

15. Що називають умовою статистичної незалежності?

16. Запишіть співвідношення для коефіцієнта кореляції та інтервалу кореляції.

17. Чи можуть бути некорельованими випадкові величини, які є функціонально залежними.

18. Дайте визначення стаціонарного випадкового процесу. В чому відмінність стаціонарних в узькому та широкому сенсі процесів?

19. Наведіть приклади стаціонарних випадкових процесів.

20. Дайте визначення нормованої кореляційної функції стаціонарного випадкового сигналу та опишіть її властивості.

21. Дайте визначення ергодичного випадкового процесу. Достатня умова ергодичності випадкового процесу стаціонарного в широкому сенсі.

22. Дайте визначення нестационарного випадкового процесу та наведіть їх приклади.

23. Сформулюйте теорему Вінера-Хинчина та запишіть співвідношення, які їй відповідають.

24. Дайте визначення ефективної ширини спектру випадкового процесу. Як вона застосовується у співвідношенні невизначеності?

Глава 5

ДИСКРЕТНІ СИГНАЛИ ТА МЕТОДИ ЇХ АНАЛІЗУ

В поточній главі розглядаються найбільш поширені методи аналізу дискретних сигналів. Дискретне перетворення Фур'є (ДПФ) є одним з основних перетворень при цифровій обробці сигналів. Воно частіш за все обчислюється швидкими методами, що дозволяє застосовувати його в цифрових аналізаторах спектру та при обробці сигналів. Зручним способом аналізу дискретних послідовностей є z -перетворення. При z -перетворенні різницеві рівняння, що описують роботу дискретної системи, перетворюються в рівняння алгебраїчні, з якими простіше виконувати необхідні дії. Для дискретних сигналів і систем z -перетворення відіграє таку ж роль, як перетворення Лапласа – для аналогових сигналів.

5.1. Моделі дискретних сигналів. Дискретизація періодичних сигналів. Дискретне перетворення Фур'є

За класифікацією, яка була проведена в главі 1, розглядалися відмінності аналогових і дискретних сигналів. Математична модель аналогового сигналу відповідає неперервній або кусково-неперервній функції, а дискретний сигнал $x_d(t)$ є послідовністю відлікових значень $\{\dots, x_{-1}, x_0, x_1, x_2, \dots\}$ сигналу $x(t)$ у відповідні моменти часу $\{\dots, t_{-1}, t_0, t_1, t_2, \dots\}$. Зазвичай часові відліки обираються рівномірно у часі з інтервалом (шагом) дискретизації, наприклад, $T = t_i - t_{i-1}$, де i – довільний номер.

Для отримання дискретного представлення $\{\dots, x_{-1}, x_0, x_1, x_2, \dots\}$ з аналогового сигналу $x(t)$ потрібно знайти скалярний добуток цього сигналу з узагальненою функцією

$\eta(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT)$, яка має назву дискретизувальної послідовності. Таким чином, маємо [17]

$$x_d(t) = (x, \eta) = \int_{-\infty}^{\infty} x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT) dt = \{ \dots, x(-T), x(0), x(T), x(2T), \dots \} \quad (5.1)$$

Виходячи з (5.1), для здійснення дискретизації сигналу виконується операція стробування, тобто обчислення добутку сигналу $x(t)$ та «гребінчастої» функції $\eta(t)$.

На практиці функція $\eta(t)$ складається з стробувальних імпульсів, які мають кінцеву тривалість. Тому послідовність часових відліків фактично представляє собою модульовану послідовність коротких імпульсів. На рис. 5.1 показано принцип отримання такої послідовності [17].

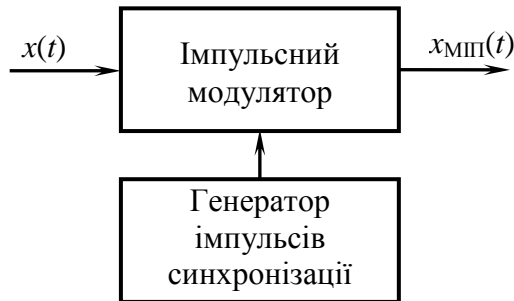


Рисунок 5.1 – Принцип отримання модульованої імпульсної послідовності

Імпульсний модулятор має два входи. На перший вхід подається вихідний аналоговий сигнал $x(t)$. При потраплянні на його другий вхід послідовності імпульсів синхронізації з періодом T виконується вимірювання миттєвих значень сигналу $x(t)$. В результаті на виході модулятора спостерігається *модульована імпульсна послідовність* $x_{МП}(t)$.

Кожен з імпульсів послідовності має площу пропорційну відповідному відліковому значенню сигналу $x(t)$. Якщо тривалість кожного імпульсу є однаковою, тоді амплітуди імпульсів пропорційні відліковому значенню. Таке перетворення аналогового неперервного сигналу має назву *амплітудно-імпульсної модуляції (АІМ)*. Навпаки, якщо амплітуди імпульсів на виході модулятора залишаються незмінними, а їх тривалість змінюється пропорційно миттєвим значенням аналогового сигналу, така модуляція називається *широтно-імпульсною (ШІМ, англ. PWM – Pulse Width Modulation)*. Обидва види модуляції використовуються на практиці.

Розглянемо як початкові дані послідовність N дискретних відліків $x(kT)$, заданих на відрізку $[0, T_1]$, $k = \overline{0, N-1}$. Математичною моделлю послідовності таких відліків є АІМ сигнал зі зсунутих за часом δ -функцій

$$x_{\text{МІП}}(t) = T \sum_{k=0}^{N-1} x(kT) \delta(t - kT), \quad (5.2)$$

яка подає динамічне представлення сигналу $x(t) = \int_{-\infty}^{\infty} x(\tau) \delta(t - \tau) d\tau$ (див. розділ 1.3).

Продовжимо подумки цей сигнал з періодом $T_1 = NT$. Оскільки дискретний сигнал (5.2) є періодичним, його спектр має бути дискретним з відстанню між гармоніками $2\pi/(NT)$. Цей дискретний періодичний сигнал можна представити рядом Фур'є

$$x_{\text{дп}}(t) = \sum_{n=-\infty}^{\infty} \dot{c}(n) e^{jn\omega_1 t}. \quad (5.3)$$

Коефіцієнти $c(n)$ ряду (5.3) обчислюються за формулою

$$\begin{aligned} \dot{c}(n) &= \frac{1}{NT} \int_0^{NT} x_{\text{МПП}}(t) e^{-jn\omega_1 t} dt = \frac{1}{NT} \int_0^{NT} \sum_{k=0}^{N-1} x(kT) \delta(t - kT) e^{-jn\omega_1 t} dt = \\ &= \frac{1}{NT} \sum_{n=0}^{N-1} x(kT) \int_0^{NT} \delta(t - kT) e^{-jn\omega_1 t} dt = \frac{1}{N} \sum_{n=0}^{N-1} x(k) e^{-jn\omega_1 kT} \end{aligned} \quad (5.4)$$

При перетвореннях в (5.4) використана фільтрувальна властивість δ -функції та позначено $x(kT) = x(k)$. Оскільки

$$\omega_1 = \frac{2\pi}{T_1} = \frac{2\pi}{NT}, \text{ остаточно маємо}$$

$$\dot{c}(n) = \frac{1}{N} \sum_{k=0}^{N-1} x(k) e^{-j \frac{2\pi}{N} kn}, \quad (5.5)$$

де $n = \overline{0, N-1}$.

Співвідношення (5.5) дозволяє обчислити комплексні амплітуди гармонік дискретного сигналу. Воно є лінійною комбінацією відліків цього сигналу і має назву *дискретного перетворення Фур'є (ДПФ)*.

Разом з ДПФ існує й *обернене дискретне перетворення Фур'є*

$$x(k) = \sum_{n=0}^{N-1} \dot{c}(n) e^{j \frac{2\pi}{N} kn}, \quad k = \overline{0, N-1}. \quad (5.6)$$

Зауважимо, що стосовно розміщення множника $1/N$ саме у виразі (5.5) немає повної єдності. У багатьох джерелах цей множник відносять до формули оберненого ДПФ, видаляючи його з формули для прямого ДПФ. В подальшому також буде застосовуватись таке представлення.

Ортогональний дискретний базис Фур'є, в якому виконується ДПФ, є системою дискретних експоненціальних функцій, заданою на дискретній часовій осі N відліками:

$$e_N(k, n) = \exp(j \frac{2\pi}{N} kn); \quad k, n = \overline{0, N-1}. \quad (5.7)$$

Система функцій (5.7) є набором N експонент з частотами кратними основній частоті $2\pi/N$, оскільки $e_N(k, n)$ періодична по k з періодом N .

Розглянемо *основні властивості ДПФ* [9, 11]:

1. Дискретне перетворення Фур'є – *лінійне перетворення*, тобто якщо послідовностям $x(k)$ і $y(k)$ з одним і тим же періодом N відповідають набори гармонік $\dot{c}_1(n)$ і $\dot{c}_2(n)$, тоді послідовності $\{ax(k)+by(k)\}$ відповідатиме спектр $\{a\dot{c}_1(n)+b\dot{c}_2(n)\}$. Ця властивість є наслідком лінійності операції підсумовування, яка використовується при обчисленні ДПФ.

2. *Властивість симетрії*, яку має спектр неперервного сигналу, зберігається і для спектру дискретного періодичного сигналу. Якщо відліки $x(k)$ – дійсні числа, тоді коефіцієнти ДПФ, номери яких розташовані симетрично відносно $N/2$, утворюють спряжені пари:

$$\dot{c}(N-n) = \frac{1}{N} \sum_{k=0}^{N-1} x(k) e^{-j\frac{2\pi}{N}k(N-n)} = \frac{1}{N} \sum_{k=0}^{N-1} x(k) e^{j\frac{2\pi}{N}kn} = \dot{c}^*(n). \quad (5.8)$$

З формули (5.8) виходить, що спектр є спряжено симетричним відносно $N/2$, тобто містить рівно таку ж кількість інформації, що і сам сигнал. Дійсно, якщо початковий сигнал представляється набором з N дійсних чисел, то його спектр представляється набором з $N/2$ комплексних чисел, кожне з яких з інформаційної точки зору еквівалентне двом дійсним. Друга половина спектру взаємно-однозначно пов'язана з першою. Можна вважати, що коефіцієнти $\dot{c}(N/2+1), \dot{c}(N/2+2), \dots, \dot{c}(N-1)$ відповідають від'ємним частотам. При вивченні спектру дійсного сигналу вони не дають нової інформації.

Гармоніка з нульовим номером (постійна складова), як випливає з (5.5), є середнім значенням усіх відліків сигналу на одному періоді

$$\dot{c}(0) = \frac{1}{N} \sum_{k=0}^{N-1} x(k) = \dot{c}(N). \quad (5.9)$$

Якщо N – парне число, тоді

$$\dot{c}\left(\frac{N}{2}\right) = \frac{1}{N} \sum_{k=0}^{N-1} x(k)(-1)^k. \quad (5.10)$$

і амплітуда гармоніки з номером $N/2$ визначається сумою відліків зі знаками, що чергуються,

$$\dot{c}\left(\frac{N}{2}\right) = \frac{1}{N} [x(0) - x(1) + \dots + x(N-2) - x(N-1)].$$

3. ДПФ кругової згортки

Візьмемо дві послідовності $\{x_1(k)\}$ і $\{x_2(k)\}$ однакової довжини N , ДПФ яких відповідно рівні $\dot{c}_1(n)$ і $\dot{c}_2(n)$. Обчислимо їх кругову згортку по одному періоду

$$y(n) = \sum_{m=0}^{N-1} x_1(m)x_2(n-m). \quad (5.11)$$

Термін «кругова згортка» введений у зв'язку з прийнятою періодичністю послідовностей $\{x_1(k)\}$ і $\{x_2(k)\}$. Завдяки цьому значення $x_2(n-m)$, якщо $(n-m) < 0$, дорівнюють $x_2(N+n-m)$, тобто фактично перший і останній члени послідовності йдуть послідовно («коло замикається»).

Знайдемо N -точкове ДПФ цієї згортки

$$\begin{aligned} s(k) &= \frac{1}{N} \sum_{n=0}^{N-1} y(n) e^{-j\frac{2\pi nk}{N}} = \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \left\{ \sum_{m=0}^{N-1} x_1(m)x_2(n-m) \right\} e^{-j\frac{2\pi nk}{N}} = \\ &= \frac{1}{N} \sum_{m=0}^{N-1} x_1(m) \left\{ \sum_{n=0}^{N-1} x_2(n-m) e^{-j\frac{2\pi nk(n-m)}{N}} \right\} e^{-j\frac{2\pi km}{N}} = \\ &= \dot{c}_1(k)\dot{c}_2(k), \quad k = 0, 1, \dots, N-1. \end{aligned} \quad (5.12)$$

При отриманні формули (5.12) врахована властивість зсуву періодичної послідовності. Таким чином, круговій згортці дискретизованих сигналів заданих на одному часовому проміжку відповідає перемножування їх спектрів.

Обчислення кругової згортки двох сигналів за допомогою ДПФ здійснюється за наступним алгоритмом:

обчислення ДПФ вихідних сигналів за формулою (5.5);

перемножування коефіцієнтів отриманих ДПФ згідно (5.12);

обчислення сигналу $y(n)$ за допомогою оберненого ДПФ (5.6) отриманої послідовності $\{s(k)\}$.

4. Рівність Парсеваля для дискретних сигналів

Визначимо значення $\sum_{n=0}^{n-1} |\dot{c}(n)|^2$, використовуючи формулу

ДПФ

$$\begin{aligned} \sum_{n=0}^{n-1} |\dot{c}(n)|^2 &= \sum_{n=0}^{N-1} \left\{ \frac{1}{N} \sum_{k=0}^{N-1} x(k) e_N^*(k, n) \frac{1}{N} \sum_{m=0}^{N-1} x^*(m) e_N(m, n) \right\} = \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \sum_{m=0}^{N-1} x(k) x^*(m) \left\{ \frac{1}{N} \sum_{n=0}^{N-1} e_N^*(k, n) e_N(m, n) \right\} = \frac{1}{N} \sum_{k=0}^{N-1} |x(k)|^2. \end{aligned} \quad (5.13)$$

При виведенні формули (5.13) використана умова ортогональності дискретних експоненціальних функцій

$$\sum_{n=0}^{N-1} e_N^*(k, n) e_N(m, n) = \begin{cases} N, & k = m \\ 0, & k \neq m \end{cases} \quad (5.14)$$

Таким чином, середня потужність сигналу $x(k)$ на N відліках дорівнює сумі потужностей його частотних компонентів.

5. Зв'язок ДПФ і спектру дискретного сигналу

Маючи один і той самий набір значень дискретного сигналу $x(k)$, можна розрахувати або *спектральну функцію*

$\dot{X}(j\omega)$ цього дискретного сигналу по формулі

$$\dot{X}(j\omega) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n}, \text{ або його ДПФ за формулою (5.5).}$$

Порівняння цих формул показує, що ДПФ є просто дискретними відліками спектральної функції дискретного сигналу, які відповідають частотам $\omega_n = n\omega_d / N$:

$$\dot{c}(n) = X\left(\frac{2\pi n}{NT}\right) = X\left(\omega_d \frac{n}{N}\right). \quad (5.15)$$

Із співвідношення (5.15) маємо важливий висновок: якщо додати до кінцевого набору відліків деяку кількість нулів, спектральна функція дискретного сигналу, природно, не зміниться, але ДПФ дасть більше число спектральних відліків, які відповідатимуть частотам більш щільно розташованим в інтервалі від нуля до частоти дискретизації.

б. Обчислення оберненого ДПФ за допомогою прямого ДПФ

Для обчислення оберненого ДПФ можна без будь-яких змін застосовувати пряме ДПФ. Обернене ДПФ N -точкової послідовності $\{\dot{c}(n)\}$, визначається наступним чином

$$x(k) = \sum_{n=0}^{N-1} \dot{c}(n) e^{j\frac{2\pi}{N}kn}, \quad k = 0, 1, \dots, N-1. \quad \text{Якщо взяти комплексне}$$

спряження від цього виразу та поділити його на N матимемо

$$\frac{1}{N} x^*(k) = \frac{1}{N} \sum_{n=0}^{N-1} \dot{c}^*(n) e^{-j\frac{2\pi}{N}kn}. \quad (5.16)$$

Права частина цього виразу є ДПФ послідовності $\{\dot{c}^*(n)\}$, яке може бути обчислене з використанням одного з алгоритмів обчислення прямого ДПФ (наприклад, одного з ШПФ). Шукану послідовність $\{x(k)\}$ можна отримати, взявши комплексно спряжений з (5.16) вираз і помножити його на N , а саме

$$x(k) = N \left[\frac{1}{N} \sum_{n=0}^{N-1} \dot{c}^*(n) e^{-j\frac{2\pi}{N}kn} \right]^*. \quad (5.17)$$

В подальшому будемо користуватися більш поширеним представленням прямого та оберненого перетворення Фур'є,

коли множник $1/N$ застосовується у співвідношенні для обчислення оберненого перетворення Фур'є.

5.2. Швидке перетворення Фур'є

Знаходження спектральних складових $c(n)$ дискретного комплексного сигналу $x(k)$ безпосередньо по формулі ДПФ вимагає N^2 комплексних множень і $N(N-1)$ комплексних додавань. Оскільки кількість обчислень, а отже, і час обчислень приблизно пропорційні N^2 , то при великих N кількість арифметичних операцій дуже велика. Тому знаходження спектру в реальному часі навіть для сучасної обчислювальної техніки є складним завданням.

З цієї причини представляють значний інтерес обчислювальні процедури, які зменшують кількість множень і складань. Основний принцип усіх цих алгоритмів полягає в розкладанні операцій обчислення ДПФ сигналу довжини N на обчислення перетворень Фур'є з меншим числом точок. Розділивши аналізований набір відліків на частини, обчислюють їх ДПФ і об'єднують результати. Такі процедури дістали назву алгоритмів швидкого перетворення Фур'є (ШПФ).

При реалізації ШПФ можливі декілька варіантів організації обчислень залежно від способу ділення послідовності відліків на частини (*проріджування за часом або по частоті*) і від того, на скільки фрагментів розбивається послідовність на кожному кроці (*основа ШПФ*). Найбільш простими і широко використовуваними є алгоритми ШПФ з основою 2, коли довжина послідовності $x(k)$ є цілою мірою числа 2, тобто $N = 2^v$, де v – ціле число.

Проаналізуємо більш детально операції ДПФ на прикладі $N=8$ [20]. У цьому випадку, ввівши загальноприйняте в літературі позначення для дискретних експоненціальних функцій

$$\tilde{e}_N(k, n) = e^{-j(2\pi/N)kn} = w_N^{nk}, \quad (5.18)$$

матимемо $w = e^{-j\frac{2\pi}{8} \cdot 1 \cdot 1} = e^{-j\frac{\pi}{4}}$, тут і далі індекс $N=8$ опущено для простоти запису.

Користуючись матричним представленням, вираз для обчислення ДПФ (вже без множника $1/N$)

$$c(n) = c_n = \sum_{k=0}^{N-1} x(k) e^{-j\frac{2\pi}{N}kn} = \sum_{k=0}^{N-1} x_k w^{kn}, \quad n, k \in \overline{0, N-1} \quad (5.19)$$

може бути представлений у вигляді добутку

$$\mathbf{C} = \mathbf{W}\mathbf{X}, \quad (5.20)$$

де матриця ДПФ прийме вигляд

$$\mathbf{W} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & w & w^2 & w^3 & w^4 & w^5 & w^6 & w^7 \\ 1 & w^2 & w^4 & w^6 & w^8 & w^{10} & w^{12} & w^{14} \\ 1 & w^3 & w^6 & w^9 & w^{12} & w^{15} & w^{18} & w^{21} \\ 1 & w^4 & w^8 & w^{12} & w^{16} & w^{20} & w^{24} & w^{28} \\ 1 & w^5 & w^{10} & w^{15} & w^{20} & w^{25} & w^{30} & w^{35} \\ 1 & w^6 & w^{12} & w^{18} & w^{24} & w^{30} & w^{36} & w^{42} \\ 1 & w^7 & w^{14} & w^{21} & w^{28} & w^{35} & w^{42} & w^{49} \end{pmatrix}, \quad (5.21)$$

і вектор $\mathbf{X} = (x_0 \ x_1 \ x_2 \ x_3 \ x_4 \ x_5 \ x_6 \ x_7)^T$, T – знак транспонування.

Проаналізуємо w^n . Точки $w^n = e^{-j\frac{\pi}{4} \cdot n}$ знаходяться на колі одиничного радіусу і являють собою періодичну послідовність. Тобто значення $n = \overline{0, N-1}$ повністю визначають будь-яку степінь при $n \geq 8$ (за умови $N = 8$) (рис. 5.2).

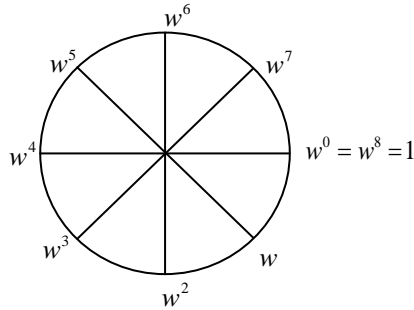


Рисунок 5.2 – Комплексна площина w

На рис. 5.2 показані числа

$$w^n = e^{-j \cdot \frac{\pi}{4} \cdot n} = \cos\left(\frac{\pi}{4} n\right) - j \cdot \sin\left(\frac{\pi}{4} n\right).$$

Звернемо також увагу на те, що числа, які відповідають точкам розташованим діаметрально протилежно на колі відрізняються лише знаком. Тому $w^4 = -w^0 = -1$, $w^5 = -w$, $w^6 = -w^2$, $w^7 = -w^3$. Тобто будь-яка ціла степінь w може бути представлена через 1 , w , w^2 , w^3 . З урахуванням цього матрицю ДПФ можна подати у вигляді

$$\mathbf{W} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & w & w^2 & w^3 & -1 & -w & -w^2 & -w^3 \\ 1 & w^2 & -1 & -w^2 & 1 & w^2 & -1 & -w^2 \\ 1 & w^3 & -w^2 & w & -1 & -w^3 & w^2 & -w \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -w & w^2 & -w^3 & -1 & w & -w^2 & w^3 \\ 1 & -w^2 & -1 & w^2 & 1 & -w^2 & -1 & w^2 \\ 1 & -w^3 & -w^2 & -w & -1 & w^3 & w^2 & w \end{pmatrix}. \quad (5.22)$$

В кожному рядку цієї матриці є або однакові або такі, що різняться за знаком члени. Ідея ШПФ застосовує описані вище властивості кругової згортки і, крім того, передбачає скорочення кількості операцій множення відповідно до тотожності

$$a \cdot b \pm a \cdot c = a \cdot (b \pm c). \quad (5.23)$$

Ліворуч у виразі (5.23) два множення, а праворуч лише одне. З урахуванням цього знайдемо складові дискретного спектру:

$$\begin{aligned} c_0 &= [(x_0 + x_4) + (x_2 + x_6)] + [(x_1 + x_5) + (x_3 + x_7)]; \\ c_1 &= [(x_0 - x_4) + (x_2 - x_6) \cdot w^2] + [(x_1 - x_5) + (x_3 - x_7) \cdot w^2] \cdot w; \\ c_2 &= [(x_0 + x_4) - (x_2 + x_6)] + [(x_1 + x_5) - (x_3 + x_7)]w^2; \\ c_3 &= [(x_0 - x_4) + (x_2 - x_6) \cdot w^2] + [(x_1 - x_5) - (x_3 - x_7) \cdot w^2] \cdot w^3; \\ c_4 &= [(x_0 + x_4) + (x_2 + x_6)] - [(x_1 + x_5) + (x_3 + x_7)]; \\ c_5 &= [(x_0 - x_4) + (x_2 - x_6) \cdot w^2] - [(x_1 - x_5) + (x_3 - x_7) \cdot w^2] \cdot w; \\ c_6 &= [(x_0 + x_4) - (x_2 + x_6)] - [(x_1 + x_5) - (x_3 + x_7)]w^2; \\ c_7 &= [(x_0 - x_4) + (x_2 - x_6) \cdot w^2] - [(x_1 - x_5) - (x_3 - x_7) \cdot w^2] \cdot w^3. \end{aligned}$$

Видно, що у круглих дужках маються лінійні комбінації елементів вектора \mathbf{X} . Кількість таких комбінацій дорівнює 8. В квадратних дужках вагові комбінації сум у круглих дужках (їх теж вісім). Останні операції передбачають утворення вагових сум з результатів у квадратних дужках.

Ці три послідовні етапи обчислення можуть бути представлені у вигляді добутку трьох порівняно простих матриць (відповідно з права на ліво). Таким чином, матриця ДПФ приймає вигляд

$$\mathbf{W} = \mathbf{W}_3 \mathbf{W}_2 \mathbf{W}_1, \quad (5.24)$$

де

$$\mathbf{W}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{pmatrix}, \quad (5.25)$$

$$\mathbf{W}_2 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & w^2 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -w^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & w^2 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -w^2 \end{pmatrix}, \quad (5.26)$$

$$\mathbf{W}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & w & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & w^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & w^3 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -w & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & w^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -w^3 \end{pmatrix}. \quad (5.27)$$

Якщо обчислити добуток цих трьох матриць отримаємо матрицю \mathbf{W} . Розкладання матриць на прості множники виду (5.24) називають *факторизацією*.

Процедуру обчислення дискретного спектру зручно подати у вигляді графа (рис. 5.3), який пояснює послідовність дій, виконуваних факторизованою матрицею.

Промисловість серійно випускає інтегральні мікросхеми, які виконують чотирьохточкове ДПФ, яке є базовою операцією. Цю мікросхему називають «метеликом». Вона дозволяє виготовляти спеціалізовані процесори ШПФ для сигналів з довільним N .

Алгоритмічна схема «метелика» і її позначення наведені на рис. 5.4.

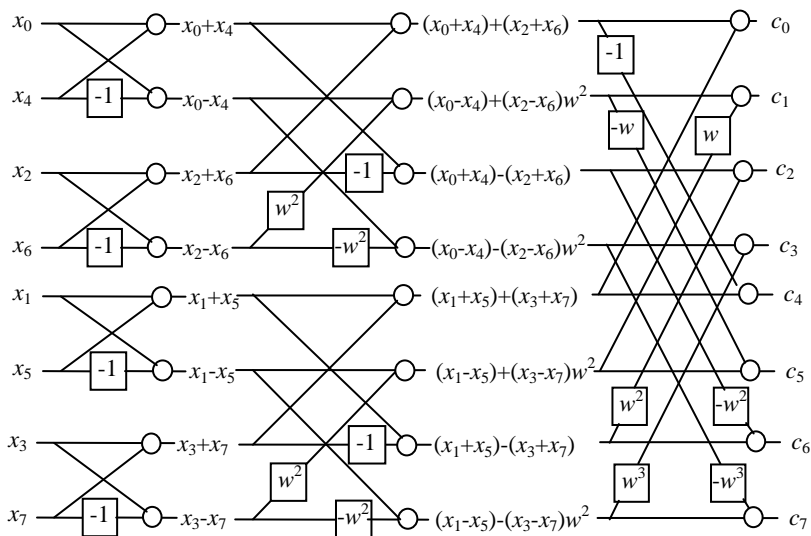


Рисунок 5.3 – Граф ШПФ. Кругом позначено підсумовування, а квадратом – множення на величину, що вказана в квадраті

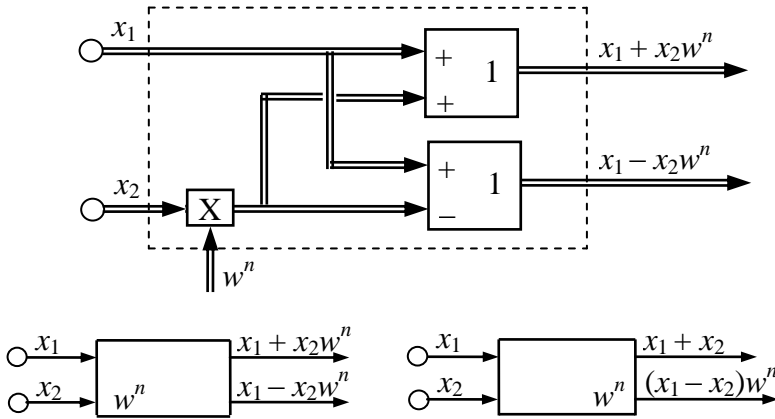


Рисунок 5.4 – Алгоритмічна схема і умовне позначення «метелика»

Якщо на її входи поступають комплексні числа x_1 і x_2 , тоді на виході формуються числа згідно з виразами $x_1 + x_2w^n$ і $x_1 - x_2w^n$, які є базовою операцією алгоритму з проріджуванням за часом. Для алгоритму з проріджуванням за частотою базова операція має вигляд: $x_1 + x_2$ і $(x_1 - x_2)w^n$. Оскільки комплексні величини являють собою пару чисел (реальна і мнма частини), тому практично «метелик» має чотири входи і чотири виходи.

Як видно зі структурної схеми восьмиточечного спеціалізованого процесора ШПФ реалізованого на «метеликах» (рис. 5.5), в ній реалізовані операції визначені факторизованою матрицею (5.24).

Формально ШПФ по основі 2 здійснюється шляхом розбивання вихідної послідовності на парні (x_0, x_2, x_4, \dots) і непарні (x_1, x_3, x_5, \dots) підпослідовності. З ними далі поступають подібним же чином. Ці процедури повторюються доти, поки не залишиться по одній парі чисел, які поступають на «метелики» першого етапу. Для восьмиточкового ШПФ

цими парами є x_0x_4 , x_2x_6 , x_1x_5 , x_3x_7 . На першому етапі операцій множення практично не проводиться (множення на ± 1). Тут формуються суми і різниці отриманих пар.

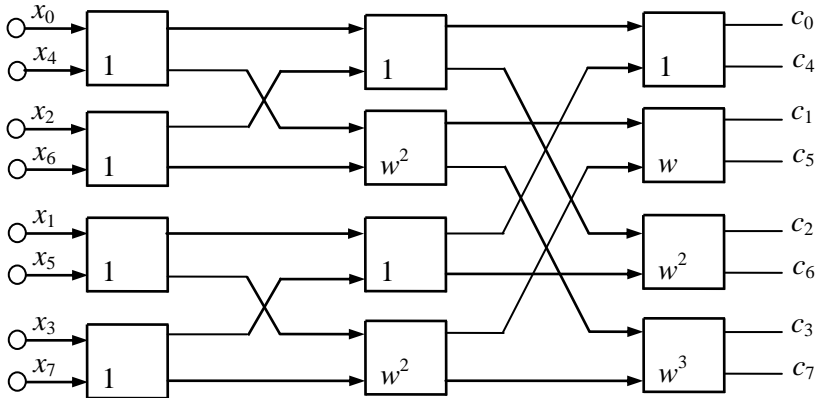


Рисунок 5.5 – Структурна схема спецпроцесора ШПФ на «метеликах»

Далі алгоритм обчислення відповідає рис. 5.5. Він складається з трьох етапів. В загальному ж випадку для N -точкового ШПФ потрібно $\log_2 N$ етапів. На кожному з них виконується $N/2$ операцій множення. Таким чином, алгоритм ШПФ потребує $\frac{N}{2} \cdot \log_2 N$ операцій множення.

Загальний алгоритм ДПФ має N^2 операцій множення. Тому алгоритм ШПФ порівняно з алгоритмом ДПФ дає вигоду по кількості операцій множення в $\frac{2N}{\log_2 N}$ разів.

В цілому, загальна кількість операцій при обчисленні ШПФ складає приблизно $\frac{N}{2} \cdot \log_2 N$ множень і $N \cdot \log_2 N$ складань комплексних чисел.

5.3. Пряме та обернене z -перетворення, його властивості

Зручним способом аналізу дискретних послідовностей є z -перетворення. При z -перетворенні різницеві рівняння, що описують роботу дискретної системи, перетворюються в рівняння алгебраїчні, з якими простіше виконувати необхідні дії. z -перетворення відіграє для дискретних сигналів і систем таку ж роль, як перетворення Лапласа – для аналогових сигналів.

Визначення z -перетворення. Дискретній послідовності відліків $\{x(n)\}$ ставиться у відповідність функція комплексної змінної z , яка визначається таким чином,

$$Z[x(n)] = X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n}. \quad (5.28)$$

Функція $X(z)$ визначена тільки для тих значень z , при яких ряд (5.28) сходиться. Якщо послідовність $x(n)$ має обмежену довжину, то $X(z)$ сходиться в z -площині скрізь, за виключенням, можливо, точок $z = 0$ або $z = \infty$.

Отримаємо z -перетворення для деяких дискретних сигналів, що часто зустрічаються на практиці.

Одиничний імпульс, який визначається як

$$u_0(n) = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases} \quad (5.29)$$

Використовуючи формулу (5.28), отримаємо:

$$X_0(z) = \sum_{n=-\infty}^{\infty} u_0(n)z^{-n} = 1. \quad (5.30)$$

Функція $X_0(z)$ сходиться в усій комплексній площині.

Одиничний стрибок, який визначається як

$$u_1(n) = \begin{cases} 1, & n \geq 0, \\ 0, & n < 0. \end{cases} \quad (5.31)$$

Використовуючи визначення z -перетворення, отримаємо

$$X_1(z) = \sum_{n=-\infty}^{\infty} u_1(n)z^{-n} = \sum_{n=0}^{\infty} 1 \cdot z^{-n}. \quad (5.32)$$

Ряд (5.32) є сумою нескінченної геометричної прогресії з першим членом 1 і знаменником z^{-1} . Як відомо, такий ряд сходиться при $|z^{-1}| < 1$, тобто при $|z| > 1$ і його сума складає

$$X_1(z) = \frac{1}{1 - z^{-1}} = \frac{z}{z - 1}. \quad (5.33)$$

Значення $z = 1$ є єдиною особливою точкою (полюсом) функції $X_1(z)$.

Експоненціальна дискретна функція

$$u_2(n) = \begin{cases} a^n, & n \geq 0, \\ 0, & n < 0. \end{cases} \quad (5.34)$$

Обчислимо її z -перетворення

$$X_2(z) = \sum_{n=-\infty}^{\infty} u_2(n)z^{-n} = \sum_{n=0}^{\infty} a^n \cdot z^{-n} = \sum_{n=0}^{\infty} (az^{-1})^n. \quad (5.35)$$

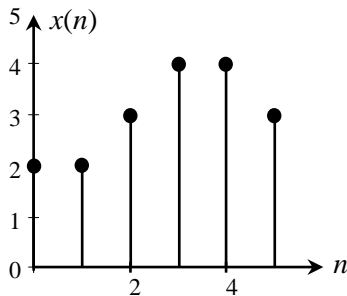
Як і у попередньому випадку ряд (5.35) є сумою геометричної прогресії з першим членом, рівним 1 і знаменником az^{-1} . Таким чином, ряд сходиться при $|az^{-1}| < 1$ тобто при $z > a$, і має особливу точку при $z = a$:

$$X_2(z) = \frac{1}{1 - az^{-1}} = \frac{z}{z - a} \quad (5.36)$$

Комплексна дискретна експонента, яка визначена співвідношенням

$$u_3(n) = \begin{cases} e^{j\omega n}, & n \geq 0 \\ 0, & n < 0 \end{cases}. \quad (5.37)$$

$$X_3(z) = \sum_{n=-\infty}^{\infty} u_3(n)z^{-n} = \sum_{n=0}^{\infty} (e^{j\omega} \cdot z^{-1})^n = \frac{1}{1 - e^{j\omega} z^{-1}} \quad (5.38)$$



причому $X_3(z)$ сходиться при $|z| > 1$, оскільки єдиною особливою точкою $X(z)$ є $z = e^{j\omega}$.

Приклад 5.1. Обчислення z -перетворення [9].

Відліковий сигнал наведено на рисунку. З використанням формули (5.28) знайти z -перетворення цієї послідовності.

Розв'язання:

Враховуючи обмеженість кількості відліків вхідної послідовності, нижня і верхня границі підсумовування в (5.28) зміняться на 0 і 5 відповідно. В результаті отримаємо

$$X(z) = \sum_{n=0}^5 x(n)z^{-n} = 2(1 + z^{-1}) + 3z^{-2} + 4(z^{-3} + z^{-4}) + 3z^{-5}.$$

Зв'язок z -перетворення з перетворенням Лапласа знайдемо, записавши аналоговий сигнал у вигляді суми добутоків дискретних відліків і набору дельта-функцій

$$x(t) = \sum_{n=-\infty}^{\infty} x(n)\delta(t - nT), \quad (5.39)$$

де T – крок дискретизації.

Перетворення Лапласа для такого сигналу дорівнює

$$F(s) = \int_0^{\infty} x(t)e^{-st} dt = \int_0^{\infty} \sum_{n=-\infty}^{\infty} x(n)\delta(t - nT)e^{-st} dt = \quad (5.40)$$

$$= \sum_{n=-\infty}^{\infty} x(n) \int_0^{\infty} \delta(t - nT)e^{-st} dt.$$

Скориставшись фільтрувальною властивістю дельта-функції, отримаємо

$$F(s) = \sum_{n=-\infty}^{\infty} x(n)e^{-snT} \quad (5.41)$$

З порівняння співвідношень (5.28) і (5.41), видно, що одна формула переходить в іншу при заміні $z = e^{sT}$. Таким чином, z -перетворення можна отримати з перетворення Лапласа шляхом переходу до нової змінної:

$$F(p) \Rightarrow X(z) \quad \text{при} \quad z = e^{sT}. \quad (5.42)$$

Сенс використання z -перетворення при аналізі дискретних сигналів витікає з наступного. Оскільки справедливе співвідношення

$$z^{-1} = e^{-sT} = e^{-j\omega T}, \quad (5.43)$$

тоді зміна фазової характеристики сигналу $\varphi(\omega) = \omega T$ означає затримку сигналу на один крок дискретизації T і відповідає затримці сигналу на один такт в z -області.

Перехід від перетворення Лапласа до z -перетворення при описанні дискретних систем потрібний з наступної причини. Дискретизація аналогового сигналу призводить до періодичності частотного спектру, тобто появи нескінченного ряду зсунутих копій спектру початкового неперервного сигналу. Очевидно, ефект дискретизації веде до появи в площині $s = \sigma + j\omega$ нескінченної конфігурації особливих точок (полюсів і нулів).

При переході від s -площини до z -площини точка $p_k = \sigma_k + j\omega_k$ відображується в точку $z_k = e^{(\sigma_k + j\omega_k)T} = e^{\sigma_k T} e^{j\omega_k T}$. Тому шлях уздовж уявної осі $j\omega$ в s -площині відображується в *одичне коло* в z -площині, оскільки на уявній осі $\sigma_k = 0$ і, отже, $|z| = 1$. При цьому ліва (стійка) смуга s -площини шириною $\Omega = 2\pi/T$ відображується всередину кола одичного радіусу z -площини. Усі наступні стійкі смуги s -площини шириною Ω , які відповідають періодичному частотному спектру дискретного сигналу, також відображуються всередину круга одичного радіусу z -площини. Тому конфігурація особливих точок (полюсів і нулів) в z -площині стає кінцевою.

Таким чином, взаємна відповідність між z -перетворенням $X(z)$ та перетворенням Лапласа $F(s)$ описується наступним чином:

$$X(z) = F\left(\frac{1}{T} \ln z\right); \quad F(s) = X(e^{sT}).$$

Подібними формулами описується зв'язок z -перетворення $X(z)$ з перетворенням Фур'є $\dot{F}(\omega)$ (при цьому немає необхідності вважати послідовність однобічною):

$$X(z) = F\left(\frac{1}{jT} \ln z\right); \quad F(j\omega) = X(e^{j\omega T}).$$

Зв'язок ДПФ з z -перетворенням

Порівнюючи формулу прямого ДПФ дискретної послідовності $x(n)$ з формулою z -перетворення, побачимо, що коефіцієнти ДПФ $c(k)$ дорівнюють значенням z -перетворення цього сигналу в N точках, рівномірно розподілених по одиничному колу z -площини. Ці коефіцієнти однозначно представляють саму послідовність, оскільки вона може бути точно відновлена за допомогою оберненого ДПФ.

Отримаємо z -перетворення послідовності через коефіцієнти ДПФ цієї послідовності

$$\begin{aligned} X(z) &= \sum_{n=0}^{N-1} x(n)z^{-n} = \frac{1}{N} \sum_{n=0}^{N-1} \left[\sum_{k=0}^{N-1} c(k)e^{j\frac{2\pi nk}{N}} \right] z^{-n} = \\ &= \frac{1}{N} \sum_{k=0}^{N-1} c(k) \sum_{n=0}^{N-1} \left(e^{j\frac{2\pi nk}{N}} z^{-1} \right) = \frac{1}{N} \sum_{k=0}^{N-1} c(k) \frac{1 - z^{-N}}{1 - e^{j\frac{2\pi k}{N}} z^{-1}}. \end{aligned} \quad (5.44)$$

Формула (5.44) показує, що z -перетворення кінцевої послідовності $x(n)$, $n = \overline{0, N-1}$, безпосередньо пов'язане з коефіцієнтами $c(k)$, $k = \overline{0, N-1}$, її ДПФ.

Властивості z-перетворення:

1. *Лінійність.*

Якщо $X_1(z)$ і $X_2(z)$ є z -перетвореннями відповідних сигналів $x_1(n)$ і $x_2(n)$, тоді сигналу $y(n) = ax_1(n) + bx_2(n)$ відповідатиме z -перетворення $Y(z) = aX_1(z) + bX_2(z)$ при будь-яких постійних a і b .

2. *Затримка (зсув послідовності).*

Якщо z -перетворення сигналу $x(n)$ дорівнює $X(z)$ або інакше: $x(n) \xrightarrow{z} X(z)$, тоді z -перетворення сигналу $y(n) = x(n - n_0)$, затриманого на n_0 тактів, буде рівне $X(z)z^{-n_0}$.

Доведення.

$$\begin{aligned} Y(z) &= \sum_{n=-\infty}^{\infty} y(n)z^{-n} = \sum_{n=-\infty}^{\infty} x(n - n_0)z^{-n} = z^{-n_0} \sum_{n=-\infty}^{\infty} x(n - n_0)z^{-n+n_0} = \\ &= z^{-n_0} \sum_{n=-\infty}^{\infty} x(n - n_0)z^{-(n-n_0)} = z^{-n_0} \sum_{k=-\infty}^{\infty} x(k)z^{-k} = z^{-n_0} X(z). \end{aligned} \quad (5.45)$$

Таким чином, при затримці сигналу на n_0 тактів необхідно помножити його z -перетворення на множник z^{-n_0} .

Зокрема, якщо сигнал $y(n) = x(n-1)$ отриманий шляхом зсуву сигналу $x(n)$ на один такт у бік запізнювання, тоді його z -перетворення $Y(z) = X(z)z^{-1}$. Отже, символ z^{-1} є *оператором одиничної затримки* (на один інтервал дискретизації) в z -області.

3. *Теорема множення на експоненту*

Якщо $x(n) \xrightarrow{z} X(z)$, тоді $a^n x(n) \xrightarrow{z} X(a^{-1}z)$.

4. *Теорема множення на n*

Якщо $x(n) \xrightarrow{z} X(z)$, тоді $nx(n) \xrightarrow{z} -z \frac{dX(z)}{dz}$.

5. *Згортка*

Введемо дискретну лінійну згортку $w(n)$, яку визначимо таким чином:

$$w(n) = \sum_{k=-\infty}^{\infty} x(k)y(n-k) = \sum_{k=-\infty}^{\infty} y(k)x(n-k). \quad (5.46)$$

Обчислимо її z -перетворення

$$\begin{aligned} W(z) &= Z[w(n)] = Z\left[\sum_{k=-\infty}^{\infty} x(k)y(n-k)\right] = \sum_{n=-\infty}^{\infty} \left[\sum_{k=-\infty}^{\infty} x(k)y(n-k)z^{-n}\right] = \\ &= \sum_{n=-\infty}^{\infty} \left[\sum_{k=-\infty}^{\infty} x(k)y(n-k)z^{-k}z^{-(n-k)}\right] = \\ &= \sum_{k=-\infty}^{\infty} \left[x(k)z^{-k} \sum_{n=-\infty}^{\infty} y(n-k)z^{-(n-k)}\right] = X(z)Y(z). \end{aligned} \quad (5.47)$$

Отже, z -перетворення лінійної згортки двох дискретних сигналів дорівнює добутку їх z -перетворень.

Різницеві рівняння зазвичай визначені при $n \geq 0$ і мають набір початкових умов. Тому при рішенні практичних задач зазвичай вводять одностороннє z -перетворення, що має вигляд

$$x(z) = \sum_{k=0}^{\infty} x(k)z^{-k} \quad (5.48)$$

Для багатьох сигналів властивості одностороннього z -перетворення аналогічні властивостям звичайного z -перетворення. Основним виключенням є властивість, пов'язана із зсувом (затримкою) сигналів. Так, затримка на один відлік, як і раніше, призводить до множення одностороннього z -перетворення на z^{-1} , але при цьому необхідно врахувати значення сигналу $n < 0$, тобто початкові умови.

б. Властивість коефіцієнтів розкладання z -перетворення в степеневий ряд

Якщо $x(n\Delta t) \xleftrightarrow{z} X(z) = q_0 + q_1z^{-1} + q_2z^{-2} + \dots$, тоді коефіцієнти q_0, q_1, q_2, \dots розкладання $X(z)$ в степеневий ряд по z дорівнюють значенням відліків функції $x(nT)$, тобто $q_n = x(nT)$.

Ця властивість отримується безпосередньо з (5.48), тому $x(0) = q_0$, $x(T) = q_1$, $x(2T) = q_2$ і т.д.

7. *Теорема початкового і кінцевого (фінального) значень*
Важливі для з'ясування поведінки дискретної послідовності $x(n)$ відповідно до відомого її z -перетворення.

Крім того, теореми початкового і кінцевого значень дозволяють контролювати обчислення оберненого z -перетворення [9].

Теорема початкового значення встановлює, що

$$x(0) = \lim_{z \rightarrow \infty} X(z).$$

Справедливість співвідношення легко довести. Запишемо (5.28) у вигляді

$$X(z) = \sum_{n=0}^{\infty} x(n)z^{-n} = x(0) + \sum_{n=1}^{\infty} x(n)z^{-n}.$$

Тоді при $z \rightarrow \infty$ сума в правій частині цього рівняння прямує до нуля, що і призводить до шуканої рівності.

Теорема кінцевого значення встановлює, що

$$x(\infty) = \lim_{z \rightarrow 1} (1 - z^{-1})X(z).$$

Цю формулу можна перевірити.

Якщо $X(z)$ має полюс (або полюси) поза одиничним колом, цьому випадку відповідає нестійка послідовність, тобто така, що зростає з часом, $x(\infty) = \infty$.

Якщо всі полюси $X(z)$ розташовані усередині одиничного кола, $x(n)$ є відліками загасаючої функції, тобто $x(\infty) = 0$.

Залишився випадок, коли полюси функції $X(z)$ лежать на одиничному колі.

Нехай полюси належать одиничному колу, але не лежать у точці, де $z = 1$. Цьому випадку відповідає функція часу, що є осцилюючою. Її значення $x(\infty)$ визначити не можна, хоча є скінченні границі, в діапазоні яких відбуваються коливання.

Ненульове значення, що встановлюється, може існувати в єдиному випадку, коли існує полюс в z -площині при $z=1$, а всі інші полюси лежать усередині одиничного кола. Тоді функцію $X(z)$ можна подати так:

$$X(z) = \frac{K}{1-z^{-1}} + G(z),$$

причому всі полюси функції $G(z)$ лежать усередині одиничного кола і $G(\infty)=0$. Як виходить з результату, отриманого при знаходженні z -перетворення одиничного стрибка, і властивості лінійності, першому доданку правої частини відповідає послідовність $x(n)=K, n \geq 0$. Тобто K – це усталене значення. Таким чином, маємо

$$x(\infty) = \lim_{z \rightarrow 1} (1-z^{-1})X(z) = K + \lim_{z \rightarrow 1} (1-z^{-1})G(z) = K.$$

Що й потрібно було довести.

Обернене z -перетворення

Відповідно до (5.48) функція $X(z)$ визначає всю нескінченну сукупність відліків $\{x(0), x(1), \dots, x(n), \dots\}$.

Помножимо обидві частини ряду (5.48) на множник z^{n-1} :

$$z^{n-1}X(z) = x(0)z^{n-1} + x(1)z^{n-2} + \dots + x(n)z^{-1} + \dots, \quad (5.49)$$

а потім обчислимо інтеграли від обох частин отриманої рівності, взявши за контур інтегрування замкнуту криву, що лежить цілком в області аналітичності і охоплює усі полюси функції $X(z)$. При цьому скористаємось відомими співвідношеннями для інтегралу по замкнутому контуру C , в середині якого лежить точка $z=0$ [21],

$$\oint_C z^n dz = \begin{cases} 2\pi j, & n = -1, \\ 0, & n \neq -1. \end{cases}$$

Видно, що інтеграли від усіх доданків правої частини (5.49), окрім доданку з номером n , будуть рівними нулю. Тому

$$x(n) = \frac{1}{2\pi j} \oint_C X(z) z^{n-1} dz = Z^{-1}[X(z)], \quad (5.50)$$

де C – замкнутий контур $C > R_1$, R_1 – радіус кола збіжності $X(z)$.

Вираз (5.50) називають оберненим z -перетворенням, воно дозволяє знайти відліки $x(n)$ по z -зображенню $X(z)$. Обернене z -перетворення існує тільки для таких функцій $X(z)$, які мають кінцеве число особливих точок (полісів), причому особливість в кожній з них є усувною.

Існують декілька методів обчислення оберненого z -перетворення. Найчастіше користуються *теоремою про лишки* (рос. вычеты), згідно якої інтеграл по замкнутому контуру від функції комплексного аргументу з точністю до множника $1/(2\pi j)$ дорівнює сумі лишків підінтегральної функції $X_0(z) = X(z)z^{n-1}$ в особливих точках (полісах p_i), що охоплюються контуром інтегрування C

$$x(n) = \frac{1}{2\pi j} \int_C X(z)z^{n-1} dz = \sum_i \operatorname{Res}_{p_i} X_0(z). \quad (5.51)$$

Обчислення лишків пов'язане з представленням функції $X_0(z)$ у вигляді

$$X_0(z) = X(z)z^{n-1} = \frac{N(z)}{\prod_{i=1}^k (z - p_i)^{m_i}}, \quad (5.52)$$

де p_i є полюсом порядку m_i .

Для знаходження лишків використовують наступні формули:

- У разі простого полюсу, тобто коли $m_i = 1$,

$$\operatorname{Res}_{p_i} X_0(z) = \lim_{z \rightarrow p_i} [(z - p_i)X_0(z)] = [(z - p_i)X_0(z)] \Big|_{z=p_i}. \quad (5.53)$$

- У разі m -кратного полюса, тобто полюсу m -го порядку,

$$\operatorname{Res}_{p_i} X_0(z) = \frac{1}{(m-1)!} \lim_{z \rightarrow p_i} \frac{d^{m-1}}{dz^{m-1}} [(z - p_i)^m X_0(z)]. \quad (5.54)$$

Приклад 5.2. Визначити по зображенню (5.36) $X(z) = z/(z - a)$ відліки сигналу $x(n)$.

Знайдемо підінтегральний вираз оберненого z -перетворення:

$$X_0(z) = X(z)z^{n-1} = \frac{z^n}{z - a}.$$

Функція $X_0(z)$ має один простий полюс в точці $p_1 = a$. Відповідно до (5.53) отримаємо послідовність

$$x(n) = \sum_{i=1} \operatorname{Res}_{p_i} X_0(z) = \lim_{z \rightarrow a} \frac{z^n}{(z - a)} (z - a) = a^n,$$

яка збігається з експоненціальною дискретною функцією (5.34).

Приклад 5.3. Знайти відліки сигналу $x(n)$ по його z -зображенню $X(z) = \frac{za}{(z - a)^2}$.

Підінтегральний вираз оберненого z -перетворення дорівнює

$$X_0(z) = X(z)z^{n-1} = \frac{z^n a}{(z - a)^2}.$$

Функція $X_0(z)$ має один двократний полюс $p_1 = a$ при $n \geq 0$. Відповідно до (5.54) отримуємо

$$\begin{aligned} x(n) &= \sum_{i=1} \operatorname{Res}_{p_i} X_0(z) = \frac{1}{(2-1)!} \lim_{z \rightarrow a} \frac{d}{dz} \left[(z - a)^2 \frac{z^n a}{(z - a)^2} \right] = \\ &= \lim_{z \rightarrow a} (naz^{n-1}) = na^n. \end{aligned}$$

Другим методом обчислення оберненого z -перетворення, який може застосовуватись на практиці, є метод розкладання функції $X(z)$ на прості дроби. Функцію $X(z)$ представляють у вигляді суми елементарних дроби

$$X(z) = \sum_i X_i(z) = \sum_i \frac{a_i}{1 - p_i z^{-1}},$$

де $X_i(z)$ – z -перетворення з одним простим полюсом.

Враховуючи, що кожен доданок $\frac{a_i}{1 - p_i z^{-1}}$ має обернене

z -перетворення виду $a_i(p_i)^n$, отримаємо

$$x(n) = \begin{cases} \sum_{i=1}^N a_i (p_i)^n, & n \geq 0, \\ 0, & n < 0. \end{cases}$$

Приклад 5.4. Обчислити обернене z -перетворення функції

$$X(z) = \frac{z}{z^2 - 3z + 2}.$$

Представимо $X(z)$ у вигляді суми простих дробів

$$X(z) = \frac{z}{(z-1)(z-2)} = \frac{z^{-1}}{(1-z^{-1})(1-2z^{-1})} = \frac{1}{(1-2z^{-1})} - \frac{1}{(1-z^{-1})}$$

Із зіставлення виду отриманих доданків з прикладами z -перетворень типових дискретних сигналів (які були розглянуті) видно, що перший доданок відповідає дискретній показовій функції $x(n) = 2^n$, а другий – одиничному стрибку з від'ємним знаком. Отже, шукана послідовність має вигляд

$$x(n) = 2^n - 1; n \geq 0.$$

Контрольні запитання

1. Дайте визначення дискретизації сигналів.
2. Дайте визначення амплітудно-імпульсної та широтно-імпульсної модуляції.
3. Чим відрізняються спектри аналогового та дискретизованого сигналів?
4. Як впливає вибір інтервалу Котельникова на спектральні характеристики дискретизованого сигналу?
5. Яку функцію буде ряд Котельникова?

6. Поясніть принцип отримання модульованої імпульсної послідовності та відмінності амплітудно-імпульсної і широтно-імпульсної видів модуляції.

7. Виведіть співвідношення для обчислення дискретного перетворення Фур'є. Обернене дискретне перетворення Фур'є.

8. Наведіть основні властивості дискретного перетворення Фур'є.

9. Поясніть сутність побудови алгоритмів швидкого перетворення Фур'є (ШПФ)

10. Дайте визначення z-перетворення.

11. Запишіть вирази одиничного імпульсу, одиничного стрибку, експоненціальної дискретної функції, комплексної дискретної експоненти.

12. Який зв'язок z-перетворення з перетворенням Лапласа?

13. Який зв'язок дискретного перетворення Фур'є з z-перетворенням

14. Які властивості має z-перетворення?

15. Як записати пряме z-перетворення дискретної послідовності?

16. Що встановлюють теореми початкового значення та кінцевого значення?

17. Як визначити імпульсну функцію дискретного ланцюга?

18. Як записати обернене z-перетворення дискретної згортки?

19. Які існують методи обчислення оберненого z-перетворення?

20. Як називаються рівняння, що описують дискретні ланцюги в часовій області?

21. Як називається дискретний фільтр з кінцевою імпульсною характеристикою?

22. Як називається дискретний фільтр з безкінцевою імпульсною характеристикою?

Глава 6

ЛІНІЙНІ СИСТЕМИ

Сучасний розвиток цифрової елементної бази та зручність використання цифрових ЗВ призводять до масового їх втілення в різних галузях вимірювань. В той же час багато базових понять і методів, які використовуються при цифровій обробці дискретних сигналів, прийшли з певними особливостями від аналогових систем. Аналогові ЗВ нині також застосовуються, а в цифрових ЗВ можуть використовуватись окремі аналогові елементи. Тому уявлення про характеристики і способи описання аналогових систем також потрібне, як і для дискретних. Це дозволить глибше зрозуміти й теорію дискретних систем, тому що багато методів аналізу аналогових і дискретних систем знаходяться в тісній спорідненості. Крім того, в основі ряду методів проектування дискретних фільтрів лежить використання аналогових прототипів, кваліфіковане застосування цих методів також вимагає знайомства з теорією аналогових систем.

Тому в даній главі стисло викладені питання аналізу проходження сигналів через лінійні аналогові і цифрові системи.

6.1. Аналогові лінійні системи

Класифікація систем.

Системи, що застосовуються для перетворення сигналів, мають різноманітні фізичні характеристики і можуть класифікуватися за різними ознаками.

Найважливішою класифікаційною ознакою є лінійність або нелінійність системи. *Лінійними називаються* системи, для яких *виконується принцип суперпозиції*: реакція на суму сигналів дорівнює сумі реакцій на ці сигнали, подані на вхід

окремо. Так, якщо зв'язок між сигналами на вході і виході системи задається оператором \mathbf{T} [10, 17]

$$\mathbf{U}_{\text{вих}}(t) = \mathbf{T}[\mathbf{U}_{\text{вх}}(t)],$$

де $\mathbf{U}_{\text{вх}}(t) = [s_{\text{вх}1}(t) \ s_{\text{вх}2}(t) \ \dots \ s_{\text{вх}m}(t)]^T$ – вхідний m -мірний векторний сигнал,

$\mathbf{U}_{\text{вих}}(t) = [s_{\text{вих}1}(t), s_{\text{вих}2}(t), \dots, s_{\text{вих}n}(t)]^T$ – вихідний n -мірний векторний сигнал, тоді для лінійної системи:

$$\mathbf{T}[\mathbf{U}_{\text{вх}1}(t) + \mathbf{U}_{\text{вх}2}(t)] = \mathbf{T}[\mathbf{U}_{\text{вх}1}(t)] + \mathbf{T}[\mathbf{U}_{\text{вх}2}(t)];$$

$$\mathbf{T}(\alpha \mathbf{U}_{\text{вх}}(t)) = \alpha \mathbf{T}(\mathbf{U}_{\text{вх}}(t)).$$

Системи, для яких *принцип суперпозиції не виконується*, називаються *нелінійними*.

Наступним критерієм класифікації систем є постійність або змінність їх характеристик в часі. Якщо довільна затримка вхідного сигналу призводить лише до такої ж затримки вихідного сигналу, не змінюючи його форми, система називається *стаціонарною*, або системою з постійними параметрами. Тобто якщо $\mathbf{U}_{\text{вих}}(t) = \mathbf{T}[\mathbf{U}_{\text{вх}}(t)]$, тоді $\mathbf{U}_{\text{вих}}(t \pm t_0) = \mathbf{T}[\mathbf{U}_{\text{вх}}(t \pm t_0)]$ при будь-якому t_0 .

Інакше система називається *нестационарною*, параметричною або системою із змінними параметрами. Два вказані способи класифікації ділять системи на чотири класи. Далі зупинимося тільки на класі лінійних стаціонарних систем. Нелінійні системи розглядаються в главі 7.

Розглянемо різні еквівалентні способи описання лінійних систем із зосередженими параметрами та переходу від одного представлення до іншого, що необхідно для правильного їх застосування на практиці.

Характеристики лінійних систем.

Для лінійних систем з постійними параметрами справедливі принципи суперпозиції і стаціонарності. Це значно спрощує аналіз проходження сигналів через такі

системи, дозволяючи використовувати характеристики, мова про які піде далі.

Лінійність і стаціонарність дозволяють легко знайти реакцію системи на будь-який вхідний сигнал, знаючи лише одну функцію – реакцію системи на подану на вхід δ -функцію. Ця реакція називається *імпульсною характеристикою* системи і позначається $h(t)$. Якщо деяка лінійна стаціонарна система описується оператором \mathbf{T} і для спрощення вважати вхідний і вихідний сигнали одновимірними, тоді $h(t) = \mathbf{T}[\delta(t)]$. Оскільки система стаціонарна, тоді аналогічний вихідний сигнал матиме місце і для зсунутої на довільний час δ -функції $\delta(t-t_0)$: $h(t-t_0) = \mathbf{T}[\delta(t-t_0)]$.

З використанням динамічного представлення (див. розд. 1.3) представимо сигнал на вході лінійної системи у вигляді його згортки з δ -функцією

$$s_{\text{вх}}(t) = \int_{-\infty}^{\infty} s_{\text{вх}}(t')\delta(t-t')dt'$$

Лінійна система під дією оператора \mathbf{T} , перетворює вхідний сигнал $s_{\text{вх}}(t)$ відносно змінної t , тобто усі функції, що входять в цей вираз і залежать від t . Вхідний сигнал $s_{\text{вх}}(t)$ при цьому перетворюється на вихідний сигнал $s_{\text{вих}}(t)$ а δ -функція $\delta(t-t')$ – в імпульсну характеристику $h(t-t')$. Функція $s_{\text{вх}}(t')$ від t не залежить і тому залишається без змін. В результаті отримаємо формулу, яка показує, що сигнал на виході лінійної системи з постійними параметрами дорівнює згортці вхідного сигналу й імпульсної характеристики системи

$$s_{\text{вих}}(t) = \mathbf{T}[s_{\text{вх}}(t)] = \mathbf{T}\left[\int_{-\infty}^{\infty} s_{\text{вх}}(t')\delta(t-t')dt'\right] = \int_{-\infty}^{\infty} s_{\text{вх}}(t')\mathbf{T}[\delta(t-t')]dt' = \int_{-\infty}^{\infty} s_{\text{вх}}(t')h(t-t')dt'. \quad (6.1)$$

Ця формула є фундаментальною в теорії лінійних систем і має назву інтеграла Дюамеля. Вона також може бути подана у вигляді

$$s_{\text{вих}}(t) = \int_{-\infty}^{\infty} s_{\text{вх}}(t-t')h(t')dt'. \quad (6.1,a)$$

Зазначимо, що, якщо вхідний і вихідний сигнали системи мають однакову розмірність, імпульсна характеристика, як і δ -функція часу, має розмірність частоти.

Якщо $h(t)$ відома, то подальші етапи знаходження вихідного сигналу системи є цілком формалізованими.

Для випадку багатомірного сигналу вводяться парціальні імпульсні характеристики $h_{ij}(t)$, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, m$, кожна з яких відображає сигнал на i -тому виході при подачі на j -й вхід δ -функції. Сукупність $h_{ij}(t)$ утворюють матрицю імпульсних характеристик

$$\mathbf{h}(t) = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1m} \\ h_{21} & h_{22} & \dots & h_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ h_{n1} & h_{n2} & \dots & h_{nm} \end{pmatrix}.$$

Формула інтеграла Дюамеля в багатомірному випадку має вигляд $\mathbf{U}_{\text{вих}}(t) = \int_{-\infty}^{\infty} \mathbf{h}(t-t')\mathbf{U}_{\text{вх}}(t')dt'$, де $\mathbf{U}_{\text{вх}}(t)$ – m -мірний вектор, $\mathbf{U}_{\text{вих}}(t)$ – n -мірний вектор.

Формування вихідного сигналу можна пояснити таким чином. Нескінченно мала «ділянка» вхідного сигналу $s_{\text{вх}}(t')$ шириною dt' породжує на виході відгук, що є імпульсною

характеристикою, помноженою на $s_{\text{вх}}(t')dt'$ і затриманою за часом на t' , тобто є $s_{\text{вх}}(t')h(t-t')$ (рис. 6.1) [11]. Щоб набути значення вихідного сигналу у момент часу t , треба скласти внески від усіх цих нескінченно малих «ділянок», тобто виконати інтегрування по t' , що і дає приведену вище формулу згортки (6.1).

Умовою фізичної реалізованості системи є важливий принцип: сигнал на виході системи, який відповідає вхідному впливові, не може з'явитися до моменту появи імпульсу на вході.

Звідси маємо просте обмеження на вид можливих імпульсних характеристик: $h(t) = 0$ при $t < 0$. Для системи, що фізично реалізується, верхня границя в інтегралі Дюамеля може бути замінена на поточний час

$$s_{\text{вих}}(t) = \int_{-\infty}^t s_{\text{вх}}(t')h(t-t')dt'.$$

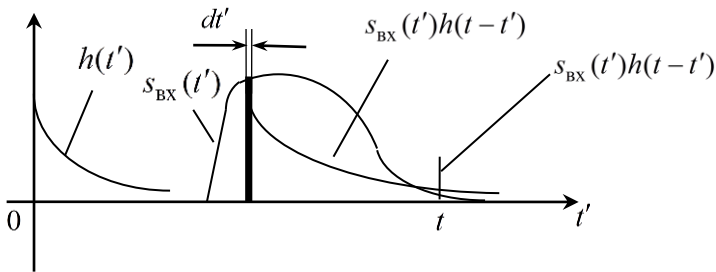


Рисунок 6.1 – Формування вихідної реакції лінійної системи

Ця формула має ясний фізичний зміст: лінійна стаціонарна система при обробці вхідного сигналу виконує операцію вагового підсумовування всіх його миттєвих значень, які існували «в минулому» при $-\infty < t' < t$. Роль вагової функції виконує при цьому імпульсна характеристика системи. Принципово важливо, що система, яка реалізується фізично, ні за яких обставин не здатна оперувати з «майбутніми» значеннями вхідного сигналу. Окрім цього,

вона має бути стійкою. Тобто, її імпульсна характеристика повинна задовольняти умові абсолютної інтегрованості

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

Перехідною характеристикою називають реакцію системи на подану на вхід функцію одиничного стрибка (функцію Хевісайда). Позначається перехідна характеристика як $g(t) = \mathbf{T}[\sigma(t)]$. Перехідна характеристика стаціонарної системи інваріантна відносно часового зсуву: $g(t-t_0) = \mathbf{T}[\sigma(t-t_0)]$.

Перехідна характеристика системи, що фізично реалізується, відрізняється від нуля лише при $t \geq 0$, як і для імпульсної характеристики $g(t) = 0$ при $t < 0$.

Оскільки δ -функція є похідною від одиничного стрибка, імпульсна і перехідна характеристики пов'язані одна з одною операціями диференціювання й інтегрування:

$$h(t) = \frac{dg(t)}{dt}; \quad g(t) = \int_{-\infty}^t h(t') dt'. \quad (6.2)$$

Скориставшись динамічним представленням довільного сигналу із застосуванням функцій Хевісайда $s(t) = s_0\sigma(t) + \int_0^{\infty} \frac{ds}{dt'} \sigma(t-t') dt'$ (розд. 1.3), отримаємо форму інтеграла Дюамеля у вигляді

$$s_{\text{вих}}(t) = \mathbf{T}[s_{\text{вх}}(t)] = s_{\text{вх}}(0)g(t) + \int_0^t \frac{s_{\text{вх}}(t')}{dt'} g(t-t') dt'. \quad (6.3)$$

Рівноцінною є й інша форма представлення сигналу на виході лінійної стаціонарної системи інтегралом Дюамеля такого виду

$$s_{\text{вих}}(t) = s_{\text{вх}}(t)g(0) + \int_0^t \frac{g(t')}{dt'} s_{\text{вх}}(t-t') dt'. \quad (6.3,а)$$

Оскільки сигнал на виході лінійної системи, як було показано вище, є згорткою вхідного сигналу й імпульсної характеристики, тому перетворення Фур'є від згортки

сигналів дає добуток їх спектрів. Спектр сигналу на виході лінійної системи в результаті має виглядає

$$\dot{S}_{\text{вих}}(\omega) = \dot{S}_{\text{вх}}(\omega) \dot{K}(\omega). \quad (6.4)$$

де $\dot{K}(\omega)$ – перетворення Фур'є імпульсної характеристики системи

$$\dot{K}(\omega) = \int_{-\infty}^{\infty} h(t) e^{-j\omega t} dt. \quad (6.5)$$

Ця функція називається *комплексним (частотним) коефіцієнтом передачі системи*, а її модуль $|\dot{K}(\omega)|$ і фаза $\varphi_K(\omega) = \arg[\dot{K}(\omega)]$ – відповідно амплітудно-частотною (АЧХ) і фазочастотною (ФЧХ) характеристиками системи.

Значення $\dot{K}(\omega)$ показує, як змінюється при проходженні через систему комплексна амплітуда синусоїди з частотою ω . Відповідно АЧХ показує у скільки разів зміниться амплітуда синусоїди, а ФЧХ – який буде отриманий нею фазовий зсув

$$\dot{K}(\omega) = |\dot{K}(\omega)| \exp[j\varphi_K(\omega)].$$

Оскільки $h(t)$ – функція дійсна, то з урахуванням властивостей перетворення Фур'є $\dot{K}(\omega) = \dot{K}^*(-\omega)$. Тому АЧХ є парною, а ФЧХ – непарною функціями частоти.

Потужність гармонічного сигналу пропорційна квадрату його амплітуди і не залежить від його фази. Тому *коефіцієнт передачі за потужністю* дорівнює квадрату модуля комплексного коефіцієнта передачі, тобто квадрату АЧХ

$$\dot{K}_{\text{пот}}(\omega) = \dot{K}(\omega) \dot{K}^*(\omega) = |\dot{K}(\omega)|^2. \quad (6.6)$$

При перетворенні сигналу лінійною системою розрізняють два види затримки. *Фазова затримка* (англ. phase delay) на частоті ω – це затримка гармонійного коливання з частотою ω , що проходить через систему. Значення фазової затримки дорівнює фазовому зсуву, що вноситься системою,

діленому на частоту гармонійного коливання, із від'ємним знаком: $\tau_{\phi}(\omega) = -\phi_K(\omega) / \omega$.

Групова затримка (англ. group delay) на частоті ω – це затримка обвідної вузько-смугового сигналу з середньою частотою ω . Групова затримка дорівнює похідній від ФЧХ системи із від'ємним знаком: $\tau_{gp}(\omega) = -d\phi_K(\omega) / d\omega$.

Взаємний спектр вихідного і вхідного сигналів лінійної системи легко знайти, виходячи з визначення взаємного спектра $\dot{S}_{12}(\omega) = \dot{S}_1(\omega)\dot{S}_2^*(\omega)$:

$$\begin{aligned} \dot{S}_{\text{вих_вх}}(\omega) &= \dot{S}_{\text{вих}}(\omega)\dot{S}_{\text{вх}}^*(\omega) = \dot{S}_{\text{вх}}(\omega)\dot{K}(\omega)\dot{S}_{\text{вх}}^*(\omega) = \\ &= |\dot{S}_{\text{вх}}(\omega)|^2 \dot{K}(\omega). \end{aligned}$$

Звідси витікає, що комплексний коефіцієнт передачі системи дорівнює відношенню взаємного спектру вихідного і вхідного сигналів до енергетичного спектру вхідного сигналу:

$$\dot{K}(\omega) = \frac{\dot{S}_{\text{вих_вх}}(\omega)}{|\dot{S}_{\text{вх}}(\omega)|^2}. \quad (6.7)$$

Взаємна кореляція між входом і виходом може бути отримана застосуванням оберненого перетворення Фур'є до формули для взаємного спектру. Вираз для ВКФ сигналів на виході і вході з урахуванням зв'язку між кореляційною функцією і спектром сигналу (3.7) має вигляд

$$B_{\text{вих_вх}}(\tau) = \int_{-\infty}^{\infty} B_{\text{вх}}(\tau')h(\tau - \tau')d\tau'.$$

Отже, ВКФ сигналів на вході і виході лінійної системи представляє собою згортку АКФ вхідного сигналу з імпульсною характеристикою системи.

Зв'язок між сигналами на вході і виході лінійного ланцюга із зосередженими параметрами може бути поданий *диференціальним рівнянням* (ДР) виду

$$\begin{aligned}
 & a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + a_{n-2} \frac{d^{n-2} y}{dt^{n-2}} + \dots + a_1 \frac{dy}{dt} + a_0 y(t) = \\
 & = b_m \frac{d^m x}{dt^m} + b_{m-1} \frac{d^{m-1} x}{dt^{m-1}} + b_{m-2} \frac{d^{m-2} x}{dt^{m-2}} + \dots + b_1 \frac{dx}{dt} + b_0 x(t),
 \end{aligned} \tag{6.8}$$

де $x(t)$ – вхідний сигнал, $y(t)$ – вихідний сигнал, a_i і b_i – постійні коефіцієнти. Таким чином, ланцюг описується наборами коефіцієнтів $\{a_i\}$ і $\{b_i\}$.

Повинна виконуватись нерівність $m \leq n$, тобто максимальний порядок похідної сигналу на вході не може перевищувати максимального порядку похідної сигналу на виході. Це пов'язано з неможливістю реалізації операції «чистого» диференціювання аналоговим ланцюгом. Значення n називається *порядком динамічної системи*.

Якщо задати конкретний вид вхідного сигналу $x(t)$, вийде лінійне неоднорідне диференціальне рівняння з постійними коефіцієнтами. Рішення цього ДУ дає вихідний сигнал $y(t)$.

Застосувавши до обох частин ДР (6.8) перетворення Лапласа, отримаємо вираз для операторного коефіцієнта передачі, або *передавальної функції динамічної системи* (англ. transfer function)

$$H(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + b_{m-2} s^{m-2} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + a_{n-2} s^{n-2} + \dots + a_1 s + a_0}, \tag{6.9}$$

де a_i і b_i – ті ж постійні коефіцієнти, що і в ДР (6.8).

Як видно з (6.9), $H(s) = \frac{Y(s)}{X(s)}$ є відношенням перетворень Лапласа сигналу на виході $Y(s)$ та на вході $X(s)$ системи.

Перетворення Лапласа є узагальненням перетворення Фур'є, при якому частота набуває комплексних значень

$s = \sigma + j\omega$. Розраховується за формулою $F(s) = \int_0^{\infty} f(t)e^{-st} dt$, яка

є подібною до формули прямого перетворення Фур'є. Тому перетворення Лапласа є лінійним і при диференціюванні сигналу в часі його перетворення Лапласа множиться на комплексну частоту s (для позначення цієї змінної в технічній літературі також часто застосовується літера p).

Комплексний коефіцієнт передачі виходить з функції передачі за Лапласом шляхом підстановки $s = j\omega$:

$$\dot{K}(\omega) = \frac{b_m (j\omega)^m + b_{m-1} (j\omega)^{m-1} + b_{m-2} (j\omega)^{m-2} + \dots + b_1 (j\omega) + b_0}{a_n (j\omega)^n + a_{n-1} (j\omega)^{n-1} + a_{n-2} (j\omega)^{n-2} + \dots + a_1 (j\omega) + a_0}.$$

Розклавши чисельник і знаменник передавальної функції за Лапласом на множники, отримаємо її в наступному вигляді

$$H(s) = k \frac{(s - z_m)(s - z_{m-1})(s - z_{m-2}) \dots (s - z_1)}{(s - p_n)(s - p_{n-1})(s - p_{n-2}) \dots (s - p_1)},$$

де $k = b_m/a_n$ – коефіцієнт підсилення (англ. gain), z_i – нулі функції передачі (англ. zero); p_i – полюси функції передачі (англ. pole). В точках нулів $H(z_i)$ також дорівнює 0, а в точках полюсів $H(p_i) \rightarrow \infty$.

В даному випадку система описується набором параметрів $\{z_i, p_i\}$, k , де нулі функції передачі можуть бути дійсними або складати комплексно-спряжені пари. Те ж відноситься й до полюсів. Коефіцієнт підсилення завжди дійсний.

Ще одним способом перетворення цієї дробово-раціональної передавальної функції є представлення її у вигляді суми простих дробів. За відсутності кратних коренів у знаменнику таке представлення має наступний вигляд

$$H(s) = \frac{r_n}{(s-p_n)} + \frac{r_{n-1}}{(s-p_{n-1})} + \frac{r_{n-2}}{(s-p_{n-2})} + \dots + \frac{r_1}{(s-p_1)} + C_0,$$

де p_i – полюси передавальної функції, а числа r_i називаються лишками (рос. вычеты), C_0 – ціла частина передавальної функції, яка відмінна від нуля тільки у разі рівності степенів поліномів чисельника і знаменника.

В даному випадку система описується набором параметрів $\{r_i\}$, $\{p_i\}$, C_0 .

Полюси передавальної функції можуть бути дійсними або складати комплексно-спряжені пари. Лишки, які відповідають комплексно-спряженим полюсам, також є комплексно-спряженими.

За наявності кратних полюсів передавальної функції розкладання на прості дроби стає складніше. Кожний m -кратний полюс p_i дає m доданків наступного виду

$$\frac{r_{i1}}{(s-p_i)} + \frac{r_{i2}}{(s-p_i)^2} + \frac{r_{i3}}{(s-p_i)^3} + \dots + \frac{r_{im}}{(s-p_i)^m}.$$

Представлення передавальної функції у вигляді суми простих дроби дозволяє обчислити імпульсну характеристику системи, оскільки кожен доданок передавальної функції виду $\frac{r_i}{(s-p_i)}$ відповідає доданку

імпульсної характеристики виду $r_i e^{p_i t}$, $t \geq 0$.

Пара комплексно-спряжених полюсів дає пару доданків імпульсної характеристики у вигляді комплексно-спряжених експонент. Сума таких доданків є косинусоїдою з амплітудою, що змінюється експоненціально

$$\begin{aligned} r_i e^{p_i t} + r_i^* e^{p_i^* t} &= 2 \operatorname{Re} \left[r_i e^{p_i t} \right] = 2 \operatorname{Re} \left[|r_i| e^{j \arg(r_i)} e^{\operatorname{Re}(p_i) t} e^{j \operatorname{Im}(p_i) t} \right] = \\ &= 2 |r_i| e^{\operatorname{Re}(p_i) t} \cos(\operatorname{Im}(p_i) t + \arg(r_i)), \end{aligned}$$

де $\arg(r_i)$ – фаза комплексного числа r_i .

Що стосується кратних полюсів, то m -кратний полюс дасть у виразі для імпульсної характеристики m доданків наступного виду

$$r_{i1}e^{p_i t} + r_{i2}te^{p_i t} + r_{i3}\frac{t^2}{2}e^{p_i t} + \dots + r_{im}\frac{t^{m-1}}{(m-1)!}e^{p_i t}.$$

Лінійна система називається *стійкою*, якщо при нульовому вхідному сигналі вихідний сигнал затухає за будь-яких початкових умов: $\lim_{t \rightarrow \infty} s_{\text{вих}}(t) = 0$ при $s_{\text{вх}}(t) = 0$. Ця вимога рівнозначна вимозі загасання імпульсної характеристики системи $\lim_{t \rightarrow \infty} h(t) = 0$.

Як показано вище, імпульсна характеристика системи в загальному випадку містить доданки виду

$$r_i \frac{t^k}{k!} e^{p_i t},$$

де p_i – полюс передавальної функції системи, r_i – відповідні їм лишки, k – цілі числа в діапазоні від нуля до значення на одиницю меншого кратності полюса p_i .

Такі доданки при $t \rightarrow \infty$ затухають, якщо дійсна частина полюса p_i є від'ємною: $\text{Re}(p_i) < 0$.

Звідси отримуємо загальну умову: лінійна система є стійкою тоді й тільки тоді, коли полюси її передавальної функції лежать в лівій півплощині комплексної площини.

Знаходження полюсів передавальної функції (6.9) в загальному випадку вимагає обчислення n коренів характеристичного рівняння

$$a_n p^n + a_{n-1} p^{n-1} + a_{n-2} p^{n-2} + \dots + a_1 p + a_0 = 0,$$

які можливо виразити через коефіцієнти рівняння при $n \leq 4$ із застосуванням кінцевого числа підсумовувань, віднімань, множень, ділень та добувань кореня. Для перевірки стійкості лінійної системи не обов'язково знаходити полюси її системної функції. Перевірку приналежності можна

здійснити, наприклад, з використанням критерію Рауса – Гурвіца для визначення приналежності всіх коренів цього рівняння (полюсів) до лівої напівплощини комплексної площини. Для цього *необхідно і достатньо*, щоб були *додатними* [10, 17, 22]:

а) коефіцієнти a_0, a_n ;

б) визначник Гурвіца

$$D_{n-1} = \begin{vmatrix} a_{n-1} & a_n & 0 & 0 & \dots & 0 & 0 \\ a_{n-3} & a_{n-2} & a_{n-1} & a_n & \dots & 0 & 0 \\ a_{n-5} & a_{n-4} & a_{n-3} & a_{n-2} & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & 0 & a_0 \end{vmatrix},$$

який будується наступним чином: по головній діагоналі зліва направо розміщуються всі коефіцієнти характеристичного рівняння від a_{n-1} до a_0 ; від кожного діагонального елемента вліво и вправо добудовуються рядки визначника таким чином, щоб індекси зростали зліва направо; на місце коефіцієнтів з індексами менше нуля або більше n ставляться нулі;

в) всі діагональні мінори $D_{n-2}, D_{n-3}, \dots, D_1 = a_{n-1}$, які отримуються послідовним викреслюванням останніх стовпця і рядка у визначнику Гурвіца.

З виконання цих умов витікає, що всі коефіцієнти характеристичного рівняння повинні бути додатними [22].

Застосування критерію Рауса – Гурвіца значно спрощує перевірку стійкості лінійної системи, але він є придатним лише для ланок із зосередженими параметрами, для яких передавальна функція є відношенням поліномів (6.9). При $n > 2$ критерій також не є наочним з точки зору зміни коефіцієнтів системи у разі дослідження властивостей системи. Більш наочними можуть виявитись критерії Михайлова або Найквіста [9].

Перетворення випадкового процесу в лінійній системі

Як відомо, випадковий процес є ансамблем реалізацій (див. гл. 4). Кожна окрема реалізація є детермінованим сигналом, і її перетворення лінійною системою аналізується за допомогою формул, приведених раніше. У цьому ж розділі розглядаються перетворення статистичних характеристик випадкового процесу. При цьому мається на увазі стаціонарний випадковий процес з нульовим математичним сподіванням.

Оскільки спектром випадкового процесу вважається *спектральна густина потужності*, він перетворюється в лінійній системі пропорційно коефіцієнту передачі за потужністю

$$W_{\text{вих}}(\omega) = W_{\text{вх}}(\omega) K_{\text{пот}}(\omega) = W_{\text{вх}}(\omega) |\dot{K}(\omega)|^2. \quad (6.10)$$

Згідно з теоремою Вінера-Хинчина, *кореляційна функція* випадкового процесу пов'язана з його спектром перетворенням Фур'є. Застосування перетворення Фур'є до цієї формули дає згортку

$$R_{\text{вих}}(\tau) = \int_{-\infty}^{\infty} R_{\text{вх}}(\tau') B_h(\tau - \tau') d\tau', \quad (6.11)$$

де $B_h(\tau)$ – результат оберненого перетворення Фур'є від коефіцієнта передавання за потужністю $|\dot{K}(\omega)|^2$. Згідно із зв'язком між кореляційними функціями і спектрами сигналів це перетворення дає кореляційну функцію імпульсної характеристики системи $B_h(\tau) = \int_{-\infty}^{\infty} h(t)h(t - \tau) dt$.

Дисперсія випадкового процесу дорівнює значенню його кореляційної функції при $\tau = 0$. Підстановка цієї величини у формулу для вихідної кореляційної функції $R_{\text{вих}}(\tau)$ дає

$$D_{\text{вих}} = R_{\text{вих}}(0) = \int_{-\infty}^{\infty} R_{\text{вх}}(\tau') B_h(\tau') d\tau'.$$

Можна розрахувати дисперсію і в частотній області. Скориставшись наведеною вище формулою для вихідного спектра, маємо

$$D_{\text{вих}} = \frac{1}{2\pi} \int_{-\infty}^{\infty} W_{\text{вх}}(\omega) |\dot{K}(\omega)|^2 d\omega. \quad (6.12)$$

Густина ймовірності випадкового процесу на виході лінійної системи у загальному випадку не піддається розрахунку простими засобами. Виключення складає окремий випадок гаусівського випадкового процесу, оскільки нормальний розподіл залишається нормальним за будь-яких лінійних перетворень. Тому нормальний випадковий процес з нульовим середнім значенням після проходження через лінійну систему збереже свою нормальність і нульове математичне сподівання, а його дисперсія може бути розрахована по одній з формул попереднього підрозділу.

Якщо вхідний випадковий процес є *білим шумом*, усі приведені раніше формули істотно спрощуються:

$$W_{\text{вих}}(\omega) = W_0 |\dot{K}(\omega)|^2; \quad R_{\text{вих}}(\tau) = W_0 B_h(\tau) = W_0 \int_{-\infty}^{\infty} h(t)h(t-\tau) dt;$$

$$D_{\text{вих}} = W_0 B_h(0) = W_0 \int_{-\infty}^{\infty} h^2(t) dt = \frac{W_0}{2\pi} \int_{-\infty}^{\infty} |\dot{K}(\omega)|^2 d\omega.$$

6.2. Дискретні лінійні системи

Для описання дискретних лінійних систем (фільтрів) подібно до аналогових широко застосовуються їх *імпульсні характеристики й системні функції (передавальні функції)*.

Імпульсною характеристикою $h(n)$ дискретного фільтра називається його реакція на одиничний імпульс $u_0(n)$ при нульових початкових умовах (подібно до аналогових систем, де вхідний вплив – δ -функція)

$$h(n) = F[u_0(n)], \text{ або } h(k, n) = h(n-k) = F[u_0(n-k)]. \quad (6.13)$$

де F – оператор перетворення дискретного фільтра. Імпульсну характеристику фільтра $h(k, n)$ зазвичай представляють сукупністю значень $\{h(0), h(1), h(2), \dots\}$.

Зазначимо, що будь-яку дискретну послідовність $x(n)$ можна представити у вигляді лінійної комбінації одиничних відліків

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)u_0(n-k). \quad (6.14)$$

Тоді сигнал на виході фільтра, виходячи з лінійності і стаціонарності такої системи, повинен представляти собою лінійну комбінацію імпульсних характеристик

$$\begin{aligned} y(n) &= F[x(n)] = F\left[\sum_{k=-\infty}^{\infty} x(k)u_0(n-k)\right] = \sum_{k=-\infty}^{\infty} x(k)F[u_0(n-k)] = \\ &= \sum_{k=-\infty}^{\infty} x(k)h(n-k) = \sum_{k=-\infty}^{\infty} x(n-k)h(k). \end{aligned} \quad (6.15)$$

Вираз (6.15) представляє собою *дискретну згортку* вхідного сигналу й імпульсної характеристики фільтра. Якщо фільтр каузальний, тобто реалізується фізично, тоді реакція на його виході не може з'явитись раніше впливу на вході (принцип часового детермінізму). Як наслідок, імпульсна характеристика такого фільтра повинна дорівнювати нулю у відлікових точках, які передують моменту подачі вхідного імпульсу, тобто $h(n) = 0$ при $n < 0$. Тому верхню границю підсумовування в формулі (6.15) може бути замінено на n

$$y(n) = \sum_{k=0}^n x(k)h(n-k) = \sum_{k=0}^n x(n-k)h(k). \quad (6.16)$$

Застосуємо третю властивість z -перетворення до рівняння дискретної фільтрації: $y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k)$. Оскільки

вихідний сигнал дискретної системи є лінійною згорткою вхідного сигналу з імпульсною характеристикою системи, тоді

$$Y(z) = X(z) \cdot H(z), \quad (6.17)$$

де $X(z), Y(z)$ – z -перетворення відповідно вхідного і вихідного сигналів системи, а $H(z)$ – z -перетворення її імпульсної характеристики

$$H(z) = \sum_{n=-\infty}^{\infty} h(n)z^{-n}. \quad (6.18)$$

Враховуючи, що $H(z) = Y(z)/X(z)$ дорівнює відношенню z -перетворень двох сигналів (вихідного і вхідного), її називають *системною або передавальною функцією*. Вона відіграє дуже важливу роль при побудові дискретних цифрових систем.

В загальному випадку дискретний фільтр може підсумовувати (з ваговими коефіцієнтами) деяку кількість вхідних відліків (в тому числі й останній) і деяку кількість попередніх вихідних відліків

$$y(k) = b_0x(k) + b_1x(k-1) + \dots + b_lx(k-l) - a_1y(k-1) - a_2y(k-2) - \dots - a_my(k-m), \quad (6.19)$$

де a_j і b_i – дійсні коефіцієнти (a_0 в цьому випадку дорівнює 1).

Застосуємо z -перетворення до обох частин різницевого рівняння

$$\begin{aligned} Y(z)[a_0 + a_1z^{-1} + a_2z^{-2} + \dots + a_mz^{-m}] &= \\ &= X(z)[b_0 + b_1z^{-1} + b_2z^{-2} + \dots + b_lz^{-l}] \end{aligned}, \quad (6.20)$$

звідки отримаємо загальний вигляд передавальної функції

$$H(z) = \frac{b_0 + b_1z^{-1} + b_2z^{-2} + \dots + b_lz^{-l}}{a_0 + a_1z^{-1} + a_2z^{-2} + \dots + a_mz^{-m}}. \quad (6.21)$$

Таким чином, передавальна (системна) функція дискретної системи, що фізично реалізується, може бути представлена у вигляді відношення поліномів по від'ємних степенях змінної z .

Дискретні лінійні системи (фільтри), які розглядаються в цій главі, є також і стаціонарними, тобто затримка вхідного сигналу призводить лише до такої ж затримки вихідного

сигналу, не змінюючи його форми. Лінійні нестационарні системи або фільтри, наприклад адаптивні, які змінюють свої параметри в залежності від статистичних властивостей вхідного сигналу розглядаються в главі 9.

Оскільки системна функція є z -перетворенням від імпульсної характеристики системи $H(z) = \sum_{n=-\infty}^{\infty} h(n)z^{-n}$,

порівняємо це співвідношення з виразом для частотного коефіцієнта передачі, який також може виражатися через значення імпульсної характеристики відповідно

$$K(j\omega) = \sum_{n=-\infty}^{\infty} h(n)e^{-j\omega nT}.$$

Очевидно, щоб отримати частотний коефіцієнт передачі (частотну характеристику) дискретної системи з його системної функції, в останній треба зробити підстановку $z = \exp(j\omega T)$.

Таким чином, маючи різницеве рівняння або структурну схему дискретної системи, не складно визначити її системну функцію і частотний коефіцієнт передачі.

Приклад 6.1. Визначити системну функцію рекурсивного фільтра другого порядку.

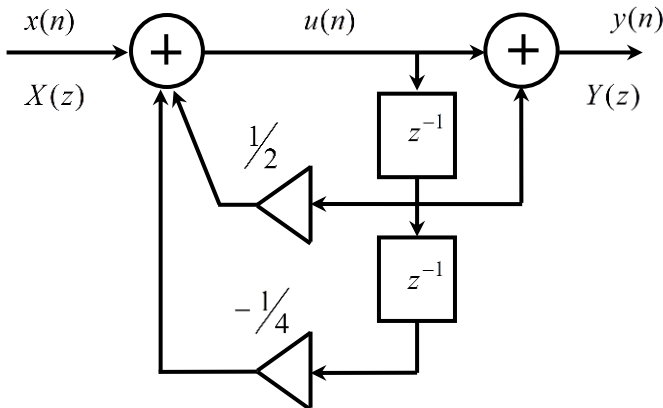


Рисунок 6.2 – Приклад рекурсивного фільтра другого порядку

Рішення. Введемо для аналізу проміжний сигнал $u(n)$ і запишемо рівняння відносно двох суматорів у формі різницевих рівнянь:

$$u(n) = x(n) + \frac{1}{2}u(n-1) - \frac{1}{4}u(n-2);$$

$$y(n) = u(n) + u(n-1).$$

Застосовуючи z -перетворення до кожного з цих рівнянь, отримуємо:

$$U(z) = X(z) + \frac{1}{2}z^{-1}U(z) - \frac{1}{4}z^{-2}U(z);$$

$$Y(z) = U(z) + z^{-1}U(z).$$

Отже,

$$U(z) = X(z) \frac{1}{1 - z^{-1}/2 + z^{-2}/4}, \quad Y(z) = U(z)(1 + z^{-1}).$$

Тоді, отримаємо $H(z) = \frac{z(z+1)}{z^2 - z/2 + 1/4}$.

Графічним представленням алгоритмів цифрової фільтрації є структурні схеми. Структурну схему дискретної системи можна скласти або по різницевому рівнянню, або за допомогою системної (передавальної) функції. Для *нерекурсивних (трансверсальних)* фільтрів рівняння дискретної фільтрації має вигляд

$$y(k) = \sum_{j=0}^m b_j x(k-j) = b_0 x(k) + b_1 x(k-1) + \dots + b_m x(k-m). \quad (6.22)$$

Цьому рівнянню відповідає передавальна функція

$$H(z) = \sum_{j=0}^m b_j z^{-j}, \quad (6.23)$$

де коефіцієнтами b_j є відліки імпульсної характеристики фільтра. Кількість використовуваних попередніх відліків m називають порядком фільтра.

Побудуємо структурну схему, що реалізує алгоритм рівнянь (6.22) і (6.23) у вигляді *прямої форми* (рис. 6.3).

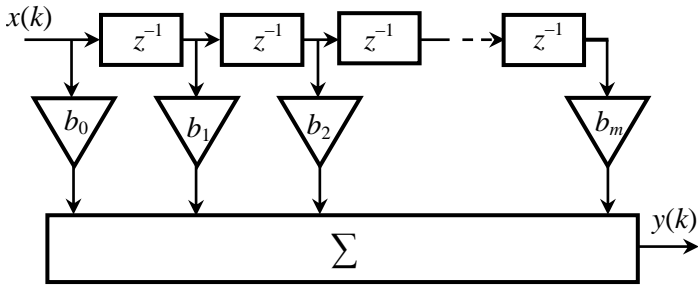


Рисунок 6.3 – Структурна схема
нерекурсивного (трансверсального) фільтра

Схема містить m елементів затримки z^{-1} на один крок дискретизації відліків сигналу, що здійснюють запам'ятовування, на час $T = 1/f_d$, $(m+1)$ помножувачів на постійні коефіцієнти b_j ($j = 0, 1, \dots, m$) і багатовходовий суматор Σ . Очевидно, що в реальному пристрої лінія затримки містить кінцеве число елементів, тому імпульсна характеристика такого фільтра є кінцевою за тривалістю (КІХ-фільтр).

Простота аналізу і реалізації, а також наочний зв'язок коефіцієнтів фільтра з відліками його імпульсної характеристики й абсолютна стійкість сприяють тому, що нерекурсивні фільтри широко застосовуються на практиці. Проте для отримання хороших частотних характеристик (наприклад, смугових фільтрів з високою прямокутністю АЧХ) потрібні нерекурсивні фільтри високого порядку – до декількох сотень і навіть тисяч.

При побудові нерекурсивних фільтрів застосовують й інші структури, наприклад, послідовні або паралельні структури, які будуть розглянуті пізніше.

Розглянемо можливі варіанти синтезу структур на прикладі рекурсивних фільтрів. Різницеве рівняння таких фільтрів в загальному випадку має вигляд

$$y(k) = \sum_{j=0}^m b_j x(k-j) - \sum_{i=0}^n a_i y(k-i), \quad (6.24)$$

а системна функція записується так

$$H(z) = \frac{\sum_{j=0}^m b_j z^{-j}}{\sum_{i=1}^n a_i z^{-i}} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}}. \quad (6.25)$$

Степені поліномів в чисельнику і знаменнику повинні задовольняти нерівність $n \geq m$. Проста структура реалізації різницевого рівняння (6.24), що відповідає прямій формі реалізації, приведена на рис. 6.4.

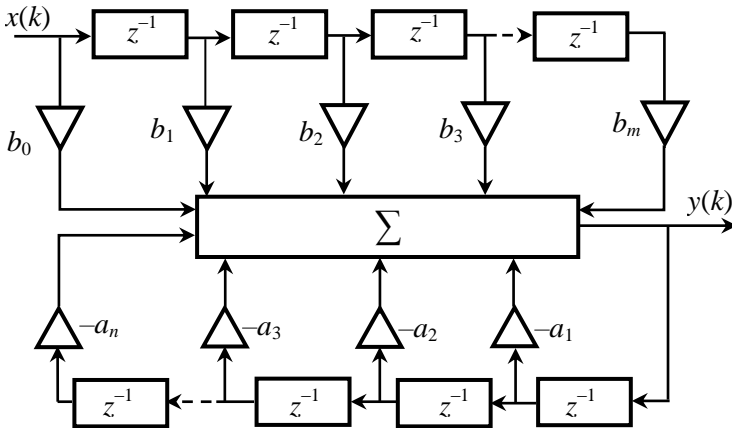


Рисунок 6.4 – Структурна схема прямої реалізації рекурсивного фільтра

Пряма форма реалізації складається з двох частин: верхня частина відображує першу суму алгоритму фільтрації і повністю відповідає структурі нерекурсивного лінійного фільтра, а нижня частина – другу суму алгоритму (6.24), яка є гілкою від’ємного зворотного зв’язку. Структурна схема містить $(n + m)$ елементів затримки z^{-1} на крок дискретизації,

$(n + m + 1)$ помножувачів на постійні коефіцієнти a_i, b_j , а також багатовходовий суматор.

Пряма форма реалізації фільтра проста, наочна, повністю відповідає системній функції (6.25) і не вимагає підвищеної розрядності ліній затримки. Проте, очевидним недоліком прямої форми є наявність великої кількості елементів окремо для нерекурсивної і рекурсивної частин. Число елементів затримки можна зменшити, реалізуючи рекурсивний фільтр в так званій канонічній формі.

Запишемо системну функцію $H(z)$ фільтра у вигляді

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{\sum_{i=1}^n a_i z^{-i}} \cdot \sum_{j=0}^m b_j z^{-j} = H_1(z) \cdot H_2(z). \quad (6.26)$$

Цифровий фільтр, що відповідає формулі (6.26), складається з двох послідовно сполучених фільтрів з передавальними функціями відповідно $H_1(z)$ і $H_2(z)$. Перший фільтр має тільки полюси, а другий – тільки нулі. Виразимо $H_1(z)$ і $H_2(z)$ за допомогою допоміжної функції $W(z)$:

$$H_1(z) = \frac{W(z)}{X(z)} = \frac{1}{\sum_{i=1}^n a_i z^{-i}}; \quad (6.27)$$

$$H_2(z) = \frac{Y(z)}{W(z)} = \sum_{j=0}^m b_j z^{-j}. \quad (6.28)$$

Співвідношенням (6.27) і (6.28) відповідає наступна пара різницьових рівнянь (у припущенні, що коефіцієнт $a_0 = 1$):

$$w(k) = x(k) - \sum_{i=1}^n a_i w(k - i); \quad (6.29)$$

$$y(k) = \sum_{j=0}^m b_j w(k - j). \quad (6.30)$$

Структура, яка реалізує ці різницеві рівняння (за умови, що $n = m = N$) показана на рис. 6.5, а. Її називають прямою формою № 2 (неканонічною). Проте, оскільки в гілках, відповідних $H_1(z)$ і $H_2(z)$, сигнал $w(k)$ затримується однаково, то для побудови фільтра досить використовувати один набір елементів затримки. Ця структура приведена на рис. 6.5, б, її називають канонічною, оскільки використовуване число елементів затримки в точності дорівнює порядку системної (передавальної) функції.

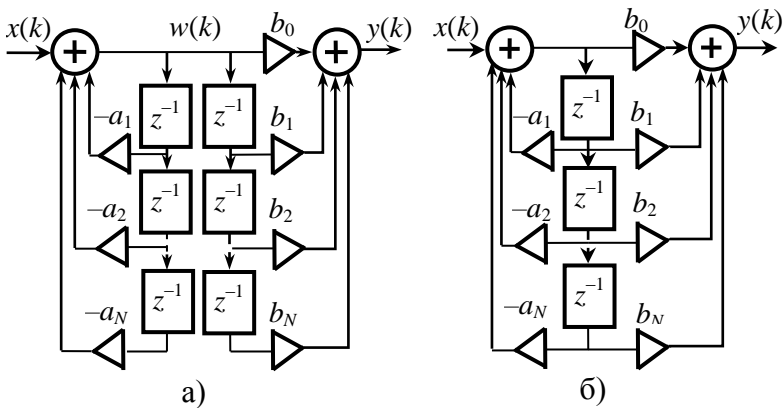


Рисунок 6.5 – Структурна схема рекурсивного цифрового фільтра: а) пряма № 2; б) канонічна

Записавши формулу (6.25) у вигляді

$$H(z) = \frac{Y(z)}{X(z)} = b_0 \prod_{i=1}^K H_i(z), \quad (6.31)$$

отримаємо, так звану, *послідовну, або каскадну* структуру побудови цифрового фільтра (рис. 6.6).

Зазвичай множники $H_i(z)$ відповідають або блокам першого порядку

$$H_i(z) = \frac{b_{0i} + b_{1i}z^{-1}}{1 + a_{1i}z^{-1}}, \quad (6.32)$$

або блокам другого порядку

$$H_i(z) = \frac{b_{0i} + b_{1i}z^{-1} + b_{2i}z^{-2}}{1 + a_{1i}z^{-1} + a_{2i}z^{-2}}. \quad (6.33)$$

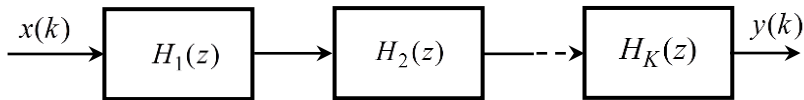


Рисунок 6.6 – Каскадна побудова цифрового фільтра

Кожен з блоків другого порядку, що утворюють послідовну форму, можна реалізувати або в прямій формі, або в канонічній формі. Функції виду (6.33) називають *біквдратними блоками*, вони є універсальною ланкою, придатною для побудови цифрового фільтра складнішої структури.

Розклавши праву частину формули (6.25) на прості дроби, отримаємо четверту структурну схему рекурсивного фільтра

$$H(z) = C + \sum_{i=1}^K H_i(z). \quad (6.34)$$

Доданки $H_i(z)$ відповідають блокам першого порядку виду

$$H_i(z) = \frac{b_{0i}}{1 + a_{1i}z^{-1}} \quad (6.35)$$

чи блокам другого порядку

$$H_i(z) = \frac{b_{0i} + b_{1i}z^{-1}}{1 + a_{1i}z^{-1} + a_{2i}z^{-2}}. \quad (6.36)$$

Структурна схема, що реалізує співвідношення (6.25) і звана *паралеельною формою*, приведена на рис. 6.7.

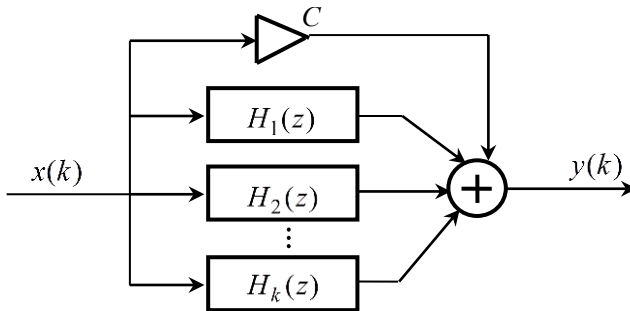


Рисунок 6.7 – Паралельна форма побудови фільтра

На практиці також розповсюдженою є форма побудови фільтра із зворотним зв'язком (рис. 6.8). Передавальні функції $H_1(z)$ і $H_2(z)$ на рис. 6.8 можуть бути результатом каскадного та (або) паралельного з'єднання кількох ланок.

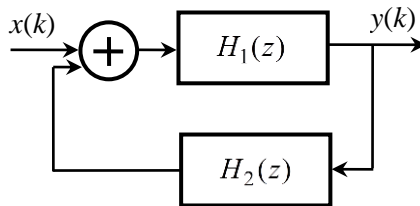


Рисунок 6.8 – Цифровий фільтр із зворотним зв'язком

Співвідношення для знаходження передавальної функції такого фільтра має вигляд

$$H(z) = \frac{H_1(z)}{1 - H_1(z) \cdot H_2(z)}.$$

Приклад 6.2. Нехай

$$H_1(z) = 1/(1 - 0,3z^{-1}); \quad H_2(z) = 0,2 + z^{-1} + z^{-2},$$

тоді для трьох форм побудови фільтра матимемо

$$H_K(z) = (0,2 + z^{-1} + z^{-2}) / (1 - 0,3z^{-1});$$

$$H_{\Pi} = (0,2 + 0,7z^{-1} + 0,7z^{-2} - 0,3z^{-3}) / (1 - 0,3z^{-1});$$

$$H_{ЗВ} = 1,25 / (1 - 1,625z^{-1} - 1,25z^{-2}).$$

Стійкість дискретних систем.

Лінійна дискретна система з постійними параметрами (стаціонарний фільтр) називається стійкою, якщо за будь-яких початкових умов і будь-якого обмеженого вхідного сигналу $x(n)$ вихідний сигнал $y(n)$ також залишається обмеженим, тобто з умови $x(n) \leq P_x < \infty$ для усіх n слідує, $y(n) \leq P_y < \infty$.

Необхідною і достатньою умовою стійкості одновимірного стаціонарного лінійного фільтра є наступна вимога до його імпульсної характеристики:

$$\sum_{n=-\infty}^{\infty} |h(n)| < \infty,$$

тобто вимога збіжності ряду, складеного з модулів відліків імпульсної характеристики фільтра.

Нерекурсивні стаціонарні лінійні фільтри мають чудову особливість: їх імпульсні характеристики $h(n)$ мають кінцеве число ненульових відліків, причому ці відліки дорівнюють коефіцієнтам b_j алгоритму фільтрації. Дійсно, відповідно до (6.16)

$$H(z) = \sum_{n=0}^{\infty} h(n)z^{-n} = \sum_{j=0}^m b_j z^{-j}.$$

Звідки видно, що

$$h(n) = \begin{cases} b_n, & n = 0, 1, \dots, m \\ 0, & n > m \end{cases} \quad (6.37)$$

Таким чином, імпульсна характеристика *нерекурсивного стаціонарного лінійного фільтра* має кінцеве число відмінних від нуля відліків, і відповідно до умови $\sum_{n=-\infty}^{\infty} |h(n)| < \infty$ такий

фільтр завжди стійкий. Саме властивість (6.37) зумовила ще одну назву таких фільтрів – фільтри з кінцевою імпульсною характеристикою (КІХ-фільтри).

Стійкість рекурсивних дискретних систем

Для рекурсивних дискретних систем використовувати критерій стійкості у формі $\sum_{n=-\infty}^{\infty} |h(n)| < \infty$ складно, оскільки необхідно підсумовувати нескінченний ряд модулів відліків імпульсної характеристики. Для дослідження рекурсивних фільтрів використовується такий критерій: якщо полюси функції лежать усередині круга одиничного радіусу z -площини, то такий фільтр стійкий. Якщо хоча б один полюс $H(z)$ розташований на одиничному колі $|z|=1$ або в зовнішній частині круга одиничного радіусу, то така $H(z)$ представляє нестійкий фільтр. Положення нулів системної функції $H(z)$ не впливає на стійкість фільтра.

Приклад 6.3. Перевірити на стійкість наступний рекурсивний фільтр другого порядку (рис. 6.2).

Рішення. Системна функція фільтра (рис. 6.2)

$$H(z) = \frac{z(z+1)}{z^2 - z/2 + 1/4}.$$

Знайдемо полюси системної функції p_1, p_2 , які є коренями характеристичного рівняння

$$p^2 - p/2 + 1/4 = 0.$$

Оскільки $p_1, p_2 = \frac{1}{4} \pm j \frac{\sqrt{3}}{4}$, тобто $|p_1| = |p_2| = 0,5 < 1$, отже, полюси системної функції $H(z)$ лежать усередині круга одиничного радіусу, і фільтр є стійким.

Примітка. Нестійкий фільтр є непрацездатним у випадку, коли сигнал $x(n)$ на його вході діє необмежено довго, оскільки при цьому сигнал $y(n)$ на виході фільтра перестане залежати від сигналу $x(n)$. Але, якщо сигнал $x(n)$ діє протягом обмеженого інтервалу часу, фільтр є працездатним і

використовується на практиці. Наприклад, цифровий інтегратор з системною функцією $H(z) = 1/(1 - z^{-1})$ (ця функція має полюс $z=1$, тобто фільтр – нестійкий) цілком працездатний і використовується на практиці, якщо вхідний сигнал $x(n)$ діє лише при $0 \leq n \leq N-1$, після чого слідує скидання і відновлення початкових умов. Цифровий інтегратор може входити до складу замкнутої стежачої системи. При цьому за рахунок зворотного зв'язку система може мати стійкість, не дивлячись на те, що інтегратор, що є динамічною ланкою системи, нестійкий.

Недоліки полюсного критерію стійкості обумовлені необхідністю визначення коренів характеристичного рівняння, що є полюсами системної функції. Аналітичних методів рішення алгебраїчних рівнянь, порядок яких вищий за четвертий, не існує. Тому знаходження полюсів системної функції високого порядку можливо лише чисельними методами.

Існують алгебраїчні і частотні критерії стійкості стаціонарних лінійних дискретних систем, що дозволяють судити про стійкість фільтра без знаходження коренів характеристичного рівняння. Наприклад, критерій Джурі, аналогічний критерію Рауса – Гурвіца для аналогових систем.

Розглянемо один з різновидів алгебраїчного критерію стійкості, названого на честь американського вченого Е. Джурі. Нехай характеристичне рівняння дискретної системи має наступний вигляд: $\alpha_m z^m + \alpha_{m-1} z^{m-1} + \dots + \alpha_1 z + \alpha_0 = 0$ [23].

Щоб визначити, чи всі корені характеристичного рівняння дискретної системи знаходяться всередині одиничного кола, будують наступну таблицю (табл. 6.1) [23], в якій $\alpha_i^j = \alpha_i^{j-1} - \alpha_{j+m-i-1} a_j$.

У таблиці Джурі перший і другий рядки – коефіцієнти характеристичного рівняння в прямому і зворотному порядку. Третій рядок отримується множенням другого рядка на $a_1 = \alpha_0/\alpha_m$ і відніманням добутку з першої. Четвертий рядок

– це записаний в зворотному порядку третій рядок. Схема повторюється до тих пір, поки в останньому рядку не залишиться єдиний елемент α_m^m .

Таблиця 6.1 – Таблиця Джурі

α_m	α_{m-1}	...	α_1	α_0	$a_1 = \alpha_0 / \alpha_m$
α_0	α_1	...	α_{m-1}	α_m	
α_m^1	α_{m-1}^1	...	α_1^1		$a_2 = \alpha_1^1 / \alpha_m^1$
α_1^1	α_2^1	...	α_m^1		
...
α_m^m					

Згідно критерію Джурі, якщо α_m і всі $\alpha_m^j, j = \overline{1, m}$ додатні, то всі корені характеристичного рівняння лежать всередині одиничного кола. Кількість від'ємних α_m^j дорівнює кількості коренів поза одиничним колом, якщо серед α_m^j немає нульових.

Для оцінки стійкості лінійної системи, яка має характеристичне рівняння $z^m + \sum_{i=1}^m a_i z^{m-i} = 0$, можливо також скористатись заміною $z = (1+w)/(1-w)$ [10]. В результаті перетворень характеристичного рівняння буде отримане рівняння того ж порядку відносно змінної w , але корені цього рівняння для стійкої системи будуть знаходитись в лівій напівплощині комплексної площини. У цьому випадку перевірка стійкості може виконуватись по відомих для аналогових систем критеріях, наприклад, Рауса – Гурвіца.

6.3. Проектування цифрових лінійних систем

Під проектуванням (синтезом) цифрової лінійної системи або цифрового фільтра розуміють вибір таких коефіцієнтів $\{a_i\}$, $\{b_j\}$ системної (передавальної) функції, при яких характеристики фільтра задовольняють заданим вимогам. Строго кажучи, в завдання проектування входить і вибір відповідної структури фільтра з урахуванням кінцевої точності обчислень. Це особливо актуально при реалізації фільтрів в апаратурному виді (у вигляді спеціалізованих ВІС або цифрових сигнальних процесорів). Тому в цілому проектування цифрового фільтра складається з наступних етапів [24]:

1. Рішення задачі апроксимації з метою визначення коефіцієнтів фільтра і системної функції, що задовольняє конкретним вимогам.

2. Вибір схеми побудови фільтра, тобто перетворення системної функції в конкретну структурну схему фільтра.

3. Оцінка ефектів квантування, тобто ефектів, пов'язаних з кінцевою точністю представлення чисел в цифрових системах, що мають кінцеву розрядність.

4. Перевірка методами моделювання чи задовольняє отриманий фільтр заданим вимогам.

В цій главі розглядається лише задача апроксимації цифрових фільтрів. Методи їх синтезу можна класифікувати за різними ознаками:

1. за типом отриманого фільтра:
 - методи синтезу фільтрів з кінцевою імпульсною характеристикою;
 - методи синтезу фільтрів з нескінченною імпульсною характеристикою;
2. за наявністю аналогового прототипу:
 - методи синтезу з використанням аналогового прототипу;

- прями методи синтезу (без використання аналогового прототипу).

На практиці КІХ-фільтрам часто віддають перевагу, для цього є наступні причини. По-перше, КІХ-фільтри забезпечують можливість точного обчислення вихідного сигналу при обмеженому вхідному по згортці, що не вимагає усікання імпульсної характеристики. По-друге, фільтри з кінцевою імпульсною характеристикою можуть мати строго лінійну ФЧХ в смузі пропускання, що дозволяє проектувати фільтри з амплітудною характеристикою, що не спотворює вхідні сигнали. По-третє, КІХ-фільтри завжди стійкі і, при введенні відповідної кінцевої затримки, фізично реалізуються. Крім того, КІХ-фільтри можуть бути реалізовані не лише за нерекурсивними схемами, але й з використанням рекурсивних форм. При нерекурсивній реалізації КІХ-фільтрів шуми округлення, пов'язані з кінцевою точністю виконання арифметичних операцій, легко мінімізувати.

Відмітимо недоліки КІХ-фільтрів:

1. Для апроксимації фільтрів, частотні характеристики яких мають гострі зрізи, потрібно імпульсну характеристику з великим числом відліків N . Тому при використанні звичайної згортки необхідно виконувати великий об'єм обчислень. Тільки розробка на основі високоефективного алгоритму БПФ методів швидкої згортки дозволила КІХ-фільтрам успішно конкурувати з НІХ-фільтрами, що мають гострі зрізи в частотній характеристиці.

2. Затримка в КІХ-фільтрах з лінійною фазовою характеристикою не завжди дорівнює цілому числу інтервалів дискретизації. У деяких застосуваннях така некротна затримка може викликати певні труднощі.

Один з варіантів проектування цифрових фільтрів пов'язаний із заданою послідовністю відліків імпульсної характеристики, які використовують для отримання і аналізу його частотної характеристики (частотного коефіцієнта передачі).

Отримаємо умову, за якої нерекурсивний фільтр має строго лінійну ФЧХ [24]. Системна функція такого фільтра має вигляд

$$H(z) = \sum_{n=0}^{N-1} h(n)z^{-n}, \quad (6.38)$$

де коефіцієнти фільтра $h(n)$ є відліками імпульсної характеристики. Перетворення Фур'є від $\{h(n)\}$ є частотною характеристикою фільтра, періодичною по частоті з періодом 2π , тобто

$$H(e^{j\omega}) = H\left[e^{j(\omega+2\pi m)}\right], \quad m = 0, \pm 1, \pm 2, \dots$$

Представимо її для дійсної послідовності $\{h(n)\}$ у вигляді $K(e^{j\omega}) = \pm |K(e^{j\omega})| e^{j\theta(\omega)}$. Отримаємо умови, за яких імпульсна характеристика фільтра забезпечуватиме строгу лінійність його фазової характеристики. Останнє означає, що фазова характеристика $\theta(\omega)$ повинна мати вигляд

$$\theta(\omega) = -\alpha\omega, \quad -\pi \leq \omega \leq \pi, \quad (6.39)$$

де α – постійна часова затримка, виражена через число інтервалів дискретизації (в цьому випадку фазова і групова затримка сигналу відповідно до їх визначення в розділі 6.1 збігаються). Запишемо частотну характеристику $K(e^{j\omega})$ у вигляді

$$K(e^{j\omega}) = \sum_{n=0}^{N-1} h(n)z^{-j\omega n} = \pm |K(e^{j\omega})| e^{-j\alpha\omega}. \quad (6.40)$$

Прирівнюючи дійсні і уявні частини, отримаємо:

$$\sum_{n=0}^{N-1} h(n) \cos(\omega n) = \pm |K(e^{j\omega})| \cos(\alpha\omega), \quad (6.41)$$

$$\sum_{n=0}^{N-1} h(n) \sin(\omega n) = \pm |K(e^{j\omega})| \sin(\alpha\omega), \quad (6.42)$$

звідки

$$\operatorname{tg}(\alpha\omega) = \frac{\sum_{n=1}^{N-1} h(n) \sin(\omega n)}{h(0) + \sum_{n=1}^{N-1} h(n) \cos(\omega n)}. \quad (6.43)$$

Існує два можливі рішення рівняння (6.43). Одне (при $\alpha = 0$) не представляє інтересу, інше відповідає випадку $\alpha \neq 0$. Перехресно помноживши члени рівняння (6.43), отримуємо

$$\sum_{n=0}^{N-1} h(n) \cos(\omega n) \sin(\alpha\omega) - \sum_{n=0}^{N-1} h(n) \sin(\omega n) \cos(\alpha\omega) = 0,$$

звідки

$$\sum_{n=0}^{N-1} h(n) \sin[(\alpha - n)\omega] = 0. \quad (6.44)$$

Оскільки рівняння (6.44) має вигляд ряду Фур'є, тоді його рішення повинне задовольняти наступним умовам:

$$\alpha = \frac{N-1}{2}, \quad (6.45)$$

$$h(n) = h(N-1-n), \quad 0 \leq n \leq N-1. \quad (6.46)$$

З умови (6.45) виходить, що для кожного N існує тільки одна часова затримка α , при якій може досягатися строга лінійність фазової характеристики фільтра. З (6.46) витікає, що при заданому α , що задовольняє умові (6.45), імпульсна характеристика повинна мати цілком певну симетрію.

Доцільно розглянути використання умов (6.45) і (6.46) окремо для випадків парного і непарного N . Якщо N непарне число, то α ціле число, тобто затримка у фільтрі дорівнює цілому числу інтервалів дискретизації. В цьому випадку центр симетрії доводиться на $(N-1)/2$ відлік. Якщо ж N – парне число, то α дробове число, і затримка у фільтрі дорівнює нецілому числу інтервалів дискретизації. Наприклад, для $N=8$ отримуємо $\alpha=3,5$, і центр симетрії імпульсної характеристики лежить посередині між двома відліками.

Значення коефіцієнтів імпульсної характеристики використовують для обчислення частотної характеристики КІХ-фільтрів. Можна показати, що для симетричної імпульсної характеристики з непарним числом відліків вираз для дійсної функції $K^*(e^{j\omega})$, що набуває додатних і від'ємних значень, має вигляд:

$$K^*(e^{j\omega}) = \sum_{n=0}^{(N-1)/2} a(n) \cos(\omega n), \quad (6.47)$$

де

$$a(0) = h[(N-1)/2], a(n) = 2h[(N-1)/2 - n], n = 1, 2, \dots, (N-1)/2.$$

Найчастіше при проектуванні КІХ-фільтра виходять з необхідної (чи бажаної) частотної характеристики з наступним обчисленням коефіцієнтів фільтра. Існують декілька методів розрахунку фільтрів з лінійними фазовими характеристиками [24]: *метод проектування за допомогою вікон, метод частотної вибірки, метод розрахунку оптимального (по Чебишеву) фільтра*. Розглянемо ідею проектування методом вікон на прикладі КІХ-фільтра нижніх частот.

Зазначимо, що, оскільки частотна характеристика будь-якого цифрового фільтра є періодичною функцією частоти, вона може бути розкладена в ряд Фур'є

$$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h(n) e^{-j\omega n}, \quad (6.48)$$

де

$$h(n) = \frac{1}{2\pi} \int_0^{2\pi} H(e^{j\omega}) e^{-j\omega n} d\omega.$$

Видно, що коефіцієнти Фур'є $h(n)$ збігаються з Фур'є коефіцієнтами імпульсної характеристики фільтра. Однак, імпульсна характеристика має нескінченну тривалість, оскільки підсумовування ведеться в нескінченних границях. Крім цього, фільтр не реалізується фізично, тому що

імпульсна характеристика починається в $-\infty$ і ніяка кінцева часова затримка не зробить фільтр фізично реалізовуваним. Таким чином, фільтр, який отримується шляхом представлення $H(e^{j\omega})$ рядом Фур'є виявляється не реалізовуваним фізично НІХ-фільтром.

Для отримання нерекурсивного фільтра заданого порядку, що реалізовується, послідовність коефіцієнтів $h(n)$ усікається – з неї вибирається центральний фрагмент потрібної довжини. Просте усікання відліків імпульсної характеристики відповідає використанню N -точкового *прямокутного вікна*, що задається ваговою функцією (N далі вважається непарним)

$$w_R(k) = \begin{cases} 1, & 0, \quad -\frac{N-1}{2} \leq k \leq \frac{N-1}{2}, \\ 0, & \text{при інших } k. \end{cases}$$

Частотна характеристика такого вікна [24]

$$W_R(e^{j\omega n}) = \sum_{n=-(N-1)/2}^{(N-1)/2} e^{-j\omega n} = \frac{\sin(\omega N/2)}{\sin(\omega/2)}$$

має пелюстковий характер подібний до функції $\sin x / x$.

Таким чином, із-за усікання відліків задана частотна характеристика фільтра спотворюється, оскільки вона є згорткою в частотній області частотних характеристик $H(e^{j\omega})$ і $W_R(e^{j\omega n})$. Це веде до прояву відомого явища Гіббса (див. розд. 2.2), тобто до появи викидів і пульсацій певного рівня до і після точки розриву частотної характеристики, що апроксимується. Наприклад, при апроксимації частотних характеристик ідеальних фільтрів низьких частот або смугового максимальна амплітуда пульсацій складає близько 9 %, причому вона не зменшується при збільшенні N . При зростанні ж N спостерігається лише зменшення ширини викиду. Тому для покращення якості апроксимації застосовують замість прямокутних вікна іншої форми.

Форма вікна (вагової послідовності кінцевої тривалості) $w(n)$ обирається таким чином, щоб шляхом модифікації коефіцієнтів $\hat{h}(n) = h(n) \cdot w(n)$ покращити збіжність ряду (6.48), причому за межами вікна $-M \leq n \leq M$ $\hat{h}(n) = 0$. Як вагові функції використовують вікна косинус на підставці (вікно Хемінга), косинус-квадрат (вікно Хана), Кайзера, Блекмена, Чебишева та ін. (рис. 6.9).

Представлене на рис. 6.9 вікно Хемінга має бічні пелюстки на рівні меншому за -43 дБ по відношенню до рівня головної пелюстки (для прямокутного вікна перша бічна пелюстка має рівень -13 дБ), однак ширина головної пелюстки на рівні -3 дБ в $1,46$ рази більша, ніж у прямокутного вікна.

Наявність значної кількості різноманітних вікон пояснюється неможливістю одночасного забезпечення виконання обох вимог до вікон. Співвідношення для їх розрахунку та приклади найбільш поширених вікон можна знайти в [25, 26].

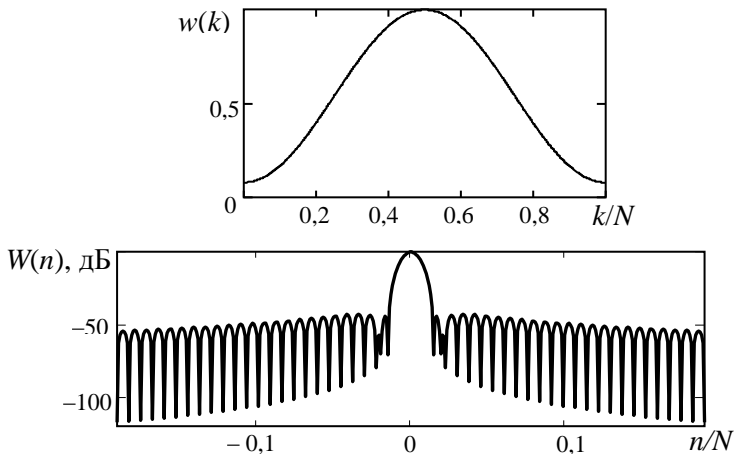


Рисунок 6.9 – Приклад вікна Хемінга

$$w(k) = 0,54 - 0,46 \cdot \cos\left(\frac{2\pi k}{N-1}\right) \text{ та його АЧХ}$$

Застосування методу ілюструється на рис. 6.10. На рис. 6.10, а показані періодична частотна характеристика $H(e^{j\omega})$ фільтра та відповідні їй коефіцієнти Фур'є $\{h(n)\}$. На рис. 6.10, б – відновлена частотна характеристика за усіченою прямокутним вікном послідовністю $\hat{h}(n) = h(n)$, де $-M \leq n \leq M$.

На рис. 6.10, в, г представлені модифіковані $\hat{h}(n) = h(n) \cdot w(n)$ трикутним вікном та вікном Хемінга та відповідна їй послідовність $g(n)$ із затримкою, яка може представляти фізично реалізовану імпульсну характеристику фільтра.

Таким чином, методом проектування КІХ-фільтрів за допомогою вікон є метод зменшення розривів вікна $w(k)$ шляхом використання вікон, відмінних від прямокутного. При цьому вагова функція (вікно) повинна мати наступні властивості:

ширина головної пелюстки частотної характеристики вікна, що містить по можливості велику частину загальної енергії, має бути якомога меншою;

енергія в бічних пелюстках частотної характеристики вікна повинна швидко зменшуватися при наближенні ω до π .

Вибір того чи іншого вікна пов'язаний з характером аналізованих даних. Необхідно враховувати динамічний діапазон корисних сигналів та наближеність джерел завад до них. Потрібен компроміс, який би забезпечував по можливості якісне розділення і спостереження сигнальних складових та зниження впливу джерел завад. Так, наприклад, якщо завади впливають в області дальніх бічних пелюсток, тоді краще обирати вікно з швидким спаданням рівня бічних пелюсток, їх же рівень поблизу сигнальної компоненти не має суттєвого значення. Навпаки, якщо поруч з потужною сигнальною компонентою потрібно спостерігати значно менш потужну, потрібне вікно з дуже швидким спаданням рівня бічних пелюсток. Для забезпечення при цьому якомога більш вузької

головної пелюстки, за відсутності віддалених компонент сигналу, рівень дальніх пелюсток може навіть зростати [25]. Якщо спектр сигналу відносно рівномірний, вікна взагалі можуть не застосовуватись.

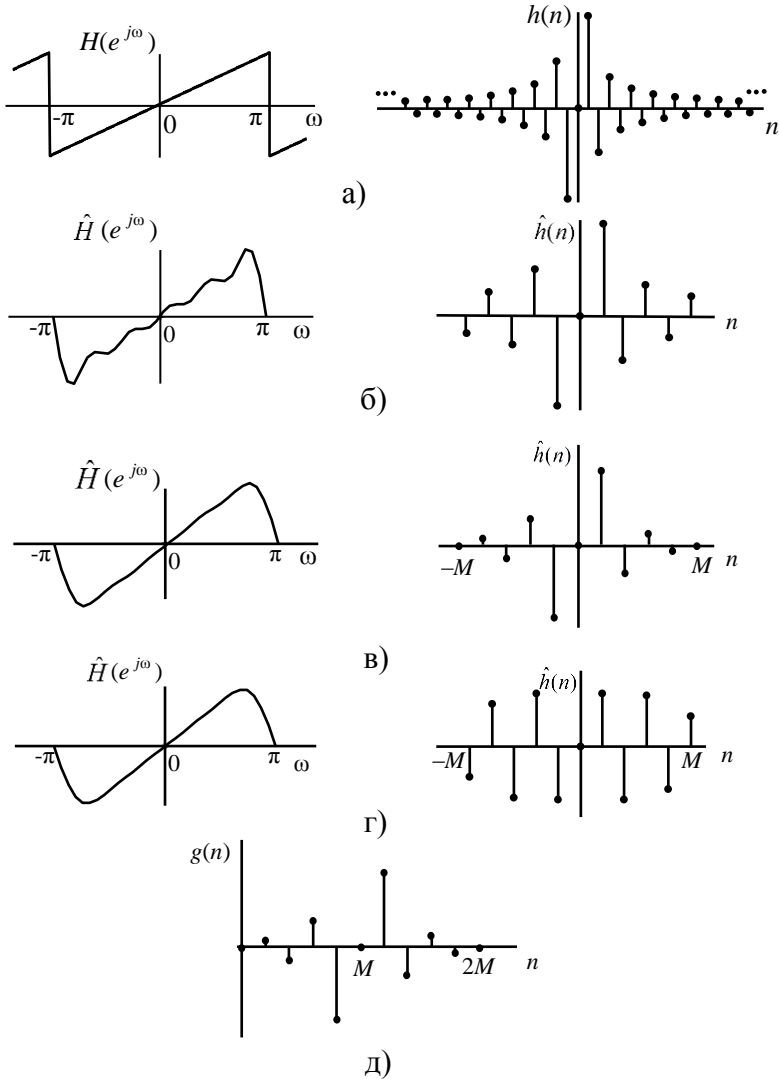


Рисунок 6.10 – Ілюстрація застосування методу вікон

Проектування фільтрів з імпульсною характеристикою нескінченної довжини.

Фільтри з нескінченною імпульсною характеристикою (НІХ-фільтри) суттєво відрізняються від КІХ-фільтрів із-за наявності зворотного зв'язку. По-перше, вони вимагають перевірки на стійкість, як і усі системи із зворотним зв'язком. По-друге, мають складнішу структуру, їх важче проектувати і аналізувати. Крім того, їх фазочастотна характеристика принципово нелінійна. Однак їх використовують, тому що вони вимагають набагато менше множень на один вихідний відлік, щоб реалізувати необхідну частотну характеристику. Це дозволяє будувати фільтри реального часу, які працюють на значно вищих частотах дискретизації, ніж КІХ-фільтри. Друга принципова перевага полягає в тому, що НІХ-фільтри на відміну від КІХ-фільтрів можуть апроксимувати задані аналогові фільтри.

Стандартні методи проектування (синтезу) НІХ-фільтрів діляться на три базові класи:

метод інваріантного перетворення імпульсної характеристики;

метод білінійного z -перетворення;

оптимізаційні методи, засновані на алгоритмах ітераційного моделювання коефіцієнтів фільтра.

Метод інваріантного перетворення імпульсної характеристики, заснований на дискретизації імпульсної характеристики аналогового прототипу. Для цього використовуємо загальну форму запису z -перетворення імпульсної характеристики НІХ-фільтрів

$$H(z) = \sum_{n=0}^{\infty} h(n)z^{-n} = \frac{\sum_{j=0}^N b_j z^{-j}}{1 + \sum_{i=1}^M a_i z^{-i}}. \quad (6.49)$$

Існують два варіанти методу інваріантного перетворення.

Варіант 1 вимагає застосування як оберненого перетворення Лапласа, так і z -перетворення. Його основні етапи (кроки) полягають в наступному:

1. Отримати передавальну функцію $H(s)$ аналогового фільтра-прототипу з необхідною частотною характеристикою.

2. По передавальній функції $H(s)$ визначити неперервну імпульсну характеристику $h(t)$, використовуючи обернене перетворення Лапласа.

3. Визначити частоту дискретизації f_d і крок дискретизації $T = 1/f_d$. Частоту дискретизації вибирають залежно від максимальної частоти аналогового фільтра-прототипу. Через проблеми накладення спектральних характеристик, властивих цьому методу, f_d повинна набагато перевершувати ширину спектру фільтрованих сигналів.

4. Підставити у вираз імпульсної характеристики замість неперервної змінної t дискретні значення з кроком T . Цим забезпечується рівність відліків дискретної імпульсної характеристики $h(n)$ значенням неперервної імпульсної характеристики в моменти часу $t = nT$.

5. Знайти z -перетворення функції $h(nT)$ і отримати системну (передавальну) функцію НХ-фільтра $H(z)$ у формі відношення поліномів від змінної z .

Примітка. Оскільки при дискретизації неперервної імпульсної характеристики частотна характеристика цифрового фільтра виявляється помноженою на коефіцієнт $1/T$, багато розробників фільтрів вважають необхідним включити множник T у вираз для $H(z)$. Це дозволяє зробити коефіцієнт передачі цифрового фільтра рівним коефіцієнту передачі прототипу. Деякі автори вважають за краще вводити множник T в дискретну імпульсну характеристику, тобто записати $h(n) = Th(nT)$.

Остаточно для $H(z)$ отримуємо:

$$H(z) = \frac{T(b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_N z^{-N})}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_M z^{-M}}, \quad (6.50)$$

і різницеве рівняння в загальній формі записується як:

$$y(n) = T[b_0 x(n) + b_1 x(n-1) + \dots + b_M x(n-N)] - a_1 y(n-1) - a_2 y(n-2) - \dots - a_N y(n-M). \quad (6.51)$$

Варіант 2 методу інваріантного перетворення імпульсної характеристики використовує інший підхід. Аналоговий фільтр-прототип математично розбивається на декілька однополюсних фільтрів, а потім кожен з цих фільтрів апроксимується однополюсним цифровим фільтром. Набір з M однополюсних фільтрів потім аналітично об'єднується в НІХ-фільтр M -го порядку, що має M полюсів.

При розрахунку фільтра цим методом необхідно виконати наступні кроки:

1. Отримати передавальну функцію $H(s)$ аналогового фільтра-прототипу у формі:

$$H(s) = \left[\sum_{k=0}^N b(k) s^k \right] / \left[\sum_{k=0}^M a(k) s^k \right]. \quad (6.52)$$

2. Обрати відповідну частоту дискретизації f_d і обчислити період (крок) дискретизації $T = 1/f_d$.

3. Виразити передавальну функцію $H(s)$ у вигляді суми однополюсних передавальних функцій. Це вимагає використання розкладання (6.52) на прості дроби виду

$$H(p) = \frac{[b_N s^N + b_{N-1} s^{N-1} + \dots + b_1 s + b_0]}{[a_M s^M + a_{M-1} s^{M-1} + \dots + a_1 s + a_0]} = \sum_{k=1}^M \frac{A_k}{(s + p_k)}, \quad (6.53)$$

де коефіцієнти A_k є константами; p_k – k -й полюс в точці $-p_k$ на s -площині. Позначимо k -у однополюсну передавальну функцію як $H_k(p)$

$$H_k(p) = A_k / (p + p_k). \quad (6.54)$$

4. Визначити імпульсну характеристику $h(t)$ аналогового фільтра з передавальною функцією виду (6.53) і записати її у формі:

$$h(t) = \sum_{k=1}^M h_k(t) = \sum_{k=1}^M A_k e^{-p_k t} \sigma(t), \quad (6.55)$$

де $\sigma(t)$ – функція Хевісайда.

5. Апроксимувати кожен однополіусний аналоговий фільтр з передавальною функцією $H_k(s)$ однополіусним цифровим фільтром з передавальною функцією $H_k(z)$. Для цього шляхом дискретизації імпульсної характеристики аналогового фільтра отримуємо імпульсну характеристику однополіусного цифрового фільтра

$$h_k(nT) = A_k^{-p_k nT} u_1(nT), \quad (6.56)$$

де $u_1(nT)$ – функція одиничного стрибка (5.31).

Знайдемо z -перетворення цього однополіусного фільтра:

$$H_k(z) = \sum_{n=0}^{\infty} h(nT) z^{-n} = A_k \sum_{n=0}^{\infty} (e^{-p_k T} z^{-1})^n = \frac{A_k}{1 - e^{-p_k T} z^{-1}}. \quad (6.57)$$

Апроксимація полягає у відображенні кожного полюса $H_k(p)$, розташованого в точці $p = -p_k$ на s -площині, в точку $z = e^{-p_k T}$ на z -площині. Іншими словами, апроксимація здійснюється за допомогою відображення, при якому використовується заміна

$$\frac{1}{p + p_k} \rightarrow \frac{1}{1 - e^{-p_k T} z^{-1}}. \quad (6.58)$$

Результуюча передавальна функція дискретного фільтра є сумою передавальних функцій однополіусних дискретних фільтрів

$$H(z) = \sum_{k=1}^M H_k(z) = \sum_{k=1}^M \frac{A_k}{1 - e^{-p_k T} z^{-1}}. \quad (6.59)$$

6. Записати вираз (6.59) у вигляді відношення двох поліномів від z . Оскільки $H(z)$ є сумою простих дробів, приводячи їх до спільного знаменника, отримаємо:

$$H(z) = Y(z) / X(z) = \frac{\sum_{k=0}^N b(k)z^{-k}}{1 - \sum_{k=1}^M a(k)z^{-k}}. \quad (6.60)$$

По аналогії з варіантом 1 з (6.60) витікає різницеве рівняння в узагальненій формі з відомими коефіцієнтами. Різницеве рівняння типу (6.60) можна реалізувати або у вигляді простої форми НІХ-фільтра, або у вигляді поліпшених структур, варіанти яких розглядалися вище.

Частотна характеристика цифрового фільтра пов'язана з частотною характеристикою аналогового прототипу так саме, як і спектр дискретизованого сигналу пов'язаний із спектром аналогового сигналу – періодичним повторенням. Тому для забезпечення відповідності частотної характеристики вихідного аналогового фільтра і інваріантного перетворення імпульсної характеристики цифрового фільтра необхідно, щоб смуга пропускання аналогового фільтра знаходилась в межах діапазону: $-\pi/T \leq \omega \leq \pi/T$. Для виконання цієї умови необхідно до початку перетворення вводити додатковий фільтр нижніх частот, що гарантує відповідне обмеження смуги пропускання аналогового фільтра.

Метод білінійного z -перетворення дозволяє синтезувати рекурсивний дискретний фільтр за частотною характеристикою аналогового прототипу. При його використанні ліва половина s -площини завжди відображується всередину одиничного кола на z -площині, тому синтез по стійкому аналоговому прототипу дає гарантовано стійкий дискретний фільтр.

При реалізації цього методу використовують просте конформне відображення s -площини в z -площину з використанням заміни

$$s \rightarrow \frac{2}{T} \cdot \frac{(1-z^{-1})}{(1+z^{-1})} = 2f_d \frac{(z-1)}{(z+1)}. \quad (6.61)$$

Можна показати, що частотні характеристики аналогового $K_a(j\omega)$ і дискретного $K_d(j\omega)$ фільтрів пов'язані одна з одною лише трансформацією частотної осі

$$K_d(j\omega) = K_a(j\omega) \cdot \left[\frac{2}{T} \operatorname{tg}(\omega T / 2) \right]. \quad (6.62)$$

На низьких частотах, коли $\omega T = 1$, тангенс приблизно дорівнює своєму аргументу і $\frac{2}{T} \operatorname{tg}(\omega T / 2) \approx \omega$, тому в області низьких частот частотні характеристики аналогового і дискретного фільтрів майже співпадають.

Різновидом метода білінійного перетворення можна вважати представлений нижче спосіб побудови різницевого рівняння для системи з заданим операторним коефіцієнтом передачі, або *передавальною функцією динамічної системи* (6.9). За цим способом для отримання дискретної передавальної функції формуючого фільтра передавальна функція по Лапласу представляється у вигляді відношення лінійних комбінацій інтегруючих ланок різного порядку. Для цього її чисельник і знаменник ділять на s^n , де n порядок передавальної функції. Після цього проводиться заміна операторів інтегрування $1/s^l$ ($n \geq l$) їх дискретними аналогами. Останні можуть мати різне представлення в залежності від способу інтегрування (табл. 6.2) [27]. Інтегратор нульового порядку є наслідком наближеного інтегрування методом прямокутників, першого порядку – методом трапецій [9].

Каскадне представлення операторної передавальної функції дозволяє обмежитись підстановками першого-другого порядку.

Таблиця 6.2 – Приклади z -перетворень операторів інтегрування

Оператор неперервного інтегрування	Оператор z -перетворення інтегрування з інтегратором	
	нульового порядку	першого порядку
$\frac{1}{s}$	$\frac{z^{-1}}{1-z^{-1}} T$	$\frac{1+z^{-1}}{1-z^{-1}} \frac{T}{2!}$
$\frac{1}{s^2}$	$\frac{z^{-1}(1+z^{-1})}{(1-z^{-1})^2} \frac{T^2}{2!}$	$\frac{1+4z^{-1}+z^{-2}}{(1-z^{-1})^2} \frac{T^2}{3!}$
$\frac{1}{s^3}$	$\frac{z^{-1}(1+4z^{-1}+z^{-2})}{(1-z^{-1})^3} \frac{T^3}{3!}$	$\frac{1+11z^{-1}+11z^{-2}+z^{-3}}{(1-z^{-1})^3} \frac{T^3}{4!}$

Приклад 6.4.

Знайти різницеві рівняння для цифрового моделювання реалізацій відліків випадкового сигналу зі спектральною густиною виду $W(\omega)$, які можуть бути сформовані лінійною системою згідно співвідношення (6.6)

$$W(\omega) = W_0 \dot{K}(\omega) \dot{K}^*(\omega) = W_0 |\dot{K}(\omega)|^2,$$

де W_0 – спектральна густина гаусівського білого шуму на вході лінійної системи, $\dot{K}(\omega)$ – комплексний коефіцієнт передачі лінійної системи, якому в результаті заміни $j\omega \rightarrow s$ відповідає передавальна функція за Лапласом

$$K(s) = K \frac{s + \beta_1}{s^2 + \alpha_2 s + \alpha_1}.$$

Припускаємо, що ця передавальна функція відповідає стійкій системі, тобто її полюси належать лівій на півплощині комплексної площини.

Розв'язання

Якщо на вхід формуючого фільтра подати дельтакорельовану гаусівську послідовність випадкових чисел з параметрами розподілу $(0, \sigma^2)$ і з дискретністю у часі T , тоді величину W_0 можна оцінити з співвідношення

$$\sigma^2 = \int_0^{\omega_0} W_0 d\omega = W_0 \omega_0,$$

де $\omega_0 = \pi/T$ – гранична частота спектра по Котельникову.

Якщо при моделюванні прийняти $\sigma^2 = 1$, тоді

$$W_0 = 1/\omega_0 = T/\pi.$$

Для отримання різницевого рівняння формування відліків випадкового сигналу спочатку від передавальної функції $K(s)$ перейдемо до z -перетворення, а потім до відповідного різницевого рівняння. Таким чином

$$K(s) = \frac{K}{\sqrt{W_0}} \frac{s + \beta_1}{s^2 + \alpha_2 s + \alpha_1} = \frac{K\sqrt{\pi}}{\sqrt{T}} \frac{\beta_1 \frac{1}{s^2} + \frac{1}{s}}{\alpha_1 \frac{1}{s^2} + \alpha_2 \frac{1}{s} + 1}.$$

Далі заміною інтегровальних інтегровальних ланок $\frac{1}{s^2}$ і $\frac{1}{s}$ їх z -перетворенням з інтегратором нульового порядку (див. табл. 6.2) отримаємо

$$\begin{aligned}
 K(z) &= \frac{K\sqrt{\pi}}{\sqrt{T}} \frac{\beta_1 \frac{z^{-1}(1+z^{-1})}{(1-z^{-1})^2} + \frac{z^{-1}}{1-z^{-1}} T}{\alpha_1 \frac{z^{-1}(1+z^{-1})}{(1-z^{-1})^2} \frac{T^2}{2} + \alpha_2 \frac{z^{-1}}{1-z^{-1}} T + 1} = \\
 &= \frac{K\sqrt{\pi}}{\sqrt{T}} \frac{(\beta_1 + T)z^{-1} + (\beta_1 - T)z^{-2}}{1 - (2 - 0,5\alpha_1 T^2 - \alpha_2 T)z^{-1} - (\alpha_2 T - 0,5\alpha_1 T^2 - 1)z^{-2}} = \\
 &= \frac{\sum_{k=1}^2 b_k z^{-k}}{1 - \sum_{k=1}^2 a_k z^{-k}},
 \end{aligned}$$

де $a_1 = (2 - 0,5\alpha_1 T^2 - \alpha_2 T)$, $a_2 = (\alpha_2 T - 0,5\alpha_1 T^2 - 1)$,

$$b_1 = \frac{K\sqrt{\pi}}{\sqrt{T}}(\beta_1 + T), \quad b_2 = \frac{K\sqrt{\pi}}{\sqrt{T}}(\beta_1 - T).$$

Звідки з урахуванням (6.60) отримаємо рівняння

$$Y(z) = b_1 X(z)z^{-1} + b_2 X(z)z^{-2} + a_1 Y(z)z^{-1} + a_2 Y(z)z^{-2},$$

з якого матимемо відповідне різницеве рівняння

$$y_n = b_1 x_{n-1} + b_2 x_{n-2} + a_1 y_{n-1} + a_2 y_{n-2},$$

де x_n, y_n – n -ні відліки послідовностей випадкових сигналів на вході та виході лінійного формуючого фільтра. При обчисленні значення y_0 попередні відліки y_{-1} і y_{-2} можуть бути прийнятими рівними нулю.

В [17] також розглядається можливість переходу до різницевого рівняння, що описує лінійну систему, безпосередньо від диференційного рівняння її аналогового прототипу. В цьому випадку похідні замінюються їх кінцево-різницевиими наближеннями. Після перетворень отриманого рівняння можна отримати співвідношення для описання рекурсивного фільтра. При цьому необхідно обов'язково перевіряти отримане рівняння на його стійкість.

Контрольні запитання

1. За якими ознаками класифікують системи? Які системи вважаються лінійними і стаціонарними? Які нелінійними?

2. Які принципи справедливі для лінійних систем з постійними параметрами?

3. Як визначається імпульсна характеристика лінійної системи? Запишіть співвідношення для обчислення сигналу на виході лінійної системи в одновимірному і багатомірному випадках з її використанням. Умова стійкості лінійної системи.

4. Реакцією на який вхідний сигнал є перехідна характеристика лінійної стаціонарної системи? Наведіть різновиди інтеграла Дюамеля для обчислення сигналу на виході системи з її застосуванням.

5. Назвіть умову фізичної реалізованості лінійної системи.

6. Що називають перехідною характеристикою?

7. Яка функція називається комплексним коефіцієнтом передачі системи?

8. Запишіть співвідношення, яким пов'язані імпульсна характеристика та комплексний коефіцієнт передачі системи.

9. Яким чином визначаються амплітудно-частотна і фазочастотна характеристики лінійної системи? Їх фізична сутність.

10. Поясніть відмінності фазової і групової затримки лінійної системи.

11. Що таке взаємний спектр вхідного і вихідного сигналів лінійної системи? Його зв'язок з комплексним коефіцієнтом передачі лінійної системи.

12. Яким співвідношенням пов'язані між собою автокореляційна функція вхідного сигналу і взаємнокореляційна функція вхідного і вихідного сигналів?

13. Операторний і комплексний коефіцієнти передачі динамічної системи та її зв'язок. Нулі і полюси.

14. Яка лінійна система називається стійкою? Як перевірити її стійкість із застосуванням критерію Рауса – Гурвіца?

15. Як обчислити імпульсну характеристику лінійної системи через нулі і лишки передавальної функції?

16. Як може бути обчислена спектральна густина потужності випадкового сигналу на виході лінійної?

17. Яка необхідна і достатня умова стійкості лінійної дискретної системи з постійними параметрами?

18. Чи може бути застосований критерій стійкості Рауса – Гурвица для дослідження на стійкість стаціонарної цифрової системи?

19. Як формується таблиця Джурі для дослідження стійкості стаціонарної цифрової системи?

20. Коли доцільно застосовувати алгебраїчні критерії стійкості для стаціонарної цифрової системи?

21. Що розуміють під проектуванням цифрової лінійної системи або цифрового фільтра?

22. Назвіть етапи проектування цифрового фільтра.

23. Як класифікуються методи синтезу цифрових фільтрів?

24. Назвіть основні недоліки фільтрів з кінцевою імпульсною характеристикою.

25. Перелічіть основні методи розрахунку фільтрів з лінійними фазовими характеристиками.

26. Які властивості повинно мати вікно при проектуванні?

27. Від чого залежить вибір необхідного вікна?

28. На які класи поділяються стандартні методи проектування НІХ-фільтрів?

29. Назвіть методи інваріативного перетворення.

30. Що можливо синтезувати за допомогою методу білінійного z -перетворення?

Глава 7

МОДУЛЬОВАНІ СИГНАЛИ І НЕЛІНІЙНІ СИСТЕМИ

7.1. Нелінійні системи. Безінерційні нелінійні перетворювачі

В попередніх главах розглядалися системи і кола, які, належали до класу лінійних систем з постійними параметрами. Перетворення гармонічних коливань в таких системах веде лише до зміни їх амплітуд і початкових фаз, тобто лінійна система з постійними параметрами не додає до спектру сигналу на виході нових частотних складових порівняно зі спектром сигналу на вході. Це є певним обмеженням щодо можливих перетворень сигналів в таких системах.

Набагато більші можливості в цьому відношенні мають нелінійні системи. В таких системах зв'язок між вхідним сигналом $x(t)$ і реакцією $y(t)$ визначається нелінійною функціональною залежністю $y(t) = f[x(t)]$. Якщо в цій залежності має місце лише безпосередній зв'язок сигналу $y(t)$ від $x(t)$ і відсутня залежність від часу, тоді нелінійна система є безінерційною. Така ідеалізація можлива за умови, якщо тривалість зміни сигналу на вході нелінійної системи значно перевищує тривалість усталення процесу в системі.

Розглянемо властивості нелінійних безінерційних систем та деякі види перетворень сигналів у таких системах [9, 17].

7.1.1. Властивості нелінійних систем

На відміну від лінійних систем, які описуються диференційними рівняннями з постійними коефіцієнтами виду (див. гл. 6)

$$a_0 \frac{d^n y(t)}{dt^n} + a_1 \frac{d^{n-1} y(t)}{dt^{n-1}} + \dots + a_{n-1} \frac{dy(t)}{dt} + a_n y(t) =$$

$$b_0 \frac{d^m x(t)}{dt^m} + b_1 \frac{d^{m-1} x(t)}{dt^{m-1}} + \dots + b_{m-1} \frac{dx(t)}{dt} + b_m x(t),$$

Для нелінійних систем потрібно застосовувати нелінійні диференційні рівняння. В таких рівняннях хоча б один з коефіцієнтів a_i є функцією змінної y або її похідних [9], тобто

$$a_i = f\left(y, \frac{dy}{dt}, \dots\right),$$

або хоча б один з коефіцієнтів b_j є функцією змінної x або її похідних, тобто

$$b_j = f\left(x, \frac{dx}{dt}, \dots\right).$$

У разі описання деякого аналогового електричного кола коефіцієнти a_i і b_j визначаються параметрами елементів схеми цього кола (опору R , індуктивності L , взаємної індуктивності M , ємності C і т. п.). Для нелінійних систем хоча б один з цих параметрів буде залежати від прикладеної до елемента напруги або струму, що проходить через нього.

Нелінійні системи мають такі основні властивості:

1. В нелінійних системах *не виконується принцип суперпозиції*, тобто відгук такої системи на вхідну дію $x_1(t) + x_2(t)$ не дорівнює $y_1(t) + y_2(t)$, де $y_1(t)$ є відгуком на вхідну дію $x_1(t)$, а $y_2(t)$ – на $x_2(t)$.

Виходячи з цього, для аналізу процесів, що відбуваються у таких системах, в загальному випадку не можна користатись методами, які основані на властивості суперпозиції сигналів.

2. У нелінійних системах відбувається перетворення спектрів сигналів, тобто у відгуку такої системи з'являються спектральні складові з частотами, яких не було в спектрі сигналу на вході.

На рис. 7.1 представлено нелінійне коло, яке включає резистивний нелінійний елемент (НЕ) R_1 та резистор R_2 . На його вході діє сигнал $u_1(t)$, сигнал на виході кола – $u_2(t)$.

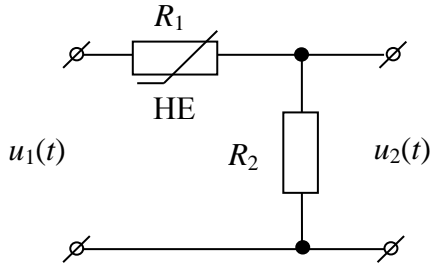


Рисунок 7.1 – Коло з резистивним нелінійним елементом

Як вже згадувалось, НЕ можуть мати не тільки активний опір (резистивні), але й бути реактивними (ємнісними або індуктивними). Основні їх статичні і динамічні характеристики пов'язані з типом НЕ (табл. 7.1) [28, 29]:

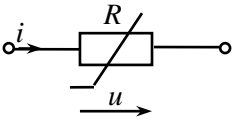
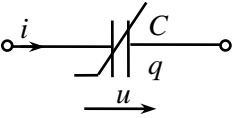
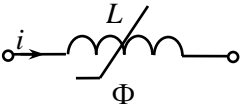
для резистивного НЕ це вольт-амперна характеристика (ВАХ) – залежність струму i від прикладеної до нього напруги u ;

для ємнісного НЕ вольт-кулонна характеристика – залежність заряду q ємності від прикладеної до неї напруги u ;

для індуктивного НЕ ампер-веберна характеристика – залежність магнітного потоку (потокосцеплення) Φ котушки індуктивності від струму i в ній.

Разом з динамічними параметрами представленими в таблиці 7.1 для аналізу впливу відносно слабких сигналів на нелінійні елементи можуть застосовуватись й обернені до них, наприклад, диференційний опір R_d .

Таблиця 7.1 – Нелінійні елементи та їх параметри

Схемне позначення НЕ	Параметри статичні	Параметри динамічні
	<p>Статична Провідність</p> $G = \frac{i}{u} = \frac{1}{R}$	<p>Диференційна провідність (крутість)</p> $G_d = S = \frac{di}{du} = \frac{1}{R_d}$
	<p>Статична ємність</p> $C = \frac{q}{u}$	<p>Диференційна ємність</p> $C_d = \frac{dq}{du}$
	<p>Статична індуктивність</p> $L = \frac{\Phi}{i}$	<p>Диференційна індуктивність</p> $L_d = \frac{d\Phi}{di}$

Для нелінійного резистивного елемента, крім цього, застосовується також середній опір (опір за першою гармонікою), який є відношенням амплітуди першої гармоніки входної дії (напруги, прикладеної до НЕ) до амплітуди першої гармоніки відгуку (струму, що проходить через нелінійний резистор), тобто

$$R_{cp} = \frac{U_{m1}}{I_{m1}}.$$

Відповідно, середня провідність нелінійного резистивного елемента (провідність за першою гармонікою) є оберненою до цього параметру

$$S_{cp} = \frac{I_{m1}}{U_{m1}}.$$

У випадку, коли U_{m1} (а отже, і I_{m1}) – величина мала, крутість ВАХ в робочій точці $S \approx S_{cp}$. Таким чином, щоб характеризувати нелінійний резистивний елемент, у загальному випадку необхідно мати такі його параметри:

опір (провідність) на постійному струмі в робочій точці R (або провідність $G = 1/R$);

диференційний опір R_d (або крутість S характеристики нелінійного елемента) у робочій точці;

середній опір R_{cp} або середню крутість S_{cp} нелінійного резистивного елемента.

Параметри, аналогічні до зазначених вище, характеризують і реактивні нелінійні елементи.

Для нелінійного конденсатора (наприклад, варикапа) застосовують також залежність диференційної ємності від положення робочої точки, яка має назву *вольт-фарадна характеристика*. Нелінійність елемента виявляється лише в залежності параметрів R_d , C_d і L_d від керувальної напруги (або струму), що визначає положення робочої точки на відповідній характеристиці НЕ. За дії слабкого сигналу такі елементи є лінійними пристроями зі змінними параметрами.

На рис. 7.2 зображена вольт-амперна характеристика резистивного безінерційного НЕ, на який діє гармонічний сигнал $u_1(t) = U_{1m} \cos(\omega_1 t + \varphi_1)$, причому показані три варіанти впливної дії [30]. Перші два (ліворуч) – відповідають випадку дії сигналу з малою амплітудою. В обох цих варіантах впливу на виході НЕ спостерігається сигнал близький за формою до гармонічного $u_2(t) = U_{2m} \cos(\omega_1 t + \varphi_1)$. Частота і початкова фаза сигналу на виході НЕ залишаються незмінними. Відмінність у перетворенні сигналів пов'язана з різним значенням крутості ВАХ у відповідних робочих точках. В робочій точці A з координатами (U_0, I_0) , яка показана на рисунку, крутість S_A , яка чисельно дорівнює тангенсу кута нахилу дотичної до кривої ВАХ в точці A , є

значно більшою, ніж у першому випадку. Тому в першому випадку (нижня крива для графіків напруг $u_2(t)$) спостерігається незначне послаблення сигналу $u_1(t)$, а в другому (в робочій точці A) – сигнал $u_1(t)$ підсилюється.

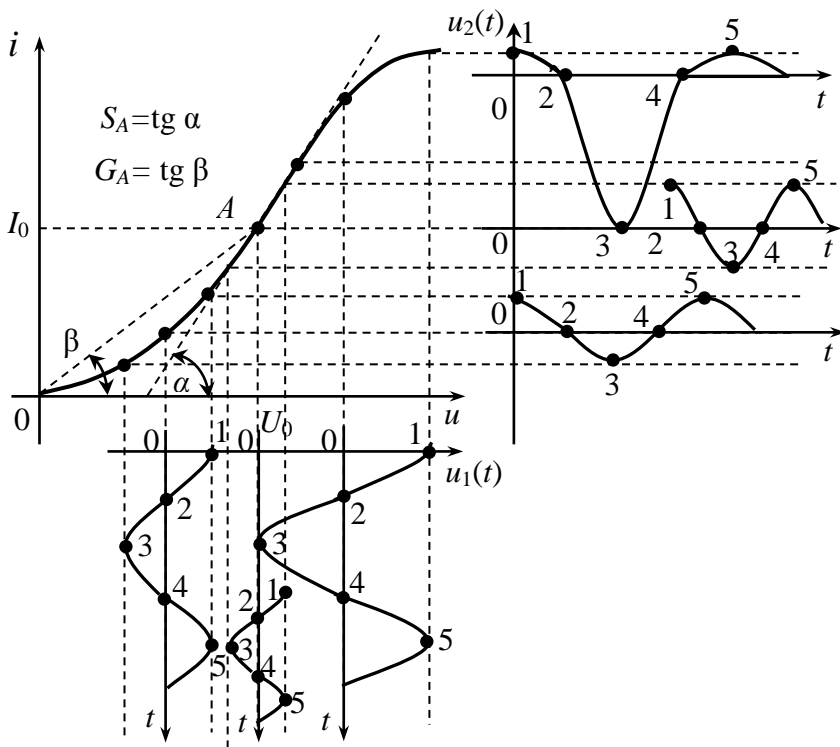


Рисунок 7.2 – Перетворення гармонічного сигналу в нелінійному елементі

Нелінійний елемент в цих випадках працює в так званому квазілінійному режимі (як би лінійному). Це означає, що діапазон зміни напруги сигналу $u_1(t)$ відповідає такій ділянці ВАХ нелінійного елементу, яка близька до лінійної. Тому відгук $u_2(t)$ має в амплітудному спектрі такі ж частоти, як і сигнал $u_1(t)$, а їх фазові спектри збігаються.

Третій варіант дії сигналу $u_1(t)$ (праворуч на рис. 7.2) відповідає випадку, коли його амплітуда значно перевищує границі лінійної ділянки ВАХ. Такий режим роботи НЕ має назву *режиму великих амплітуд*. Як видно з рис. 7.2, нижня напівхвиля гармонічного коливання $u_1(t) = U_{1m} \cos \omega_1 t$, що подається на вхід кола (рис. 7.1), залишається на лінійній ділянці, тоді як верхня – діє на суттєво нелінійній ділянці. Тому сигнал $u_2(t) = iR$ не є гармонічним, хоча він залишається періодичним і матиме лінійчатий спектр.

Для обчислення спектру сигналу на виході нелінійного кола застосовують різні *способи апроксимації ВАХ* нелінійного елемента. Найбільш поширеними є кусково-лінійна, поліноміальна та з використанням трансцендентних функцій (показникових, тригонометричних та ін.) [9, 17].

Кусково-лінійна апроксимація передбачає заміну окремих ділянок ВАХ нелінійного елемента відрізками прямих ліній. Найпростіша з них має вигляд

$$i(u) = \begin{cases} 0, & u \leq U_{\text{відс}}; \\ S(u - U_{\text{відс}}), & u > U_{\text{відс}}, \end{cases}$$

де S – крутість ВАХ нелінійного елемента;

$U_{\text{відс}}$ – напруга відсічки.

Вона може застосовуватись для апроксимації вхідної характеристики біполярного транзистора, тобто залежності струму бази від напруги між базою і емітером, або ВАХ напівпровідникового діоду (рис. 7.3) [17].

На рис. 7.3 зображено ВАХ НЕ при кусково-лінійній апроксимації та часові діаграми напруги на вході і виході нелінійної системи (рис. 7.1). Сигнал $u_1(t)$ описується виразом

$$u_1(t) = U_3 + U_{1m} \cos \omega_1 t,$$

де U_3 – початок відліку амплітудного значення гармонічного впливу (напруга зсуву);

θ – кут відсічки (половина інтервалу, протягом якого має місце сигнал відгуку, $\omega_1 t_2 = \theta$ – відповідає точці „2” на

графіку $u_1(\omega t)$). Порядок розрахунку цього кута представлено в [17, 30].

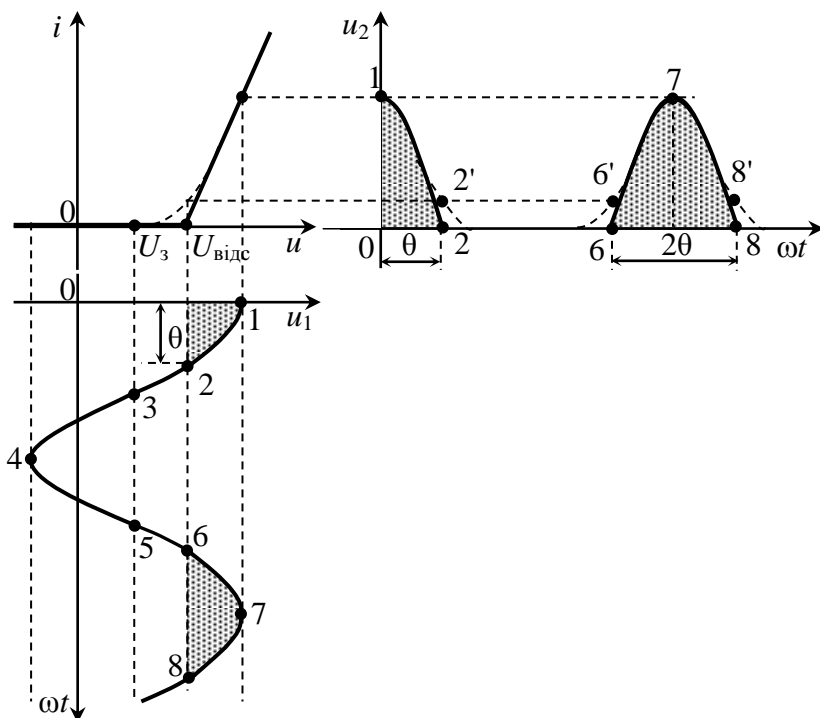


Рисунок 7.3 – Визначення відгуку нелінійного кола при кусково-лінійній апроксимації ВАХ НЕ

На графіку ВАХ пунктирною лінією показано ділянку кривої, що має помітну відмінність від лінійної залежності. Тому при виборі кусково-лінійної або іншого виду апроксимації потрібно оцінювати ступінь можливих спотворень та їх припустимість. В даному випадку сигнал на виході нелінійної системи є періодичним, тому для оцінки ступеня спотворень сигналу за рахунок неточності апроксимації може бути застосоване порівняння його розкладань в ряд Фур'є для різних законів апроксимації.

Наприклад, може бути задано відсоток припустимого відхилення в енергії гармонік за рахунок апроксимації ВАХ.

Вибір найбільш придатної апроксимації ВАХ може бути виконаний мінімізацією її відхилення від експериментальної кривої за критерієм середньоквадратичного наближення, найбільшого відхилення або за ймовірнісним критерієм, які розглядались в главі 1.

Очевидно, що представлена кусково-лінійна апроксимація краще відповідатиме випадку, коли амплітуда сигналу на вході буде великою.

Оскільки $\omega_1 t_2 = \theta$, тоді, як видно для точки „2” на графіку $u_1(\omega t)$ (рис. 7.3), можна записати:

$$U_3 + U_{1m} \cos \theta = U_{\text{відс}},$$

звідки

$$\cos \theta = \frac{U_{\text{відс}} - U_3}{U_{1m}}$$

і

$$\theta = \arccos \frac{U_{\text{відс}} - U_3}{U_{1m}}.$$

Тому при кусково-лінійній апроксимації ВАХ НЕ (рис. 7.3) коефіцієнти розкладання в ряд Фур'є ряду періодичної послідовності імпульсів відгуку $u_2(t)$ мають вид

$$U_{mk} = R_2 I_{mk} = R_2 S U_{1m} \gamma_k(\theta),$$

де коефіцієнти $\gamma_k(\theta)$ є функціями Берга [17]:

$$\gamma_0(\theta) = \frac{1}{\pi}(\sin \theta - \theta \cos \theta); \quad \gamma_1(\theta) = \frac{1}{\pi}(\theta - \theta \sin \theta \cos \theta); \quad \dots;$$

$$\gamma_k(\theta) = \frac{2}{\pi} \frac{\sin k\theta \cos \theta - k \cos k\theta \sin \theta}{k(k^2 - 1)},$$

де $k = 2, 3, \dots$.

Графіки залежностей перших чотирьох функцій Берга представлені на рис. 7.4.

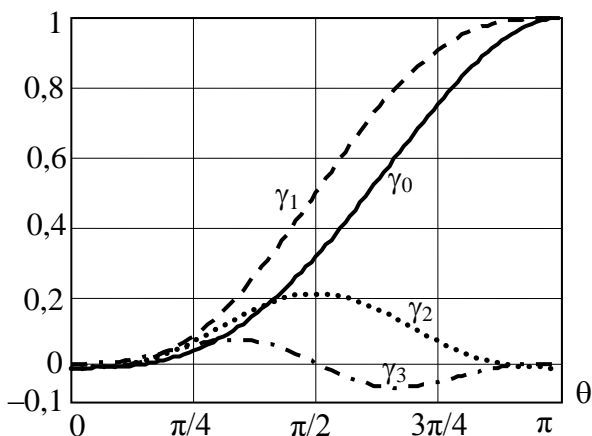


Рисунок 7.4 – Графіки чотирьох перших функцій Берга

Апроксимація ВАХ з використанням *трансцендентних функцій*, а саме показникових, часто застосовується для представлення нелінійних характеристик напівпровідникових пристроїв. Наприклад [17], характеристика напівпровідникового діода може бути представлена у вигляді

$$i(u) = I_3 \left[\exp(u/u_T) - 1 \right], \text{ якщо } u > 0,$$

де I_3 – зворотний струм насичення, u_T – температурний потенціал, який дорівнює 25 мВ для кремнієвих приладів при стандартній температурі 300 К.

Така апроксимація є достатньо вірною при струмах до кількох міліампер. При великих значеннях струму експоненціальна залежність переходить в лінійну, що пов'язано із впливом об'ємного опору напівпровідникового матеріалу.

У разі впливу на НЕ з експоненціальною характеристикою напруги $u_1(t) = U_3 + U_{1m} \cos \omega_1 t$ для знаходження спектрального сигналу на його виході може бути застосоване

відоме з математики співвідношення [31]

$$\exp(m \cdot \cos x) = I_0 + 2 \sum_{n=1}^{\infty} I_n(m) \cos nx,$$

де $I_k(m)$ – модифікована функція Бесселя k -го порядку.

Із застосуванням цієї формули можна отримати вираз для обчислення струму в нелінійному колі

$$i(u) = I_3 \left[\exp(U_3/u_T) I_0(U_{1m}/u_T) - 1 \right] + 2I_3 \exp(U_3/u_T) \sum_{n=1}^{\infty} I_n(U_{1m}/u_T) \cos n\omega_1 t.$$

Іншим поширеним способом апроксимації ВАХ нелінійних елементів є *поліноміальна апроксимація*.

Така апроксимація є зручною з точки зору порівняно простого обчислення спектрального складу сигналу на виході нелінійної системи при гармонічному вхідному сигналі.

Наприклад, ВАХ НЕ поблизу робочої точки U_3 апроксимована поліномом

$$i = a_0 + a_1(u - U_3) + a_2(u - U_3)^2 + \dots + a_n(u - U_3)^n + \dots, \quad (7.1)$$

тоді відгук на сигнал $u_1(t) = U_3 + U_{1m} \cos \omega_1 t$, що визначається як $u_2(t) = i R_2$, можна записати як

$$u_2(t) = R_2 \left[a_0 + a_1(U_{1m} \cos \omega_1 t) + \dots + a_n(U_{1m} \cos \omega_1 t)^n + \dots \right].$$

Якщо покласти $R = 1$ Ом та скористатись тригонометричними формулами [31]:

$$\cos^2 x = \frac{1}{2} + \frac{1}{2} \cos 2x;$$

$$\cos^3 x = \frac{3}{4} \cos x + \frac{1}{4} \cos 3x;$$

$$\cos^4 x = \frac{3}{8} + \frac{4}{8} \cos 2x + \frac{1}{8} \cos 4x;$$

$$\cos^5 x = \frac{10}{16} \cos x + \frac{5}{16} \cos 3x + \frac{1}{16} \cos 5x$$

і т. д., в результаті перетворень одержимо:

$$u_2(t) = U_{20} + U_{2m1} \cos \omega_1 t + \\ + U_{2m2} \cos 2\omega_1 t + \dots + U_{2mn} \cos n\omega_1 t + \dots,$$

що відповідає ряду Фур'є

$$u_2(t) = U_{20} + \sum_{n=1}^{\infty} U_{2mn} \cos(n\omega_1 t + \varphi_k), \quad (7.2)$$

де, як показано в [31], довільний n -й коефіцієнт можна обчислити за формулою

$$U_{2mn} = \sum_{k=0}^{\infty} \frac{(2k+n)!}{2^{2k+n-1} k!(k+n)!} a_{2k+n} U_{1m}^{2k+n}, \quad (7.3)$$

Таким чином, відгук нелінійного кола на гармонічний сигнал в режимі великих амплітуд є періодичним сигналом. Якщо цей сигнал подати на вхід аналізатора спектра, то на екрані будемо спостерігати амплітудний спектр цього сигналу, аналогічний наведеним у главі 2.

Розглянемо *перетворення бігармонічних коливань в нелінійній системі*. Вхідний сигнал $u_1(t)$ при цьому містить два гармонічних коливання з різними частотами ω_1 й Ω ($\omega_1 > \Omega$) [17]:

$$u_1(t) = U_3 + U_{m1} \cos \omega_1 t + U_{m2} \cos \Omega t. \quad (7.4)$$

Спектр цього сигналу представлений на рис. 7.5.

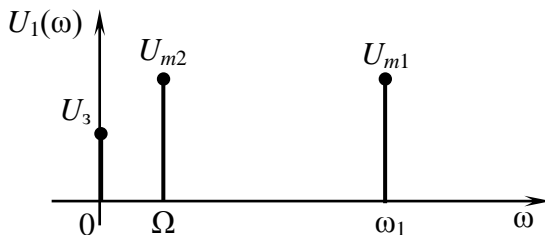


Рисунок 7.5 – Спектр бігармонічного сигналу

Бігармонічний сигнал подається на вхід нелінійного кола (рис. 7.1), ВАХ нелінійного елемента в якому описується квадратичною функцією

$$i(u) = a_0 + a_1(u - U_3) + a_2(u - U_3)^2.$$

В цьому випадку струм нелінійного кола визначиться як:

$$i = a_0 + a_1 U_{m1} \cos \omega_1 t + a_1 U_{m1} \cos \Omega t + a_2 (U_{m1} \cos \omega_1 t)^2 + 2a_2 U_{m1} U_{m2} \cos \omega_1 t \cdot \cos \Omega t + a_2 (U_{m2} \cos \Omega t)^2.$$

Виконавши перетворення за тригонометричними формулами для добутку косинусів, одержимо:

$$i = a_0 + \frac{1}{2} a_2 (U_{m1}^2 + U_{m2}^2) + a_1 U_{m1} \cos \omega_1 t + a_1 U_{m1} \cos \Omega t + \frac{1}{2} a_2 U_{m1}^2 \cos 2\omega_1 t + \frac{1}{2} a_2 U_{m2}^2 \cos 2\Omega t + a_2 U_{m1} U_{m2} \cos(\omega_1 - \Omega)t + a_2 U_{m1} U_{m2} \cos(\omega_1 + \Omega)t.$$

Тоді напругу на виході нелінійного кола можна подати у вигляді

$$u_2(t) = i R_2 = U_{20} + U_{m\Omega} \cos \Omega t + U_{m2\Omega} \cos 2\Omega t + U_{m3} \cos(\omega_1 - \Omega)t + U_{m\omega} \cos \omega_1 t + U_{m5} \cos(\omega_1 + \Omega)t + U_{m2\omega} \cos 2\omega_1 t. \quad (7.5)$$

Спектр відгуку (7.5) представлено на рис. 7.6. Порівняння спектрів вхідного впливу (рис. 7.5) і відгуку рис. 7.6 вказує на суттєву трансформацію спектру. В спектрі з'явилися нові складові з так званими комбінаційними частотами, тобто частоти яких є сумою, різницею та подвійними значеннями частот бігармонічного сигналу.

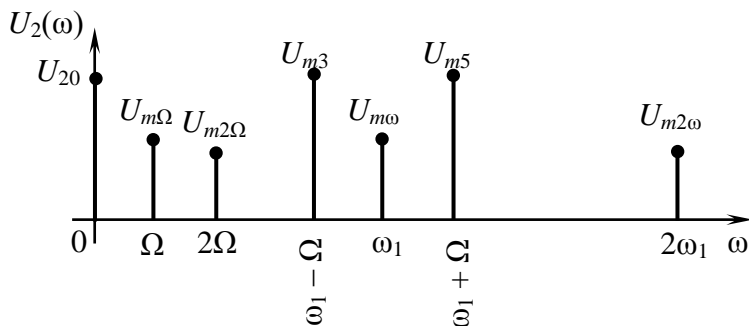


Рисунок 7.6 – Амплітудний спектр відгуку нелінійного кола при бігармонічному впливі

Якщо $\omega_1 \gg \Omega$, і низькочастотний сигнал має деяку смугу частот $\Delta\Omega$, тоді *комбінаційні частоти* відгуку будуть містити комбінації $(\omega_1 + \Delta\Omega)$ і $(\omega_1 - \Delta\Omega)$. Тобто спостерігається перенесення смуги частот з однієї частоти-носія на інші більш високі частоти.

При збільшенні кількості гармонічних складових у вхідному сигналі до M кількість комбінаційних частот у сигналі на виході НЕ може суттєво зрости. Комбінаційні частоти визначаються співвідношенням [17]

$$\omega_{\text{ком}} = |n_1\omega_1 + n_2\omega_2 + \dots + n_M\omega_M|, \quad (7.6)$$

де n_1, n_2, \dots, n_M – довільні цілі числа, як додатні, так і від’ємні, а також нуль.

Комбінаційні частоти групують шляхом об’єднання всіх частот, для яких

$$|n_1| + |n_2| + \dots + |n_k| + \dots = N, \quad (7.7)$$

де N – *порядок комбінаційної частоти*.

Піднесення у степінь N гармонічного сигналу з частотою ω_1 , як було показано при отриманні співвідношення (7.2), дає складову з максимальною частотою $N\omega_1$. Спектр сигналу на виході нелінійної системи може включати гармоніки з

частотами згідно (7.6) при $N=1,2,\dots,N_{\max}$, де N_{\max} – максимальне значення порядку полінома, яким апроксимується ВАХ нелінійного елемента. Наприклад, при збудженні нелінійної системи, яка апроксимується поліномом з найбільшим степенем $N=3$, двома гармоніками з частотами ω_1 і ω_2 будуть можливі наступні комбінаційні частоти [17]:

$$N=1 - \omega_1, \omega_2;$$

$$N=2 - 2\omega_1, 2\omega_2, \omega_1 + \omega_2, \omega_1 - \omega_2;$$

$$N=3 - 3\omega_1, 3\omega_2, 2\omega_1 + \omega_2, 2\omega_1 - \omega_2, 2\omega_2 + \omega_1, 2\omega_2 - \omega_1.$$

Крім цього, якщо N – парне число, тоді в спектрі спостерігаються комбінаційні частоти парних порядків: $N, N-2, N-4, \dots, 0$. Відповідно для непарного N складові в спектрі матимуть комбінаційні частоти непарних порядків: $N, N-2, N-4, \dots, 1$.

Ця властивість нелінійних елементів застосовується для перенесення модуляції сигналу з однієї частоти на іншу. Нелінійні елементи застосовуються також для ділення і множення частоти сигналів, демодуляції сигналів, обмеження їх амплітуд та ін. Відомий метод вимірювання комбінаційних спотворень немодульованих сигналів застосовується, наприклад, для вимірювання нелінійних спотворень дуже малого рівня в підсилювачах та інших майже лінійних пристроях [32]. Для цього на вхід пристрою, який досліджується, подається бігармонічний сигнал (можлива й сума трьох гармонічних напруг), і на виході виконуються вимірювання комбінаційних частот. Перевагою цього методу порівняно з методами вимірювання гармонічних спотворень є його більша чутливість, завдяки більшій потужності комбінаційних складових.

Якщо ВАХ нелінійного елемента має ділянку з від'ємною крутістю, тоді цей елемент має від'ємний опір (на рис. 7.7 ділянка ab). Як видно з рисунку, в точці A : динамічна провідність $S_A = \text{tg } \beta < 0$, в той час як статична – $G_A = \text{tg } \alpha > 0$.

Двополюсники з характеристиками подібними до зображеної на рис. 7.7, наприклад тунельні діоди, знаходять застосування у генераторах коливань різної частоти і форми [9].

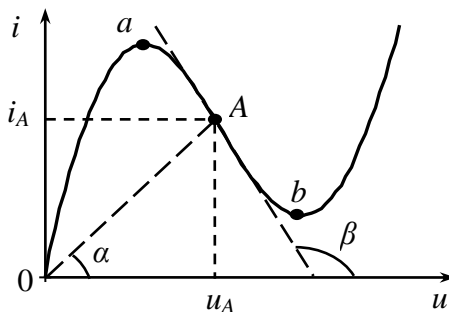


Рисунок 7.7 – Ілюстрація ВАХ з ділянкою, яка має від’ємний динамічний опір

Загалом, нелінійні опори надають електричним колам властивості, принципово недосяжні в лінійних колах, наприклад, за їх допомогою здійснюють стабілізацію струму, стабілізацію напруги, підсилення постійної напруги тощо.

7.1.2. Вплив стаціонарних випадкових сигналів на безінерційні нелінійні системи

Безінерційна нелінійна система представляє собою сполучення резистивних (активних) елементів, один з яких є нелінійним із заданою ВАХ. Інерційними є такі кола, які мають у своєму складі реактивні елементи (індуктивності або ємності), нелінійні кола до того ж мають у своєму складі також активний або реактивний НЕ. Реальну систему з нелінійними елементами у загальному випадку можна представити як сполучення нелінійних безінерційних елементів з лінійними інерційними колами, що значно ускладнює визначення ймовірнісних і статистичних

характеристик сигналу і шуму на її виході порівняно з безінерційними колами. Нелінійні кола можуть бути розізнаними або із зворотним зв'язком.

При дослідженні ймовірнісних характеристик випадкових сигналів на виході нелінійної системи можуть бути сформульовані наступні задачі: знаходження закону розподілу миттєвих значень або обвідної сигналу; обчислення окремих моментів розподілів (наприклад, математичного сподівання і дисперсії); знаходження моментних функцій, перш за все кореляційної функції (або спектра середньої потужності) сигналу на виході нелінійної системи.

Аналітичне вирішення перелічених вище задач можливе лише в окремих порівняно простих випадках, які, однак, мають практичне значення. Далі обмежимося розглядом проходження випадкових сигналів лише через безінерційні нелінійні кола.

Як і для лінійних кіл (див. главу 6), знаходження густини розподілу сигналу на виході нелінійної системи у загальному випадку не вдається виконати простими засобами.

Нехай на вхід безінерційного НЕ з вольт-амперною характеристикою $y = f(x)$ діє випадковий сигнал $x(t)$ з відомою густиною ймовірності $p_x(x, t)$. Потрібно знайти густину ймовірності $p_y(y, t)$ значень сигналу $y(t)$ на його виході.

Для вирішення цієї задачі можуть бути використані методи «безпосереднього» розрахунку або метод характеристичних функцій.

Почнемо з розгляду методу «безпосереднього» розрахунку. Спочатку вважатимемо функцію $y = f(x)$. В цьому випадку існує однозначна обернена функція $x(y) = f^{-1}(y)$. Завдяки такому зв'язку, ймовірність потрапляння випадкової величини ξ в інтервал $(x, x + dx)$ –

$P(x < \xi < x + dx)$ буде дорівнювати ймовірності потрапляння випадкової величини на виході НЕ η у відповідний інтервал $(y, y + dy)$

$$P(x < \xi < x + dx) = P(y < \eta < y + dy),$$

або ж

$$p_x(x, t) dx = p_y(y, t) dy. \quad (7.8)$$

З рівняння (7.8) виходить

$$p_y(y, t) = p_x(x, t) \left| \frac{dx}{dy} \right|.$$

Наявність модулю похідної легко пояснюється, оскільки густина ймовірності не може бути від'ємною. Підстановка замість x оберненої функції $x(y)$ дає

$$p_y(y, t) = p_x[x(y), t] \left| \frac{d}{dy} x(y) \right|. \quad (7.9)$$

Якщо функція $x(y)$ буде багатозначною останнє співвідношення потрібно застосовувати для кожної монотонної гілки, підсумовуючі отримані результати. Наприклад, у випадку двозначної функції $x(y)$ (див. рис. 7.7) отримаємо

$$p_y(y, t) = p_x[x_1(y), t] \left| \frac{d}{dy} x_1(y) \right| + p_x[x_2(y), t] \left| \frac{d}{dy} x_2(y) \right|. \quad (7.10)$$

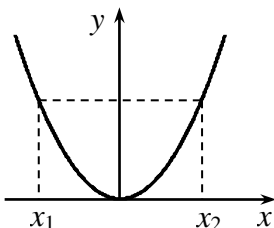


Рисунок 7.7 – Двозначна ВАХ нелінійного елемента

В загальному випадку безінерційного перетворення співвідношення для знаходження n -мірної густини ймовірності матиме вигляд

$$p_y(y_1, y_2, \dots, y_n) = p_x[x_1, x_2, \dots, x_n]J,$$

$$\text{де } x_1 = x(y_1) = f^{-1}(y_1), x_2 = f^{-1}(y_2), \dots, x_n = f^{-1}(y_n),$$

J – якобіан переходу, який з урахуванням незалежності у цьому випадку поточного значення сигналу на виході НЕ від попередніх можна представити у наступному вигляді

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \dots & \frac{\partial x_1}{\partial y_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \dots & \frac{\partial x_n}{\partial y_n} \end{vmatrix} = \begin{vmatrix} \frac{dx_1}{dy_1} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{dx_n}{dy_n} \end{vmatrix} = \prod_{i=1}^n \frac{dx_i}{dy_i} = \left(\frac{dx}{dy}\right)^n$$

Для спрощення в подальшому при записі одновимірної густини випадкового процесу змінну t показувати не будемо.

Розглянемо *приклад* застосування співвідношення (7.10) при обчисленні густини ймовірності випадкового сигналу на виході нелінійного елемента з ВАХ на рис. 7.7. В цьому випадку $y = x^2$. Нехай випадковий сигнал, що діє на вході НЕ, має нормальний закон розподілу

$$p_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{x^2}{2\sigma_x^2}}.$$

Обернена функція $x(y)$ у випадку $y > 0$ є двохзначною $x = \pm\sqrt{y}$, тому потрібно скористатись співвідношенням (7.10).

Похідну оберненої функції

$$\left| \frac{d}{dy} x(y) \right| = \frac{1}{2} \frac{1}{\sqrt{y}}$$

підставимо у вираз (7.10) і отримаємо

$$p_y(y) = p_x(\sqrt{y}) \frac{1}{2\sqrt{y}} + p_x(-\sqrt{y}) \frac{1}{2\sqrt{y}}, y > 0,$$

$$p_y(y) = 0, y < 0.$$

Підстановка виразу для $p_x(x)$ у перше співвідношення дасть

$$p_y(y) = \frac{1}{\sqrt{2\pi}\sigma_x} \frac{1}{2\sqrt{y}} \left[e^{-\frac{(\sqrt{y})^2}{2\sigma_x^2}} + e^{-\frac{(-\sqrt{y})^2}{2\sigma_x^2}} \right] =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_x\sqrt{y}} e^{-\frac{y}{2\sigma_x^2}}, y > 0.$$

Тому остаточно отримаємо

$$p_y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_x\sqrt{y}} e^{-\frac{y}{2\sigma_x^2}}, & y > 0; \\ 0, & y < 0. \end{cases}$$

В деяких випадках ВАХ безінерційних НЕ достатньо точно апроксимуються кусково-неперервними кривими (рис. 7.8).

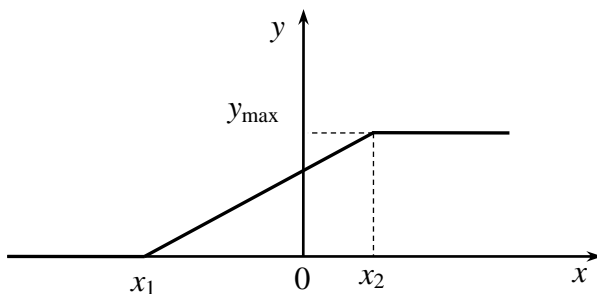


Рисунок 7.8 – Апроксимація ВАХ нелінійного елемента кусково-неперервними лінійними функціями

У випадку на рис. 7.8 кожна з окремих ділянок є лінійною функцією. Розглянемо спочатку приклад обчислення закону розподілу сигналу на окремій ділянці з лінійним перетворенням вхідного сигналу.

Приклад. Нехай характеристика перетворення *лінійна* $y = f(x) = ax + b$, де a і b – будь-які скінченні дійсні постійні величини. Густина ймовірності $p_x(x)$ значень випадкового процесу $\xi(t)$ на вході перетворювача відома. Знайти густину ймовірності $p_y(y)$ значень процесу $\eta(t)$ на виході перетворювача.

Розв'язання

Оскільки характеристика $y = ax + b$ забезпечує однозначний зв'язок x з y , для розв'язання задачі застосуємо співвідношення (7.9). Обернена функція $x(y)$ і похідна $\frac{dx}{dy}$ матимуть вигляд:

$$x = \frac{y-b}{a} \quad \text{і} \quad \frac{dx}{dy} = \frac{1}{a}.$$

Після підстановки цих значень в (7.9) одержимо

$$p_y(y) = p_x\left(\frac{y-b}{a}\right) \frac{1}{|a|}. \quad (7.11)$$

Для ділянок ВАХ, де сигнал обмежується, порівняно просто отримати одновимірні закони розподілів випадкових сигналів на виході НЕ, не застосовуючи співвідношення (7.11).

Розглянемо *приклад* знаходження закону розподілу сигналу на виході нелінійної системи з ВАХ НЕ на рис. 7.8.

Випадковий сигнал на вході НЕ (рис. 7.8) має одновимірний нормальний закон розподілу з математичним сподіванням, яке дорівнює нулю, і дисперсією σ_x^2

$$p_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{x^2}{2\sigma_x^2}}.$$

Потрібно знайти густину ймовірності сигналу на виході НЕ.

Розв'язок

Позначимо границі трьох ділянок, з яких складається ВАХ нелінійного елемента: $(-\infty, x_1]$, (x_1, x_2) і $[x_2, \infty)$, і розглянемо перетворення закону розподілу сигналу на кожній з них. При цьому, як видно з рисунку 7.8, сигнал на виході НЕ змінюється в границях $[0, y_{\max}]$.

Позначимо ймовірності потрапляння вхідного сигналу на ділянки з обмеженням як:

$$Q_1 = \int_{-\infty}^{x_1} p_x(x) dx \text{ і } Q_2 = \int_{x_2}^{\infty} p_x(x) dx.$$

Тоді, враховуючи властивість густини розподілу ймовірності щодо рівності одиниці при інтегруванні у нескінченних границях, ймовірність потрапляння вхідного сигналу на ділянку (x_1, x_2) складе: $1 - Q_1 - Q_2$.

$$\text{Для цієї ділянки } y(x) = \frac{y_{\max}}{x_2 - x_1} x + y(0) = ax + b,$$

$$\text{де } a = \frac{y_{\max}}{x_2 - x_1} \text{ і } b = y(0) = \frac{y_{\max}}{x_2 - x_1} |x_1|.$$

З урахуванням введених позначень та попереднього прикладу для ділянки (x_1, x_2) , густина ймовірності сигналу на виході НЕ матиме нормальний закон

$$p_y(y) = \frac{1}{\sqrt{2\pi}|a|\sigma_x} e^{-\frac{(y-b)^2}{2a^2\sigma_x^2}}.$$

Для двох інших ділянок з використанням функції Дірака матимемо $p_y(y) = Q_1\delta(y)$ для $x \in (-\infty, x_1]$ і

$p_y(y) = Q_2 \delta(y - y_{\max})$ для $x \in [x_2, \infty)$ відповідно. Поєднаємо отримані для кожної з ділянок результати в остаточний вираз для густини ймовірності сигналу на виході НЕ

$$p_y(y) = \begin{cases} Q_1 \delta(y), & y = 0; \\ \frac{1}{\sqrt{2\pi}|a|\sigma_x} e^{-\frac{(y-b)^2}{2a^2\sigma_x^2}}, & 0 < y < y_{\max}; \\ Q_2 \delta(y - y_{\max}), & y = y_{\max}. \end{cases}$$

Математичне сподівання і дисперсія випадкового процесу на виході нелінійного кола

Найпростіші статистичні характеристики (середнє значення і дисперсію) стаціонарного випадкового процесу на виході безінерційного нелінійного кола можна визначити двома способами.

Якщо методами, розглянутими вище, знайдено густину ймовірності $p_y(y)$ процесу $\eta(t)$ на виході безінерційного НЕ, можна скористатися відомими формулами для обчислення математичного сподівання випадкової величини

$$m_y = \int_{-\infty}^{\infty} y p_y(y) dy$$

та її дисперсії

$$D_y = \sigma_y^2 = \int_{-\infty}^{\infty} (y - m_y)^2 p_y(y) dy = \int_{-\infty}^{\infty} y^2 p_y(y) dy - m_y^2.$$

Якщо густину ймовірності $p_y(y)$ знайти не вдається, але відома густина ймовірності $p_x(x)$ процесу $\xi(t)$ на вході нелінійного кола з характеристикою $y = f(x)$, для знаходження математичного сподівання можна скористатись співвідношенням

$$M\{f(x)\} = m_y = \int_{-\infty}^{\infty} f(x)p_x(x)dx$$

і дисперсії

$$D_y = \sigma_y^2 = \int_{-\infty}^{\infty} f^2(x)p_x(x)dx - m_y^2.$$

В деяких випадках обчислення густини ймовірності або числових характеристик на виході НЕ може бути виконано простіше із застосуванням апарату характеристичних функцій.

За визначенням характеристична функція є математичним сподіванням функції $e^{j\omega x}$, тобто

$$\theta_x(j\omega) = M\{e^{j\omega x}\} = \int_{-\infty}^{\infty} e^{j\omega x} p_x(x) dx. \quad (7.12)$$

Як видно з (7.12), характеристична функція є перетворенням Фур'є. Однак, на відміну від його визначення, наприклад при обчисленні спектральної густини сигналу, в показнику степені комплексної експоненти стоїть знак плюс.

Відповідне (7.12) обернене перетворення має вигляд

$$p_x(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \theta_x(j\omega) e^{-j\omega x} d\omega. \quad (7.13)$$

Для знаходження характеристичної функції $\theta_x(j\omega)$ або за відомою $\theta_x(j\omega)$ густини розподілу $p_x(x)$ можуть бути використані таблиці перетворень по Фур'є.

У разі перетворення $y = f(x)$ випадкової функції $\xi(t)$ характеристична функція процесу $\eta(t)$ на виході безінерційного НЕ може бути обчислена за формулою

$$\theta_y(j\omega) = M\{e^{j\omega f(x)}\} = \int_{-\infty}^{\infty} e^{j\omega f(x)} p_x(x) dx. \quad (7.14)$$

Скориставшись співвідношенням (7.13) по відношенню до $\theta_y(j\omega)$, отримаємо густину ймовірності $p_y(y)$ випадкової функції на виході НЕ.

Властивості характеристичної функції багато в чому повторюють властивості перетворення Фур'є. Наприклад, для лінійного перетворення $y = f(x) = ax + b$ характеристична функція

$$\theta_y(j\omega) = M \left\{ e^{j\omega(ax+b)} \right\} = e^{bj\omega} \theta_x(ja\omega).$$

Для нормованої випадкової величини $\frac{x - m_x}{\sigma_x}$

характеристична функція матиме вигляд

$$M \left\{ e^{j\omega \frac{x - m_x}{\sigma_x}} \right\} = e^{-\frac{m_x j\omega}{\sigma_x}} \theta_x \left(\frac{j\omega}{\sigma_x} \right).$$

За відомою характеристичною функцією доволі просто визначати початкові моменти розподілу випадкової величини. Виконавши диференціювання $\theta_x(j\omega)$ по ω

$$\frac{d^k \theta_x(j\omega)}{d\omega^k} = j^k \int_{-\infty}^{\infty} x^k e^{j\omega x} p_x(x) dx$$

та прирівнявши $\omega = 0$, знайдемо

$$\frac{d^k \theta_x(0)}{d\omega^k} = j^k m_k,$$

де m_k – k -й початковий момент випадкової функції $\xi(t)$. Звідси отримуємо співвідношення для обчислення початкових моментів за відомою $p_x(x)$ характеристичною функцією

$$m_k = \frac{1}{j^k} \frac{d^k \theta_x(0)}{d\omega^k}. \quad (7.15)$$

Для знаходження центральних моментів випадкових величин (як на вході так і на виході НЕ) користуються логарифмом характеристичної функції (або кумулянтною функцією)

$$\psi(j\omega) = \ln \theta(j\omega). \quad (7.16)$$

Оскільки перша та друга похідні функції (7.16)

$$\psi'(j\omega) = \frac{\theta'(j\omega)}{\theta(j\omega)}; \quad \psi''(j\omega) = \frac{\theta(j\omega)\theta''(j\omega) - [\theta'(j\omega)]^2}{\theta^2(j\omega)},$$

і, як видно з (7.12), $\theta(0) = 1$, тому

$$\psi'(0) = \theta'(0); \quad \psi''(0) = \theta''(0) - [\theta'(0)]^2.$$

Оскільки за формулою (7.15) $\theta'(0) = jm_1$ і $\theta''(0) = -m_2$, тому

$$\psi''(0) = -m_2 - [jm_1]^2 = -(m_2 - m_1^2).$$

З урахуванням наведених вище співвідношень для обчислення дисперсії випадкової величини D_y отримаємо

$$\psi''(0) = -M_2 = -D,$$

де M_2 – другий центральний момент розподілу випадкової величини, який дорівнює її дисперсії D .

Таким чином дисперсія випадкового процесу може бути обчислена через другу похідну кумулянтної функції

$$M_2 = D = -\psi''(0) = -[\ln \theta(0)]''. \quad (7.17)$$

Кореляційна функція випадкового сигналу на виході безінерційного нелінійного елемента може бути обчислена, якщо відомий зв'язок $y = f(x)$ випадкового процесу на виході НЕ $\eta(t)$ з випадковим процесом на його вході $\xi(t)$. Крім того, задано двовимірну густину ймовірності $p_{x_2}(x_1, x_2, t_1, t_2)$ випадкового процесу $\xi(t)$. Потрібно знайти кореляційну функцію $R_y(t_1, t_2)$ випадкового процесу $\eta(t)$ та її математичне сподівання.

В загальному вигляді її можна визначити за формулою

$$R_y(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 y_2 p_{y_2}(y_1, y_2, t_1, t_2) dy_1 dy_2 - \quad (7.18)$$

$$- m_y(t_1) m_y(t_2),$$

де $p_{y_2}(y_1, y_2, t_1, t_2)$ – двовимірна густина ймовірності випадкового процесу $\eta(t)$; $y_1 = y(t_1)$ і $y_2 = y(t_2)$;

$m_y(t)$ – математичне сподівання випадкового процесу $\eta(t)$ на виході НЕ.

Щоб скористатись цим співвідношенням, необхідно спочатку знайти густину ймовірності $p_{y_2}(y_1, y_2, t_1, t_2)$ і математичне сподівання $m_y(t)$. Однак, якщо коло безінерційне і його характеристика $y = f(x)$ однозначна, виконується рівність $p_x(y)dy = p_y(x)dx$. У цьому разі вираз (7.18) можна подати у вигляді

$$R_y(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1) f(x_2) p_{x_2}(x_1, x_2, t_1, t_2) dx_1 dx_2 - m_y(t_1) m_y(t_2), \quad (7.19)$$

тобто представити через густину ймовірності вхідного процесу.

У (7.19) $x_1 = x(t_1)$ і $x_2 = x(t_2)$. Якщо процеси $\xi(t)$ і $\eta(t)$ стаціонарні, принаймні, у широкому розумінні, і центровані (7.19) набуває вигляду

$$R_y(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1) f(x_2) p_{x_2}(x_1, x_2, \tau) dx_1 dx_2, \quad (7.20)$$

де $x_1 = x(t)$ і $x_2 = x(t + \tau)$.

Якщо ж, крім цього, процеси є ергодичними тоді

$$R_y(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} f[x(t)] f[x(t + \tau)] dt. \quad (7.21)$$

Обчислення кореляційних функцій на виході НЕ за формулами (7.19) – (7.21), за винятком окремих випадків, є досить складною задачею через складність знаходження інтегралів навіть у разі стаціонарних випадкових процесів.

Тепер можна сформулювати, у чому полягає відмінність в задачі знаходження кореляційних функцій на виході лінійних і нелінійних систем.

У лінійних системах для розв'язання задачі стосовно стаціонарних процесів достатньо знати кореляційну функцію сигналу на вході та імпульсну характеристику (або передавальну функцію) системи (див. главу 6).

У нелінійних системах для розв'язання задачі необхідно знати, крім характеристики $y = f(x)$, *двовимірну густину ймовірності* (або двовимірну функцію розподілу ймовірностей) вхідного сигналу. Цим і пояснюється те, що подібні *нелінійні задачі значно складніші за лінійні*.

В деяких випадках більш зручним для обчислення кореляційної функції на виході нелінійної безінерційної системи може бути застосування методу характеристичних функцій. В цьому методі виконується заміна закону розподілу його характеристичною функцією, а нелінійної функції $f(x)$ – її зображенням по Фур'є:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega) e^{j\omega x} d\omega; \quad (7.22)$$

$$S(j\omega) = \int_{-\infty}^{\infty} f(x) e^{-j\omega x} dx.$$

В результаті такої заміни і підстановки формули (7.22) в формулу для обчислення ковариаційної функції

$$K_y(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1) f(x_2) p_{x_2}(x_1, x_2, t_1, t_2) dx_1 dx_2, \quad \text{яка є}$$

першим доданком в (7.19), отримаємо:

$$\begin{aligned}
 K_y(t_1, t_2) &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(j\omega_1) e^{j\omega_1 x_1} d\omega_1 \times \\
 &\times \int_{-\infty}^{\infty} S(j\omega_2) e^{j\omega_2 x_2} d\omega_2 p_{x_2}(x_1, x_2, t_1, t_2) dx_1 dx_2 = \\
 &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(j\omega_1) S(j\omega_2) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{x_2}(x_1, x_2, t_1, t_2) e^{j(\omega_1 x_1 + j\omega_2 x_2)} \times \\
 &\times dx_1 dx_2 d\omega_1 d\omega_2.
 \end{aligned}$$

Оскільки

$$\begin{aligned}
 &\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{x_2}(x_1, x_2, t_1, t_2) e^{j(\omega_1 x_1 + \omega_2 x_2)} dx_1 dx_2 = \\
 &= M \left\{ e^{j\omega_1 x_1} e^{j\omega_2 x_2} \right\} = \theta_2(j\omega_1, j\omega_2, t_1, t_2),
 \end{aligned}$$

де $\theta_2(j\omega_1, j\omega_2, t_1, t_2)$ – двовимірна характеристична функція випадкового процесу з густиною ймовірності $p_{x_2}(x_1, x_2, t_1, t_2)$. Тому остаточно ковариаційна функція може бути обчислена за формулою

$$K_y(t_1, t_2) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(j\omega_1) S(j\omega_2) \theta_2(j\omega_1, j\omega_2, t_1, t_2) d\omega_1 d\omega_2. \tag{7.23}$$

Приклади рішення задач, пов'язаних з обчисленням кореляційних (ковариаційних) функцій, можна знайти, наприклад, у [9, 17, 33].

Отже, ми досить докладно проаналізували окремий випадок перетворення *одновимірної* густини ймовірності $p_x(x)$ значень випадкового процесу $\xi(t)$ в *одновимірну* густину ймовірності $p_y(y)$ процесу $\eta(t)$, що є відгуком безінерційної нелінійної системи на дію коливання $\xi(t)$. Однак, одновимірні густини ймовірності в загальному

випадку не дають вичерпного опису випадкових процесів (див. главу 6). Наприклад, щоб визначити кореляційну функцію і спектральну густину середньої потужності процесу на виході, необхідно знати *двовимірну* густину ймовірності. У деяких випадках потрібні ймовірнісні характеристики *більш високих порядків*.

7.2. Модульовані сигнали

При створенні каналів зв'язку інформаційно-вимірювальних систем для ефективного передавання низькочастотних сигналів на великі відстані виникає потреба у перенесенні їх на більш високі. При цьому також дуже часто в одному й тому ж каналі зв'язку необхідно передавати декілька сигналів одночасно. Якщо різні сигнали займають смуги частот, що не перекриваються, можливим є, так зване, частотне розділення. Частіш за все вимагається, щоб випромінюваний сигнал був *вузькосмуговим*, тобто ефективна ширина його спектру була набагато менше його центральної частоти $\Delta f \ll f_0$. Тому з метою передавання вимірювальний сигнал перетворюється так, щоб вимоги до смуги частот каналу передавання виконувались, а сам вхідний сигнал можна було б відновити. Подібні перетворення застосовуються також в аналізаторах спектру сигналів.

Вирішення вказаної проблеми досягається шляхом використання *модуляції* (англ. modulation), тобто процесу зміни в часі одного або декількох параметрів сигналу переносника з метою передавання інформації каналом зв'язку. Її суть полягає в наступному. Формується деяке електромагнітне коливання, призначене для створення радіочастотного сигналу зміною одного чи декількох параметрів цього коливання (радіохвилі) – *коливання-носії* (англ. carrier) (в ДСТУ 3254-95. Радіозв'язок. Терміни та визначення. застосовується термін *несівне коливання* або просто *несівна*, інколи зустрічається термін *носійне*

коливання). Параметр цього коливання змінюється в часі пропорційно сигналу на вході. Цей сигнал називають модулювальним (*англ. modulating signal*), а результуюче коливання з параметрами, що змінюються в часі, – *модульованим сигналом* (*англ. modulated signal*). Зворотний процес – виділення модулювального сигналу з модульованого коливання – називається *демодуляцією* (*англ. demodulation*) [11].

На рисунку 7.9 представлена функціональна схема модулятора, де $u(t)$ – вхідний модулювальний сигнал (сигнал, що модулює); $s(t)$ – вихідний сигнал (модульований сигнал); $s_0(t)$ – сигнал переносник. В процесі *модуляції* параметри коливання-носія $s_0(t)$ змінюються під впливом вхідного сигналу $u(t)$ для отримання модульованого сигналу $s(t)$ [30].

Коливання-носії частіш за все є гармонічним

$$s_0(t) = A_0 \cos(\omega_0 t + \varphi_0) = A_0 \cos \theta(t), \quad (7.24)$$

де A_0 – амплітуда коливання, $\omega_0 = 2\pi f_0$ – кутова частота, φ_0 – початкова фаза, $\theta(t)$ – повна фаза.

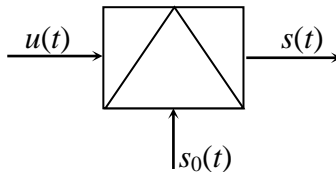


Рисунок 7.9 – Функціональна схема модулятора

Основними видами модуляції є амплітудна (АМ) (*англ. amplitude modulation*) і кутова (КМ) (*англ. angle modulation*). Якщо змінним параметром є амплітуда сигналу переносника (S_0), тоді сигнал буде *амплітудно-модульований* ($s_{\text{АМ}}(t)$). При зміні повної фази $\theta(t)$ отримаємо сигнал з *кутовою модуляцією* ($s_{\text{КМ}}(t)$). Кутова модуляція поділяється на два тісно пов'язані між види: частотна модуляція (ЧМ) (*англ. frequency modulation*) і фазова модуляція (ФМ)

(англ. phase modulation). Сигнали з такими видами модуляції позначатимемо відповідно як $s_{\text{ЧМ}}(t)$ і $s_{\text{ФМ}}(t)$

На практиці зустрічаються й змішані види модуляції: амплітудно-фазова, амплітудно-частотна та інші.

7.2.1. Амплітудно-модульовані сигнали

При *амплітудній модуляції* у відповідності до модульовального сигналу змінюється амплітуда колювання-носія

$$s_{\text{АМ}}(t) = A(t) \cos(\omega_0 t + \varphi_0).$$

Для реалізації АМ до модульовального сигналу зазвичай додають постійну складову. Це робить його однополярним і дозволяє запобігти спотворенню обвідної двополярного сигналу при його демодуляції

$$A(t) = A_0 (1 + m s_M(t)),$$

де m – коефіцієнт амплітудної модуляції (коефіцієнт глибини модуляції), практично його значення приймають в межах 0,4...0,6.

$s_M(t)$ – сигнал, що модулює, ($-1 < s_M(t) \leq 1$);

Остаточно АМ сигнал можна записати в наступному виді

$$s_{\text{АМ}}(t) = A_0 (1 + m s_M(t)) \cos(\omega_0 t + \varphi_0). \quad (7.25)$$

Спектральну структуру АМ сигналу зручно дослідити для випадку *однотональної АМ*, коли модульовальний сигнал є гармонічним

$$s_M(t) = A_M \cos(\Omega t + \Phi_0),$$

з частотою $\Omega \ll \omega_0$.

В цьому випадку

$$s_{\text{АМ}}(t) = A_0 (1 + m \cos(\Omega t + \Phi_0)) \cos(\omega_0 t + \varphi_0). \quad (7.26)$$

Коефіцієнт модуляції $m = A_M / A_0$ може бути обчислений за результатами вимірювання (наприклад, за допомогою

осцилографа) максимальної A_{\max} і мінімальної A_{\min} амплітуд сигналу: $m = (A_{\max} - A_{\min}) / (A_{\max} + A_{\min})$. В окремих випадках вимірюють відносні коефіцієнти модуляції вгору $m_{\text{в}} = (A_{\max} - A_0) / A_0$ та вниз $m_{\text{н}} = (A_0 - A_{\min}) / A_0$ [11, 17].

На рис. 7.10 приведені епюри напруг для різних значень коефіцієнта амплітудної модуляції, які дають уяву про вигляд сигналу при порівняно неглибокій (б), глибокій модуляції (в) та при його перемодуляції (г).

Якщо коефіцієнт амплітудної модуляції лежить в діапазоні $0 \dots 1$, обвідна модульованого сигналу повторює форму модульовального сигналу $s_M(t)$ (рис. 7.10, б, в).

Перемодуляція має місце при $m > 1$, при цьому амплітудна обвідна спотворюється (рис. 7.10, г). Проте й цей режим в деяких випадках також є корисним на практиці.

З метою аналізу спектрального складу такого коливання розкриємо дужки у виразі (7.26) й виконаємо тригонометричні перетворення

$$\begin{aligned} s_{\text{AM}}(t) &= A_0 \cos(\omega_0 t + \varphi_0) + A_0 m \cos(\Omega t + \Phi_0) \cos(\omega_0 t + \varphi_0) = \\ &= A_0 \cos(\omega_0 t + \varphi_0) + \frac{A_0 m}{2} \cos((\omega_0 + \Omega)t + \varphi_0 + \Phi_0) + \\ &\quad + \frac{A_0 m}{2} \cos((\omega_0 - \Omega)t + \varphi_0 - \Phi_0). \end{aligned} \quad (7.27)$$

Результат перетворення показує, що однотональний АМ сигнал складається з трьох гармонічних складових. Одна з них є коливанням-носієм з частотою ω_0 , а дві інші – бічні коливання з частотами $\omega_0 - \Omega$ і $\omega_0 + \Omega$ (їх називають нижньою і верхньою бічними частотами). Амплітуда коливання-носія дорівнює A_0 і не залежить від рівня модульовального сигналу. Амплітуди ж коливань бічних частот дорівнюють $A_0 m / 2$ і є пропорційними коефіцієнту модуляції.

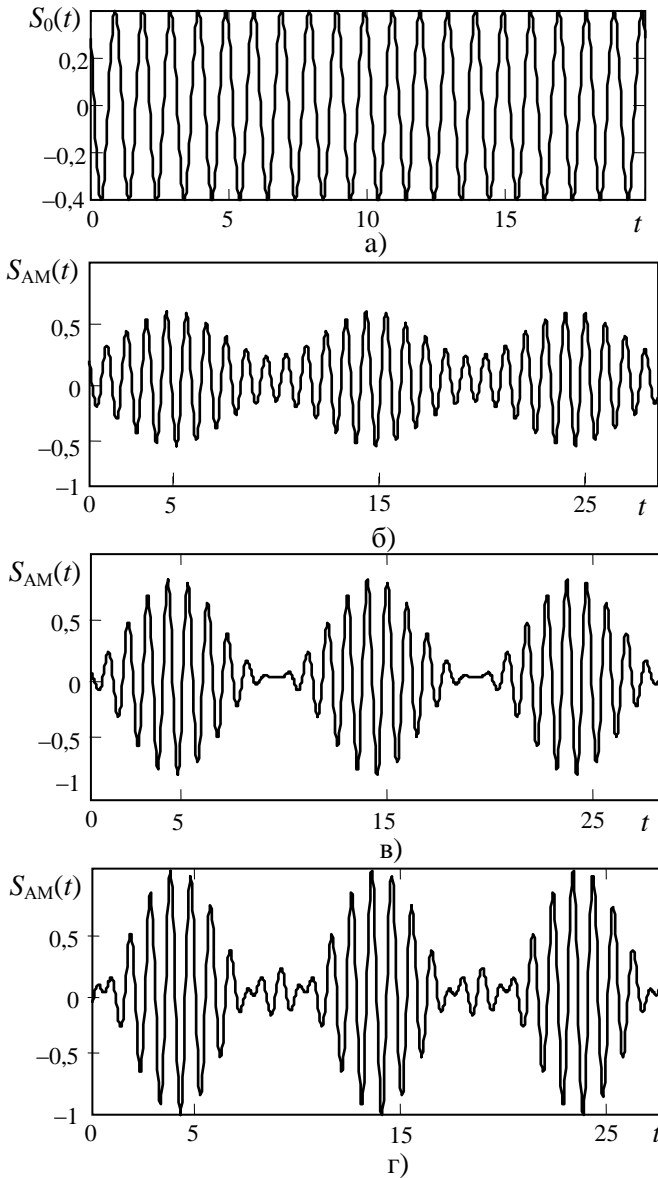


Рисунок 7.10 – Однотональний АМ сигнал: а) $m = 0$ (коливання-носії); б) $m = 0,5$; в) $m = 1$; г) $m = 1,5$

Для коливання з частотою $\omega_0 + \Omega$ початкові фази коливання-носія і модульовального сигналу додаються, а для коливання з частотою $\omega_0 - \Omega$ – віднімаються. Амплітудний і фазовий спектри однотонального АМ сигналу показані на рис. 7.11. Цей спектр дискретний, але в загальному випадку він не є розкладанням в ряд Фур'є. Це має місце лише тоді, коли частоти спектральних складових мають загальний дільник, тобто співвідношення ω_0/Ω є раціональним дробом.

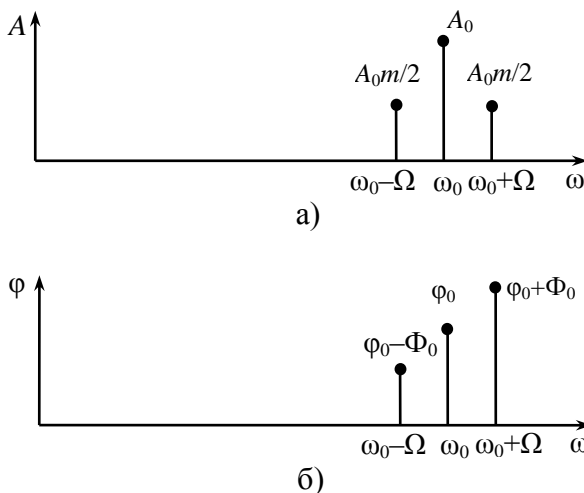


Рисунок 7.11 – Спектр однотонального АМ сигналу:
а) амплітудний; б) фазовий

Видно також, що ширина спектру однотонального АМ сигналу в два рази перевищує частоту модульовального сигналу: $\Delta\omega = 2\Omega$.

Векторна діаграма (рис. 7.12) пояснює яким чином з трьох гармонічних складових з постійними амплітудами складається сигнал із змінною амплітудою \dot{S}_{AM} .

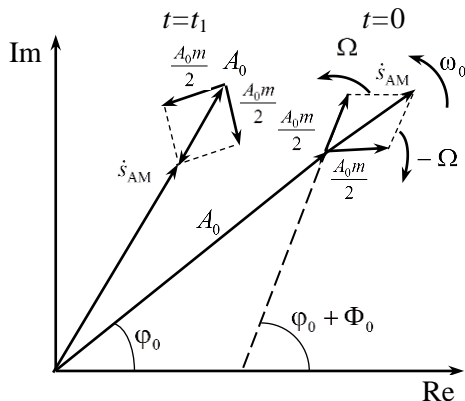


Рисунок 7.12 – Векторна діаграма однотонового АМ сигналу для двох фіксованих моментів часу $t=0$ і $t=t_1$

Для побудови векторної діаграми кожне з трьох гармонічних коливань представляється як дійсна частина комплексної експоненти, тобто

$$s_{AM}(t) = \text{Re} \left[e^{j\omega_0 t} \left(A_0 e^{j\varphi_0} + \frac{A_0 m}{2} e^{j(\Omega t + \varphi_0 + \Phi_0)} + \frac{A_0 m}{2} e^{j(-\Omega t + \varphi_0 - \Phi_0)} \right) \right].$$

При складанні коливань з різними частотами постійно змінюються їх взаємні фазові співвідношення. Вектор коливання-носія обертається з кутовою швидкістю ω_0 . Вектори бічних коливань при цьому обертаються в різних напрямках з кутовою швидкістю Ω відносно вектора коливання-носія. В результаті сумарний вектор бічних коливань виявляється або спрямованим як і вектор коливання-носія (синфазним), збільшуючи його амплітуду, або протифазним з ним (тоді амплітуда сигналу зменшується). Зауважимо, що до момент часу t_1 на діаграмі рис. 7.12 вектор коливання-носія встиг зробити кілька обертів, тоді як вектори бічних коливань зробили поворот на кути $\pm\Omega t_1$ (кути не

позначені на діаграмі). Поточне значення АМ сигналу є проекцією результуючого вектору на дійсну вісь.

Середня потужність коливання-носія, тобто гармонічного сигналу за відсутності модульовального сигналу, дорівнює

$$P_0 = \frac{A_0^2}{2}.$$

Амплітуда модульованого сигналу змінюється в межах від мінімальної [30]

$$S_{AM \min} = A_0 - \frac{mA_0}{2} - \frac{mA_0}{2} = A_0(1-m)$$

до максимальної

$$S_{AM \max} = A_0 + \frac{mA_0}{2} + \frac{mA_0}{2} = A_0(1+m).$$

Зі зміною амплітуди S_{AM} змінюється й середня потужність сигналу [30]:

– поблизу до максимальної амплітуди потужність

$$P_{\max} = \frac{S_{AM \max}^2}{2} = \frac{A_0^2(1+m)^2}{2} = P_0(1+m)^2;$$

– в околиці мінімальної S_{AM} потужність

$$P_{\min} = \frac{S_{AM \min}^2}{2} = \frac{A_0^2(1-m)^2}{2} = P_0(1-m)^2.$$

Максимальна потужність при $m = 1$, становить

$$P_{\max} = 4P_0.$$

Тепер знайдемо середню потужність АМ сигналу. Оскільки гармонічні коливання різних частот є некорельованими, тому їх потужності можна складати. Застосувавши сказане до подання однотонального АМ сигналу (7.27) отримаємо співвідношення

$$P_{\text{cp}} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s_{AM}^2(t) dt = \frac{A_0^2}{2} + \frac{A_0^2 m^2}{4}.$$

Перший доданок не залежить від коефіцієнта модуляції і являє собою потужність немодульованого коливання-носія. Корисна потужність, яка зосереджена в бічних коливаннях, представлена другим доданком.

Коефіцієнт корисної дії (ККД) амплітудної модуляції, визначається як відношення потужності бічних коливань до загальної середньої потужності сигналу [11].

$$\eta_{AM} = \frac{A_0^2 \frac{m^2}{4}}{A_0^2 \left(\frac{1}{2} + \frac{m^2}{4} \right)} = \frac{m^2}{m^2 + 2}.$$

Графік залежності коефіцієнта корисної дії (ККД) від коефіцієнта амплітудної модуляції m наведено на рис. 7.13.

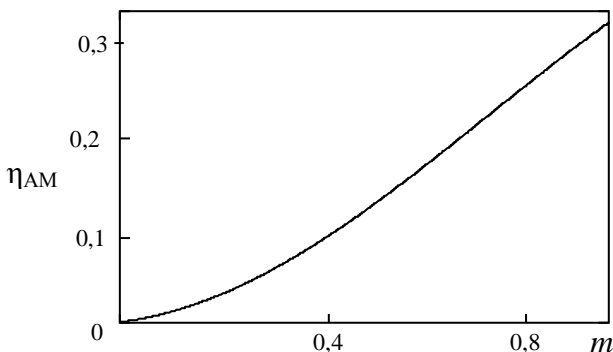


Рисунок 7.13 – Залежність ККД від коефіцієнта амплітудної модуляції

В [9], як показник ефективності використання потужності при передаванні АМ сигналу, пропонується застосовувати відношення середніх потужностей в обох бічних коливаннях до потужності коливання-носія

$$(P_{cp} - P_0) / P_0 = m^2 / 2.$$

При $m = 1$ частка потужності обох бічних коливань складає 50 % від потужності коливання-носія. Але як ККД цей показник використовувати не можна (при видаленні з сигналу коливання-носія – маємо нескінченність).

Амплітудні модулятори частіш за все будуються за принципом перетворення спектру суми двох сигналів (модулювального та гармонічного коливання-носія) в безінерційному нелінійному пристрої. Цей принцип можна реалізувати, наприклад, із застосуванням нелінійного транзисторного підсилювача. Резонансний контур на виході підсилювача налаштовується на частоту коливання-носія. Вольтамперна характеристика НЕ може бути апроксимована кусково-лінійною кривою, як це показано на рис. 7.3. Напруга зсуву U_3 при додаванні низькочастотної модулювальної напруги змінюється в часі. Це призводить до зміни в часі кута відсічки коливання-носія. В результаті амплітуда першої гармоніки, на частоту якої налаштований вихідний контур в колі колектора, також змінюється за законом модулювального коливання. Тому з колектора транзистора через роздільний конденсатор отримується АМ сигнал. Якщо на вході (на базі транзистора) модулятора діє напруга

$$u_1(t) = U_3 + U_{\text{тнос}} \cos \omega_0 t + U_{\text{тмод}} \cos \Omega t,$$

де $U_{\text{тнос}}$ – амплітуда коливання-носія,

$U_{\text{тмод}}$ – амплітуда модулювального гармонічного коливання, тоді поблизу частоти ω_0 в коливальному контурі протікатиме струм, який на вихідному опорі дозволить отримати АМ сигнал

$$u_2(t) = a_1 U_{\text{тнос}} \cos \omega_0 t + a_2 U_{\text{тнос}} U_{\text{тмод}} \cos(\omega_0 + \Omega)t + \\ + a_2 U_{\text{тнос}} U_{\text{тмод}} \cos(\omega_0 - \Omega)t.$$

З урахуванням (7.27) коефіцієнт АМ матиме значення

$$m = 2a_2 U_{\text{тмод}} / a_1.$$

Історично АМ була першим практично освоєним видом модуляції. Однак низький ККД і ширина спектра, що вдвічі

перевищує ширину спектра модульовального сигналу, обмежили сфери застосування АМ у зв'язку. В даний час АМ застосовується для радіомовлення на порівняно низьких частотах (в діапазонах довгих, середніх і коротких хвиль) і для передавання зображення в телевізійному мовленні.

З метою поліпшення характеристик АМ розроблено кілька її модифікацій, які розглянемо далі.

Очевидно, що для підвищення ККД амплітудної модуляції доцільно видалити «зайве» коливання-носія, все ж таки відмовившись від додавання постійної складової до модульовального сигналу. Такий спосіб називається АМ з придушеною частотою-носієм (АМ-ПН, англійський термін – amplitude modulation with suppressed carrier, АМ-SC) або балансною АМ: $s_{AM}(t) = s_M(t) \cos(\omega_0 t + \varphi_0)$. Тоді, для однотонального модульовального сигналу (без урахування початкових фаз коливань) при $s_M(t) = m \cos \Omega t$ маємо:

$$\begin{aligned} s_{AM}(t) &= mU_m \cos \Omega t \cos \omega_0 t = \\ &= \frac{mU_m}{2} \{ \cos [(\omega_0 + \Omega)t] + \cos [(\omega_0 - \Omega)t] \}. \end{aligned}$$

На рис. 7.14 побудовано графік такого сигналу. На рисунку спостерігаємо биття двох гармонічних сигналів з однаковими амплітудами з верхньою і нижньою бічними частотами.

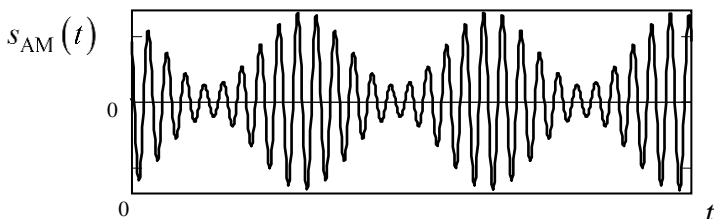


Рисунок 7.14 – Однотональна АМ з придушеною частотою-носієм

Енергетичний вигравш при цьому, звичайно, великий (згідно введеного визначення, ККД стає рівним 100%).

Структурна схема балансного модулятора (БМ) представлена на рис. 7.15 [17].

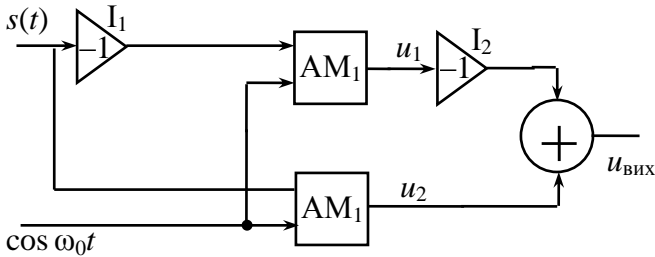


Рисунок 7.15 – Структурна схема балансного модулятора

Тут гармонійне коливання-носій з частотою ω_0 синфазно підводиться до нижніх входів двох однакових амплітудних модуляторів AM_1 і AM_2 . Модулювальний сигнал $s(t)$ надходить на модулятор AM_1 через інвертор I_1 , коефіцієнт передачі якого дорівнює -1 . Тому на виходах модуляторів будуть отримані сигнали:

$$u_1(t) = A[1 - Ms(t)] \cos \omega_0 t ;$$

$$u_2(t) = A[1 + Ms(t)] \cos \omega_0 t ,$$

де A – постійний коефіцієнт.

Інвертор I_2 змінює знак сигналу u_1 на протилежний, тому вихідний сигнал

$$u_{\text{вых}}(t) = u_2 - u_1 = 2AMs(t) \cos \omega_0 t$$

являє собою добуток модулювального коливання і коливання-носія, тобто є балансно-модульованим коливанням.

Ширина спектра АМ сигналу з придушеною частотою-носієм така ж, як у випадку звичайної АМ (адже придушена середня частота, а бічні частоти залишилися на місці).

Таким чином, амплітудна модуляція з придушеною частотою-носієм має певні енергетичні переваги в порівнянні зі звичайною АМ. Однак цей спосіб модуляції не отримав

широкого розповсюдження, що пов'язано з проблемами, які виникають при демодуляції сигналу.

Двосмугова АМ з придушеною частотою-носієм має переваги перед звичайною АМ тільки в сенсі. Ширина спектра сигналу при цьому залишається рівною подвоєній частоті модульовального сигналу.

Однак, оскільки спектри двох бічних смуг АМ сигналу є дзеркальним відображенням одна одної, вони несуть одну й ту саму інформацію. Тому одну з бічних смуг можна видалити. Отримана модуляція називається *односмуговою* (англійський термін – *single side band, SSB*).

Залежно від того, яка бічна смуга зберігається, говорять про односмугову модуляцію з використанням верхньої або нижньої бічної смуги. Формування однополосного сигналу пояснюється графіками спектрів (рис. 7.16).

По суті, при односмуговій модуляції відбувається просто зсув спектра сигналу в околиці частоти-носія.

Очевидно, що ширина спектра односмугового сигналу дорівнює ширині спектра модульовального сигналу. Таким чином, спектр односмугового сигналу виявляється в два рази вузким, ніж при звичайній АМ.

На відміну від попередніх випадків, щоб встановити зв'язок між модульованим і модульовальним сигналами зручно скористатись перетворенням Гільберта і поняттям аналітичного сигналу (див. розділ 1.2).

Для АМ сигналу $u(t) = U_m \left[1 + \sum_n m_n s(t, n) \right] \cos \omega_0 t$ його квадратурне доповнення визначається перетворенням Гільберта і дорівнює $\tilde{u}(t) = U_m \left[1 + \sum_n m_n s(t, n) \right] \sin \omega_0 t$.

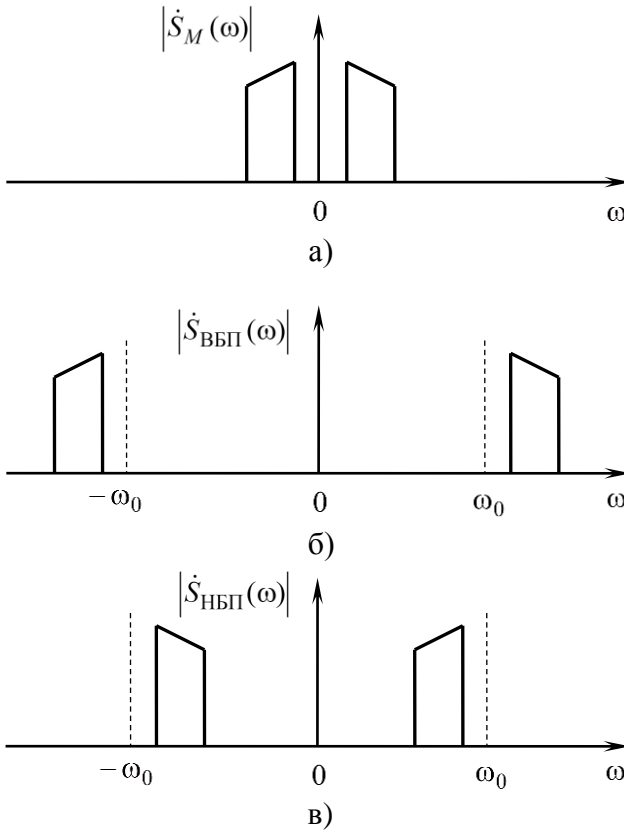


Рисунок 7.16 – Односмугова модуляція:

а) спектр модулювального сигналу; б) спектр односмугового сигналу з верхньою бічною смугою; в) те ж саме з нижньою бічною смугою

Обвідна аналітичного сигналу

$$\begin{aligned}
 |\dot{z}(t)| &= \sqrt{u^2(t) + \tilde{u}^2(t)} = \\
 &= \sqrt{U_m^2 \left[1 + \sum_n m_n s(t, n) \right]^2} (\cos^2 \omega_0 t + \sin^2 \omega_0 t) =
 \end{aligned}$$

$$= U_m \left[1 + \sum_n m_n s(t, n) \right].$$

Для знаходження однобічного АМ сигналу за довільного модульовального сигналу з обмеженою смугою частот спочатку потрібно сформувавши аналітичний сигнал, що має односторонній спектр [17]. Множення цього сигналу на $\exp(j\omega_0 t)$ зсуває його односторонній спектр на ω_0 у бік зростання й утворює односторонній спектр односмугового сигналу з верхньою бічною смугою. Для переходу від аналітичного сигналу знову до дійсного, потрібно взяти дійсну частину. Формування сигналу з нижньою бічною смугою описується аналогічно, тільки множити аналітичний сигнал потрібно на $\exp(-j\omega_0 t)$ (тоді його спектр зсунеться вліво, в область від'ємних частот, і займе положення нижньої бічної смуги). Математично це виглядає так

$$\begin{aligned} s_{SSB}(t) &= \operatorname{Re} \left[(s_M(t) + j\tilde{s}_M(t)) \exp(\pm j\omega_0 t) \right] = \\ &= s_M(t) \cos \omega_0 t \mp \tilde{s}_M(t) \sin \omega_0 t. \end{aligned}$$

Знак «мінус» в цьому виразі відповідає видаленню верхньої бічної смуги, «плюс» – нижньої.

Отже, односмуговий сигнал можна представити сумою двох АМ сигналів, коливання-носії яких мають однакові частоти та відносний фазовий зсув 90° . Амплітудними функціями цих АМ сигналів є модульовальний сигнал і його квадратурне доповнення. Залежно від того, складаються ці два АМ сигнали або віднімаються (а точніше, яка з двох частот-носіїв випереджає іншу по фазі), формується односмуговий сигнал з верхньою або нижньою бічною смугою.

У випадку додаткового усунення коливання-носія ККД АМ сигналу зростає ще більше. Амплітудна обвідна односмугового сигналу не схожа на модульовальний низькочастотний сигнал, що ілюструється на прикладі модульовального сигналу з двох гармонік $s_M(t) = \cos(\omega_1 t + \varphi_1) + \cos(\omega_2 t + \varphi_2)$. В цьому прикладі

сигнали з верхньою та нижньою бічною смугою з усунутою частотою-носієм:

$$s_B(t) = \cos[(\omega_0 + \omega_1)t + \varphi_1] + \cos[(\omega_0 + \omega_2)t + \varphi_2];$$

$$s_H(t) = \cos[(\omega_0 - \omega_1)t - \varphi_1] + \cos[(\omega_0 - \omega_2)t - \varphi_2].$$

Обвідні таких сигналів мають суттєві відмінності, що легко перевірити, побудувавши їх графіки.

Односмугова модуляція з різним рівнем коливання-носія – від повного придушення до повного збереження – широко застосовується в техніці радіозв'язку.

7.2.2. Сигнали з кутовою модуляцією

При фазовій модуляції модулювальний сигнал визначає початкову фазу коливання-носія: $\varphi(t) = ks_M(t)$. Тому сигнал з фазовою модуляцією (ФМ; англійський термін – phase modulation, PM) має вигляд

$$s_{\text{ФМ}}(t) = A \cos(\omega_0 t + ks_M(t)),$$

де k – коефіцієнт пропорційності, $s_M(t)$ – закон модуляції.

Позначимо повну фазу коливання: $\psi(t) = \omega_0 t + ks_M(t)$, також, як і у випадку АМ, прийемо $-1 \leq s_M(t) \leq 1$. При $s_M(t) = 1$ коефіцієнт k дорівнює максимальній девіації фази вгору і при $s_M(t) = -1$ – максимальній девіації фази вниз.

Кругова частота коливання за визначенням є швидкістю зміни фази. Для коливань з кутовою модуляцією вона змінюється у часі і називається *миттєвою частотою* (англ. instantaneous frequency), яка є похідною від повної фази за часом

$$\omega(t) = \frac{d\psi}{dt} = \omega_0 + k \frac{ds_M}{dt}. \quad (7.28)$$

Отже, у разі фазової модуляції змінюється не тільки

початкова фаза, але й миттєва частота коливання.

Відповідно, повна фаза може бути знайдена шляхом інтегрування миттєвої частоти

$$\Psi(t) = \int_{-\infty}^t \omega(\tau) d\tau. \quad (7.29)$$

У випадку *частотної модуляції* (ЧМ; англійський термін – frequency modulation, FM) модульовальний сигнал лінійно пов'язаний з миттєвою частотою коливання

$$\omega(t) = \frac{d\Psi}{dt} = \omega_0 + k s_M(t).$$

Доданок у вигляді константи ω_0 робить коливання високочастотним. Повна фаза знаходиться згідно (7.29) шляхом інтегрування $\omega(t)$

$$\Psi(t) = \omega_0 t + k \int_{-\infty}^t s_M(\tau) d\tau + \varphi_0.$$

де φ_0 – початкова фаза коливання в момент $t=0$, яка в загальному випадку може бути довільною. Нарешті, сам ЧМ сигнал має наступний вигляд

$$s_{\text{ЧМ}}(t) = A \cos \left[\omega_0 t + k \int_{-\infty}^t s_M(\tau) d\tau + \varphi_0 \right].$$

Як видно, початкова фаза коливання при частотній модуляції змінюється пропорційно інтегралу від модульовального сигналу: $\varphi(t) = k \int s_M(\tau) d\tau + \varphi_0$.

Фазову і частотну модуляції об'єднують під загальною назвою «кутова модуляція» саме, завдяки тому, що певній зміні початкової фази коливання відповідатиме й зміна миттєвої частоти (і навпаки). Це дозволяє зробити висновки [11]:

за формою коливання з ЧМ не можна визначити це сигнал з ФМ чи ЧМ, для цього потрібно також знати й форму модульовального сигналу;

якщо модульований сигнал пропустити через ідеальний диференціальний пристрій, з виходу якого подати на частотний модулятор, вийде ФМ сигнал (верхня гілка на рис. 7.17) і, навпаки, при проходженні модульовального сигналу через ідеальний інтегрувальний пристрій та далі через фазовий модулятор, отримаємо ЧМ сигнал (нижня гілка на рис. 7.17).

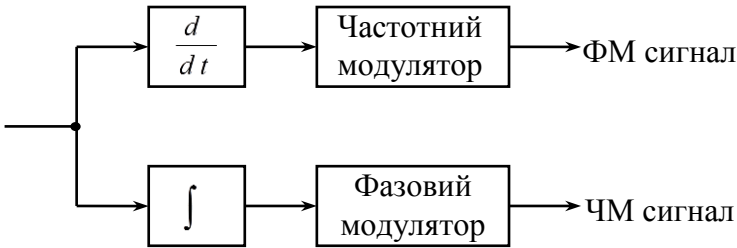


Рисунок 7.17 – Взаємозв’язок фазової і частотної модуляції

Узагальнені дані, які показують зв’язок модульовального сигналу з різними характеристиками модульованого колювання при ФМ і ЧМ, зведені в таблицю 7.2 [11].

Теоретичний розрахунок спектру сигналу з КМ в загальному випадку не має простого рішення. Тому для його оцінки, як і у випадку АМ сигналу, розглянемо *гармонічну кутову модуляцію*. Початкова фаза колювання матиме закон

$$\varphi(t) = \beta \sin(\Omega t + \Phi_0) + \varphi_0,$$

де коефіцієнт β – *індекс кутової модуляції* (англ. modulation index), Φ_0 – початкова фаза модульовального гармонічного сигналу. Повна фаза отримується додаванням лінійного доданка $\omega_0 t$

$$\Psi(t) = \omega_0 t + \beta \sin(\Omega t + \Phi_0) + \varphi_0.$$

Таблиця 7.2 – Характеристики модульованого коливання при фазовій і частотній модуляції

Параметр	ФМ	ЧМ
Початкова фаза	$\varphi(t) = ks_M(t)$	$\varphi(t) = k \int s_M(\tau) d\tau + \varphi_0$
Повна фаза	$\psi(t) = \omega_0 t + ks_M(t)$	$\psi(t) = \omega_0 t + k \int s_M(\tau) d\tau + \varphi_0$
Миттєва частота	$\omega(t) = \omega_0 + k \frac{ds_M}{dt}$	$\omega(t) = \omega_0 + ks_M(t)$

Видно, що індекс β визначає найбільше відхилення повної фази від лінійного закону, тобто девіацію фази.

Нарешті, сам сигнал з гармонічною КМ має вигляд

$$s_{\text{КМ}}(t) = A \cos \left[\omega_0 t + \beta \sin(\Omega t + \Phi_0) + \varphi_0 \right]. \quad (7.30)$$

Закон зміни миттєвої частоти в цьому випадку також буде гармонічним

$$\omega(t) = \frac{d\Psi}{dt} = \omega_0 + \beta \Omega \sin(\Omega t + \Phi_0).$$

Як видно з отриманої формули, максимальне відхилення від середнього значення ω_0 , тобто девіація частоти (*англ.* frequency deviation), становить $\beta \Omega$. Вона позначається ω_d . Відповідно індекс КМ є відношенням девіації частоти до частоти модульовального сигналу: $\beta = \omega_d / \Omega$.

Отже, при гармонічній КМ як початкова фаза, так і миттєва частота змінюються по гармонічному закону. Однак, при ФМ індекс β , тобто девіація фази, не залежить від частоти модульовального сигналу і є характеристичним параметром ФМ модуляції, тоді як при ЧМ девіація частоти є прямо пропорційною Ω : $\beta = \text{const}$, $\omega_d = \beta \Omega$. Тому відмінності між ЧМ та ФМ починають виявлятися при зміні частоти модульовального сигналу Ω .

Зазвичай при ЧМ девіація частоти є фіксованим характеристичним параметром, який не залежить від частоти

модульовального сигналу, $\omega_d = \text{const}$. Тому в цьому випадку індекс модуляції обернено пропорційний Ω : $\beta = \omega_d / \Omega$.

Розглянуті параметри ФМ і ЧМ, як функції частоти Ω , ілюструються на рис. 7.18.

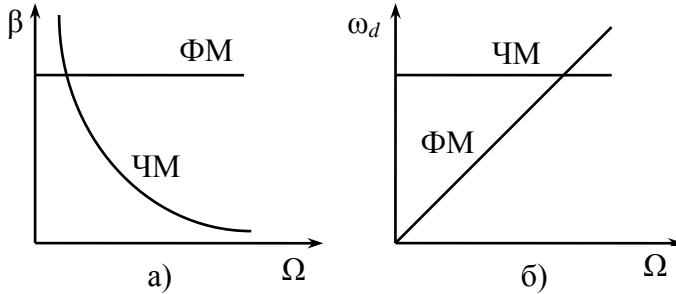


Рисунок 7.18 – Залежності індексу модуляції (а) і девіації частоти (б) від частоти гармонічного модульовального сигналу при ФМ і ЧМ

Для розрахунку спектру сигналу з гармонічною КМ представимо сигнал у вигляді дійсної частини комплексної експоненти:

$$\begin{aligned}
 s_{\text{КМ}}(t) &= \text{Re} \left\{ A \exp \left[j\omega_0 t + j\varphi_0 + j\beta \sin(\Omega t + \Phi_0) \right] \right\} = \\
 &= \text{Re} \left\{ A \exp \left[j\omega_0 t + j\varphi_0 \right] \exp \left[j\beta \sin(\Omega t + \Phi_0) \right] \right\}.
 \end{aligned}
 \tag{7.31}$$

Далі скористаємось поданням виразу $\exp(j\beta \sin x)$ у вигляді ряду Фур'є:

$$\exp(j\beta \sin x) = \sum_{k=-\infty}^{\infty} J_k(\beta) \exp(jkx),$$

де $J_k(\beta)$ – функція Бесселя першого роду порядку k від аргументу β . Підстановка цього співвідношення в (7.31), дає

$$s_{\text{КМ}}(t) = \text{Re} \left\{ A \exp(j\omega_0 t + j\varphi_0) \sum_{k=-\infty}^{\infty} J_k(\beta) \exp[jk(\Omega t + \Phi_0)] \right\} =$$

$$= A \sum_{k=-\infty}^{\infty} J_k(\beta) \cos[(\omega_0 + k\Omega)t + \varphi_0 + k\Phi_0].$$

Як видно, спектр сигналу містить нескінченну кількість складових з частотами $\omega_0 + k\Omega$, $k = 0, \pm 1, \pm 2, \dots$. Амплітуда k -ї складової дорівнює $A J_k(\beta)$, тобто пропорційна функції Бесселя k -го порядку, аргументом якої є індекс модуляції β . В теорії функцій Бесселя також доводиться, що функції з додатними і від'ємними індексами пов'язані співвідношенням

$$J_{-k}(\beta) = (-1)^k J_k(\beta).$$

Тому амплітудний спектр сигналу з КМ симетричний відносно частоти-носія. Функції Бесселя мають коливальний характер із поступовим спаданням (рис. 7.19).

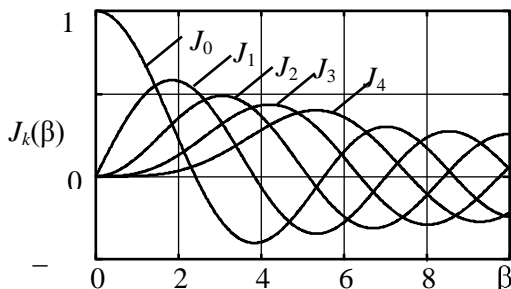
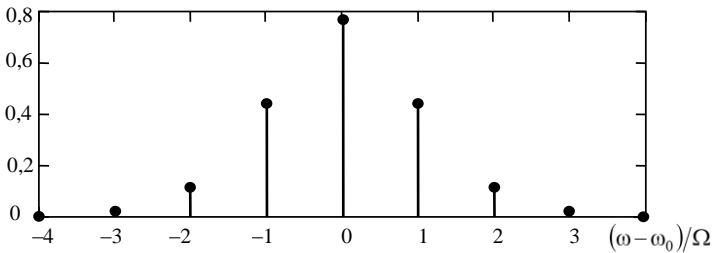
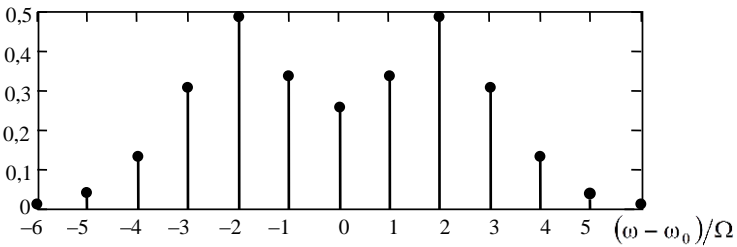


Рисунок 7.19 – Приклади функцій Бесселя

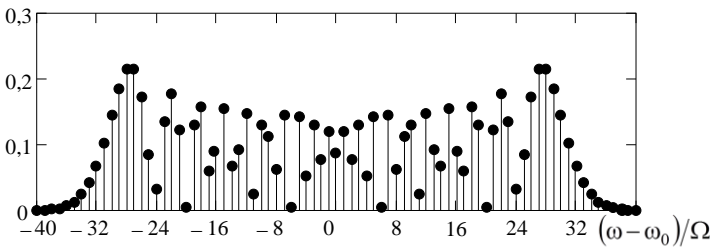
Завдяки цьому, в спектрі КМ-коливання при віддаленні від частоти-носія амплітуди складових зменшуються немонотонно. Ці особливості ілюструються прикладами амплітудних спектрів сигналів з гармонічною КМ при індексах модуляції $\beta = 1, 3$ і 30 (рис. 7.20). При побудові спектрів вважалось, що $\omega_0 \gg \Omega$.



а)



б)



в)

Рисунок 7.20 – Амплітудні спектри сигналів з гармонічною КМ при індексах модуляції 1 (а), 3 (б) і 30 (в)

Таким чином, навіть при гармонічній модуляції спектр КМ-сигналу має необмежений за шириною лінійчатий спектр, сусідні складові якого відстають одна від одної на величину Ω . На практиці при достатньо великих значеннях індексу модуляції β потужність складових, номер яких перевищує β , значно менша за потужність складових більш наближених до ω_0 (рис. 7.20, в). За цієї умови ефективну ширину спектру

сигналу з гармонічною КМ можна наближено обчислити за формулами $\Delta\omega_{\text{ef}} \approx 2\beta\Omega$ або $\Delta\omega_{\text{ef}} \approx 2\omega_d$.

Окремо розглянемо випадок, коли індекс модуляції малий ($\beta \ll 1$) і можна застосувати наближення:

$$\cos(\beta \sin(\Omega t + \Phi_0)) \approx 1;$$

$$\sin(\beta \sin(\Omega t + \Phi_0)) \approx \beta \sin(\Omega t + \Phi_0).$$

Спочатку, застосувавши тригонометричні перетворення, представимо (7.30) у вигляді

$$s_{\text{KM}}(t) = A \cos(\omega_0 t + \varphi_0) \cos(\beta \sin(\Omega t + \Phi_0)) - A \sin(\omega_0 t + \varphi_0) \sin(\beta \sin(\Omega t + \Phi_0)).$$

Далі з урахуванням представлених вище наближень отримаємо

$$s_{\text{KM}}(t) \approx A \cos(\omega_0 t + \varphi_0) - A \sin(\omega_0 t + \varphi_0) \beta \sin(\Omega t + \Phi_0).$$

Якщо подати останній доданок у вигляді піврізниці косинусів, отримаємо представлення сигналу у вигляді суми трьох складових

$$s_{\text{KM}}(t) \approx A \cos(\omega_0 t + \varphi_0) + \frac{A\beta}{2} \cos[(\omega_0 + \Omega)t + \varphi_0 + \Phi_0] - \frac{A\beta}{2} \cos[(\omega_0 - \Omega)t + \varphi_0 - \Phi_0]. \quad (7.31)$$

Представлення сигналу з КМ (7.31) майже збігається з (7.27) для АМ сигналу з гармонічною модуляцією. Три складові мають однакові частоти, їх амплітуди розраховуються аналогічно (тільки замість коефіцієнта амплітудної модуляції m у формулі фігурує індекс кутової модуляції β). Однак, принциповою відмінністю є наявність знака «мінус» перед одним з доданків, які відповідають бічним частотам. Саме вона перетворює амплітудну модуляцію в кутову.

Амплітудний і фазовий спектри сигналу з КМ при малому індексі модуляції (її ще називають вузькосмуговою ЧМ)

показані на рис. 7.21 (пунктирна лінія демонструє фазу верхньої бічної частоти для випадку АМ).

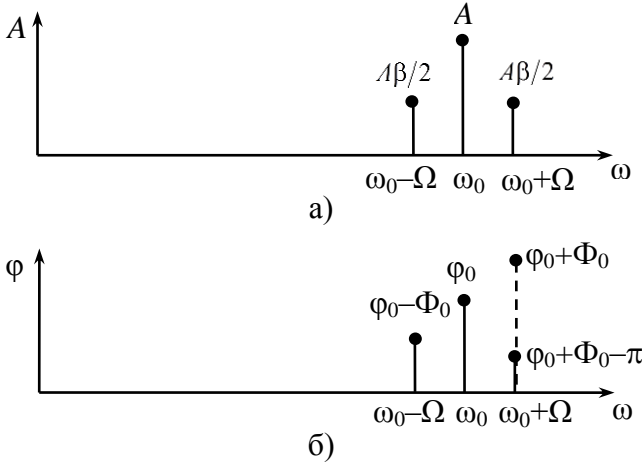


Рисунок 7.21 – Амплітудний і фазовий спектри сигналу з гармонічною КМ при $\beta \ll 1$

На рис. 7.22 представлено векторну діаграму сигналу з КМ для двох моментів часу (для більшої наочності при побудові обрано досить велике значення індексу модуляції β). Вона є аналогічною випадку гармонічної АМ (рис. 7.12), але один з векторів бічної частоти різниться лише знаком перед одним з доданків.

Як видно з рис. 7.22, наявність від'ємного знаку у коливання бічної частоти призводить до того, що амплітуда сигналу з гармонічною КМ при $\beta \ll 1$ є практично незмінною, а змінюється лише його фаза. Таким чином, для перетворення сигналу з гармонічною АМ в сигнал з гармонічною КМ, достатньо змінити на 180° початкову фазу однієї з бічних частот ($\beta \ll 1$).

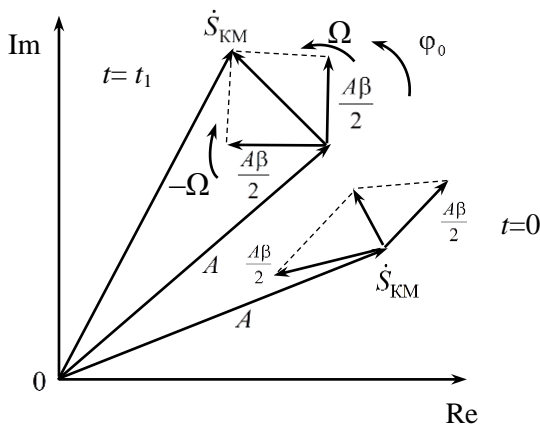


Рисунок 7.22 – Векторна діаграма сигналу з гармонічною КМ при $\beta \ll 1$ для двох моментів часу

Судячи з діаграм рис. 7.12 і рис. 7.22, іншим способом перетворити АМ в ЧМ є зміна фази складової частоти-носія на 90° .

Саме за таким принципом функціонує, так званий, модулятор Армстронга (рис. 7.23) [17]. Тут на один вхід суматора прикладається сигнал u_1 , що надходить з балансного модулятора БМ. На другий вхід суматора подається немодульований сигнал u_2 з виходу фазообертача, який змінює фазу гармонічного коливання-носія на 90° в бік запізнювання. Таким чином, сигнал на виході даного модулятора

$$u_{\text{вих}}(t) = U_{m1} s(t) \cos \omega_0 t + U_{m2} \sin \omega_0 t ,$$

де U_{m1} та U_{m2} – деякі постійні амплітуди.

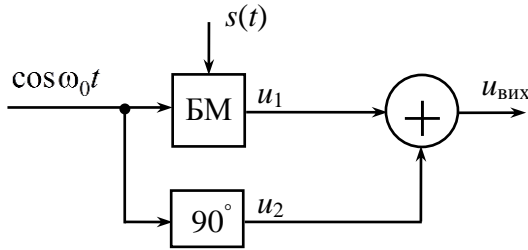


Рисунок 7.23 – Модулятор Армстронга

Не складно показати, що тангенс фази такого сигналу може бути обчислений як $\operatorname{tg}\varphi(t) = U_{m1}s(t)/U_{m2}$, при цьому амплітуда $A(t) = \sqrt{[U_{m1}s(t)]^2 + U_{m2}^2}$ є змінною у часі. Для того, щоб фаза $\varphi(t)$ лінійно залежала від модульовального сигналу $s(t)$, а амплітуда була практично постійною необхідно виконання умови $U_{m2} \gg U_{m1}$, тобто модулятор Армстронга має працювати при малих індексах модуляції і відповідно малої девіації частоти. Для збільшення девіації частоти можуть додатково застосовуватись множники частоти.

Отже, залежно від значення індексу модуляції β можна навести дві формули для розрахунку ефективної ширини спектру КМ-сигналу:

при $\beta \ll 1$, як і у випадку АМ сигналу, ширина спектру дорівнює подвоєній частоті модулюючого сигналу $\Delta\omega_{\text{еф}} = 2\Omega$;

при $\beta \gg 1$ ширина спектру дорівнює подвоєній девіації частоти: $\Delta\omega_{\text{еф}} = 2\omega_d$.

На відміну від АМ, у разі кутової модуляції отримати аналітичний вираз для спектра при довільному модульовальному сигналі не вдається. Навіть для модульовального сигналу з суми гармонік з частотами Ω_1, Ω_2 при малих парціальних індексах модуляції β_1, β_2 в спектрі

модульованого сигналу спостерігаються складові не тільки з частотами $\omega_0 \pm \Omega_1, \omega_0 \pm \Omega_2, \omega_0 \pm 2\Omega_1, \omega_0 \pm 2\Omega_2$, але й складові з комбінаційними частотами $\omega_0 \pm \Omega_1 \pm \Omega_2$ з чотирма можливими сполученнями знаків перед Ω_1, Ω_2 .

В загальному випадку, коли КМ здійснюється групою низькочастотних коливань з частотами $\Omega_1, \Omega_2, \dots, \Omega_N$ і парціальними індексами модуляції $\beta_1, \beta_2, \dots, \beta_N$ відповідно, спектральне представлення сигналу має вигляд [17]

$$s_{\text{КМ}}(t) = A \sum_{k=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \dots \sum_{k=-\infty}^{\infty} J_k(\beta_1) J_k(\beta_2) \dots J_k(\beta_N) \times \\ \times \cos(\omega_0 + k_1\Omega_1 + k_2\Omega_2 + \dots + k_N\Omega_N)t.$$

7.2.3. Модульовані сигнали з кількома змінними параметрами

Як вже згадувалось раніше, існують види модуляції, коли одночасно змінюються кілька параметрів коливання-носія, наприклад, одночасно можуть змінюватись амплітуда і початкова фаза гармонічного коливання

$$s(t) = A(t) \cos(\omega_0 t + \varphi(t)).$$

Таку модуляцію можна назвати амплітудно-фазовою і застосувати для одночасного передавання двох сигналів.

Однак модульовальні сигнали в цьому випадку виявляються «нерівноправними», оскільки вони модулюють істотно різні параметри коливання-носія. Можна зробити ситуацію більш «симетричною», якщо розкрити косинус суми

$$s(t) = A(t) \cos[\varphi(t)] \cos \omega_0 t - A(t) \sin[\varphi(t)] \sin \omega_0 t.$$

В такому вигляді сигнал є сумою двох АМ коливань. Їх коливання-носії $\cos \omega_0 t$ і $\sin \omega_0 t$ мають зсув фази 90° по відношенню одна до одної, а от амплітудні функції дорівнюють $A(t) \cos[\varphi(t)]$ і $-A(t) \sin[\varphi(t)]$. Позначивши їх

як $a(t)$ і $b(t)$, замість змінних амплітуди і початкової фази отримаємо нову пару модулювальних сигналів і відповідно подання розглянутого сигналу

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

Таке представлення називається квадратурним (*англ.* quadrature), а даний спосіб модуляції – *квадратурною модуляцією* (КАМ, англійський термін – quadrature amplitude modulation, QAM). Модулювальні сигнали при цьому виявляються повністю рівноправними.

Оскільки КАМ сигнал є сумою двох АМ сигналів, його *спектр*, скориставшись загальною формулою для спектру АМ сигналу, можна записати як суму чотирьох бічних смуг (див. гл. 2)

$$\begin{aligned} \dot{S}(\omega) = & \frac{1}{2} \dot{A}(\omega + \omega_0) + \frac{1}{2} \dot{A}(\omega - \omega_0) - \\ & - \frac{1}{2} j \dot{B}(\omega + \omega_0) + \frac{1}{2} j \dot{B}(\omega - \omega_0). \end{aligned}$$

Оскільки спектри модулювальних сигналів $a(t)$ і $b(t)$ зазвичай займають одну й ту саму смугу частот, тоді вони будуть перекриватись і після зсуву в область частоти-носія. Однак, розділити квадратурні складові при прийомі сигналу можливо, оскільки синусне коливання-носії додатково множиться на $\pm j$.

Для передачі цифрової інформації різноманітними каналами зв'язку у цифровому вигляді використовуються способи модуляції, які забезпечують передавання послідовності символів (*англ.* symbol), кожен з яких є цілим числом з деякої фіксованої кінцевої множини. Ці символи надходять від джерела інформації з частотою $f_T = 1/T$, де T – період їх слідування. Ця частота називається символною швидкістю (*англ.* symbol rate). На практиці часто використовується двійкова (*англ.* binary) послідовність

символів, коли кожне з чисел може мати одне з двох значень – 0 або 1.

Послідовність переданих символів є дискретним сигналом. Оскільки символи приймають значення з кінцевої множини, цей сигнал фактично є й квантованим. Тому, відповідно до визначень, введених в главі 1, його можна назвати цифровим сигналом. Далі розглянемо питання, пов'язані з перетворенням цього цифрового сигналу в аналоговий модульований сигнал.

Типовий підхід до здійснення *передавання дискретної послідовності символів* полягає в наступному. Кожному з можливих значень символу зіставляється деякий набір параметрів коливання-носія. Ці параметри підтримуються постійними протягом інтервалу T , тобто до приходу наступного символу. Фактично це означає перетворення послідовності чисел $\{n_k\}$ в ступінчастий сигнал $s_n(t)$ з використанням кусково-неперервної інтерполяції

$$s_n(t) = f(n_k), \quad kT \leq t < (k+1)T,$$

де f – деяка функція перетворення. Отриманий сигнал $s_n(t)$ далі використовується звичайним способом як модульовальний.

Спосіб модуляції, коли параметри коливання-носія змінюються стрибкоподібно, *називається маніпуляцією* (англ. keying). Залежно від того, які саме параметри змінюються, розрізняють амплітудну (АМн), фазову (ФМн), частотну (ЧМн) і квадратурну (КАМн) маніпуляції. Крім того, при передаванні цифрової інформації може використовуватись коливання-носій відмінне за формою від гармонічного. Так, при використанні в якості коливання-носія послідовності прямокутних імпульсів можливі амплітудно-імпульсна (АІМ), широтно-імпульсну (ШІМ) (див. підрозділ 5.1) і фазово-імпульсну (ФІМ, англ. PPM – Pulse Position Modulation) види модуляції, де ШІМ і ФІМ є різновидами часово-імпульсної (ЧІМ, англ. PTM – Pulse Time Modulation).

Найпростіші приклади формування двохпозиційної (бінарної) маніпуляції представлені на рис. 7.24.

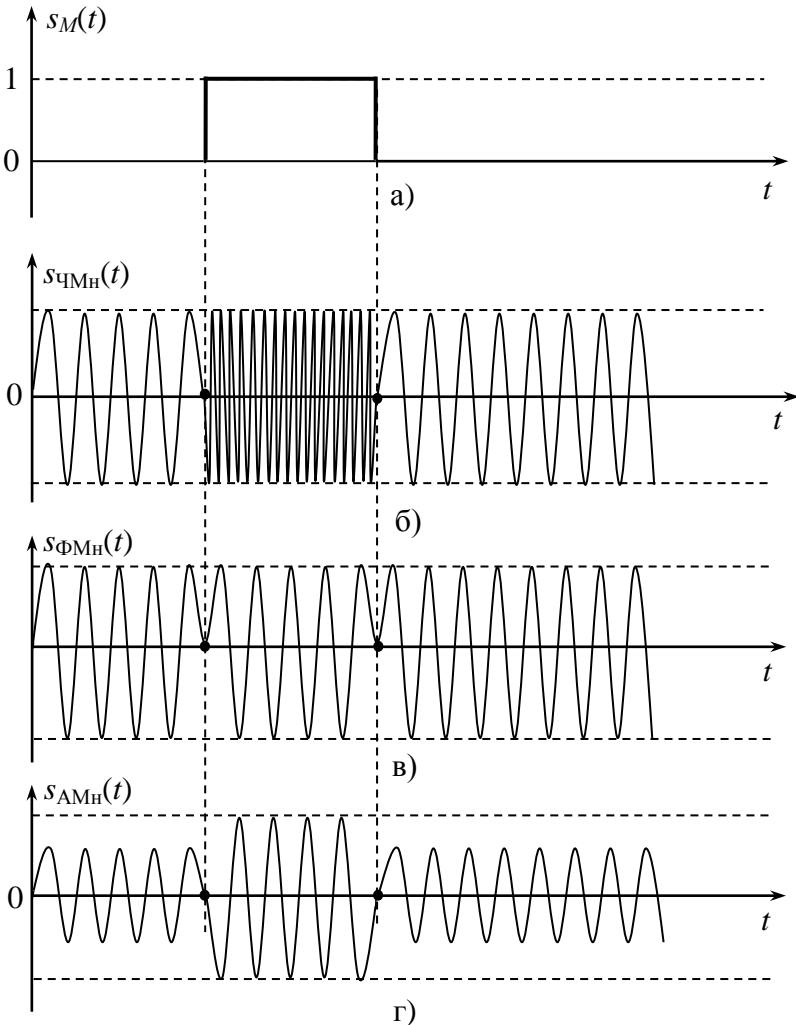


Рисунок 7.24 – Приклади маніпульованих сигналів:
 а) модульований бінарний сигнал; б) частотно-маніпульований; в) фазо-маніпульований;
 г) амплітудно-маніпульований

При частотній маніпуляції (англійський термін – frequency shift keying, FSK) кожному можливому значенню переданого символу протягом кожного символного інтервалу зіставляється своя частота (рис. 7.24, б – випадок передавання символів 0 і 1).

Прийом ЧМн сигналу зазвичай здійснюється кореляційним методом, тобто обчисленням взаємної кореляції прийнятого сигналу з коливаннями-зразками (опорними сигналами), які є гармонічними коливаннями з використовуваними для маніпуляції частотами. На виході обирається той символ, частота якого виявляється максимально корельованою з вхідним сигналом [11].

Кореляційний прийом може бути когерентним або некогерентним. Когерентний алгоритм може використовуватись, якщо відома початкова фаза коливання. Опорними сигналами при цьому є дійсні синусоїди з потрібними частотами і початковими фазами. Для прийняття рішення про прийнятий символ порівнюються отримані результати обчислення кореляційних сум.

Якщо початкові фази коливань невідомі, застосовують некогерентний кореляційний прийом. Модуль обчислюваної при цьому комплексної ВКФ не залежить від початкових фаз сигналів, проте завадостійкість такого алгоритму дещо гірша, ніж у випадку когерентного прийому.

Для підвищення завадостійкості ЧМн послілки, що відповідають різним символам, бажано обрати некорельованими. Якщо початкові фази послілок є нульовими, ЧМн сигнали для символів 0 і 1 можна записати так:

$$s_0(t) = A \cos(\omega_0 t), \quad s_1(t) = A \cos(\omega_1 t), \quad 0 \leq t \leq T.$$

Їх ВКФ при нульовому зсувові у часі дорівнює

$$B_{10}(0) = \int_0^T s_0(t)s_1(t) dt = A^2 \int_0^T \cos(\omega_0 t) \cos(\omega_1 t) dt = \\ = \frac{A^2 \sin(\omega_1 + \omega_0)T}{2(\omega_1 + \omega_0)} + \frac{A^2 \sin(\omega_1 - \omega_0)T}{2(\omega_1 - \omega_0)}.$$

Якщо $(\omega_1 + \omega_0)T \gg 1$, перший доданок значно менше другого, тому їм можна знехтувати, відповідно

$$B_{10}(0) \approx \frac{A^2 \sin(\omega_1 - \omega_0)T}{2(\omega_1 - \omega_0)}.$$

Її значення дорівнює нулю при $(\omega_1 - \omega_0)T = \pi k$, де k – ціле не рівне нулю число. Таким чином, мінімальне значення різниці між частотами маніпуляції, при якому посліди, що відповідають різним символам, виявляються некорельованими, становить

$$\Delta\omega_{\min} = \frac{\pi}{T}, \Delta f_{\min} = \frac{1}{2T} = \frac{f_T}{2}.$$

Бінарна ЧМн, частоти якої обрані згідно цьому співвідношенню, має назву *мінімальної частотної маніпуляції* (МЧМн, англійський термін – minimum shift keying, MSK).

При *фазовій маніпуляції* (ФМн; англійський термін – phase shift keying, PSK) стрибкоподібно змінюється фаза коливання-носія. Графік сигналу з двохпозиційною фазовою маніпуляцією представлено на рис. 7.24, в.

На практиці фазова маніпуляція використовується при невеликому числі можливих значень початкової фази – як правило, 2, 4 або 8. Крім того, при прийомі сигналу складно виміряти абсолютне значення початкової фази; значно простіше визначити відносний фазовий зсув між двома сусідніми символами. Тому зазвичай використовується фазорізницева маніпуляція (синоніми – диференціальна фазова маніпуляція, відносна фазова маніпуляція; англійський термін – differential phase shift keying, DPSK).

Як буде показано далі, *амплітудна маніпуляція* (АМн;

англійський термін – amplitude shift keying, ASK), при якій стрибкоподібно змінюється амплітуда коливання-носія, як і ФМн, є окремими випадками квадратурної маніпуляції. Приклад АМн сигналу наведено на рис. 7.24, г.

При *квадратурній маніпуляції* (КАМн; англійський термін – quadrature amplitude shift keying, QASK) кожному з можливих значень дискретного символу C_k , ставиться у відповідність пара величин – амплітуди синфазної і квадратурної складових або, що еквівалентно, амплітуда і початкова фаза коливання-носія, тобто

$$C_k \rightarrow (a_k, b_k), \quad s(t) = a_k \cos(\omega_0 t) + b_k \sin(\omega_0 t), \quad kT \leq t \leq (k+1)T,$$

або

$$C_k \rightarrow (A_k, \varphi_k), \quad s(t) = A_k \cos(\omega_0 t + \varphi_k), \quad kT \leq t \leq (k+1)T.$$

Параметри аналогового коливання, зіставлені дискретному символу C_k , зручно представляти комплексними числами в алгебраїчній $(a_k + jb_k)$ або експоненціальній $A_k \exp(j\varphi_k)$ формі. Сукупність цих комплексних чисел для всіх можливих значень дискретного символу називається сигнальним сузір'ям (англ. constellation) або «сигнально-кодуючою конструкцією» (СКК). Далі застосовується термін «сузір'я».

При поданні дискретного символу комплексним числом \dot{C}_k сигнал з КАМн можна записати наступним чином:

$$s(t) = \operatorname{Re}(\dot{C}_k \exp(-j\omega_0 t)), \quad kT \leq t \leq (k+1)T.$$

На практиці використовуються сузір'я, що містять від чотирьох точок до кількох тисяч. На рис. 7.25 показані деякі сузір'я, використовувані в модемах, призначених для передавання даних телефонними лініями.

Графік сигналу з КАМн є не дуже наочним через змішаний (амплітудно-фазовий) характер модуляції та невеликі зміни амплітуди і фази при переході від символу до символу.

При квадратурній маніпуляції одночасно змінюються і амплітуда і початкова фаза коливання-носія, тому амплітудна і фазова маніпуляція є окремими випадками квадратурної – потрібно лише використовувати відповідні сузір'я. Приклади цих сузір'їв показані на рис. 7.26 [11].

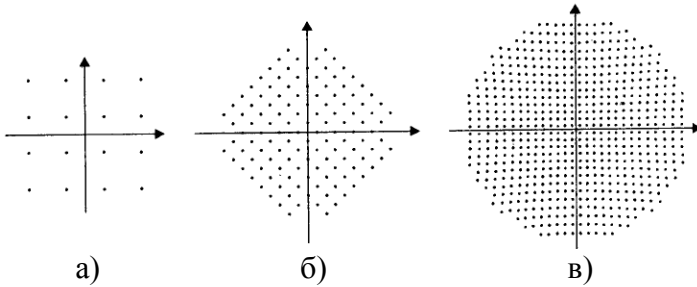


Рисунок 7.25 – Приклади сузір'їв при квадратурній маніпуляції [11]:

а) 16 точок (протокол V.32, швидкість 9600 біт / с); б) 128 точок (протокол V.32bis, швидкість 14400 біт/с); в) 640 точок (протокол V.34, швидкість 28800 біт/с)

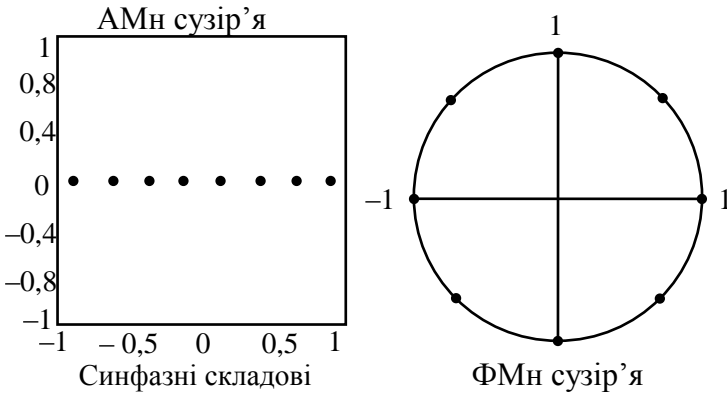


Рисунок 7.26 – Сузір'я, що відповідають 8-позиційній маніпуляції: а) амплітудній; б) фазовій

Для порівняльної оцінки перешкодостійкості сузір'їв КАМн, ФМн і КАМн сигналів при однаковій кількості символів N може бути застосоване відношення мінімальної відстані між точками на комплексній площині до середньоквадратичного рівня сигналу

$$\sigma = \sqrt{\frac{1}{N} \sum_{k=1}^{N-1} |\dot{C}_k|^2}.$$

За цим показником найбільш перешкодостійким виявляється КАМн сигнал, що є цілком очікуваним, оскільки, на відміну від ФМн і КАМн сигналів, зміна двох параметрів дозволяє більш рівномірно розподілити їх в комплексній площині.

7.3. Демодуляція сигналів

Для виділення на приймальному боці з високочастотного модульованого коливання корисного (модульовального) сигналу потрібно виконати обернену до модулювання операцію демодуляції (*англ.* demodulation) або детектування. Детектори зазвичай будуються шляхом об'єднання нелінійних елементів з відповідними лінійними частотно-селективними колами. Розглянемо теорію роботи і приклади побудови найважливіших видів детекторів [17].

Принцип детектування АМ сигналів розглянемо на прикладі ідеального детектора АМ коливання. Якщо на його вході діє АМ коливання $u_{\text{вх}}(t) = U_{\text{твх}}(1 + M \cos \Omega t) \cos \omega_0 t$, тоді на виході буде отримано пропорційний переданому повідомленню низькочастотний сигнал $u_{\text{вих}}(t) = U_{\text{твих}} \cos \Omega t$. Ефективність роботи детектора прийнято оцінювати коефіцієнтом детектування

$$k_{\text{дет}} = U_{\text{твих}} / MU_{\text{твх}}, \quad (7.32)$$

який є відношенням амплітуди низькочастотного сигналу на виході до «розмаху» зміни амплітуди високочастотного

сигналу на вході.

Детектування може бути виконане, якщо подати АМ сигнал на безінерційний нелінійний елемент і здійснити подальшу фільтрацію низькочастотних складових спектра. Прикладом такого сполучення є схема, так званого, колекторного детектора. Він представляє собою транзисторний пристрій з навантаженням у вигляді паралельного RC -кола [17] (рис. 7.27).

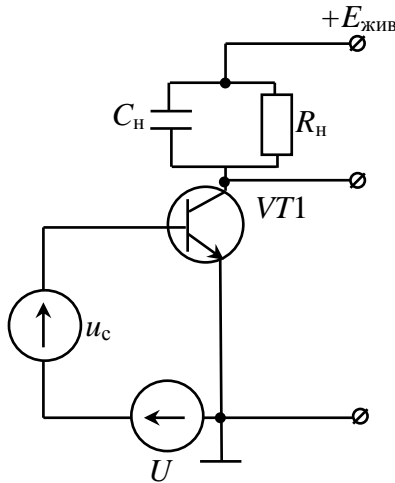


Рисунок 7.27 – Схема колекторного детектора

Для того, щоб навантажувальне коло було частотним фільтром, який придушує високочастотні спектральні складові, потрібне виконання нерівностей

$$1/(\omega_0 C_H) \ll R_H, \quad 1/(\Omega C_H) \gg R_H. \quad (7.33)$$

Нерівності (7.33) забезпечують для сигналу з частотою модуляції Ω практично резистивне навантаження детектора R_H , тоді як модуль опору навантаження і відповідно коефіцієнт передачі системи на частоті ω_0 нехтовно малий.

Нехай вхідна напруга на базі транзистора

$$u_{BE}(t) = U_0 + U_{mвх} (1 + M \cos \Omega t) \cos \omega_0 t ,$$

причому амплітуда $U_{mвх}$ досить велика для того, щоб можна було скористатись кусково-лінійною апроксимацією ВАХ нелінійного елемента. Припустимо також для простоти, що $U_0 = U_n$ і кут відсічки струму $\theta = 90^\circ$ незалежно від зміни в часі амплітуди вхідного сигналу. Процеси в колекторному детекторі ілюструються графіками рис. 7.28.

Послідовність імпульсів колекторного струму виявляється промодульованою по амплітуді; нульова складова струму повільно (з частотою Ω) змінюється за часом, причому

$$\begin{aligned} I_{OK}(t) &= SU_{mвх}(1 + M \cos \Omega t) \gamma_0(90^\circ) = \\ &= 0,318SU_{mвх}(1 + M \cos \Omega t). \end{aligned}$$

Вихідна напруга детектора

$$u_{ввх}(t) = E_{пит} - I_{OK}(t)R_n = 0,318 \cdot SR_n U_{mвх}(1 + M \cos \Omega t), \quad (7.34)$$

звідки коефіцієнт детектування

$$k_{дет} = 0,318SR_n. \quad (7.35)$$

Як видно з (7.34) і (7.35), амплітуди сигналів на вході і на виході пов'язані прямою пропорційністю. Тому такий режим роботи детектора, називається лінійним. Його відмінна риса – відсутність спотворень переданого повідомлення.

Квадратичне детектування. Розглянемо окремо важливий для додатків випадок детектування слабких сигналів, коли вольт-амперна характеристика повинна бути апроксимована степеневою залежністю виду

$$i(u) = a_0 + a_1(u_{вх} - U_0) + a_2(u_{вх} - U_0)^2 + \dots \quad (7.36)$$

Обмежимося лише виписаними в (7.36) членами і припустимо, що на детектор подано напругу АМ сигналу разом з постійним зсувом

$$u_{вх}(t) = U_0 + U_{mвх}(1 + M \cos \Omega t) \cos \omega_0 t. \quad (7.37)$$

Підставивши (7.37) в (7.36), виявимо серед різноманітних комбінаційних коливань, які наявні в струмі i_k , низькочастотну складову

$$i_{\text{НЧ}}(t) = a_2 U_{\text{мвх}}^2 M \cos \Omega t + \frac{1}{4} a_2 U_{\text{мвх}}^2 M^2 \cos 2\Omega t. \quad (7.38)$$

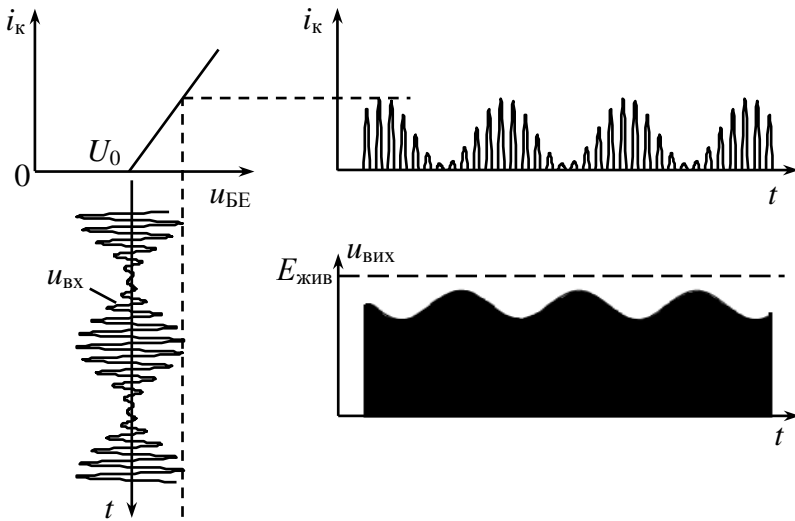


Рисунок 7.28 – Осцилограми струмів і напруг в колекторному детекторі

Завдяки фільтрувальній дії навантажувального RC -кола вихідний сигнал визначається саме цим струмом:

$$u_{\text{вих}}(t) = E_{\text{пит}} - a_2 R_{\text{н}} U_{\text{мвх}}^2 M \cos \Omega t - \frac{1}{4} a_2 R_{\text{н}} U_{\text{мвх}}^2 M^2 \cos 2\Omega t. \quad (7.39)$$

Корисний ефект детектування в (7.39) пропорційний величині $U_{\text{мвх}}^2$, тому детектування АМ сигналів з малими амплітудами є квадратичним. Наявність в (7.39) доданка, пропорційного $\cos 2\Omega t$ вказує, що квадратичне детектування супроводжується спотвореннями переданого повідомлення. В [17] введено коефіцієнт нелінійних спотворень $k_{\text{нл}}$, який є відношенням амплітуд вихідних коливань з частотами 2Ω і Ω , та з (7.39) знайдено, що $k_{\text{нл}} = M/4$. Нелінійні спотворення

виявляються значними при глибокій амплітудній модуляції на вході. Тому в радіоприймальних пристроях бажано, щоб амплітуда коливання-носія АМ сигналу, що подається на детектор, становила кілька вольт. При цьому реалізується режим лінійного детектування і нелінійні спотворення є незначними.

На практиці широко застосовується також *діодний детектор АМ сигналів*, який особливо придатний для роботи з сигналами великого рівня. Такий детектор утворений послідовним з'єднанням діода і паралельного RC -кола, яке виконує роль частотного фільтра. Параметри RC -кола обирають за умовами (7.33).

Принципова схема діодного амплітудного детектора з послідовним включенням діода та навантаження зображена на рис. 7.29.

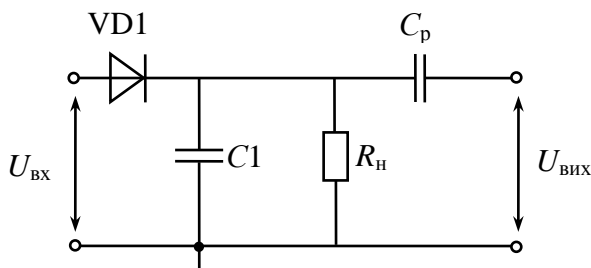


Рисунок 7.29 – Діодний амплітудний детектор

Напруга вхідного сигналу проміжної частоти ділиться між ємністю діода і ємністю навантаження, пропорційно величинам їх опорів. Для детектування необхідно щоб більша частина вхідної напруги була прикладена до діода. Для цього необхідно, щоб ємність діода була менше ємності навантаження.

З огляду на нелінійність ВАХ діода через нього протікає періодично пульсуючий струм, спектр якого може бути представлений рядом Фур'є.

Постійна складова струму протікає через резистор R_n , а високочастотні складові струму замикаються через ємність навантаження на корпус.

Постійна напруга на резисторі пропорційна напрузі обвідної вхідного сигналу і є корисним сигналом на виході амплітудного детектора.

Далі вважатимемо, що діод має кусково-лінійну ВАХ з нульовою початковою напругою

$$i(u) = \begin{cases} 0, & u < 0, \\ Su, & u \geq 0. \end{cases}$$

Для нормальної роботи детектора необхідно, щоб опір резистора навантаження R_n значно перевищував опір діода в прямому напрямку, тобто $SR_n \gg 1$. Нехай на вхід детектора поданий немодульований гармонічний сигнал $u_{вх}(t) = U_{мвх} \cos \omega_0 t$. Конденсатор заряджається через відкритий діод набагато швидше, ніж розряджається через високоомний резистор навантаження. Тому осцилограма вихідного сигналу являє собою пилкоподібну криву з малою відносною висотою зубців. Середній рівень вихідної напруги близький до амплітуди вхідного сигналу. Таким чином, діод більшу частину періоду виявляється запертим.

Нехтуємо зазначеною непостійністю вихідного сигналу і вважаємо, що $U_{вих}$ – постійна величина. Зауважимо далі, що напруга $U_{вих}$ прикладена до діоду в зворотному напрямку, і для нього є напругою зсуву $U_0 = -U_{вих}$. Коефіцієнт детектування даного пристрою

$$k_{дет} = U_{мвих} / MU_{мвх} = \cos \theta$$

може бути зроблений близьким до одиниці, оскільки $U_{вих} \approx U_{мвх}$, а це означає, що кут відсічки струму досить малий.

Кут відсічки знаходять із співвідношення

$$-U_0 = I_0 R_H = S U_{mвх} \gamma_0(\theta) R_H,$$

звідки слідує трансцендентне рівняння

$$\cos \theta = (SR_H / \pi) (\sin \theta - \theta \cos \theta)$$

або

$$\operatorname{tg} \theta - \theta = \pi / SR_H. \quad (7.40)$$

При $SR_H \gg 1$ корінь цього рівняння наближається до нуля, так що з (7.40) випливає формула для розрахунку коефіцієнта детектування

$$k_{\text{дет}} = \cos \left[\sqrt[3]{3\pi / SR_H} \right]. \quad (7.41)$$

Взаємодія сигналу і перешкоди в амплітудному детекторі. Припустимо, що на вході ідеального лінійного детектора АМ сигналу з відомим коефіцієнтом детектування $k_{\text{дет}}$ присутня сума корисного однотонального АМ колювання і немодульованого колювання перешкоди:

$$u_{\text{вх}}(t) = U_{mc} (1 + M \cos \Omega t) \cos \omega_0 t + U_{мп} \cos \omega_{п} t,$$

частоти ω_0 і $\omega_{п}$ в загальному випадку різні.

Сигнал на виході детектора пропорційний фізичній обвідній колювання $u_{\text{вх}}(t)$. Для його обчислення зручно скористатись поняттям спряженого сигналу $\hat{u}_{\text{вх}}(t)$, який матиме вигляд:

$$\hat{u}_{\text{вх}}(t) = U_{mc} (1 + M \cos \Omega t) \sin \omega_0 t + U_{мп} \sin \omega_{п} t,$$

звідки

$$\begin{aligned} U_{\text{вих}}(t) &= k_{\text{дет}} \sqrt{u_{\text{вх}}^2 + \hat{u}_{\text{вх}}^2} = k_{\text{дет}} U_{мп} [1 + (U_{mc} / U_{мп})^2 \times \\ &\times (1 + 2M \cos \Omega t + M^2 \cos^2 \Omega t) + \\ &+ 2(U_{mc} / U_{мп}) (1 + M \cos \Omega t) \cos(\omega_0 - \omega_{п}) t]^{1/2}. \end{aligned} \quad (7.42)$$

Розглянемо випадок, коли корисний сигнал значно слабкіший за перешкоду, тобто $U_{mc} / U_{мп} \ll 1$. Будемо цікавитись корисною складовою вихідного колювання, яка змінюється в часі пропорційно переданому повідомленню

Розкладаючи радикал, що входить в (7.42), в ряд за степенями малого параметра $U_{mc}/U_{мп}$ переконаємось, що дана спектральна складова створює на виході коливання

$$U_{пол}(t) = k_{дет} U_{mc} (U_{mc}/U_{мп}) M \cos \Omega t. \quad (7.43)$$

Видно, що із зростанням амплітуди перешкоди відбувається придушення корисного сигналу.

Шкідлива дія перешкоди проявляється також у тому, що на виході детектора може виникнути велика кількість комбінаційних коливань через нелінійну взаємодію сигналу і перешкоди.

Фазове детектування. Відомо багато схем фазових детекторів – пристроїв для демодуляції коливань з повною фазою $\psi(t) = \omega_0 t + \varphi(t)$, яка є модульованою по фазовому куту. Фазовий детектор виробляє напругу пропорційну різниці фаз вхідних коливань, одне з яких не модульоване є опорним (формується допоміжним зовнішнім джерелом), а інше модульоване – сигнальним. Робота таких детекторів заснована на нелінійній взаємодії цих коливань.

Нехай, наприклад, до нелінійного безінерційного двухполюсника з ВАХ виду $i(u) = a_0 + a_1 u + a_2 u^2$ прикладена сума двох напруг:

$$u(t) = u_1 + u_2 = U_{m1} \sin[\omega_0 t + \varphi(t)] + U_{m2} \cos \omega_0 t.$$

Через квадратичний доданок ВАХ в струмі буде присутня складова, яка описує нелінійну взаємодію коливань:

$$\begin{aligned} i_{вз}(t) &= 2a_2 U_{m1} U_{m2} \sin[\omega_0 t + \varphi(t)] \cos \omega_0 t = \\ &= 2a_2 U_{m1} U_{m2} \sin \varphi(t) + a_2 U_{m1} U_{m2} \sin[2\omega_0 t + \varphi(t)]. \end{aligned} \quad (7.44)$$

Другому доданку в останній частині формули (7.44) відповідає високочастотний сигнал з середньою частотою $2\omega_0$ який легко придушується лінійним ФНЧ (наприклад, RC-колом). Перший доданок в (7.44) описує низькочастотний струм

$$i_{нч}(t) = a_2 U_{m1} U_{m2} \sin \varphi(t) \approx a_2 U_{m1} U_{m2} \varphi(t), \quad (7.45)$$

який наближено пропорційний переданому повідомленню $\varphi(t)$ за умови, що девіація фази (індекс модуляції) сигналу на вході детектора досить мала.

При створенні фазових детекторів виникає необхідність забезпечувати жорстку стабілізацію фази опорного генератора.

Розглянемо одну зі схем фазового детектора, в якому для формування вихідної напруги використовуються обидві півхвилі сумарної напруги, такий фазовий детектор називають балансным (рис. 7.30) [34].

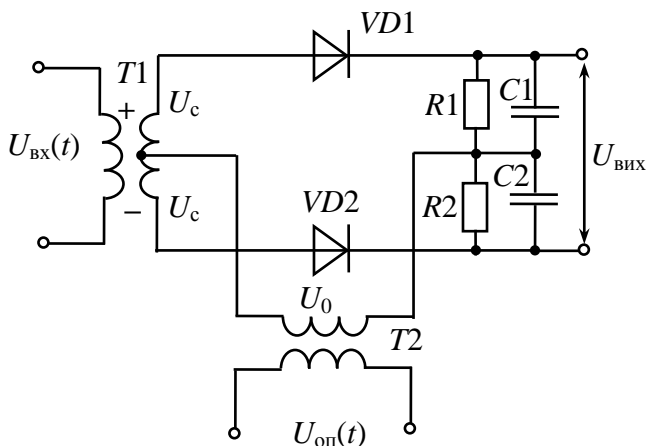


Рисунок 7.30 – Схема балансного фазового детектора

Він складається з двох діодних амплітудних детекторів включених по балансній схемі. На їх входи за допомогою резонансних трансформаторів подаються сигнальна і опорна напруги. При цьому одна з напруг, наприклад, сигнальна діє на перший і другий амплітудні детектори в протифазі, а інша – опорона, прикладена синфазно.

Згідно рис. 7.30 до діода $VD1$ прикладена напруга

$$u_{д1}(t) = U_0 \cos \omega_0 t + U_c \cos [\omega_0 t + \varphi(t)],$$

а до діода $VD2$

$$u_{д2}(t) = U_0 \cos \omega_0 t - U_c \cos [\omega_0 t + \varphi(t)].$$

Амплітуди цих напруг $U_{д1}$ і $U_{д2}$ легко знайти, побудувавши векторну діаграму (рис. 7.31).

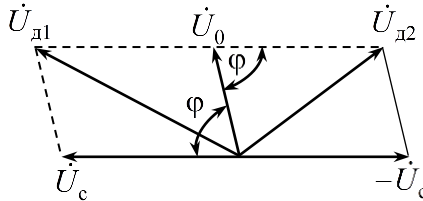


Рисунок 7.31 – Векторну діаграма напруг в баланському фазовому детекторі

З рис. 7.31 видно, що амплітуди сигналів на діодах:

$$U_{д1} = \sqrt{U_0^2 + U_c^2 + 2U_0U_c \cos \varphi};$$

$$U_{д2} = \sqrt{U_0^2 + U_c^2 - 2U_0U_c \cos \varphi}.$$

Тому сигнал на виході фазового детектора

$$U_{вих} = k_{дет} (U_{д1} - U_{д2}).$$

Після підстановки в цю формулу співвідношень для $U_{д1}$ і $U_{д2}$ та введення позначення $m = U_c / U_0$ отримаємо

$$U_{вих} = k_{дет} U_0 \left(\sqrt{1 + m^2 + 2m \cos \varphi} - \sqrt{1 + m^2 - 2m \cos \varphi} \right).$$

Практичне значення мають два випадки: $m = 1$; $m \ll 1$.

Для $m = 1$ отримаємо

$$U_{вих} = 2k_{дет} U_c \left(\cos \frac{\varphi}{2} - \sin \frac{\varphi}{2} \right).$$

У разі $m \ll 1$

$$U_{вих} = 2k_{дет} U_0 m \cos \varphi.$$

Амплітудно-фазові характеристики графічно

представлені на рисунку 7.32.

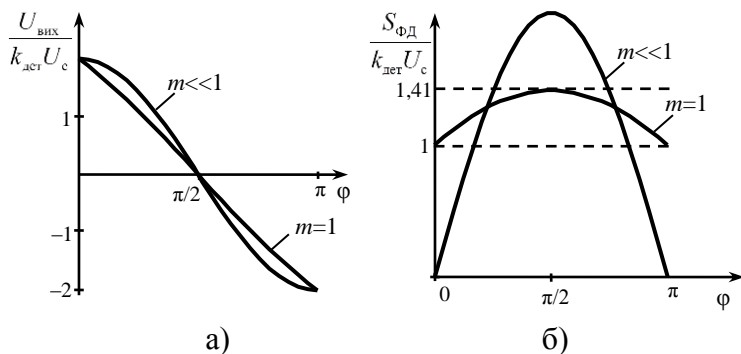


Рисунок 7.32 – Графіки амплітудно-фазових характеристик (а) та їх крутизни (б)

Крутизна амплітудно-фазової характеристики визначається наступними виразами:

для випадку $m = 1$

$$S_{\text{ФД}} = \left| \frac{dU_{\text{вих}}}{d\varphi} \right| = k_{\text{дет}} U_c \left(\cos \frac{\varphi}{2} + \sin \frac{\varphi}{2} \right);$$

для випадку $m \ll 1$

$$S_{\text{ФД}} = 2k_{\text{дет}} U_c \sin \varphi.$$

Амплітудно-фазова характеристика у випадку $m = 1$ відрізняється підвищеною лінійністю (рис. 7.32, а). Однак максимальна крутизна амплітудно-фазової характеристики вище при $m \ll 1$ (рис. 7.32, б). Амплітудно-фазова характеристика переходить через 0 при різниці фаз порівнюваних коливань рівній $\pi/2$, маючи при цьому максимальну крутизну, тому початковий фазовий зсув порівнюваних коливань зазвичай обирається рівним $\pi/2$.

Частотне детектування. При частотній модуляції, як відомо, корисне повідомлення пропорційне відхиленню миттєвої частоти сигналу від частоти-носія. Пристрій, напруга

на виході якого пропорційна відхиленню частоти від деякого опорного значення, називається частотним детектором.

Частотну модуляцію можна перетворити в неглибоку амплітудну модуляцію, якщо подати вхідний сигнал на лінійний частотний фільтр налаштований таким чином, щоб в розкладанні АЧХ

$$|K(j\omega)| = |K(j\omega_0)| + |K(j\omega_0)|' (\omega - \omega_0) + \dots$$

коефіцієнт $|K(j\omega_0)|'$ був відмінний від нуля. Тоді, якщо частота вхідного сигналу $\omega(t) = \omega_0 + \Delta\omega \cos \Omega t$, отримаємо на виході фільтра сигнал зі складною амплітудно-кутовою модуляцією (рис. 7.33). Миттєва амплітуда змінної складової цього сигналу матиме вигляд

$$U_{\text{мвих}}(t) = B_0 |K(j\omega_0)|' \Delta\omega \cos \Omega t, \quad (7.46)$$

де B_0 – постійний коефіцієнт. Тобто сигнал (7.46) повторює за формою передане повідомлення.

Остаточна обробка сигналу проводиться звичайним АМ детектором, включеним на виході фільтра.

Розглянутий спосіб частотного детектування має ряд недоліків – високі вимоги до якості обмеження можливої паразитної АМ на вході фільтра, а також недостатня лінійність характеристики детектування.

Якість роботи частотного детектора характеризується його амплітудно-частотною характеристикою (АЧХ), що представляє собою залежність вихідної напруги U_d від частоти вхідних коливань при постійній амплітуді (рис. 7.34). Робоча ділянка знаходиться між точками A та B .

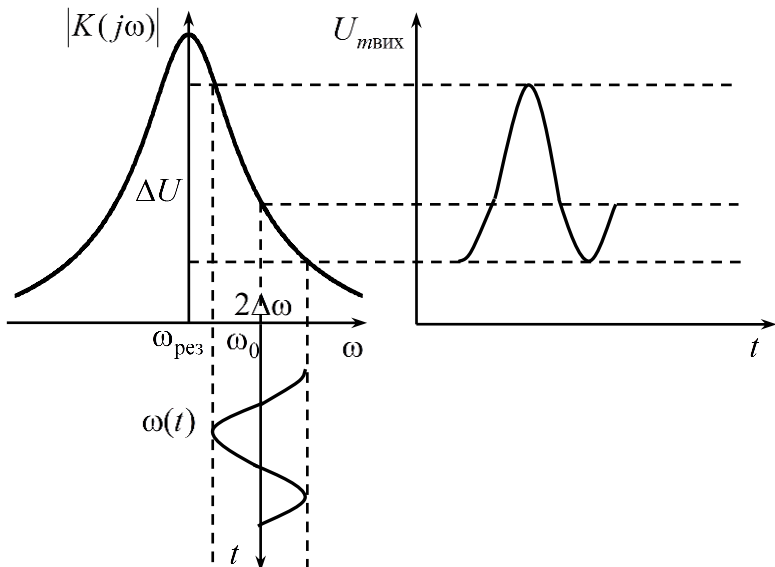


Рисунок 7.33 – Принцип перетворення зміни частоти сигналу в амплітудні зміни

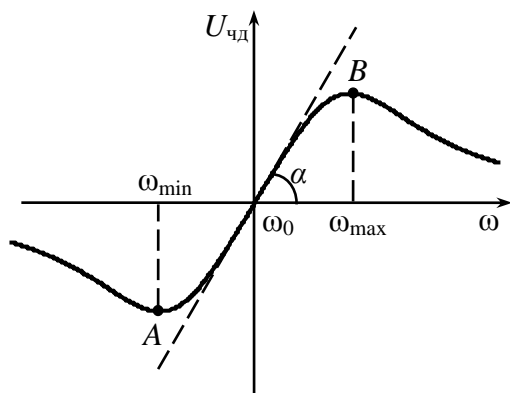


Рисунок 7.34 – АЧХ частотного детектора

Частота ω_0 при якій $U_{чд}=0$ називається перехідною частотою. За допомогою АЧХ частотного детектора

визначаються такі його параметри:

- крутизна АЧХ

$$S_{\text{чд}} = \left. \frac{dU_{\text{д}}}{d(\Delta\omega)} \right|_{\Delta\omega=0};$$

- смуга пропускання

$$P_{\text{чд}} = \omega_{\text{max}} - \omega_{\text{min}}.$$

Для отримання АЧХ подібних до представленої на рис. 7.34 частотні детектори зазвичай будуються за балансними схемами.

Схема частотного детектора із взаємно розузгодженими паралельними контурами зображена на рис 7.35 [35].

Контури $L3$, $C4$ і $L4$, $C5$ включені в колекторні кола підсилювачів та симетрично розузгоджені відносно частоти переходу. Вихідна напруга $U_{\text{чд}}$ дорівнює різниці напруг, що створюються струмами діодів $VD1$ та $VD2$ на опорах навантажень $R3$ та $R4$. Якщо частота сигналу на вході дорівнює ω_0 , тоді за рахунок симетрії у відхиленні частоти налаштування контурів сигнал $U_{\text{чд}}=0$ і АЧХ детектора наближається до показаної на рис. 7.34. Недоліком такого детектора, як видно з рис. 7.34, є порушення лінійності робочої ділянки при наближенні до точок A і B .

Для досягнення достатньо великої протяжності майже лінійної ділянки АЧХ в схемах з коливальними контурами необхідне їх ретельне налаштування та забезпечення відповідної добротності контурів.

Кращі результати забезпечують так звані частотно-фазові частотні детектори, в яких застосовується перетворення ЧМ сигналу в ФМ сигнал за допомогою лінійного частотно-селективного фільтра з подальшим фазовим детектуванням. При такому методі демодуляції фазочастотна характеристика селективного вузькосмугового кола в малій околиці частоти має вигляд

$$\varphi_K(\omega) = \varphi_K(\omega_0) - t_{\text{гр}}(\omega - \omega_0), \quad (7.46)$$

де $t_{\text{гр}}$ – час групової затримки.

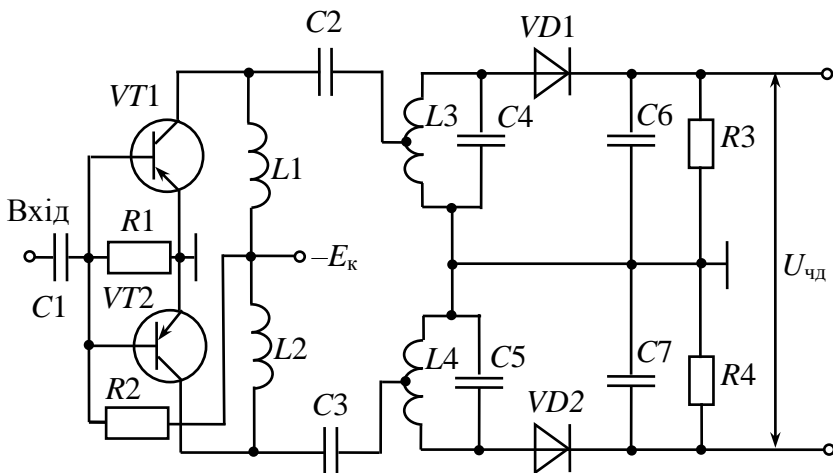


Рисунок 7.35 – Схема резонансного частотного детектора на розузгоджених контурах

Якщо $\omega(t) = \omega_0 + \Delta\omega \cos \Omega t$, тоді вузькосмуговий сигнал на виході фільтра має повну фазу

$$\psi(t) = \omega_0 t + \varphi_K(\omega_0) - \Delta\omega t_{\text{гр}} \cos \Omega t,$$

тобто справді є ФМ сигналом. Структурна схема частотно-фазового частотного детектора представлена на рис. 7.36, який може бути виконаний на інтегрованій схемі [28].

Як фазосувне коло ФЗК використовується коливальний контур, як фазовий детектор ФД – схему збігу СЗ. Нехай на вхід підсилювача-обмежувача амплітуди ПОА подається гармонічний сигнал $s(t)$, з виходу ПОА сигнал U_1 потрапляє на СЗ та на вхід ФЗК з фазочастотною характеристикою показаною на рис. 7.37, а [28].

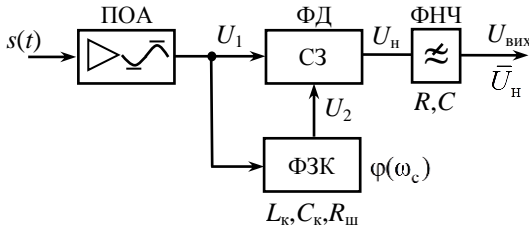


Рисунок 7.36 – Структурна схема частотно-фазового частотного детектора

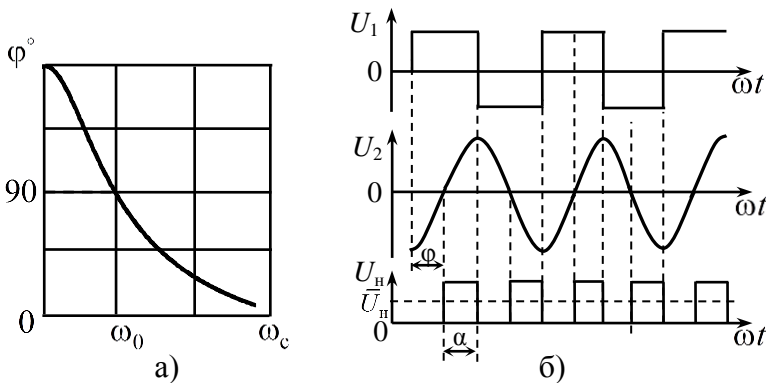


Рисунок 7.37 – Фазочастотна характеристика ФЗК (а) та епюри напруг (б) в колах частотного детектора

Струм у навантаженні має місце лише у разі збігу полярностей напруг U_1 та U_2 . В цьому випадку на навантаженні утворюється послідовність прямокутних імпульсів $U_H(t)$ тривалістю α та середнім значенням $\overline{U_H(t)}$, що визначається фазовим зсувом $\varphi(\omega_c)$, який в межах лінійного відрізка ФЧХ повторює закон ЧМ вхідного сигналу (рис. 7.37, б). Вихідний ФНЧ виділяє закон модуляції $\overline{U_H(t)}$. Побудовані за таким принципом частотні детектори широко застосовуються, завдяки можливості їх реалізації на інтегральних схемах і простоті регулювання.

Частотні детектори лічильного типу, вихідна напруга якого є функцією кількості імпульсів, що надходять на вхід в одиницю часу, застосовуються в більшості найбільш точних вимірювачів девіації частоти, хоча їх реалізація має певні труднощі на високих проміжних частотах [28, 32]. Структурна схема частотного детектора лічильного типу представлена на рис. 7.38 [32].



Рисунок 7.38 – Структурна схема частотного детектора лічильного типу

Принцип його роботи полягає в перетворенні ЧМ сигналу в послідовність імпульсів, модульованих за частотою слідування, з наступним виділенням з неї напруги пропорційної зміні частоти ЧМ сигналу за допомогою ФНЧ.

В представленій схемі (рис. 7.38) ЧМ сигнал спочатку зазнає впливу симетричного двобічного амплітудного обмеження в схемі формування, далі диференціюється з виділенням імпульсів однієї полярності та фільтрується ФНЧ. Для підвищення чутливості імпульси подаються не безпосередньо на ФНЧ, а через тригер Шмітта, на виході якого утворюються імпульси модульовані за частотою слідування з таким же законом, що й вхідні імпульси, але значно більші за площею, ніж однополярні імпульси, отримані в результаті диференціювання.

Для демодуляції КМ сигналів в реальному масштабі часу методами цифрової обробки сигналів застосовується квадратурна обробка. Цій обробці зазвичай передуює перетворення частоти сигналу на проміжну частоту, хоча з розвитком цифрової техніки частота дискретизації сигналу все більше наближується до радіочастоти колювання-носія. При

квадратурній обробці вхідний сигнал множиться на два опорні коливання, зсув по фазі між якими складає 90° [11]:

$$\begin{aligned} y_I(t) &= S_{\text{КМ}}(t) \cos \omega_0 t = A \cdot \cos(\omega_0 t + \varphi(t)) \cos \omega_0 t = \\ &= \frac{A}{2} \cos \varphi(t) + \frac{A}{2} \cos(2\omega_0 t + \varphi(t)); \end{aligned}$$

$$\begin{aligned} y_Q(t) &= S_{\text{КМ}}(t) \sin \omega_0 t = A \cdot \cos(\omega_0 t + \varphi(t)) \sin \omega_0 t = \\ &= -\frac{A}{2} \sin \varphi(t) + \frac{A}{2} \sin(2\omega_0 t + \varphi(t)). \end{aligned}$$

Кожен з результатів множення містить два доданки. Один з них – низькочастотний сигнал (косинус або синус початкової фази), інший – високочастотний КМ-сигнал з частотою-носієм $2\omega_0$. Низькочастотні складові виділяються за допомогою ФНЧ:

$$y'_I(t) = \frac{A}{2} \cos \varphi(t); \quad y'_Q(t) = -\frac{A}{2} \sin \varphi(t).$$

Подальші дії залежать від виду кутової модуляції. Для демодуляції ФМ необхідно обчислити фазу отриманої пари квадратурних складових

$$\begin{aligned} s_{\text{ФМ}}(t) &= -\arg(y'_I(t) + jy'_Q(t)) = -\arg\left(\frac{A}{2} \cos \varphi(t) - j\frac{A}{2} \sin \varphi(t)\right) = \\ &= -\arg\left(\frac{A}{2} \exp(-j\varphi(t))\right) = \varphi(t). \end{aligned}$$

Для демодуляції ЧМ сигналу отриману фазову функцію необхідно продиференціювати

$$\begin{aligned} s_{\text{ЧМ}}(t) &= \frac{ds_{\text{ФМ}}}{dt} = -\frac{d}{dt} \arg(y'_I(t) + jy'_Q(t)) = \\ &= -\frac{d}{dt} \arctg \frac{y'_Q(t)}{y'_I(t)} = \frac{\frac{dy'_I}{dt} y'_Q(t) - \frac{dy'_Q}{dt} y'_I(t)}{y'^2_I(t) + y'^2_Q(t)}. \end{aligned}$$

Структурна схема такого демодулятора показана на рис. 7.39 [11].

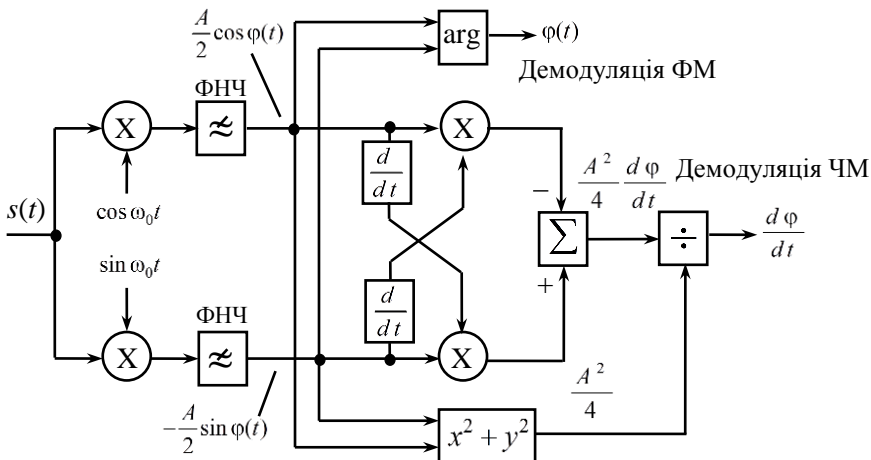


Рисунок 7.39 – Квадратурна обробка сигналу з кутовою модуляцією

Перевагою цієї схеми є те, що при високій частоті-носії вхідні блоки (генератор опорного колювання і помножувачі) можуть бути виконані в аналоговому виді.

Демодуляція сигналу з квадратурною модуляцією. Як й інші різновиди сигналів з КМ, квадратурно-модульований сигнал може демодулюватись шляхом множення на опорне колювання. Проте оскільки КАМ сигнал є сумою двох АМ сигналів, то і опорних колювань повинно бути два – із зсувом фаз на 90° :

$$y_1(t) = s_{\text{КАМ}}(t) \cos \omega_0 t = (a(t) \cos \omega_0 t + b(t) \sin \omega_0 t) \cos \omega_0 t = \\ = \frac{1}{2} a(t) + \frac{1}{2} a(t) \cos 2\omega_0 t + \frac{1}{2} b(t) \sin 2\omega_0 t,$$

$$y_Q(t) = s_{\text{КАМ}}(t) \sin \omega_0 t = (a(t) \cos \omega_0 t + b(t) \sin \omega_0 t) \sin \omega_0 t = \\ = \frac{1}{2} b(t) + \frac{1}{2} a(t) \sin 2\omega_0 t + \frac{1}{2} b(t) \cos 2\omega_0 t.$$

Результат кожного множення містить три доданки. Один з них є низькочастотним модульовальним коливанням, заданим функцією $a(t)$ або $b(t)$ із зменшеним удвічі рівнем. Інші два доданки утворюють КАМ сигнал з частотою-носієм $2\omega_0$. Тому корисні складові легко виділяються шляхом пропускання результатів множення через ФНЧ. Структура демодулятора показана на рис. 7.40, а.

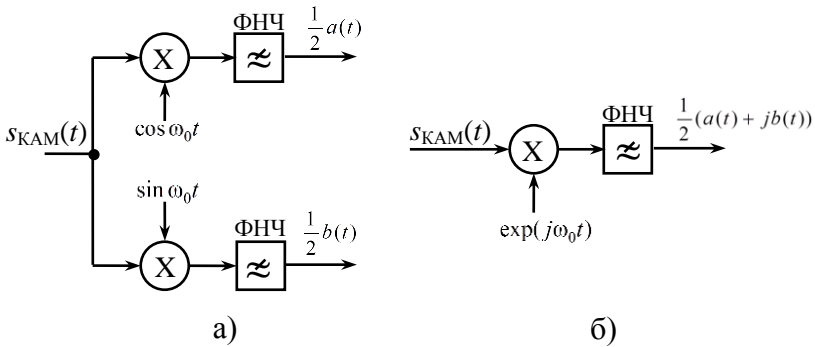


Рисунок 7.40 – Демодуляція КАМ-сигналу:
а) в дійсному і б) комплексному представленні

Якщо не обмежуватись дійсними сигналами, можна представити демодуляцію КАМ дещо простіше (рис. 7.40, б). При цьому потрібно всього один канал, а опорним коливанням є комплексна експонента $\exp(j\omega_0 t)$:

$$y(t) = s_{\text{КАМ}}(t) \exp(j\omega_0 t) = (a(t) \cos \omega_0 t + b(t) \sin \omega_0 t) \exp(j\omega_0 t) = \\ = \frac{1}{2}(a(t) + jb(t)) + \frac{1}{2}(a(t) - jb(t)) \exp(j2\omega_0 t).$$

Результат комплексного множення містить низькочастотний доданок, що представляє обидві модульовальні функції, і високочастотний комплексний доданок, що містить множник $\exp(j2\omega_0 t)$. Низькочастотний доданок, як завжди, виділяється за допомогою ФНЧ. Дійсна і уявна частини комплексного вихідного

сигналу відповідають двом вихідним сигналам дійсного демодулятора.

При демодуляції важливо забезпечити високу точність формування частоти і початкової фази опорного коливання. За наявності фазової помилки $\Delta\varphi$ на виходах демодулятора будуть отримані лінійні комбінації модульовальних функцій:

$$\operatorname{Re}\left((a(t) + jb(t))\exp(j\Delta\varphi)\right) = a(t)\cos\Delta\varphi - b(t)\sin\Delta\varphi,$$

$$\operatorname{Im}\left((a(t) + jb(t))\exp(j\Delta\varphi)\right) = a(t)\sin\Delta\varphi + b(t)\cos\Delta\varphi,$$

За наявності помилки по частоті фазовий зсув змінюється лінійно у часі, так що досить підставити в попередню формулу $\Delta\varphi = \Delta\omega t$:

$$\operatorname{Re}\left((a(t) + jb(t))\exp(j\Delta\omega t)\right) = a(t)\cos\Delta\omega t - b(t)\sin\Delta\omega t,$$

$$\operatorname{Im}\left((a(t) + jb(t))\exp(j\Delta\omega t)\right) = a(t)\sin\Delta\omega t + b(t)\cos\Delta\omega t.$$

Таким чином, сигнали на виході демодулятора виявляються амплітудно-модульованими з частотою биття.

Аналогічно демодулюється сигнал з квадратурною маніпуляцією. Як і у випадку аналогової квадратурної модуляції – сигнал множиться на два коливання-носія, що зсунуті по фазі одне відносно одного на 90° , а результати множення пропускаються через ФНЧ (рис. 7.40, а). На виходах цих ФНЧ будуть отримані аналогові сигнали синфазної і квадратурної складових. Далі ці сигнали дискретизуються з частотою, яка дорівнює символній швидкості. Пари відліків синфазної і квадратурної складових утворюють комплексне число, і найближча до цього числа точка використуваного сузір'я (а точніше – відповідний цій точці інформаційний символ) видається як вихідний результат.

Графік розміщення прийнятих точок на комплексній площині називається діаграмою розсіяння (*англ.* scatter plot), яка ілюструє якість (перешкодостійкість) прийому, що визначається ступенем розкиду точок поблизу очікуваних їх положень в сузір'ї (див. рис. 7.25).

Контрольні запитання

1. Які нелінійні системи вважаються безінерційними?
2. Перелічить різновиди нелінійних елементів та їх основні властивості.
3. Якими параметрами характеризуються нелінійні елементи?
4. Яку назву має залежність диференційної ємності від положення робочої точки?
5. В яких випадках може застосовуватись кусково-лінійна апроксимація ВАХ нелінійного елемента? Її переваги та обмеження.
6. Як виконується вибір найбільш придатної апроксимації ВАХ? Перелічить найбільш поширені види апроксимувальних функцій.
7. Як змінюються спектри гармонічного та бігармонічного сигналу при проходженні нелінійної системи? Які значення мають комбінаційні частоти?
8. Як визначаються одновимірна та багатовимірна функції розподілу ймовірностей на виході безінерційного нелінійного кола?
9. Якими способами можна обчислювати математичне сподівання та дисперсію випадкового сигналу на виході безінерційного нелінійного кола? Запишіть формули для їх обчислення.
10. Опишіть способи обчислення кореляційної функції випадкового сигналу на виході безінерційного нелінійного кола? В чому їх принципова відмінність порівняно з її обчисленням на виході лінійних систем?
11. За яких умов може бути обчислена кореляційна функція випадкового сигналу на виході безінерційного нелінійного елемента?
12. Поясніть сутність процесів модуляції і демодуляції.
13. Які основні види модуляції ви знаєте?

14. В яких межах зазвичай задається коефіцієнт амплітудної модуляції? За яких умов виникає перемодуляція та які її наслідки?

15. Які є різновиди АМ, в чому їх переваги та недоліки?

16. Як визначається коефіцієнт корисної дії амплітудної модуляції? Обчисліть їх значення для різних видів АМ.

17. За яким принципом частіш за все будуються амплітудні модулятори?

18. Що таке кутова модуляція? Як розрізнити сигнали з фазовою і частотною модуляцією?

19. За яким принципом функціонує модулятор Армстронга?

20. Приведіть формули для розрахунку ефективної ширини спектру КМ-сигналу при малих і великих індексах модуляції.

21. Охарактеризуйте види модуляції, в яких одночасно змінюються кілька параметрів коливання-носія?

22. Як називається спосіб модуляції, при якому сигнал має представлення: $s(t) = a(t) \cos \omega_0 t + b(t) \sin \omega_0 t$? В чому його відмінність від сигналу з амплітудно-фазовою модуляцією?

23. Дайте визначення маніпуляції сигналу. Які види маніпульованих сигналів ви знаєте?

24. Опишіть принцип детектування АМ сигналів і принципову схему амплітудного детектора.

25. Як може вплинути перешкода на корисний сигнал в амплітудному детекторі?

26. Перелічіть різновиди частотних і фазових детекторів. За якими принципами вони будуються?

27. Зобразіть варіанти структурних схем демодуляторів КАМ і КАМн сигналів та запишіть співвідношення, на яких ґрунтується їх побудова.

28. Як виглядають «сузір'я» КАМн сигналів, та яким чином за їх допомогою можна оцінити перешкодостійкість таких сигналів?

Глава 8

ВЕЙВЛЕТ-АНАЛІЗ СИГНАЛІВ

8.1. Поняття про спектрально–часовий аналіз

Спектрально-часовий аналіз (СЧАН) – це метод дослідження нестационарних випадкових сигналів, що полягає в оцінці „потокowego” спектра частини сигналу, що відповідає ковзному часовому вікну деякої малої довжини [36]. При подібній оцінці виходить залежність спектра потужності сигналу, як від частоти, так і від положення (середини або правого краю) вікна. Ця залежність візуалізується у вигляді або двовимірних ліній рівнів, або тривимірних рельєфів, які називають СЧАН–діаграмами. Максимальним значенням такої залежності відповідають періоди часу збільшення потужності коливань у тих або інших частотних смугах. СЧАН часто застосовується для виявлення нестационарних сигналів у дискретних у часі протяжних випадкових процесах.

Будується спектрограма в такий спосіб. Часове вікно довжиною в задану кількість відліків ряду зміщується уздовж ряду зліва направо із заданим кроком. Для кожного кроку розраховується спектр потужності фрагмента ряду, що потрапив у вікно. Одержаний спектр потужності розгортається уздовж осі ординат зі значенням абсциси, відповідної до положення середини вікна. У результаті виходить залежність оцінки спектра потужності від двох параметрів у вигляді рельєфу, а саме, від частоти й від положення часового вікна. Слідкуючи за хребтами на рельєфі можна одержати уявлення про те, як інтенсивність прояву тієї або іншої циклічності змінювалася за часом (рис. 8.1).

Зіставляючи звичайне й віконне перетворення Фур’є можна дійти висновку про те, що останнє розв’язало всі проблеми з аналізом і синтезом функцій і сигналів. Але принцип невизначеності Гейзенберга, відомий з курсу фізики, говорить про те, що неможливо одночасно одержати високе

часове і частотне відрізнєння. Вибираючи вікно з малою шириною за часом, одержимо високе часове відрізнєння і низьке частотне, вибираючи вікно з великою шириною за часом, одержимо високе частотне відрізнєння і низьке часове. Віконне перетворєння Фур'є оперує з вікнами однакової ширини й тому дане протиріччя для нього нерозв'язне [36].

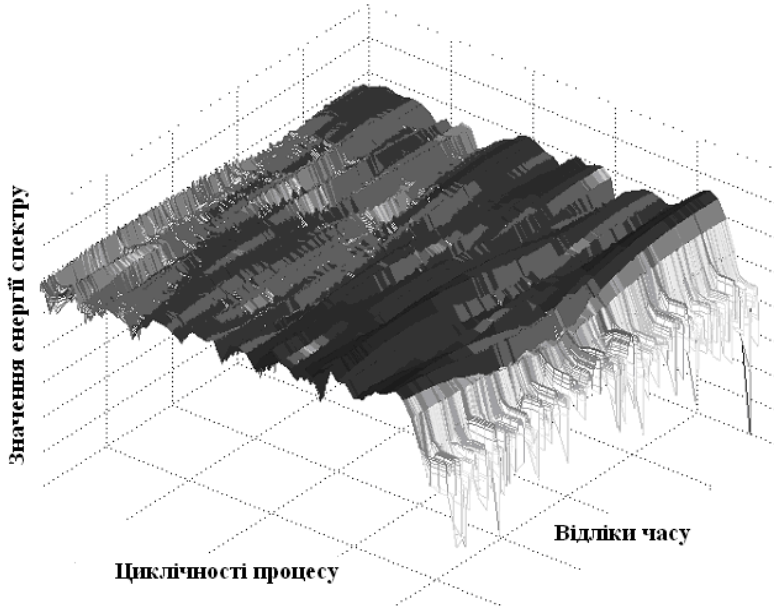


Рисунок 8.1 – Тривимірна СЧАН–діаграма

І, крім того, у ньому використовується єдина базисна функція – синусоїда – з усіма властивими їй недоліками й, виходить, недоліками перетворєння Фур'є:

- незастосовність до аналізу нестационарних сигналів, у яких певні частотні компоненти існують тільки в певні проміжки часу або коли параметри сигналів змінюються в часі;
- перетворєння Фур'є навіть для однієї заданої частоти

вимагає знання сигналу не тільки в минулому, але й у майбутньому, що є теоретичною абстракцією;

- в умовах практично неминучого обмеження числа гармонік або спектра коливань точне відновлення сигналу після прямого та зворотного перетворень Фур'є теоретично і практично неможливо через наявність ефекту Гібса;
- базисною функцією при розкладанні в ряд Фур'є є гармонійне коливання, яке математично визначене в нескінченному інтервалі часу та має незмінні в часі параметри;
- при чисельному інтегруванні в нескінченних межах при прямому та зворотному перетворенні відповідно в часовий і в частотній області виникають значні обчислювальні труднощі;
- істотні особливості сигналу у вигляді розривів або піків викликають несуттєві зміни частотного образу сигналу у всьому нескінченному інтервалі частот, які „розмазуються” по всій частотній осі, що робить їхнє виявлення в спектрі практично неможливим;
- навіть уможливно зрозуміло, що плавна синусоїдальна базисна функція в принципі не може представляти перепади сигналів з нескінченною крутістю, хоча на практиці подібні імпульсні сигнали застосовуються дуже широко;
- єдине пристосування до визначення швидких змін сигналів – це різке збільшення кількості гармонік, які, однак, впливають на форму сигналу й за межами локальних особливостей сигналу;
- по складу вищих гармонік спектра практично неможливо оцінити місце розташування особливостей на часовій залежності сигналу і їх характер;
- для нестационарних сигналів при прямому та зворотному перетворенні Фур'є суттєво збільшується обчислювальна складність.

8.2. Загальні відомості про вейвлет–перетворення

Ідея вейвлет-перетворення. На підставі поняття про векторний простір загальноприйнятим підходом до аналізу сигналів $s(t)$ стало їхнє представлення у вигляді зваженої суми простих складових – базисних функцій $\psi_k(t)$, помножених на коефіцієнти C_k [37]:

$$s(t) = \sum_k C_k \psi_k(t).$$

Оскільки базисні функції $\psi_k(t)$ передбачаються заданими як функції цілком певного виду, то тільки коефіцієнти C_k містять інформацію про конкретний сигнал. Таким чином, можна говорити про можливість представлення довільних сигналів на основі рядів з різними базисними функціями.

Наприклад, в ряді Фур'є $s(t) = \sum_k C_k e^{jtk}$ використовуються в якості базисної функції синусоїди $e^{jtk} = \cos(tk) + j\sin(tk)$, $j = \sqrt{-1}$. Вони гранично локалізовані в частотній області й взагалі не локалізовані в часовій області [38].

Прямо протилежний приклад – ряд Шенона $s(t) = \sum_k C_k \delta_k(t)$, у якому використовується в якості базисної

функції функція Кронекера $\delta_k(t) = \begin{cases} 1, & k = t \\ 0, & k \neq t \end{cases}$, яка добре

локалізована в часі, але не несе ніякої інформації про частоту й тому погано пристосована для представлення повільно мінливих сигналів на заданому відрізку часу і, тим більше, періодичних сигналів.

В обох наведених випадках базисні функції мають принциповий недолік – вони не здатні адаптуватися до локальних змін сигналів.

Таким чином, виникла нагальна потреба в створенні класу функцій вільних від зазначених недоліків. Ними стали вейвлети.

Поняття про вейвлети. Термін „вейвлет” (англ. wavelet) у перекладі з англійської мови означає „коротка або маленька хвиля”. У побут його ввів сейсмолог Морле, тому що подібні хвилі поширюються в об’ємі Землі при тектонічних процесах. Термін „вейвлет” вдалий через відбиття їм суті процесу, тому що більшість вейвлетів має часові залежності з яскраво вираженим коливальним компонентом, як у хвилях. Вейвлети займають проміжне положення між наведеними вище базисними функціями – синусоїдою та функцією Кронекера.

Досить грубо можна уявити вейвлети як деякі хвильові функції, здатні здійснювати перетворення Фур’є не по всій часовій осі, а локально по місці свого розташування. Для цього, природно, крім зміни „середньої частоти” вейвлети повинні переміщатися до того місця сигналу або функції, у якому повинне здійснюватися „локальне перетворення Фур’є” і масштабуватися. Така інтерпретація вейвлетів дуже спрощена й дає про них тільки саме грубе уявлення, яке часом навіть принципово помилкове через те, що переважна більшість вейвлетів не є модульованою по амплітуді синусоїдальною хвилею й властивості вейвлетів принципово та сильно відрізняються від властивостей синусоїди.

Базисними функціями вейвлетів можуть бути всілякі функції, у тому числі такі, що нагадують модульовані синусоїди, функції зі стрибками рівня і т.д. Це забезпечує різне представлення сигналів з локальними стрибками та розривами наборами вейвлетів того або іншого типу і відкриває простір у доборі найбільш підходящих вейвлетів, виходячи з умов розв’язуваних задач.

Слід зазначити, що майже всі вейвлети не мають аналітичного представлення у вигляді однієї формули, але можуть задаватися ітераційними вираженнями, які легко обчислюються.

Вейвлети характеризуються своїм часовим і частотним

образами. Часовий образ визначається функцією часу $\psi(t)$. А частотний образ визначається її Фур'є-образом $\widehat{\psi}(\omega) = F(\omega)$, який задає огинальну спектра вейвлета. При цьому Фур'є-образ визначається вираженням $F(\omega) = \int_{-\infty}^{\infty} \widehat{\psi}(t) e^{-i\omega t} dt$.

Із цього виразу видно, що неможлива гарна локалізація вейвлетів одночасно в часовій області й у частотній області. Але було б помилкою вважати цю властивість недоліком. У порівнянні з віконним перетворенням Фур'є саме гарна часова локалізація вейвлетів дає засіб для підвищення частоти тих спектральних складових вейвлетів, які потрібні для представлення локальних короткочасних особливостей сигналу.

Визначення вейвлета. Будь-яка локалізована \mathbf{R} -функція $\psi \in L^2(\mathbf{R})$ називається \mathbf{R} -вейвлетом або просто вейвлетом, якщо для неї існує функція $\psi^* \in L^2(\mathbf{R})$ (її пара, двійник) така, що сімейства $\{\psi_{ik}\}$ і $\{\psi_{ik}^*\}$, побудовані відповідно до

$$\psi_{j,k}(t) = a_0^{-j/2} \psi(a_0^{-j}t - k)$$

та

$$\psi_{j,k}^*(t) = a_0^{-j/2} \psi^*(a_0^{-j}t - k),$$

є парними базисами функціонального простору $L^2(\mathbf{R})$.

Кожний, визначений подібним чином вейвлет ψ , незалежно від того, ортогональний він чи ні, дозволяє будь-яку функцію $f \in L^2(\mathbf{R})$ представити у вигляді ряду

$$f(t) = \sum_{j,k=-\infty}^{\infty} c_{jk} \psi_{jk}(t),$$

коефіцієнти якого визначаються інтегральним вейвлет – перетворенням f відносно ψ^* .

Вейвлет–двійник ψ^* – єдиний і сам є \mathbf{R} -вейвлетом. Пара (ψ, ψ^*) симетрична в тому розумінні, що ψ у свою чергу є двійником для ψ^* .

Якщо \mathbf{R} -вейвлет ψ має властивість ортогональності, то $\psi^* \equiv \psi$, і $\{\psi_{ik}\} \equiv \{\psi_{ik}^*\}$ – ортогональний базис.

Для багатьох практичних цілей досить, щоб вейвлет ψ мав властивість напівортогональності, тобто щоб його базис Рисса $\{\psi_{ik}\}$ задовольняв умові $\langle \psi_{ik}, \psi_{lm} \rangle = 0$ при $j, k, l, m \in I$.

\mathbf{R} -вейвлет називається не ортогональним, якщо він не є напівортогональним вейвлетом. Однак, будучи \mathbf{R} -вейвлетом, він має двійника, і пари (ψ, ψ^*) дає можливість сформуувати сімейства $\{\psi_{ik}\}$ й $\{\psi_{ik}^*\}$, що задовольняють умові біортогональності $\langle \psi_{ik}, \psi_{lm} \rangle = \delta_{il} \delta_{km}$, що $j, k, l, m \in I$ й дозволяють побудувати повноцінний ряд по вейвлетам і реконструкційну формулу.

З необхідністю мати зворотне вейвлет–перетворення (або реконструкційну формулу) зв’язана більшість обмежень, що накладаються на вейвлет.

Ознаки вейвлета.

Локалізація. Вейвлет–перетворення на відміну від перетворення Фур’є використовує локалізовану базисну функцію. Вейвлет повинен бути локалізований і в часовому просторі, і по частоті.

Нульове середнє. $\int_{-\infty}^{\infty} \psi(t) dt = 0$. Часто виявляється

необхідним, щоб не тільки нульовий, але й усі перші m моментів дорівнювали нулю:

$$\int_{-\infty}^{\infty} t^m \psi(t) dt = 0.$$

Такий вейвлет називається вейвлетом m -го порядку. Вейвлети, що мають більше число нульових моментів дозволяють, ігноруючи найбільш регулярні поліноміальні складові сигналу, аналізувати дрібномасштабні флуктуації та особливості високого порядку.

Обмеженість. $\int |\psi(t)|^2 dt < \infty$. Оцінка гарної локалізації та обмеженості може бути записана у вигляді $|\psi(t)| < (1 + |t|^n)^{-1}$ або $|\hat{\psi}(\omega)| < (1 + |k - \omega_0|^n)^{-1}$, тут ω_0 – домінантна частота вейвлета, число n повинне бути можливо більшим.

Автомодельність базису. Характерною ознакою базису вейвлет-перетворення є його самоподоба. Усі вейвлети даного сімейства $\psi_{ab}(t)$ мають однакове число осциляцій з базисним вейвлетом $\psi(t)$, оскільки отримані з нього за допомогою масштабних перетворень і зсувів. Завдяки цьому вейвлет-перетворення з успіхом застосовується для аналізу фрактальних сигналів.

Пряме безперервне вейвлет-перетворення (ПБВП) – означає розкладання (декомпозицію) довільного вхідного сигналу по базису у вигляді сукупності хвильових пакетів – вейвлетів, які характеризуються чотирма основними властивостями:

- мають вигляд коротких, локалізованих у часі або в просторі хвильових пакетів з нульовим значенням

інтеграла $\int_{-\infty}^{\infty} \psi(t) dt = 0,$

- мають можливість зсуву за часом,
- здатні до масштабування, тобто до стиску й розтягання,

– мають обмежений або локальний частотний спектр.

Якщо базис ортогональний, то аналіз суттєво полегшується, з'являється можливість відновлення (реконструкції) сигналу та реалізації алгоритмів швидких вейвлет-перетворень.

Але існують і не ортогональні вейвлети (майже базисні вейвлети), які, проте, корисні в завданнях аналізу та виявлення локальних особливостей сигналів і функцій.

Умовно кажучи, при ПБВП сигнал представляється сукупністю хвильових пакетів – вейвлетів, утворених на основі деякої вихідної (базової) функції $\psi_0(t)$. Ця сукупність, різна в різних частинах часового інтервалу визначення сигналу і коректується множниками, що мають вид складних часових функцій, та представляє сигнал з тим або іншим ступенем деталізації.

Кількість використовуваних при розкладанні сигналу вейвлетів задає *рівень розкладання* або *рівень декомпозиції* сигналу. За нульовий (вищий) рівень вважається сам сигнал, а наступні рівні розкладання утворюють дерево розкладання (декомпозиції) або вейвлет-дерево. Точність представлення сигналу в міру переходу на більш низькі рівні розкладання знижується, але з'являється можливість вейвлет-фільтрації сигналів, усунення шумів із суміші сигнал+шум, стиск (компресія) сигналів, тобто стає можливою їхня вейвлет-обробка.

Також однією з основних ідей вейвлет-представлення сигналів є розбивка наближення до сигналу на *дві складові* – грубу або *апроксимуючу* й уточнену або, *деталізуючу* – з наступним уточненням цих складових ітераційним методом. Кожний крок такого уточнення відповідає певному рівню декомпозиції (розкладання) або реставрації (відновлення) сигналу. Ці операції можливі як у часовій, так і в частотній області представлення сигналів вейвлетами.

Також в основі ПБВП лежить використання двох безперервних та інтегрованих по всій осі t (або x) функцій:

- вейвлет-функція $\psi(t)$ з нульовим значенням інтегралу $\int_{-\infty}^{\infty} \psi(t) dt = 0$, що визначає деталі сигналу, що й породжує деталізуючі коефіцієнти;
- масштабуюча (апроксимуюча) або скейлинг-функція $\varphi(t)$ з одиничним значенням інтегралу $\int_{-\infty}^{\infty} \varphi(t) dt = 1$, що визначає грубе наближення сигналу, що й породжує апроксимуючі коефіцієнти.

Апроксимуючі функції $\varphi(t)$ властиві не всім вейвлетам, а тільки ортогональним.

Функція $\psi(t)$ створюється на основі тієї або іншої базисної функції $\psi_0(t)$, яка, як і $\psi(t)$, визначає тип вейвлета. Базисна функція $\psi_0(t)$ повинна задовольняти всім тим вимогам, які ставляться до $\psi(t)$, а саме:

- зсув по осі часу $t - \psi_0(t-b)$ при $b \in \mathbf{R}$,
- масштабування $-\frac{1}{\sqrt{a}} \psi_0\left(\frac{t}{a}\right)$ при $a > 0$ й $a \in \mathbf{R}^+ - \{0\}$.

Параметр a задає ширину функції, а параметр b – його положення на осі t .

Ці властивості відбиваються при наступному записі функції $\psi(t)$:

$$\psi(t) \equiv \psi(a, b, t) = a^{-1/2} \psi_0\left(\frac{t-b}{a}\right).$$

Для заданих a і b функція $\psi(t)$ і є вейвлет. Вейвлети, позначувані подібним чином, називають ще батьківськими або материнськими вейвлетами, оскільки вони породжують цілий ряд вейвлетів певного роду.

Вейвлети є дійсними функціями часу t й коливаються навколо осі t (або осі x , якщо вейвлети – функції простору). Параметр b задає положення вейвлетів, а параметр a – його масштаб. Коефіцієнт $a^{-1/2}$ забезпечує нормування енергії при зміні масштабу.

Стосовно до сигналів як функціям часу параметр $b \in \mathbf{R}$ задає положення вейвлета на осі часу, а параметр a – задає його масштабування по осі часу. Оскільки параметр масштабу a реально може бути тільки додатним і його не можна брати рівним 0, те й покладається, що $a \in \mathbf{R}^+ - \{0\}$, а на практиці $-\{0\}$ опускають і не пишуть.

У частотній області малі значення a відповідають високим частотам, а більші – низьким частотам. Таким чином, операція завдання вікна, використовувана у віконному перетворенні Фур'є, закладена в самій базисній функції вейвлетів, а це й створює передумови їх пристосування до сигналів, які можуть бути представлені сукупністю вейвлетів.

Нехай енергія сигналу $s(t)$ дорівнює $\int_R s^2(t) dt$ й вона $< \infty$.

Пряме безперервне вейвлет-перетворення сигналу $s(t)$ задається по формальній аналогії з перетворенням Фур'є, відповідно до наведеної раніше ідеї вейвлет-перетворення, шляхом обчислення вейвлет-коефіцієнтів по формулі:

$$C(a, b) = \langle s(t), \psi(a, b, t) \rangle = \int_{-\infty}^{\infty} s(t) a^{-1/2} \psi\left(\frac{t-b}{a}\right) dt.$$

З урахуванням обмеженої області визначення сигналів і $a, b \in \mathbf{R}, a \neq 0$:

$$C(a,b) = \int_R s(t) a^{-1/2} \psi\left(\frac{t-b}{a}\right) dt.$$

Таким чином, вейвлет-коефіцієнти визначаються інтегральним значенням скалярного добутку сигналу на вейвлет-функцію заданого виду, а ПБВП можна розглядати як розкладання сигналу $s(t)$ по всіх можливих зсувах і по розтяганнях (стисканнях) сигналу $s(t)$ або деякої довільної функції. При цьому параметри a й b можуть приймати будь-які значення в межах областей їх визначення.

Слід зазначити кілька важливих обставин.

Практично неможливо вести обчислення з безперервними вейвлетами. На практиці доводиться задавати дискретні значення a й b .

Безперервне вейвлет-перетворення має велику надмірність, що веде до невиправдано великих витрат машинного часу. Але, крім того, надмірність ще може бути й небезпечна при аналізі результатів вейвлет-перетворення. При відображенні спектрограм штатними засобами систем комп'ютерної математики локальні особливості сигналу можуть відображатися по-різному. Це може привести до помилкової інтерпретації рентгенівських або томографічних знімків, до помилок геологорозвідки й т.п.

Зворотне безперервне вейвлет-перетворення (ЗБВП) – здійснюється по формулі реконструкції в часовій області, яка має цілий ряд форм, що залежать від математичного стилю запису та від визначення областей існування сигналу.

Один з варіантів:

$$f = C_{\psi}^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dadb}{a^2} \langle f, \psi^{a,b} \rangle \psi^{a,b},$$

де

$$\psi^{a,b}(x) = |a|^{-1/2} \psi\left(\frac{x-b}{a}\right) \quad \text{і}$$

$$C_\psi = 2\pi \int_{-\infty}^{\infty} d\xi |\widehat{\psi}(\xi)|^2 |\xi|^{-1} = 2\pi \int_{-\infty}^{\infty} \frac{|\widehat{\psi}(\xi)|^2}{|\omega|^{-1}} d\omega$$

Тут трохи незвичний запис інтеграла та параметрів масштабу й зсуву. Параметри зсуву й стискання (розтягання) міняються безупинно уздовж \mathbf{R} із обмеженням $a \neq 0$. Вважається, що $C_\psi < \infty$, оскільки в протилежному випадку вираження для f не має змісту.

Якщо ψ є дійсною функцією, то справедливий наступний вираз:

$$C_\psi = 2\pi \int_0^{\infty} d\xi |\widehat{\psi}(\xi)|^2 |\xi|^{-1} = 2\pi \int_{-\infty}^0 d\xi |\widehat{\psi}(\xi)|^2 |\xi|^{-1} < \infty.$$

У цьому випадку:

$$f = C_\psi^{-1} \int_0^{\infty} \frac{da}{a^2} \int_{-\infty}^{\infty} db \langle f, \psi^{a,b} \rangle \psi^{a,b}.$$

Існують і інші варіанти запису ЗБВП. Практичний інтерес представляють лише ті формули реконструкції, які реально використовуються в системах комп'ютерної математики при прямому та зворотному безперервних вейвлет-перетвореннях. Наприклад,

$$s(t) = \frac{1}{C_\psi} \int_{\mathbf{R}^+ - \{0\}} \int_{\mathbf{R}} C(a,b) a^{-1/2} \psi\left(\frac{t-b}{a}\right) \frac{dadb}{a^2}.$$

У теорії вейвлет-перетворень доведено, що не всі

вейвлети є ортогональними, тому за допомогою перетворень не завжди вдається відновити будь-який сигнал у цілому. Однак також доведено, що довільний сигнал з відмінним від нуля інтегральним значенням може бути відновлений за допомогою вейвлетів з нульовим інтегральним значенням, якщо вираження для C_{ψ} належить із деякими обмеженнями до норми $L^1(\mathbf{R})$.

Логічно припустити, що вейвлет-перетворення тільки на основі деталізуючої ортогональної вейвлет-функції $\psi(t)$ здатне відновити, принаймні, тонкі деталі часової залежності сигналу $s(t)$. Для відновлення повної форми сигналу потрібно вдаватися до допомоги ще однієї часової функції – апроксимуючої $\phi(t)$. Більш докладно це буде розглянуто в кратномасштабному аналізі.

Вейвлети, що допускають розкладання сигналу і які забезпечують його повну реконструкцію, є ортогональними вейвлетами. Вони є найбільш важливим типом вейвлетів, здатних повноцінно наближати довільні функції та сигнали.

Однак не всі типи вейвлетів гарантують точну реконструкцію сигналів у цілому або навіть таку взагалі. Проте, застосування й таких вейвлетів може бути корисно для виявлення тонких особливостей сигналів або зображень, які добре узгодяться з певними типами вейвлетів.

Діадне вейвлет-перетворення безперервних сигналів. Безперервне вейвлет-перетворення вимагає значних обчислювальних витрат. Тому для практичного його застосування необхідна дискретизація значень a і b . Щоб уникнути надмірності вейвлет-перетворення можна задавати дискретні значення a і b на деякій безлічі $Z = \{\dots -1, 0, 1, 2, \dots\}$, які дорівнюють $a = 2^j$ та $b = k \cdot 2^j$, де j та k – цілі числа. Можливі й інші способи завдання дискретних значень a і b . Параметр j , іноді позначуваний як m , називається параметром

масштабу.

Описана дискретизація є найпоширенішою, а сітка дискретизації називається *діадичною*. Відповідно й це перетворення прийнято називати *діадним вейвлет-перетворенням*.

Важливою особливістю діадного вейвлет-перетворення є виключення перекриття носіїв вейвлетів, тобто усунення надмірності в ході вейвлет-перетворення.

Виділення діадного вейвлет-перетворення в окремий клас виправдане ще й тим, що й безперервне вейвлет-перетворення на практиці використовується в тому або іншому дискретному варіанті, але при цьому не накладається особливих обмежень на характер дискретизації параметрів a і b . Дискретизація може бути іншою, наприклад, часто використовується рівномірний ряд значень a і b .

На відміну від дискретного вейвлет-перетворення безперервних сигналів, що має більшу надмірність, діадне перетворення такої надмірності не має. Це означає, що компактні носії вейвлетів щільно укладаються на осі t (або осі x), не перекриваючись.

Таким чином, при дискретних значеннях a і b вейвлет - функція може бути представлена у вигляді:

$$\psi_{j,k}(t) = a_0^{-j/2} \psi(a_0^{-j}t - k).$$

Умова, що множина $\{\psi_{j,k}\}_{j,k} \in \mathbb{Z}$ утворює ортонормовану систему, означає, що будь-яка функція $f \in L^2(\mathbf{R})$ може бути представлена рядом:

$$f = \sum_{j,k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k},$$

і цей вираз можна розглядати як накладення вейвлетів $\psi_{j,k}$,

що створює функцію f .

Зовні останній вираз нагадує розкладання в ряд Фур'є для функції $f \in L^2(\mathbf{R})$ по експонентних функціях. Але експонентна функція має глобальний носій, і всі члени розкладання Фур'є залежать від значення f в точці x .

Вейвлети мають компактний носій або експонентний порядок убуття на нескінченності, і тільки члени цього розкладання, що відповідають $\psi_{j,k}$ із $j=2^{-k}$ в околиці точки x , вносять найбільший вклад у функцію в цій околиці. Тому останнє розкладання локальне.

Таким чином, **пряме дискретне вейвлет-перетворення** (ПДВП) зводиться до вже відомого обчислення коефіцієнтів $C(a,b)$

$$C(a,b) = \int_{\mathbf{R}} s(t) a^{-1/2} \psi\left(\frac{t-b}{a}\right) dt$$

с підстановкою дискретних значень a і b .

І тоді остаточно:

$$C(j,k) = d_{j,k} = \int_{\mathbf{R}} s(t) a^{-j/2} \psi(a_0^{-j}t - k) dt,$$

де $C(j,k) = d_{j,k}$ – деталізуючі коефіцієнти для вейвлет-декомпозиції сигналу k рівня, причому коефіцієнти дискретні, тобто обчислюються для заданих j і k .

Зворотнє дискретне вейвлет-перетворення (ЗДВП) для безперервних сигналів задається формулою

$$s(t) = \frac{1}{C_\psi} \sum_{j \in \mathbf{Z}} \sum_{k \in \mathbf{Z}} C(j,k) a_0^{-j/2} \psi(a_0^{-j}t - k).$$

При нормуванні базових функцій у частотній області домагаються того, що $C_\psi = 1$. Тоді остаточно формула реконструкції сигналу буде мати вигляд:

$$s(t) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} C(j, k) \psi_{j, k}(t).$$

Відомо, що для ортогональних вейвлетів можливо точне відновлення сигналу після прямого та зворотного дискретного вейвлет-перетворення з використанням додаткової апроксимації сигналу за допомогою ϕ функції. В іншому випадку відновлення дає близький до вихідного сигналу $s(t)$ наближений сигнал, причому близькість розуміється в змісті забезпечення мінімуму середньоквадратичної похибки відновлення.

У підсумку видно, що діадне вейвлет-перетворення має ієрархічну структуру у вигляді дерева. При фіксованому індексі відрізнення j вейвлети мають однакові масштаби й лише дискретно зсуюються в часі. При збільшенні відрізнення на 1 величина зсуву зменшується вдвічі, і вейвлети стискаються.

Діадне вейвлет-перетворення часто називають дискретним. Але подібне формулювання не зовсім коректне. Вірніше говорити про діадне перетворення як про різновид безперервного вейвлет-перетворення, що дозволяє усунути надмірність останнього.

Варто згадати ще один особливий тип вейвлетів – **фрейми**. Вейвлет-фрейми використовують кратне двом масштабування, але безперервні зсуви. Таким чином, вони зберігають надмірність властиву безперервному перетворенню, але в набагато меншому ступені, ніж при цьому перетворенні. Тобто вейвлет-фрейми займають проміжне положення між діадним і безперервним вейвлет-перетворенням. Але слід зазначити, що в системах комп'ютерної математики вони застосування не знайшли.

Кратномасштабний аналіз і синтез. Тепер зрозуміло, що можливе представлення сигналу на основі підсумовування його грубого представлення з детальними представленнями у його різних місцях. Для цього є ортогональні вейвлети. Їх можна створити, ґрунтуючись на представленні простору сигналів V у вигляді вкладених підпросторів V_j , що відрізняються один від одного тільки перемасштабуванням незалежної змінної. Заснований на цьому аналіз і називають кратномасштабним або багатомасштабним аналізом.

Він базується на наступних вихідних передумовах:

- простір сигналів V може бути розбитий на ієрархічно вкладені підпростори V_j , які не перетинаються й об'єднання яких у межі дає $L^2(\mathbf{R})$;
- для будь-якої функції $s(t) \in V_j$ її стисла версія належить простору V_{j-1} ;
- існує така функція $\phi(x) \in V_0$, для якої її зсув $\phi_{0,j}(t) = \phi(t-k)$ при $k \in Z$ утворює ортонормований базис простору V_0 .

Оскільки функції $\phi_{0,k}(t)$ утворюють ортонормований базис простору V_0 , то функції

$$\phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j}t - k)$$

утворюють ортонормований базис простору V_j .

Таким чином, функції $\phi_{j,k}$ називаються масштабуючими саме тому, що вони створюють свої масштабовані версії в просторі сигналу. Їх ще називають батьківськими або материнськими вейвлетами. При цьому сигнал $s(t)$ може бути представлений безліччю послідовних наближень $s_j(t)$ у

підпросторах V_j . Тому j й називається масштабним коефіцієнтом і завжди можна сказати, що сигнал $s(t)$ є межа апроксимації $s_j(t) \in V_j$ при $j \rightarrow -\infty$, тобто $s(t) = \lim_{j \rightarrow -\infty} s_j(t)$.

Відповідно до цього при більших j виходять грубі наближення сигналу, а при малих – точні. Наближення сигналу відповідає ітераційній формулі:

$$s(t) = \sum_j C(j, k) \varphi_{j, k}(t),$$

при цьому

$$\varphi_{0,0}(t) = 2 \sum_k h_k \varphi(2t - k),$$

де h_k – деяка послідовність.

У підсумку сума приблизної та детальної компонент дає вихідний сигнал з тим або іншим наближенням. У якості початкового значення часто беруть $j = 0$, але можна починати ітераційний процес наближення й з інших значень j .

Як приклад можна привести передачу зображення, коли спочатку передається грубий образ зображення у вигляді функції $\varphi(t)$, а потім воно уточнюється за допомогою вейвлет-коефіцієнтів і приймає звичний вид. Причому процес уточнення завжди можна зупинити, якщо, наприклад, досягнута максимальна роздільна здатність системи відображення і більше немає необхідності в уточненні зображення або похибка перетворення досягалася рівня похибки машинних розрахунків.

Таким чином, у загальному випадку реконструкція сигналу на n -тому рівні відрізнення j_n задається вираженням:

$$s(t) = \sum_{k=-\infty}^{\infty} a_{j_n, k} \varphi_{j_n, k}(t) + \sum_{j=j_n}^{\infty} \sum_{k=-\infty}^{\infty} d_{j, k} \psi_k(t),$$

де a – коефіцієнти апроксимації, d – коефіцієнти деталізації.

З теорії кратномасштабного аналізу відомо, що в ході прямого й зворотного діадного вейвлет-перетворення можливо повне відновлення сигналу, якщо для цілих k існують такі коефіцієнти $\{h_k\}$, які задовольняють розв'язанню рівняння масштабування або рівняння уточнення:

$$\varphi\left(\frac{t}{2}\right) = \sqrt{2} \sum_k h_k \varphi(t-k).$$

Представлення результатів вейвлет-перетворення. Одновимірне перетворення Фур'є дає також одномірну інформацію про відносний внесок (амплітуду) різних часових масштабів (частот). Результатом вейвлет-перетворення одновимірного ряду є двовимірний масив амплітуд вейвлет-перетворення – значень коефіцієнтів $W(a,b)$. При аналізі комплексного одновимірного сигналу або при використанні комплексного аналізуючого вейвлета в результаті вейвлет-перетворення виходять двовимірні масиви значень модуля коефіцієнтів і фази:

$$W(a,b) = |W(a,b)| \exp[i\Phi(a,b)],$$

тобто два двовимірні масиви – для модуля та фази.

Розподіл цих значень у просторі (a,b) (часовий масштаб, часова локалізація) подає інформацію про еволюцію відносного внеску компонент різного масштабу в часі й називається спектром коефіцієнтів вейвлет-перетворення, (частотно-) масштабно-часовим спектром або вейвлет-спектром (time-scale spectrum, або wavelet spectrum на відміну від signal spectrum перетворення Фур'є).

Спектр $W(a,b)$ одновимірного сигналу являє собою поверхню в тривимірному просторі. Способи візуалізації цієї інформації можуть бути різними. Замість зображення

поверхонь часто представляють їхні проекції на площину ab з ізолініями або ізорівнями, що дозволяють простежити зміну інтенсивності амплітуд вейвлет-перетворення на різних масштабах і в часі, а також картини ліній локальних екстремумів цих поверхонь (т.зв. *skeleton*), що чітко виявляють структуру аналізованого процесу. Термін „скелет” або „скелетон” добре відбиває характер картин ліній локальних екстремумів.

У випадках, коли необхідно показати дуже широкий діапазон масштабів, візуалізація результатів у логарифмічних координатах $(\log a, b)$ переважніше, чим у лінійних координатах.

У системах комп’ютерної математики широке поширення одержало відображення вейвлет-спектрів з колірним розфарбуванням. Чистому гармонічному сигналу із частотою 50 Гц відповідають яскраві смуги, де модуль деякого коефіцієнта великий (рис. 8.2).

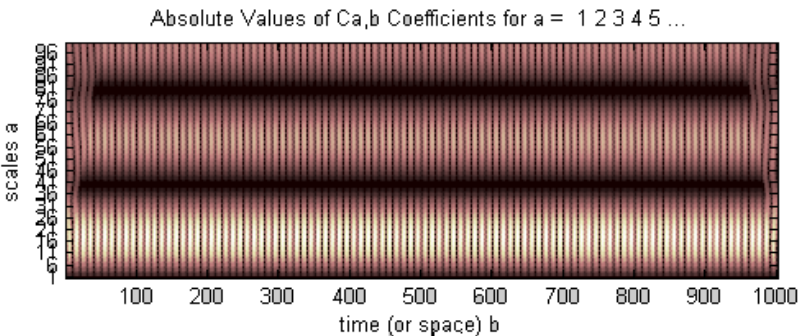


Рисунок 8.2 – Вейвлет-спектр гармонічного сигналу

Локальними особливостям негладких сигналів (наприклад, стикованим послідовно в часі синусоїдам), відповідають смуги, що виходять із точки, де перебуває особливість (рис. 8.3).

Пікам сигналів відповідає згущення світлих областей

вейвлет спектрограм, а западаючим – згущення темних областей.

Чим різкіше виражена особливість сигналу, тим сильніше вона виділяється на спектрограмі й тим вище рівень вейвлет-коефіцієнтів. Вейвлет-спектрограми чітко виділяють такі особливості сигналу, як невеликі розриви, зміни знаків першої та другої похідних, зміна частоти складових сигналу в часі і т.д. Це саме ті особливості сигналу, які погано виділяються на Фур'є-спектрі.

Відмінності між Фур'є і вейвлет-спектрограмами носять принциповий характер. Фур'є-спектрограми зручні для виявлення періодичної стаціонарної компоненти сигналу. Якщо такою компонентою є гармонічний сигнал, то для його представлення достатня всього одна гармоніка ряду Фур'є.

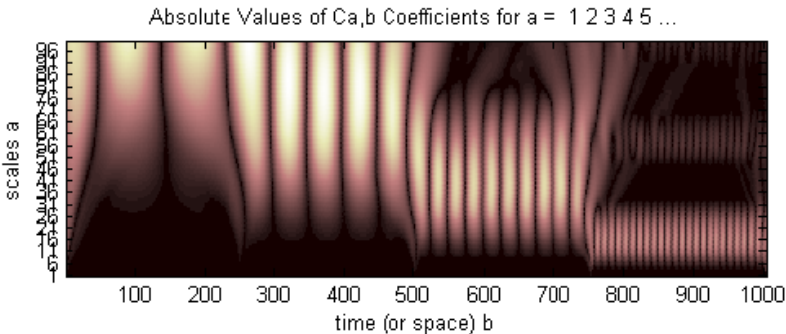


Рисунок 8.3 – Вейвлет-спектр сигналу з особливістю: послідовно стиковані в часі синусоїди із частотами 5, 10, 20 і 50 Гц

При вейвлет-аналізі навіть гармонічного сигналу прийдеться використовувати багато вейвлетів, що перекривають усю вісь часу. Це може здатися невиправданим ускладненням для даного часткового випадку, коли сигнал – проста синусоїда. Але якщо сигнал має локальні особливості, якщо він не періодичний і не стаціонарний, то вейвлет-аналіз

стає саме тим інструментом для його дослідження, до якого класичний Фур'є-аналіз не пристосований. Більші рівні вейвлет- коефіцієнтів характерні тільки для області, що розташовується поблизу тієї або іншої локальної особливості функції або сигналу. Це дозволяє легко визначати місце розташування локальних особливостей, що практично неможливо при Фур'є-спектральному аналізі, тобто вейвлет-спектрограми дають уявлення про досліджуваний процес у амплітудно-часовій області.

8.3. Властивості вейвлет-перетворення

Коефіцієнти вейвлет-перетворення містять комбіновану інформацію про те, що аналізує вейвлет й аналізований сигнал. Незважаючи на це, вейвлет-аналіз дозволяє одержати й об'єктивну інформацію про аналізований сигнал, тому що деякі властивості вейвлет-перетворення не залежать від вибору аналізуючого вейвлета. Незалежність від аналізатора робить ці прості властивості перетворення дуже важливими.

Основними елементарними властивостями вейвлет-перетворення функції $f(t)$ є: (при цьому будемо використовувати позначення $W[f] = W(a, b)$).

Лінійність.

$$W[\alpha f_1(t) + \beta f_2(t)] = \alpha W[f_1] + \beta W[f_2] = \alpha W_1(a, b) + \beta W_2(a, b).$$

Звідси, зокрема, впливає, що вейвлет-перетворення векторної функції є вектор з компонентами, що представляють собою вейвлет-перетворення кожної з компонентів аналізованого вектора окремо.

Інваріантність щодо зсуву.

$$W[f(t - b_0)] = W(a, b - b_0).$$

Із цієї властивості випливає комутативність диференціювання, зокрема, $\partial_t W[f] = W[\partial_t f]$ (тут $\partial_t = \partial/\partial t$). Разом з першою властивістю це означає перестановлюваність і для похідних векторного аналізу.

Інваріантність щодо розтягання (стискування).

$$W\left[f\left(\frac{t}{a_0}\right)\right] = \frac{1}{a_0} W\left(\frac{a}{a_0}, \frac{b}{a_0}\right).$$

Ця властивість дозволяє, зокрема, визначити наявність і характер особливостей аналізованої функції.

Крім трьох описаних незалежних від вибору аналізуючого вейвлета елементарних властивостей вейвлет-перетворення має ще декілька важливих і корисних властивостей.

Диференціювання.

$$W\left[\partial_t^m f\right] = (-1)^m \int_{-\infty}^{\infty} f(t) \partial_t^m \left[\left(\psi_{ab}^*(t)\right)\right] dt.$$

Таким чином, щоб проігнорувати, наприклад, великомасштабні поліноміальні складові та проаналізувати особливості високого порядку або дрібномасштабні варіації функції f , байдуже, чи диференціювати потрібне число раз аналізуючий вейвлет або саму функцію. Це дуже корисна властивість, якщо врахувати, що часто функція f являє собою ряд чисел, а аналізуючий вейвлет заданий формулою.

Для вейвлет-перетворення існує **аналог теореми Парсеваля** й виконується рівність

$$\int f_1(t) f_2^*(t) dt = C_{\psi}^{-1} \iint W_1(a,b) W_2^*(a,b) \frac{dadb}{a^2},$$

звідки випливає, що енергія сигналу може бути обчислена через амплітуди (коефіцієнти) вейвлет-перетворення подібно тому, як вона обчислюється через значення коефіцієнтів перетворення Фур'є:

$$E_f = \int f^2(t) dt = \frac{1}{2\pi} \int |A(\omega) - jB(\omega)|^2 d\omega.$$

Частотно-часова локалізація. Вірніше було б говорити про масштабно-часову локалізацію. Перетворення Фур'є та ряди Фур'є є прекрасним математичним апаратом для фізичної інтерпретації процесів при аналізі сигналів, що їх характеризують. Однак іноді вони виявляються недостатньо ефективними. Так, щоб одержати спектральну інформацію на обраній частоті, необхідно мати й минулу, і майбутню часову інформацію; до того ж перетворення Фур'є не враховує, що частота може еволюціонувати згодом. Воно, наприклад, не відрізняє сигнал, що представляє собою суму двох синусоїд з різними частотами, від сигналу, що складається з тих же синусоїд, що йдуть послідовно одна за іншою.

Для одержання високочастотної інформації з гарною точністю важливо витягати її з відносно малих часових інтервалів, а не із усього сигналу; і навпаки – низькочастотну спектральну інформацію витягати з відносно протяжних часових інтервалів сигналу.

Частина описаних труднощів знімається при використанні віконного перетворення Фур'є. Однак нескінченно осцилююча базисна функція (синусоїда) не дозволяє одержувати по-справжньому локалізовану інформацію.

Елементом базису вейвлет-перетворення є добре локалізована функція, що швидко прагне до нуля поза невеликим інтервалом, що дозволяє провести „локалізований спектральний аналіз”. Іншими словами, вейвлет-перетворення автоматично має рухливе частотно-часове вікно, вузьке на малих масштабах і широке на більших масштабах.

Розглянемо параметри частотно-часового вікна вейвлет-перетворення.

Уведемо функцію $\eta(\omega) = \hat{\psi}(\omega + \langle\omega\rangle)$. Вона є віконною функцією із центром у нулі й радіусом $\Delta\hat{\psi}$. Скориставшись

рівністю $\langle f, g \rangle = \hat{f} \cdot \hat{g} / (2\pi)$, можна записати вейвлет-перетворення для Фур'є-образа \hat{f} у вигляді:

$$W(a, b) = |a|^{-1/2} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{jb\omega} \eta^* \left(a \left[\omega - \frac{\langle \omega \rangle}{a} \right] \right) d\omega.$$

Якщо відволіктися від фазового зсуву, то стане очевидним, що останнє перетворення дає про спектр $\hat{f}(\omega)$ сигналу $f(t)$ локалізовану інформацію з „частотним вікном”

$$[\text{win}_\omega] = \left[\frac{\langle \omega \rangle}{a} - \frac{1}{a} \Delta_{\hat{\psi}}, \frac{\langle \omega \rangle}{a} + \frac{1}{a} \Delta_{\hat{\psi}} \right].$$

Частотна локалізація відбувається близько центру вікна в $\langle \omega \rangle / a$ із шириною вікна, рівної $2\Delta_{\hat{\psi}} / a$.

Помітимо, що відношення центральної частоти до ширини вікна,

$$\frac{\langle \omega \rangle}{a} \left(\frac{2\Delta_{\hat{\psi}}}{a} \right)^{-1} = \frac{\langle \omega \rangle}{2\Delta_{\hat{\psi}}}$$

не залежить від місця розташування центральної частоти, а частотно-часове вікно $[\text{win}_t] \cdot [\text{win}_\omega]$, що має площу $4\Delta_\psi \Delta_{\hat{\psi}}$, звужується при високій центральній частоті $\langle \omega \rangle / a$ й розширюється при низкій (рис. 8.4, а).

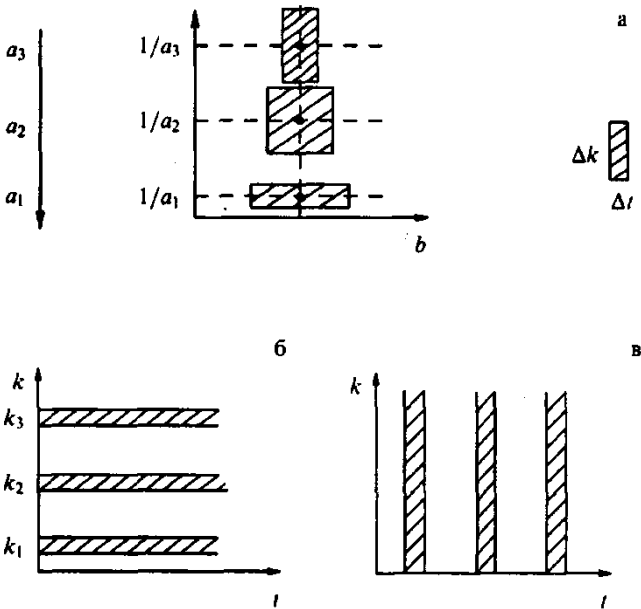


Рисунок 8.4 – Частотно-часова локалізація перетворень с різними аналізаторами:

а) вейвлети; б) синусоїди; в) функції Кронекера

Для порівняння на рис. 8.4 показана локалізація в частотно-часовому просторі перетворень із іншими аналізуючими функціями: перетворення Фур'є (рис. 8.4, б) і перетворення Шеннона, аналізуючою функцією якого є функція Кронекера (рис. 8.4, в).

З рис. 8.4 видно, що перетворення Фур'є добре локалізує частоту, але без часового відрізнення; перетворення Шеннона не має частотну локалізацію при відмінній часовій; вейвлет-перетворення має рухливе вікно, локалізоване близько обраного моменту часу, що розширюється з ростом масштабу, що і є найбільш бажаним при одержанні спектральної інформації. Визначається це тим, що базисними функціями перерахованих вище перетворень є відповідно: локалізуюча

тільки частоту синусоїдальна хвиля, локалізуюча тільки момент часу функція Кронекера та досить добре локалізуюча і масштаб, і момент часу вейвлет-функція. Вейвлет-перетворення являє собою як би безперервний банк віконних перетворень Фур'є з різними вікнами для кожної частоти.

Ця властивість вейвлет-перетворення дає йому велику перевагу при аналізі сигналів, тому що швидкі варіації сигналів (високочастотні характеристики) добре локалізовані, а для виявлення повільно мінливих характеристик досить гарного низькочастотного відрізнення. Вейвлет-перетворення, що має рухливе частотно-часове вікно, однаково добре виявляє низькочастотні і високочастотні характеристики сигналів.

Наявність кута впливу або достовірності. Можна відзначити, що значення аналізованої функції в точці t_0 впливає на значення коефіцієнтів перетворення в зростаючих з масштабом часових діапазонах у так званому куті впливу (рис. 8.5, а). Кут впливу виглядає інакше, якщо масштаб міняється не лінійно, а за іншим законом (логарифмічному, статечному).

У свою чергу, коефіцієнт $W(a, b)$ у точці (a_0, b_0) залежить від значень частини ряду з обумовленого тим же кутом впливу часового діапазону (діапазону інтегрування) близько b_0 (рис. 8.5, б), і діапазон цей тим більше, чим більше масштаб a_0 , тобто високочастотна (або дрібномасштабна) інформація обчислюється на основі малих відрізків ряду, а низькочастотна – на основі великих.

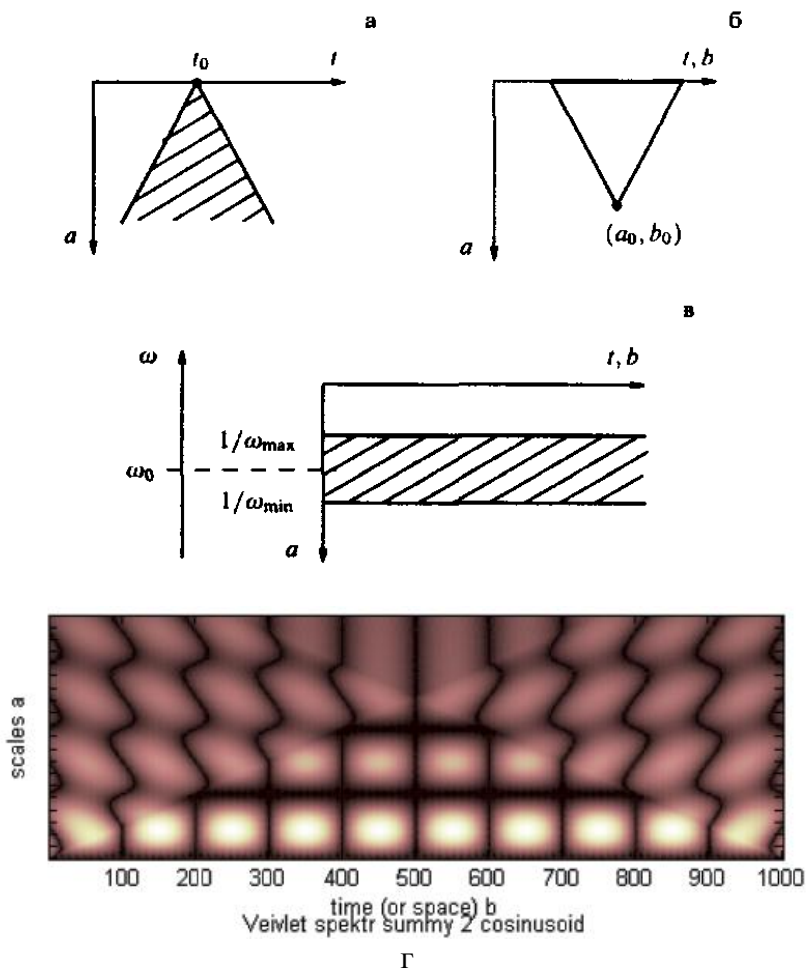


Рисунок 8.5 – Кут впливу або достовірності (а, б), смуга впливу або достовірності (в), зображення кута достовірності на вейвлет спектрі суми двох синусоїд (г)

Максимальний кут впливу (кут достовірності) відзначає область достовірності – область, поза якої коефіцієнти $W(a, b)$ обчислюються вже по відрізках, що виходять за межі ряду (по доповненому ряду). Оскільки аналізований ряд завжди скінченний, і необхідно одержати максимум

інформації, отримані поза максимальним кутом впливи (з деякою помилкою) наближені значення коефіцієнтів часто зберігаються. Для зменшення помилки ряд доповнюється з урахуванням його поведінки (середнім значенням, відомим часовим ходом і т.д.).

Вейвлет-перетворення може бути записане також через образи Фур'є сигналу $\hat{f}(\omega)$ й вейвлета $\hat{\psi}(\omega)$. Легко показати, що вплив компоненти Фур'є $\hat{f}(\omega_0)$ випробовують коефіцієнти $W(a,b)$, що перебувають у горизонтальній смузі $\omega_{\min} < a\omega_0 < \omega_{\max}$ (рис. 8.5, в); у свою чергу, на коефіцієнт $W(a,b)$ у точці (a_0, b_0) впливають усі компоненти Фур'є $\hat{f}(\omega)$ сигналу, для яких $\omega_{\min} < a_0\omega < \omega_{\max}$.

Значення кута впливу й ширина смуги впливу залежать від базисного вейвлета. Так, наприклад, МНАТ-вейвлет добре локалізований у часі та має вузький енергетичний спектр. Ця влада якості означає, зокрема, що коефіцієнти $W(a,b)$ залежать від малого інтервалу області частот вейвлета, тобто ширина смуги впливу $(\omega_{\min}, \omega_{\max})$ невелика.

8.4. Класифікація та приклади вейвлетів

Вибір вейвлетів досить великий. Широко використовуються близько півтора десятків базових типів вейвлетів і безліч варіантів для ряду базових типів вейвлетів [39]. Однак необгрунтоване застосування того або іншого типу вейвлета може привести до неотримання очікуваного результату.

Достаток класів вейвлетів і їх конкретних реалізацій у кожному класі дозволяє легко використовувати для розв'язання кожного практичного завдання найбільш підходящий тип вейвлета. У цей час створення нових типів вейвлетів раціонально тільки при розв'язанні нових і досить

специфічних завдань і за умови, що для цього не вдається підібрати підходящий тип вейвлета із числа описаних.

Тому виникає необхідність провести класифікацію вейвлетів по певних ознаках. Це дозволить підбирати найбільш підходящі типи вейвлетів для розв'язання конкретних завдань обробки сигналів і зображень. Але й подібний вибір треба розглядати як попередню операцію. При розв'язанні складних завдань в області обробки сигналів і зображень необхідно застосування хоча б декількох типів вейвлетів з наступним порівнянням результатів і вибором найкращих з них.

Грубі вейвлети або передвейвлети.

До „грубих” вейвлетів або передвейвлетів відносяться вейвлети Гауссового типу, вейвлети Морле та вейвлет „мексиканський капелюх”.

Вони мають мінімум властивостей, якими повинні володіти вейвлети, забезпечуючи повноцінні можливості у техніці перетворення сигналів:

- симетричність,
- функція ψ задається явно (має аналітичне вираження);
- функція φ в них відсутня (не має аналітичного вираження);
- аналіз не є ортогональним;
- функція ψ не має компактного носія;
- можливість реконструкції не гарантована;
- можлива безперервна декомпозиція;
- швидкі алгоритми перетворень і точна реконструкція неможливі.

Нескінченні регулярні вейвлети.

До нескінченних регулярних вейвлетів належать вейвлети Мейера.

Вони мають наступні властивості:

- мають функцію φ і їх аналіз ортогональний;
- функції ψ й φ не визначені явно;

- функції ψ й φ не мають компактного носія;
- вейвлети є симетричними та регулярними в нескінченності, тобто нескінченно диференціюючими;
- швидкий алгоритм перетворень не підтримується.
У цих вейвлетів можливі наступні методи аналізу:
- безперервні перетворення;
- дискретні перетворення, але побудувати для них фільтри з кінцевою імпульсною характеристикою неможливо.

Ще один вейвлет цього типу – дискретний вейвлет Мейера. Його властивості:

- можлива апроксимація фільтром з кінцевою імпульсною характеристикою;
- підтримка безперервного та дискретного перетворень.

Ортогональні вейвлети з компактим носієм.

До цих вейвлетів відносяться вейвлети Добеши, симлети та койфлети.

Їхні основні властивості:

- функція φ є й аналіз відноситься до ортогонального типу;
- функції ψ та φ мають деяке число моментів зникнення (нульових моментів);
- функції ψ та φ мають компактний носій;
- можливі безперервні перетворення та дискретні перетворення із застосуванням швидкого вейвлет-перетворення;
- забезпечується принципова можливість реконструкції сигналів і функцій;
- можлива апроксимація фільтром з кінцевою імпульсною характеристикою.

Деякі труднощі: недостатня періодичність (слабка регулярність).

Специфічні властивості:

- вейвлети Добеши несиметричні;
- вейвлети симлети близькі до симетричних;
- вейвлети койфлети: відсутність симетрії, функцій ψ та φ ,

наявність моментів зникнення (кілька нульових моментів).

Біортогональні парні вейвлети з компактним носієм.

До них відносяться B -сплайнові біортогональні вейвлети.

Вони мають наступні властивості:

- функція ϕ , ψ , і аналіз відноситься до біортогонального типу;
- обидві функції ψ та ϕ для декомпозиції і реконструкції мають компактний носій;
- функції ψ та ϕ для декомпозиції мають моменти зникнення (нульові моменти);
- функції ψ та ϕ для реконструкції можуть мати періодичність.

Можливі види аналізу – безперервне перетворення й дискретне перетворення з використанням алгоритму швидкого вейвлет-перетворення.

Найбільш суттєві переваги:

- симетрія з фільтрами,
- бажані властивості для розкладання та відновлення розділені.

Найбільші труднощі: відсутність ортогональності.

Комплексні вейвлети.

До комплексних відноситься досить більша група вейвлетів: Гауса, Морле, Шеннона та частотні B -сплайнові вейвлети.

Вони мають мінімальні властивості:

- функція ϕ відсутня;
- аналіз не ортогональний;
- функція ψ не має компактного носія;
- властивості реконструкції не гарантуються;
- можливий аналіз типу безперервної комплексної декомпозиції;

Основні хороші властивості: симетрія, ψ має явне вираження.

Труднощі застосування: швидкий алгоритм і

реконструкція неможливі.

Приклади вейвлетів. У системах комп'ютерної математики найбільше часто використовуються наступні вейвлети: Хаара, Добеши, симлети (симетричні вейвлети Добеши), койфлети (вейвлети Добеши з досить гарною симетрією), Мейєра, дискретна апроксимація вейвлету Мейєра, Гауса, комплексний вейвлет Гауса, Морле, комплексний вейвлет Морле, Шеннона, біортогональний, зворотний біортогональний, мексиканський капелюх, частотний B -сплайновий і деякі інші. Варіанти представлення різних вейвлетів за допомогою системи комп'ютерної математики наведені на рис. 8.6 – 8.8.

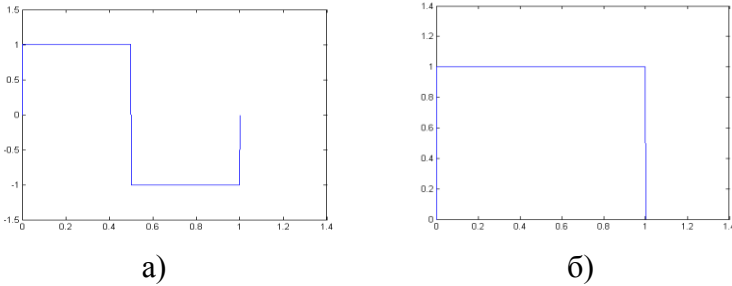


Рисунок 8.6 – Вейвлет Хаара:
а) вейвлет; б) масштабуюча функція

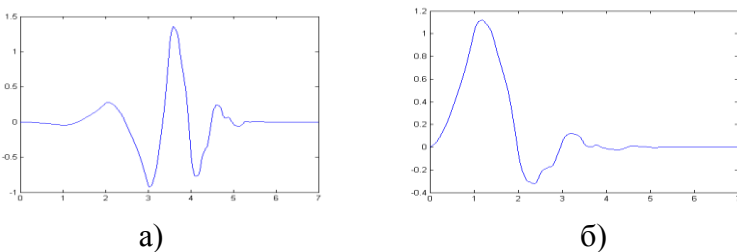


Рисунок 8.7 – Вейвлет Добеши (db4):
а) вейвлет; б) масштабуюча функція

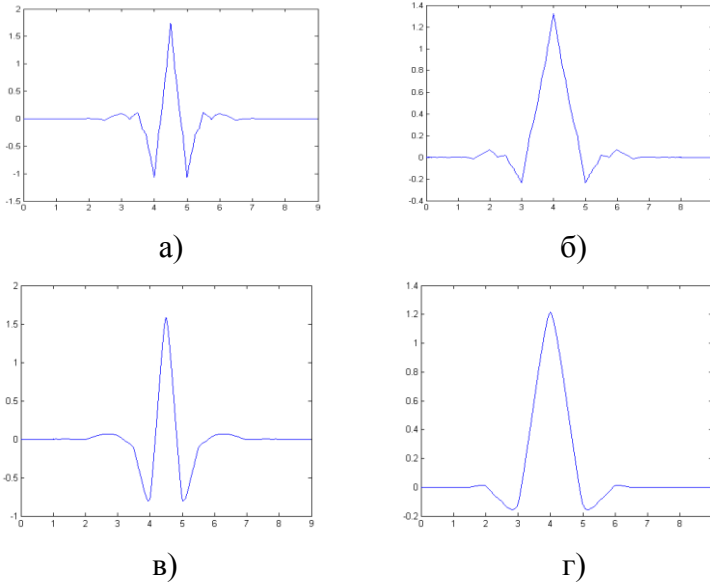


Рисунок 8.8 – Біортогональні вейвлети (bior4.4):
 а) вейвлет розкладання; б) масштабуюча функція розкладання;
 в) вейвлет відновлення; г) масштабуюча функція відновлення

8.5. Можливості вейвлет-аналізу

Маючи вейвлет-спектри, можна обчислити корисні характеристики досліджуваного процесу та проаналізувати багато його властивостей [40]. Розглянемо докладніше деякі можливості аналізу особливостей сигналу і його енергетичних характеристик.

Аналіз локальної регулярності. Розглянемо деякі наслідки властивості масштабною інваріантності.

Вейвлет-перетворення так улаштоване, що $W(a, t)$ – регулярна функція навіть при нерегулярній $f(t)$. Уся інформація про можливу особливість $f(t)$ (локалізація t_0 ,

інтенсивність s , показник α) укладена в асимптотичній поведінці коефіцієнтів $W(a, t_0)$ при малих a . Якщо коефіцієнти на малих масштабах розходяться, то f має особливість у t_0 й показник сингулярності α визначається нахилом залежності $\log|W(a, t_0)|$ до $\log a$. Якщо вони, навпаки, близькі до нуля в околиці t_0 на малих масштабах, то f в точці t_0 регулярна.

Описана властивість часто й успішно використовується при аналізі фрактальних і мультифрактальних сигналів. Базис перетворення самоподібний і коефіцієнти перетворення масштабуються з тим же показником, що й аналізована функція:

$$W(\lambda a, t_0 + \lambda b) \approx \lambda^{\alpha(t_0)} W(a, t_0).$$

Звідси легко одержати шкальний (масштабний) показник $\alpha(t_0)$, який пов'язаний із фрактальною розмірністю множини. Аналіз мультифрактальної множини дозволяє визначити спектр показників і спектр розмірностей.

Можна відзначити, що аналіз локальної регулярності в деякому змісті універсальний – він не залежить від вибору аналізуючого вейвлета.

Енергетичні характеристики. З існування для вейвлет-перетворення аналога рівності Парсеваля випливає, що в просторі дійсних функцій повна енергія сигналу f може бути записана через амплітуди вейвлет-перетворення у вигляді

$$E_f = \int f^2(t) dt = C_{\psi}^{-1} \iint W^2(a, b) \frac{dadb}{a^2}.$$

Щільність енергії сигналу $E_W(a, b) = W^2(a, b)$

характеризує енергетичні рівні (рівні збудження) досліджуваного сигналу $f(t)$ в просторі $(a, b) = (\text{масштаб}, \text{час})$.

Вейвлет спектр може відбивати той факт, що енергія нерівномірно розподілена по масштабах – можуть виявитися виділені масштаби, він може демонструвати стаціонарну або нестаціонарну структуру аналізованого процесу з елементами квазіперіодичності з частотами, що еволюціонують і з діапазонами локальних періодичностей на різних масштабах.

Локальний спектр енергії. Однією з основних особливостей вейвлет-перетворення є можливість одержувати локалізовані характеристики та вивчати локальні властивості процесів. Природа вейвлет-перетворення така, що термін „локальний енергетичний спектр” має право на існування.

Знаючи щільність енергії $E_W(a, b)$, можна за допомогою вікна визначити локальну щільність енергії в точці b_0 (або t_0):

$$E_{\xi}(a, t_0) = \int E_W(a, b) \xi\left(\frac{b-t_0}{a}\right) dt.$$

Віконна функція ξ „підтримує” діапазон близько t_0 й задовольняє рівності

$$\int \xi(b) db = 1.$$

Якщо в якості ξ вибрати функцію Кронекера, то локальний спектр енергії буде мати вигляд:

$$E_{\delta}(a, t_0) = W^2(a, t_0).$$

Ця характеристика дозволяє проаналізувати часову динаміку передачі енергії процесу по масштабах – обмін енергією між складовими процесу компонентами різного масштабу в будь-який заданий момент часу.

Глобальний спектр енергії. Повна енергія розподілена по

масштабах відповідно до глобального спектра енергії коефіцієнтів вейвлет-перетворення

$$E_W(a) = \int W^2(a, b) db = \int E_W(a, b) db.$$

Його називають також шкалограмою (*англ.* scalogram) або дисперсією вейвлет-перетворення.

На рис. 8.9, а представлений спектр потужності $E_F(\omega)$ потоку іонів сонячного вітру та шкалограма E_W – глобальний спектр енергії коефіцієнтів вейвлет-перетворення того ж сигналу. Спектри досить добре узгодяться, але спектр, обчислений за коефіцієнтами вейвлет-перетворення, – набагато більше гладка крива.

Причина полягає в тому, що вейвлет-спектр енергії сигналу E_W відповідає згладженому спектру потужності E_F . Це можна показати, виразивши спектр енергії $E_W(a)$ через спектр енергії сигналу в просторі Фур'є $E_F(\omega) = |\hat{f}(\omega)|^2$:

$$E_W(a) = \int E_F(\omega) |\hat{\psi}(a\omega)|^2 d\omega.$$

Легко бачити, що шкалограма E_W відповідає спектру потужності E_F , згладженому на кожному масштабі спектром Фур'є аналізуючого вейвлета.

Вейвлет-перетворення, що надає як би часове розгорнення спектра, дозволяє одержати й більш локалізовану енергетичну інформацію. На рис. 8.9, б представлено чотири шкалограми $E_W(a)$, одержані не по всій довжині ряду, а окремо для кожної четвертої його частини. На тому ж рисунку для порівняння наведений глобальний спектр (на рис. 8.9, а він показаний залежно від частоти, на рис. 8.9, б шкалограми показані залежно від масштабу, який росте нагору). Енергія

сигналу визначається через спектр енергії рівністю

$$E_f = C_{\psi}^{-1} \int E_W(a) \frac{da}{a^2}.$$

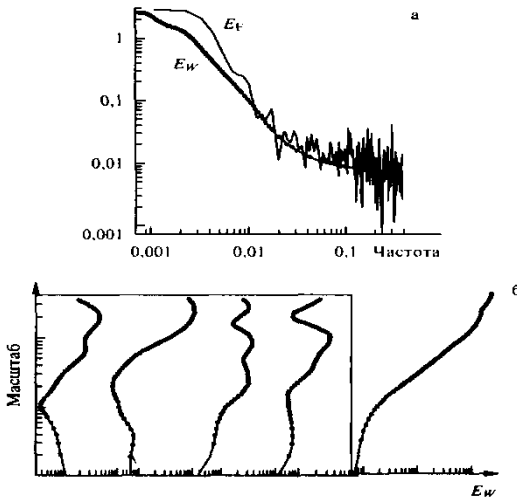


Рисунок 8.9 – Енергетичні спектри потоку іонів сонячного вітру: а) спектр потужності E_f і шкалограма E_W залежно від частоти; б) шкалограми для чотирьох часових діапазонів (у рамці) і повна шкалограмма залежно від масштабу.

Таким чином, величина E_f пропорційна площі під кривій $E_W(a)/a^2$, а шкалограмма відбиває відносний внесок різних масштабів у повну енергію та виявляє розподіл енергії процесу по масштабах.

Міра локальної перемежаємості – міра локальних відхилень від середнього поля спектрів на кожному масштабі; вона дозволяє визначити ступінь нерівномірності розподілу енергії по масштабах (кутовими дужками позначене усереднення):

$$I_W(a, t) = \frac{E_W(a, t)}{\langle E_W(a, t) \rangle_t}.$$

Рівність $I_W(a, t) = 1$ при всіх a і t означає, що енергія розподілена рівномірно й усі локальні спектри енергії однакові; $I_W(a, t_0) = \alpha$ означає, що внесок компоненти масштабу a в точці t_0 в α раз перевершує усереднений по всіх t .

Міра контрастності – дозволяє визначати навіть малі зміни в сигналі, коли необхідно, наприклад, виявити структурованість слабого сигналу або слабкі варіації на тлі великої структури (вбудовані структури):

$$C_W(a, t) = \frac{E_W(a, t)}{E'_W(a, t)},$$

$$\text{де } E'_W(a, t) = \int_{a'=0}^{a'=a} E_W(a', t) da'.$$

8.6. Вейвлет-обробка сигналів

Фур'є-обробка сигналів складається із двох основних етапів. Спочатку здійснюється розкладання сигналу $f(x)$, тобто знаходяться коефіцієнти Фур'є c_n . Вони аналізуються і, при необхідності, змінюються. Потім проводиться відновлення сигналу $f(x)$ по, може бути, зміненим коефіцієнтам c_n . Аналогічна ситуація й у випадку вейвлет-обробки [39].

Нехай $\{x_n\}$, $n \in Z$ – деякий сигнал. З метою аналізу відповідно до **частотного підходу до вейвлет-аналізу**

виділимо його низькочастотну $\{x_n^H\}$ та високочастотну $\{x_n^G\}$ частини.

При розкладанні сигналу одержуємо обидві компоненти $x(t) = x^H(t) + x^G(t)$.

Низькочастотна компонента $x^H(t)$ сигналу після фільтрації відрізняється від функції $x(t)$ всього сигналу тим, що її перетворення Фур'є $\hat{x}(\omega)$ має носій на проміжку $[-\pi/2, \pi/2]$. При розкладанні $A(\omega) = H(\omega)X(\omega)$ фільтр h із послідовності $\{x_n\}$ робить послідовність $\{a_n\}$ середніх арифметичних

$$a_n = \sum_{k \in \mathbb{Z}} h_k x_{n-k} = \frac{1}{2} x_n + \frac{1}{2} x_{n-1},$$

згладжуючи коливання послідовності $\{x_n\}$, тобто низькочастотний фільтр згладжує коливання.

Високочастотна компонента $x^G(t)$ виділяється поза проміжком $[-\pi/2, \pi/2]$ за допомогою фільтра g . Його передатна функція пригнічує низькі частоти (близькі до нуля) і пропускає високі. Тому в області значень дія цього фільтра $D(\omega) = G(\omega)X(\omega)$ дає новий сигнал

$$d_n = \sum_{k=0}^1 g_k x_{n-k} = \frac{x_n - x_{n-1}}{2},$$

який має сенс першої кінцево-різницевої похідної сигналу x , (точніше, її половини). Новий сигнал $\{d_n\}$ відбиває коливання послідовності $\{x_n\}$.

Таким чином, при аналізі сигналу одержимо

$$x_n = a_n + d_n.$$

Це означає, що сигнал $\{x_n\}$ розкладений на згладжену частину $\{a_n\}$ та високочастотну компоненту $\{d_n\}$.

По теоремі Котельникова-Шеннона функція $x^H(t)$ може бути відновлена по її парних значеннях $x^H(2n) = x_{2n}^H$. Отже, значення $x_{2n+1}^H = x^H(2n+1)$ з непарними номерами повністю визначаються парними елементами $\{x^H(2n)\}$. Тому при аналізі непарні значення з послідовності $\{x_n^H\}$ можуть бути без шкоди вилучені (децимація $2\downarrow$). Звичайно так і робиться з метою економії. Високочастотна компонента $x^G(t)$ сигналу після фільтрації також відповідно до теореми Котельникова-Шеннона може бути піддана операції децимації.

Вейвлет-розкладання сигналу проводиться по викладеній вище схемі, але багаторазово.

До сигналу $x = \{x_n\}$ застосовується низькочастотний фільтр розкладання $\{h_n\}$ з децимацією для одержання згладженої складової $A = \{A_k\}$. Потім застосовується високочастотний фільтр розкладання $\{g_n\}$ з децимацією для одержання деталей $D = \{D_k\}$, які були загублені при згладжуванні. Коефіцієнти $A = \{A_k\}$ називаються коефіцієнтами апроксимації першого рівня розкладання й позначаються $A_1 = \{A_{1,k}\}$. Коефіцієнти $D = \{D_k\}$ називаються, коефіцієнтами деталізації першого рівня розкладання й позначаються $D_1 = \{D_{1,k}\}$.

Далі процедура розкладання застосовується до набору коефіцієнтів A_1 і виходять коефіцієнти другого рівня

розкладання $A_2 = \{A_{2,m}\}$ й $D_2 = \{D_{2,m}\}$ (рис. 8.10).

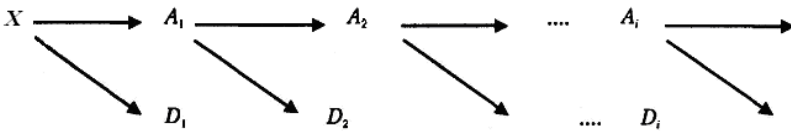


Рисунок 8.10 – Схема вейвлет-розкладання сигналу

Після кожного кроку кількість i апроксимуючих коефіцієнтів A і деталізуючих коефіцієнтів D зменшується у два рази.

Повторюючи необхідне число j раз процедуру розкладання сигналу, одержимо замість сигналу $x = \{x_n\}$ серію коефіцієнтів

$$(A_j, D_j, D_{j-1}, D_{j-2}, \dots, D_1).$$

По цьому набору коефіцієнтів можна будувати вейвлет-спектрограму сигналу для оцінки його особливостей.

Вейвлет-відновлення. Крім процедури аналізу (розкладання), повинна бути ще передбачена й процедура синтезу (відновлення, реконструкції) сигналу $\{x_n\}$ по його компонентах.

З рівняння $x_n = a_n + d_n$ також випливає, що подібним нескладним образом може бути здійснене відновлення (реконструкція) сигналу $\{x_n\}$.

З теорії фільтрів відомо, що при відновленні сигналів в операції підсумовування коефіцієнтів фільтрів беруть участь тільки парні коефіцієнти. Таким чином, відновлення сигналу $x(t)$ по низькочастотних компонентах $x^H(t)$ і

високочастотних компонентах $x^G(t)$ проводиться в наступному порядку:

- додаються непарні нульові значення в масиви $\{x^H(t)\}$ й $\{x^G(t)\}$ (операція зворотна децимації $2\uparrow$),
- до отриманих масивів $\{x^H(t)\}$ і $\{x^G(t)\}$ застосовується фільтрація за допомогою фільтрів відновлення $\{\tilde{h}_n\}$ й $\{\tilde{g}_n\}$

Таким чином, при вейвлет-розкладанні використовуються сполучені фільтри $\bar{H}(\omega)$ й $\bar{G}(\omega)$ вейвлетів, а для вейвлет-відновлення використовуються фільтри $H(\omega)$ й $G(\omega)$ вейвлетів $\varphi(x)$ і $\psi(x)$.

Таким чином, **вейвлет-аналіз** складається з наступних кроків:

- вибирається досить високий рівень відрізнення j , такий, що апроксимація $P_j(x)$ досить точно відбиває функцію $f(x)$. Практично тут проблем не виникає, оскільки функція $f(x)$ звичайно задається масивом значень. Ці значення й беруться за початкове наближення $P_j(f)$;
- вибирається глибина розкладання N й знаходяться коефіцієнти розкладання $(A_N, D_N, D_{N-1}, D_{N-2}, \dots, D_1)$, ці коефіцієнти аналізуються та змінюються, якщо буде потреба;
- відновлюється функція $f(x)$, з використанням можливо змінених коефіцієнтів розкладання:

$$f(x) \approx \sum_{k \in \mathbb{Z}} a_{j-N,k} \Phi_{j-N,k}(x) + \sum_{k \in \mathbb{Z}} d_{j-N,k} \Psi_{j-N,k}(x) + \dots + \sum_{k \in \mathbb{Z}} d_{j-1,k} \Psi_{j-1,k}(x).$$

Отже, для вейвлет-перетворення функції $f(x)$ необхідно обчислити серію коефіцієнтів $(A_N, D_N, D_{N-1}, D_{N-2}, \dots, D_1)$, причому кожний коефіцієнт знаходиться інтегруванням

$$a_{j-N,k} = \left\langle f, \Phi_{j-N,k} \right\rangle = \int_R f(x) \bar{\Phi}_{j-N,k}(x) dx,$$

$$d_{j-m,k} = \left\langle f, \Psi_{j-m,k} \right\rangle = \int_R f(x) \bar{\Psi}_{j-m,k}(x) dx,$$

де $m = 1, 2, \dots, N$, тому звідси й виникає проблема обчислення великої кількості інтегралів з необхідною точністю. Слід також ураховувати, що при високому рівні відрізнення, носії функцій $\Phi_{j,k}(x)$ і $\Psi_{j,k}(x)$ стають малими порядку $1/2^j$.

Швидке вейвлет-перетворення дозволяє розв'язати цю проблему. Алгоритм дає можливість обчислювати коефіцієнти вейвлет-розкладання без інтегрування, використовуючи алгебраїчні операції на основі згортки:

$$a_{j-1,k} = \sum_n \bar{h}_n a_{j,n+2k} \quad \text{і} \quad d_{j-1,k} = \sum_n \bar{g}_n a_{j,n+2k}$$

Ці рівності забезпечують швидкі алгоритми обчислення вейвлет-коефіцієнтів (каскадні або пірамідальні алгоритми, алгоритми Малла). Термін „швидкі” означає не тільки, що використовуються більш швидкі алгебраїчні процедури, але й те, що при кожному перетворенні загальне число нових коефіцієнтів не збільшується у два рази, а залишається колишнім.

Вейвлет-пакетна обробка. При застосуванні звичайного алгоритму Малла на кожному кроці „відрізається” половина низькочастотної частини діапазону сигналу. Реалізація цього

алгоритму виходить із загальної уяви про більшу інформаційність низькочастотної частини спектра сигналу, що для багатьох видів сигналу (наприклад, цифрових) не завжди справедливо. Існують процедури розкладання не тільки апроксимуючих коефіцієнтів A_j , але і деталізуючі коефіцієнтів D_j . Тоді коефіцієнти розкладання зображуються у вигляді дерева, нагорі якого перебуває аналізований сигнал $x = \{x_n\}$. Усі гілки, що йдуть уліво, указують на апроксимуючі коефіцієнти, а всі гілки, що йдуть вправо, указують на деталізуючі коефіцієнти попереднього вузла. Вузли, у яких перебувають коефіцієнти розкладання, зручно нумерувати так, як показано на рис. 8.11 або на рис. 8.12.

Звичайне вейвлет-розкладання та відповідні коефіцієнти виділені стовщеними лініями ліворуч на рис. 8.11. Вейвлет-коефіцієнтів досить для відновлення сигналу, тому можна розкласти високочастотні коефіцієнти деталей або відмовитися від цього. Таким чином, з'являється набагато більше можливостей для розкладання – від „мінімального” вейвлет-розкладання до повного пакетного розкладання на всіх рівнях.

Схемі на рис. 8.11 можна дати тлумачення стосовно до вейвлетів. У цьому випадку відбувається заміна вейвлета $\psi(t)$ на два нові вейвлета:

$$\psi_1(t) = \sum h_n \psi(t-n) \quad \text{і} \quad \psi_2(t) = \sum g_n \psi(t-n),$$

і так далі. Нові вейвлети теж локалізовані в часі (просторі), але на вдвічі більш широкому відрізку, чим вихідний вейвлет. Цьому бінарному дереву розкладання відповідає набір підпросторів з базисами, побудованими аналогічним образом. Функції, що породжують ці базиси, називаються *вейвлет - пакетами*.

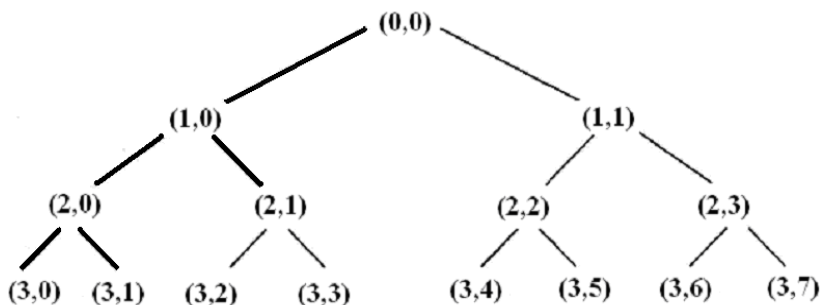


Рисунок 8.11 – Схема пакетного розкладання сигналу

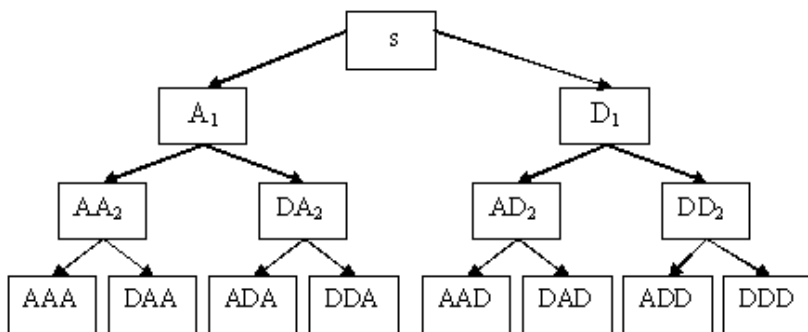


Рисунок 8.12 – Структура вейвлет-представлення сигналу для пакетних вейвлетів

Перетворення за допомогою вейвлет-пакетів є адаптивним вейвлет-перетворенням, оскільки воно легко пристосовується до особливостей сигналу й може успішно використовуватися для компресії (стискання) сигналів, очищення їх від шумів. Зокрема, є можливість відсікання деяких гілок деталізуючих коефіцієнтів дерева.

Перевагою, а якоюсь мірою й недоліком вейвлет-пакетів і адаптивних алгоритмів їх реалізації є відсутність необхідності в навчанні системи (характерному, наприклад, для систем на основі нейронних мереж) і навіть в оцінці статистичних характеристик сигналів. Усе, що потрібно, – це ввести оцінку

вартості вейвлет-коефіцієнтів, мірою якої може служити ентропія – концентрація кількості вейвлет-коефіцієнтів, що потрібні для опису сигналу з деякою заданою точністю (або погрішністю).

Існує багато варіантів базисів розкладання сигналу. З великої кількості представників потрібно вибрати той варіант базису, розкладання по якому представляє сигнал найбільше ефективно. Під „ефективним” мається на увазі те, що сигнал повинен бути представлений невеликою кількістю коефіцієнтів розкладання. Тобто базис для розкладання повинен бути таким, що більші коефіцієнти були б сконцентровані на невеликій кількості елементів вейвлет-пакетного базису, а більша частина коефіцієнтів була б близька до нуля.

У цьому випадку в якості *критерію для вибору* найефективнішого або *найкращого базису* для даного сигналу також використовується критерій мінімальності ентропії. Ентропія характеризує „розмазаність” сигналу або концентрацію кількості вейвлет-коефіцієнтів, що вимагаються для опису сигналу з деякою заданою точністю. Ентропія повинна мати властивість адитивності (об’єднання) стосовно сигналів.

Існує кілька варіантів визначення ентропії.

Розглянемо, як приклад, ентропію Шенона. Нехай s – сигнал і s_i – його значення. Вважаємо сигнал нормалізованим

$\sum_i s_i^2 = 1$. Ентропія Шенона визначається формулою

$$E(s) = - \sum_i s_i^2 \log_2(s_i^2) \text{ при наступному допущенні: } 0 \cdot \log 0 = 0$$

. Використання суми забезпечує властивість адитивності стосовно об’єднання масивів, що задають сигнали.

Кращий базис – це той базис, який дає найменшу кількість ентропії. Ентропія, як задано вище, дає критерій того, скільки ефективних компонентів необхідно, щоб представити сигнал у певному базисі. Наприклад, якщо в

деякому базисі при розкладанні одержуємо всі нульові коефіцієнти, крім одного (тобто сигнал збігається з базисною функцією), тоді ентропія досягає свого мінімального значення – нуля. З іншого боку, якщо в деякому базисі коефіцієнти розкладання всі однаково важливі й $s_i = 1/\sqrt{N}$, де N є довжиною сигналу, ентропія в цьому випадку має максимальне значення, рівне $\log_2 N$. Будь-яке інше розкладання перебуває між цими двома крайніми випадками. Таким чином, чим менше ентропія, тем менше істотних коефіцієнтів потрібно для представлення сигналу.

Практично для вибору оптимального базису критерій ентропії використовується в такий спосіб: якщо при розкладанні коефіцієнтів деякого вузла сума ентропій, отриманих при розкладанні компонентів, менше, чим ентропія коефіцієнтів у вихідному вузлі, то розкладання застосовується, а якщо ні, то коефіцієнти разом з базисними функціями залишаються без зміни.

Двовимірна вейвлет-обробка.

Двовимірні вейвлети. На практиці необхідно аналізувати сигнали (т.зв. зображення), які являють собою функції двох змінних. Тоді багатомасштабний аналіз повинен будуватися в просторі $L^2(\mathbf{R}^2)$ функцій двох змінних. Існує кілька способів узагальнити викладені раніше міркування на функції двох змінних. Найпростіший і широко розповсюджений шлях – це тензорний добуток одномірних вейвлетів [37]. Він заснований на тому факті, що простір $L^2(\mathbf{R}^2)$ є тензорним добутком просторів $L^2(\mathbf{R})$ функцій однієї змінної

$$L^2(\mathbf{R}^2) = L^2(\mathbf{R}) \otimes L^2(\mathbf{R}).$$

Це означає, що лінійні комбінації функцій виду $f(x) \cdot g(y)$ утворюють щільну множину у просторі $L^2(\mathbf{R}^2)$ функцій двох змінних.

За аналогією з одномірним кратномасштабним аналізом, нехай функції $\varphi(x)$ та $\psi(x)$ породжують ортогональний кратномасштабний аналіз простору $L^2(\mathbf{R})$:

$$\dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots$$

Зсуви $\varphi_{0,n}(x) = \varphi(x-n)$, $n \in \mathbf{Z}$ функції $\varphi(x)$ утворюють ортонормований базис простору V_0 . Інші простори $V_j \in$ масштабованими версіями V_0 , $f(x) \in V_0 \Leftrightarrow f(2^j x) \in V_j$. Зокрема, V_j має в якості ортонормованого базису систему функцій $\varphi_{j,k}(x) = \sqrt{2^j} \varphi(2^j x - k)$.

Простір вейвлетів W_0 визначається як ортогональне доповнення до V_0 в підпросторі V_1 . Тоді $V_1 = V_0 \oplus W_0$ – пряма сума ортогональних підпросторів. У просторі W_0 ортонормований базис утворюють зсуви $\psi_{0,n}(x) = \psi(x-n)$, $n \in \mathbf{Z}$ функції $\psi(x)$. У випадку кожного $j \in \mathbf{Z}$, маємо $V_j \subset V_{j+1}$ й $V_{j+1} = V_j \oplus W_j$.

Простори вейвлетів $W_j \in$ масштабованими версіями W_0 , зокрема W_j має ортонормований базис із функцій $\psi_{j,k}(x) = \sqrt{2^j} \psi(2^j x - k)$.

У двовимірному випадку визначимо простір V_0^2 як тензорний добуток просторів V_0 функцій однієї змінної:

$$V_0^2 = V_0 \otimes V_0.$$

Інакше кажучи, простір V_0^2 є замикання лінійної оболонки, породженої зсувами $\varphi_{0,k,n}(x, y) = \varphi(x-k)\varphi(y-n)$, $k, n \in \mathbf{Z}$.

Ці функції утворюють ортонормований базис простору V_0 .
Простори $V_j^2 = V_j \otimes V_j$ є масштабованими версіями простору V_0^2 :

$$f(x, y) \in V_0^2 \Leftrightarrow f(2^j x, 2^j y) \in V_j^2.$$

Тоді ми одержуємо послідовність замкнених вкладених друг у друга підпросторів $\dots \subset V_{-1}^2 \subset V_0^2 \subset V_1^2 \subset \dots$, що мають аналогічні властивості.

Оскільки $V_1 = V_0 \oplus W_0$, то для тензорного добутку маємо наступне ортогональне розкладання:

$$\begin{aligned} V_1^2 &= V_1 \otimes V_1 = (V_0 \oplus W_0) \otimes (V_0 \oplus W_0) = \\ &= (V_0 \otimes V_0) \oplus (V_0 \otimes W_0) \oplus (W_0 \otimes V_0) \oplus (W_0 \otimes W_0). \end{aligned}$$

Тут $V_0 \otimes V_0 = V_0^2$. Тому інші три простори в сумі утворюють простір W_0^2 двовимірних вейвлетів. При цьому:

- простір $V_0 \otimes W_0$ породжений зсувами функції $\psi^H(x, y) = \varphi(x)\psi(y)$,
- простір $W_0 \otimes V_0$ породжений зсувами функції $\psi^V(x, y) = \psi(x)\varphi(y)$,
- простір $W_0 \otimes W_0$ породжений зсувами функції $\psi^D(x, y) = \psi(x)\psi(y)$.

Позначимо ці простори в такий спосіб:

$$W_0^H = V_0 \otimes W_0, W_0^V = W_0 \otimes V_0, W_0^D = W_0 \otimes W_0.$$

Їхня сума дає весь простір вейвлетів W_0^2 . У підсумку одержуємо ортогональне розкладання

$$V_1^2 = V_0^2 \otimes W_0^2 = V_0^2 \oplus W_0^H \oplus W_0^V \oplus W_0^D.$$

Відповідне розкладання виходить для будь-якого $j \in \mathbf{Z}$

$$V_{j+1}^2 = V_j^2 \otimes W_j^2 = V_j^2 \oplus W_j^H \oplus W_j^V \oplus W_j^D.$$

Системи функцій, які утворюють ортонормовані базиси в

зазначених вище просторах, будуть мати вигляд:

$$\varphi_{j,k,n}(x, y) = 2^j \varphi(2^j x - k) \varphi(2^j y - n);$$

$$\psi_{j,k,n}^H(x, y) = 2^j \varphi(2^j x - k) \psi(2^j y - n);$$

$$\psi_{j,k,n}^V(x, y) = 2^j \psi(2^j x - k) \varphi(2^j y - n);$$

$$\psi_{j,k,n}^D(x, y) = 2^j \psi(2^j x - k) \psi(2^j y - n).$$

Двовимірне вейвлет-перетворення. Для системи функцій, які утворюють чотири ортонормованих базиса, відповідно, і вейвлет-коефіцієнти сигналу $f(x, y)$ будуть чотирьох типів.

Коефіцієнти

$$cA_1 = \langle f, \varphi_{j,k,n} \rangle = 2^j \int_{\mathbf{R}^2} f(x, y) \overline{\varphi(2^j x - k) \varphi(2^j y - n)} dx dy$$

називаються апроксимуючими.

Інші коефіцієнти – деталізуючі трьох типів.

При знаходженні деталізуючих вейвлет-коефіцієнтів функції $f(x, y)$

$$cH_1 = \langle f, \psi_{j,k,n}^H \rangle = 2^j \int_{\mathbf{R}^2} f(x, y) \overline{\varphi(2^j x - k) \psi(2^j y - n)} dx dy,$$

по змінній x відбувається інтегрування з масштабуючою функцією $\varphi(x)$, а це, як відомо, викликає згладжування сигналу $f(x, y)$ по змінній x . Інтегрування по змінній y проводиться за участю вейвлета $\psi(y)$ – це дає коефіцієнти, які враховують зміни функції по змінній y . У результаті виходить сигнал cH_1 , у якому більше відбиті зміни по вертикалі. Будуть виділятися горизонтальні смуги. Тому такі коефіцієнти cH_1 називаються горизонтальними.

Аналогічний зміст і назву мають і інші вейвлет-коефіцієнти.

Коефіцієнти cV_1 – вертикальні,

$$cV_1 = \left\langle f, \psi_{j,k,n}^V \right\rangle = 2^j \int_{R^2} f(x, y) \overline{\psi(2^j x - k) \varphi(2^j y - n)} dx dy,$$

по змінній y відбувається інтегрування з масштабуючою функцією $\varphi(y)$, а це викликає згладжування сигналу $f(x, y)$ по змінній y . Інтегрування по змінній x проводиться за участю вейвлета $\psi(x)$ – це дає коефіцієнти, які враховують зміни функції по змінній x . У результаті виходить сигнал cV_1 , у якому більше відбиті зміни по горизонталі. Будуть виділятися вертикальні смуги.

Коефіцієнти cD_1 – діагональні,

$$cD_1 = \left\langle f, \psi_{j,k,n}^D \right\rangle = 2^j \int_{R^2} f(x, y) \overline{\psi(2^j x - k) \psi(2^j y - n)} dx dy,$$

інтегрування по змінній x проводиться за участю вейвлета $\psi(x)$ – це дає коефіцієнти, які враховують зміни функції по змінній x . Інтегрування по змінній y проводиться за участю вейвлета $\psi(y)$ – це дає коефіцієнти, які враховують зміни функції по змінній y . У результаті виходить сигнал cD_1 , у якому більше відбиті зміни як по осі x , так і по осі y , тобто по діагоналі. Будуть виділятися діагональні смуги. При цьому згладжування по обом осям відсутнє.

Таким чином, схема розкладання сигналу S може бути зображена у вигляді:

$$S \rightarrow (cA_1, cH_1, cV_1, cD_1) \rightarrow (cA_2, cH_2, cV_2, cD_2, cH_1, cV_1, cD_1) \rightarrow \dots$$

Практично вхідний сигнал задається матрицею S . При розкладанні цього сигналу виходять зазначені чотири типи коефіцієнтів у такий спосіб.

Наприклад, горизонтальні деталізуючі коефіцієнти cH_1 виходять згортокою рядків \mathbf{S} матриці з низькочастотним фільтром вейвлета та децимацією, а потім стовпців отриманої матриці з високочастотним фільтром і децимацією.

Інші коефіцієнти розкладання виходять аналогічним образом.

Оскільки масив початкових коефіцієнтів двовимірний, то природно графічно схему розкладання сигналу \mathbf{S} зобразити у вигляді плоскої картини (рис. 8.13):

Слід зазначити, що ця схема відбиває й зміну розмірів масивів коефіцієнтів розкладання. Якщо сигнал \mathbf{S} має розмірність $n \times n$, то коефіцієнти першого рівня розкладання cA_1, cH_1, cV_1, cD_1 будуть мати розмірність $n/2 \times n/2$ і т.д.

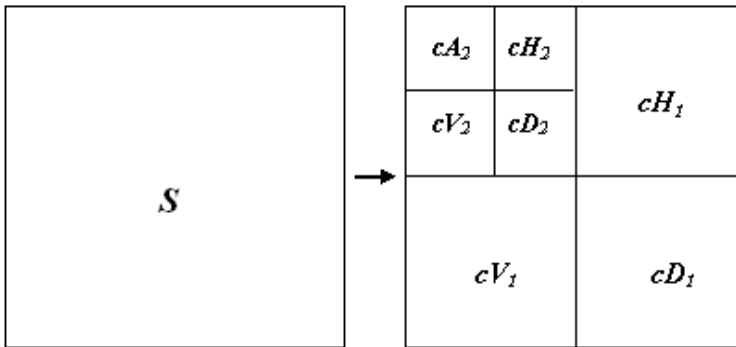


Рисунок 8.13 – Схема розкладання двовимірного сигналу

Двовимірне вейвлет-розкладання можна застосовувати не тільки до апроксимуючих коефіцієнтів cA_j , але й до всіх інших. Тоді з'являються пакетні вейвлети. Їхню схему розкладання прийнято зображувати у вигляді дерева, аналогічного дереву пакетних одновимірних розкладань.

8.7. Приклади вейвлет-обробки

При здійсненні *прямого дискретного вейвлет-перетворення* обчислюються апроксимуючі коефіцієнти, cA та деталізуючі коефіцієнти cD першого рівня вейвлет-перетворення сигналу, заданого вектором s . Коефіцієнти одержуються згорткою сигналу s з фільтром розкладання нижніх частот LoD для апроксимації та з фільтром розкладання високих частот HiD для деталізації, а потім супроводжуються двійковою децимацією. Обчислення прямого вейвлет-перетворення можливо як із застосуванням звичайних алгоритмів, так і за допомогою швидких алгоритмів. Швидкі алгоритми можливі для вейвлетів: Хаара, Добеши, симлетів, койфлетів, біортогональних, дуальних біортогональних, дискретних апроксимацій вейвлетів Мейєра.

Як приклад представлений однорівневий дискретний одновимірний вейвлет-аналіз гармонічного сигналу за допомогою двох вейвлетів – Хаара та Добеши. Сигнал представлений на рис. 8.14. Коефіцієнти cA та cD для вейвлета Хаара представлені на рис. 8.15, коефіцієнти cA та cD для вейвлета Добеши – на рис. 8.16.

Видна істотна різниця в коефіцієнтах, отриманих з використанням різних вейвлетів.

При здійсненні *зворотного однорівневого дискретного одновимірного вейвлет-перетворення* по відомих апроксимуючих коефіцієнтах cA та деталізуючих коефіцієнтах cD першого рівня обчислюються значення сигналу s з використанням прямого обчислення або фільтрів реконструкції LoR , HiR , тобто відновлення проводиться у зворотному порядку. У коефіцієнти cA та cD додаються нулі на непарні місця, і далі проводиться згортка отриманих масивів з фільтрами реконструкції LoR та HiR або пряме обчислення. Можна здійснити відновлення й по одному типу коефіцієнтів – тільки по апроксимуючих коефіцієнтах або тільки по деталізуючих коефіцієнтах.

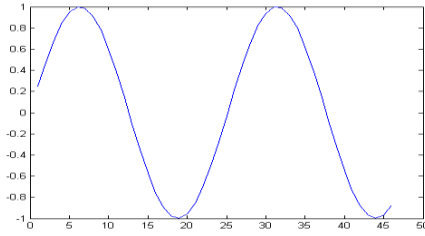
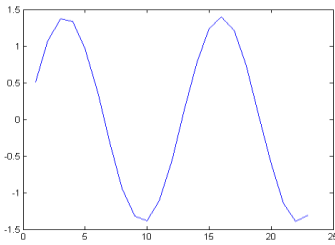
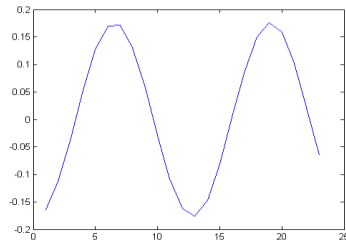


Рисунок 8.14 – Гармонічний сигнал



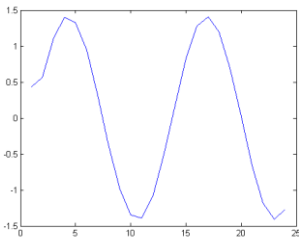
а)



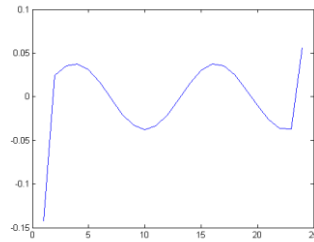
б)

Рисунок 8.15 – Коефіцієнти вейвлет-перетворення гармонічного сигналу за допомогою вейвлету Хаара (haar):

а) апроксимуючі cA ; б) деталізуючі cD



а)



б)

Рисунок 8.16 – Коефіцієнти вейвлет-перетворення гармонічного сигналу за допомогою вейвлету Добеши (db2):

а) апроксимуючі cA ; б) деталізуючі cD

Як приклад (рис. 8.17, 8.18) представлено однорівневе дискретне одновимірне вейвлет-відновлення гармонічного сигналу за коефіцієнтами, отриманими у попередньому прикладі, і оцінена якість відновлення, причому показником якості відновлення є максимальна сумарна абсолютна помилка.

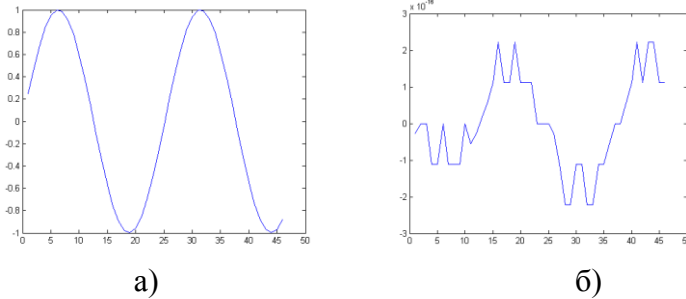


Рисунок 8.17 – Гармонічний сигнал (а), відновлений зворотним однорівневим дискретним одновимірним вейвлет-перетворенням за допомогою вейвлета Хаара і графік (б) його абсолютної помилки відновлення

Максимальна сумарна абсолютна помилка відновлення склала $8.3969\text{e-}016$.

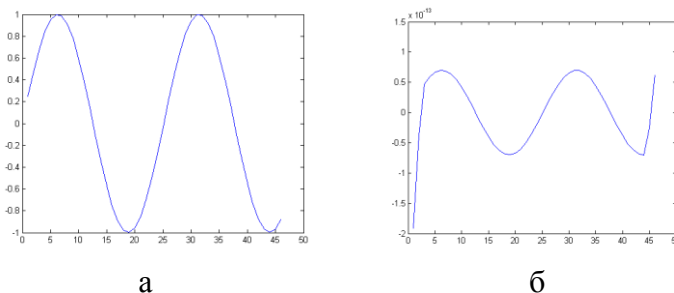


Рисунок 8.18 – Гармонічний сигнал (а), відновлений зворотним однорівневим дискретним одновимірним вейвлет-перетворенням за допомогою вейвлета Добеші (db2) і графік (б) його абсолютної помилки відновлення

Максимальна сумарна абсолютна помилка відновлення склала $3.8974e-013$.

У системах комп'ютерної математики передбачене також відображення коефіцієнтів одновимірного дискретного вейвлет-розкладання у вигляді спектрограм з кольоровим або яскравим відображенням значень вейвлет-коефіцієнтів. Для цього випадку для більшої наочності сигнал у вигляді суми двох синусоїд розкладемо до 5 рівня. Спектр його апроксимуючих коефіцієнтів буде виглядати, як показано на рис. 8.19, а, а спектр його деталізуючих коефіцієнтів – на рис. 18.9, б.

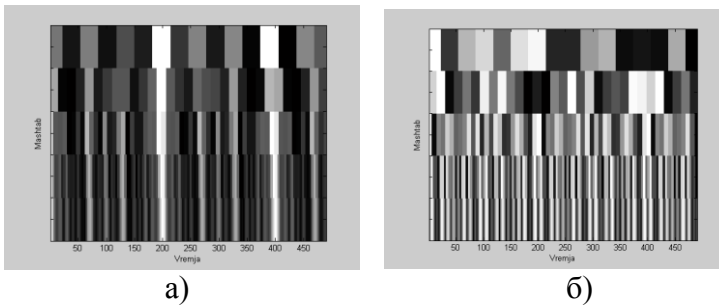


Рисунок 8.19 – Відображення коефіцієнтів одновимірного дискретного вейвлет-розкладання у вигляді спектрограм з яскравим відображенням значень вейвлет-коефіцієнтів

При *однорівневому прямому дискретному двовимірному вейвлет-перетворенні* для вхідної матриці S обчислюються апроксимуючі коефіцієнти cA та деталізуючі коефіцієнти першого рівня трьох типів: горизонтальні cH , вертикальні cV і діагональні cD , з використанням для цього тензорного добутку $w(x)w(y)$ вейвлета $w(x)$. Усі отримані коефіцієнти утворюють матриці розміром у два рази менше, чим вихідна матриця S .

Для прикладу відобразимо результат однорівневого прямого дискретного двовимірного вейвлет-перетворення

спутникового знімка у двох варіантах:

а) у вигляді яскравого зображення та тривимірної поверхні;

б) у вигляді яскравого зображення з нормованими коефіцієнтами розкладання.

Перший варіант відображення наведений на рис. 8.20 – 8.25.



Рисунок 8.20 – Вихідний супутниковий знімок для перетворення

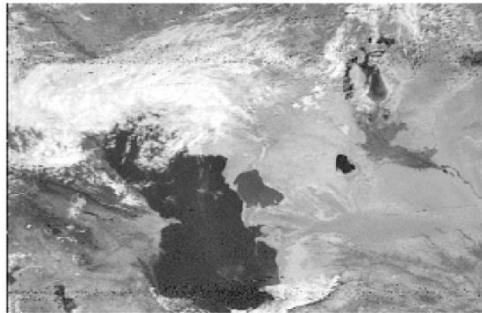


Рисунок 8.21 – Яскраве зображення апроксимуючих коефіцієнтів cA_1 однорівневого прямого дискретного двовимірного вейвлет-перетворення

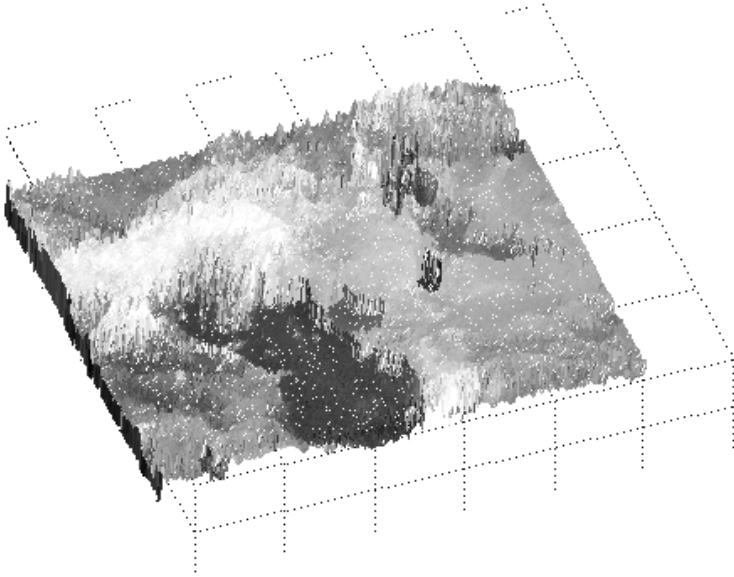


Рисунок 8.22 – Амплітудне зображення апроксимуючих коефіцієнтів sA_1 однорівневого прямого дискретного двовимірного вейвлет-перетворення

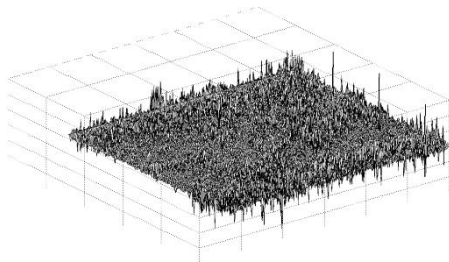
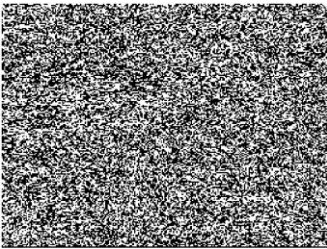


Рисунок 8.23 – Яскраве (ліворуч) та амплітудне (праворуч) зображення деталізуючих коефіцієнтів sH_1 однорівневого прямого дискретного двовимірного вейвлет-перетворення

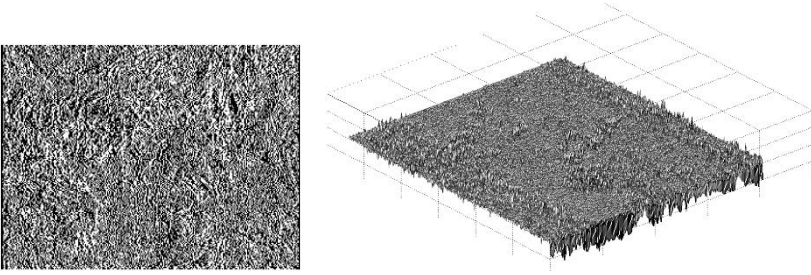


Рисунок 8.24 – Яскраве (ліворуч) та амплітудне (праворуч) зображення коефіцієнтів cV_1 однорівневого прямого дискретного двовимірного вейвлет-перетворення

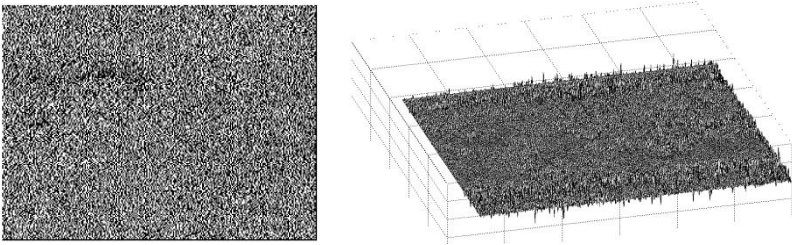


Рисунок 8.25 – Яскраве (ліворуч) і амплітудне (праворуч) зображення деталізуючих коефіцієнтів cD_1 однорівневого прямого дискретного двовимірного вейвлет-перетворення

Другий варіант відображення наведений на рис. 8.26.

На рис. 8.20 наведений вихідний супутниковий знімок району Каспійського моря. Спостерігається суша, ріки, велика водна поверхня, гори, хмарні системи. Угорі та унизу знімка добре видна прийнята високочастотна імпульсна радіозавада. На рис. 8.21, на якому зображені апроксимуючі коефіцієнти розкладання cA , картина явно згладжена за рахунок роботи масштабуючого фільтра (φ), усі контури обрублені, завади згладжені. На всіх трьох амплітудних зображеннях деталізуючих коефіцієнтів (рис. 8.23 – 8.25) чітко

локалізується в просторі імпульсна завада та видні високочастотні деталі зображення.

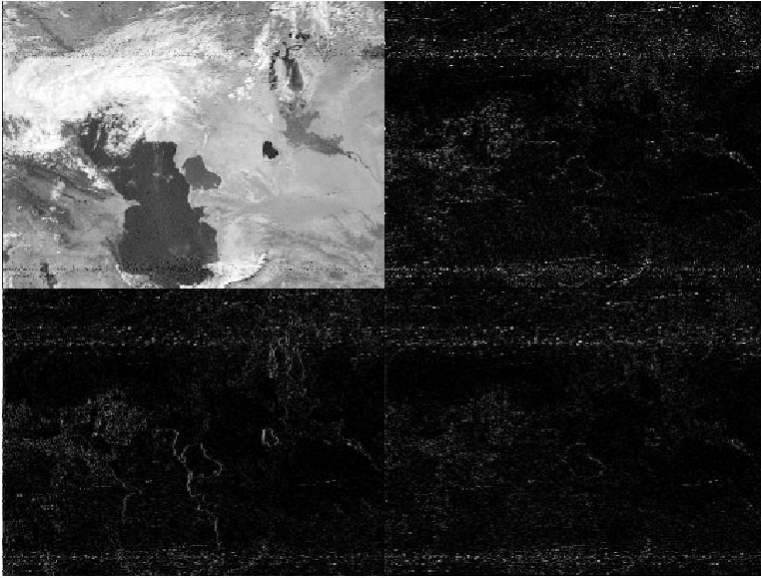


Рисунок 8.26 – Коефіцієнти однорівневого прямого дискретного двовимірного вейвлет-перетворення (ліворуч угорі – cA_1 , праворуч угорі – cH_1 , ліворуч унизу – cV_1 , праворуч унизу – cD_1)

На рис. 8.26 видно, що коефіцієнти cH_1 більшою мірою відбивають горизонтальні елементи зображення (праворуч угорі), коефіцієнти cV_1 більшою мірою відбивають вертикальні елементи рисунка (ліворуч унизу), а коефіцієнти cD_1 краще представляють діагональні лінії зображення (праворуч унизу). Особливо добре це видно на контурах затоки Кара-Богаз-Гол.

Контрольні запитання

1. У чому полягає спектрально – часовий аналіз?
2. У чому полягає ідея вейвлет – перетворення?
3. Яке визначення вейвлета?
4. Які ознаки вейвлета?
5. У чому полягає пряме безперервне вейвлет – перетворення?
6. У чому полягає зворотне безперервне вейвлет – перетворення?
7. У чому полягає діадне вейвлет – перетворення?
8. У чому полягає пряме дискретне вейвлет – перетворення?
9. У чому полягає зворотне дискретне вейвлет – перетворення?
10. Що таке вейвлет – фрейми?
11. Що таке кратномасштабний аналіз і синтез?
12. На яких передумовах базується кратномасштабний аналіз?
13. Яким чином представляються результати вейвлет – перетворення?
14. Які властивості вейвлет – перетворення?
15. У чому полягає властивість частотно – часової локалізації?
16. Що таке кут впливу або достовірності?
17. Як класифікуються вейвлети?
18. У чому полягають можливості вейвлет – аналізу?
19. У чому полягає частотний підхід до вейвлет – аналізу?
20. У чому полягає вейвлет – розкладання і вейвлет – відновлення?
21. Який зміст вейвлет – аналізу?
22. Що таке алгоритм Малла?
23. Що таке вейвлет – пакетна обробка?
24. Що використовується як критерій вибору якнайкращого базису для обробки даного сигналу?

25. Яким чином здійснюється двовимірна вейвлет – обробка?

Практичні завдання

1. У системі комп'ютерної математики здійснити моделювання двох синусоїд різної частоти таких, що містять від 5 до 50 коливань. Виконати їх складання і стикування. Побудувати вейвлет – спектри: а) окремих синусоїд, б) їх суми, в) стикованих синусоїд, застосувавши пряме безперервне вейвлет – перетворення. Параметри спектру встановити самостійно. Проаналізувати отриманий результат, зробити висновки.

2. Засобами операційної системи комп'ютера здійснити запис звуку власного голосу тривалістю 2 – 3 секунди у вигляді файлу звукового формату і засобами системи комп'ютерної математики побудувати вейвлет – спектр записаного звуку, застосувавши безперервне вейвлет – перетворення. Параметри спектру встановити самостійно. Проаналізувати отриманий результат, зробити висновки.

3. У системі комп'ютерної математики побудувати вейвлет – спектр будь-якого зображення, узятого в одному з графічних форматів. Параметри спектру встановити самостійно. Зображення повинне містити низькочастотні і високочастотні просторові складові. Здійснити низькочастотну фільтрацію зображення шляхом відкидання деталізуючих коефіцієнтів і відновлення зображення тільки по його апроксимуючих коефіцієнтах. Проаналізувати отриманий результат, зробити висновки.

Глава 9

ОПТИМАЛЬНІ І АДАПТИВНІ ЛІНІЙНІ СИСТЕМИ

В цій главі розглядаються питання оптимальної лінійної обробки сигналів у вимірювальному каналі. Критерій оптимальності визначається або потребою вимірювання окремих параметрів сигналу в умовах перешкод найкращим чином, або ж – відтворенням форми сигналу з найменшим рівнем спотворень.

Окрім систем обробки сигналів з незмінними незалежно від результатів попередніх вимірювань характеристиками (вони є розімкнутими системами), розглядаються також адаптивні системи, які враховують попередні результати для певної зміни характеристик системи з метою її оптимізації до умов вимірювання, тобто діють за принципом замкнутого контуру (зворотного зв'язку).

9.1. Оптимальна лінійна обробка сигналів

Задачу оптимальної обробки вимірювального сигналу відомої форми можна сформулювати наступним чином. Вхідний сигнал $s_0(t)$ в суміші з перешкодою $n(t)$ надходить на засіб вимірювання (ЗВ):

$$s(t) = s_0(t) + n(t). \quad (9.1)$$

У ЗВ $s(t)$ перетворюється в деякий інший сигнал $y(t)$ однозначно пов'язаний з $s(t)$

$$y(t) = \mathbf{T}[s(t)], \quad (9.2)$$

де $\mathbf{T}[\bullet]$ – оператор перетворення лінійної стаціонарної системи обробки сигналів.

Сигнал $y(t)$ в подальшому застосовується для вимірювання невідомих інформативних параметрів сигналу $s_0(t)$.

Лінійна стаціонарна система, що виконує обробку суми сигналу і перешкоди деяким найкращим чином, називається *оптимальним лінійним фільтром*.

Проблема оптимальної обробки суми відомого за формою сигналу і перешкоди виникає, наприклад, при застосуванні локаційного методу вимірювання. В цьому випадку прийнятий корисний сигнал $s_0(t)$ є масштабною копією зондувального сигналу $s_{\text{пер}}(t - \tau)$ із затримкою в часі τ , тобто [17]

$$s_0(t) = A s_{\text{пер}}(t - \tau), \quad (9.3)$$

де постійне число $A \ll 1$.

Амплітуда сигналу $s_0(t)$ може бути досить малою у порівнянні з ефективною напругою шуму на вході приймача.

Подібна задача виникає також в приймальному пристрої цифрового каналу зв'язку інформаційно-вимірювальної системи при передаванні радіоканалом кодованого вимірювального сигналу.

Приймальний пристрій в цих випадках перш за все встановлює факт присутності корисного сигналу (який, наприклад, при цифровому зв'язку відповідає нулю або одиниці) в прийнятому коливанні. У випадку імпульсної локації може бути виміряний час затримки сигналу τ , який є пропорційним відстані до об'єкта.

В цих випадках не потрібно зберігати форму корисного сигналу. Навпаки, бажано його обробити так, щоб на виході фільтра в деякий момент часу спостерігався значний «сплеск» миттєвих значень сигналу. При постановці задачі оптимізації, як правило, припускають, що сигнал перешкоди є стаціонарним гаусівським процесом (4.12), який характеризується малою ймовірністю великих викидів. Такою перешкодою, наприклад, є власний шум в приймальному каналі. Тому, якщо сигнал на виході системи в деякий момент часу істотно перевищує ефективну напругу шуму, ймовірність присутності корисного сигналу на вході приймача є великою.

Завдяки лінійності перетворення, сигнал $y(t)$ на виході фільтра складатиметься з суми двох сигналів $y_s(t)$ і $y_n(t)$. Перший з них є лінійним перетворенням корисного сигналу $s_0(t)$, а другий – перешкоди $n(t)$. Складову $y_s(t)$ можна подати у вигляді оберненого перетворення Фур'є

$$y_s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}_0(\omega) \dot{H}(\omega) e^{j\omega t} d\omega, \quad (9.4)$$

де $\dot{S}_0(\omega)$ – спектр корисного сигналу, $\dot{H}(\omega)$ – комплексна частотна характеристика фільтра (частотний коефіцієнт передачі).

Тоді миттєва потужність корисного сигналу на виході фільтра має вигляд

$$p_s(t) = |y_s(t)|^2 = \frac{1}{4\pi^2} \left| \int_{-\infty}^{\infty} \dot{S}_0(\omega) \dot{H}(\omega) e^{j\omega t} d\omega \right|^2. \quad (9.5)$$

Якщо припустити, що перешкода $n(t)$ є стаціонарним шумом з постійною спектральною густиною потужності N В²/Гц, середня потужність шуму на виході складе

$$W_n = \frac{N}{2\pi} \int_{-\infty}^{\infty} |\dot{H}(\omega)|^2 e^{j\omega t} d\omega. \quad (9.6)$$

З урахуванням (9.5) і (9.6) відношення сигнал/шум (за потужністю) можна представити у вигляді [41]

$$\rho(t) = \frac{p_s(t)}{W_n} = \frac{\left| \int_{-\infty}^{\infty} \dot{S}_0(\omega) \dot{H}(\omega) e^{j\omega t} d\omega \right|^2}{2\pi N \int_{-\infty}^{\infty} |\dot{H}(\omega)|^2 e^{j\omega t} d\omega}. \quad (9.7)$$

Завдання пошуку оптимального фільтра зводиться, таким чином, до визначення такої функції $\dot{H}(\omega)$, яка при заданій функції $\dot{S}_0(\omega)$ забезпечує максимальне значення $\rho(t)$ для деякого моменту часу t . Цей момент часу відповідає

максимальному перевищенню сигналу над шумом; в інші моменти рівень сигналу зменшується, а шуму – зберігається (оскільки випадковий процес є стаціонарним).

Далі до (9.7) буде застосовано нерівність Буняковського – Шварца, яка для двох довільних комплексних функцій $\dot{A}(x)$ і $\dot{B}(x)$ має вигляд

$$\left| \int_a^b \dot{A}^*(x)\dot{B}(x) dx \right|^2 \leq \int_a^b |\dot{A}(x)|^2 dx \int_a^b |\dot{B}(x)|^2 dx \text{ при}$$

$b > a$ [42].

Перетворивши за цим співвідношенням чисельник формули (9.7), отримаємо

$$\rho(t) \leq \frac{\int_{-\infty}^{\infty} |\dot{S}_0(\omega)|^2 d\omega \int_{-\infty}^{\infty} |\dot{H}(\omega)|^2 d\omega}{2\pi N \int_{-\infty}^{\infty} |\dot{H}(\omega)|^2 d\omega} = \frac{1}{2\pi N} \int_{-\infty}^{\infty} |\dot{S}_0(\omega)|^2 e^{j\omega t} d\omega.$$

Таким чином, при будь-якій частотній характеристиці фільтра $\dot{H}(\omega)$ відношення сигнал/шум $\rho(t)$ не може перевищити максимального значення

$$\rho_{\max} = \frac{1}{2\pi N} \int_{-\infty}^{\infty} |\dot{S}_0(\omega)|^2 d\omega = \frac{E_s}{N}, \quad (9.8)$$

де E_s – згідно рівності Парсеваля (2.17) – енергія сигналу.

На практиці спектральна потужність перешкоди часто задається однобічною N_0 , тоді $N = N_0/2$ і відношення сигнал/шум $\rho_{\max} = 2E_s/N_0$.

Величина ρ_{\max} може бути отримана, якщо частотна характеристика фільтра має вигляд [41, 42]

$$\dot{H}(\omega) = c\dot{S}_0^*(\omega)e^{-j\omega t_0} = c|\dot{S}_0(\omega)|e^{-j[\omega t_0 + \varphi(\omega)]}, \quad (9.9)$$

де $\dot{S}_0^*(\omega) = \dot{S}_0(-\omega) = |\dot{S}_0(\omega)| e^{-j\varphi(\omega)}$ – функція, комплексно-спряжена зі спектром сигналу $s_0(t)$, c – довільна постійна, t_0 – параметр, що має розмірність часу.

Переконаємось в цьому безпосередньою підстановкою співвідношення (9.9) в рівність (9.7)

$$\rho(t) = \frac{\left| \int_{-\infty}^{\infty} \dot{S}_0(\omega) c \dot{S}_0^*(\omega) e^{-i\omega t_0} e^{j\omega t} d\omega \right|^2}{2\pi N \int_{-\infty}^{\infty} |c \dot{S}_0^*(\omega) e^{-i\omega t_0}|^2 d\omega} = \frac{|c|^2 \left| \int_{-\infty}^{\infty} |\dot{S}_0(\omega)|^2 e^{i\omega(t-t_0)} d\omega \right|^2}{2\pi N |c|^2 \int_{-\infty}^{\infty} |\dot{S}_0^*(\omega)|^2 d\omega}.$$

Якщо $t = t_0$, з урахуванням (9.8) отримаємо

$$\rho(t_0) = \frac{\left(\int_{-\infty}^{\infty} |\dot{S}_0(\omega)|^2 d\omega \right)^2}{2\pi N \int_{-\infty}^{\infty} |\dot{S}_0(\omega)|^2 d\omega} = \frac{1}{2\pi N} \int_{-\infty}^{\infty} |\dot{S}_0(\omega)|^2 d\omega = \rho_{\max}.$$

Таким чином, частотна характеристика оптимального фільтра (9.9) забезпечує отримання максимального пікового значення сигналу на виході лінійної стаціонарної системи в момент часу $t = t_0$. Вона цілком визначається спектром сигналу, і може бути подана його АЧХ і ФЧХ:

$$\begin{aligned} |\dot{H}(\omega)| &= |c| |\dot{S}_0(\omega)|; \\ \psi(\omega) &= -[\varphi(\omega) + \omega t_0]. \end{aligned} \tag{9.10}$$

З співвідношень (9.10) видно, що АЧХ оптимального фільтра з точністю до постійного множника збігається з амплітудним спектром сигналу, а ФЧХ визначається фазовим спектром сигналу $\varphi(\omega)$ і лінійною функцією частоти ωt_0 . Таким чином, частотна характеристика оптимального фільтра «узгоджена» зі спектром сигналу, а такий фільтр має назву – *узгоджений фільтр* (УФ).

Фаза сигналу на виході узгодженого фільтра дорівнює $\theta(t) = \omega t + \varphi(\omega) + \psi(\omega) = \omega t + \varphi(\omega) - \varphi(\omega) - \omega t_0 = \omega(t - t_0)$, тому $\theta(t_0) = 0$. Це означає, що в момент t_0 , всі гармонічні складові сигналу мають однакову фазу, тому в цей момент виникає «сплеск» сигналу на виході УФ. Спектральні ж складові завади на виході фільтра мають випадкові фази, завдяки чому УФ максимізує відношення сигналу до завади (типу білого шуму) на його виході. Крім цього, оскільки АЧХ узгодженого фільтра повторює форму амплітудно-частотного спектру корисного сигналу, УФ найкращим чином пропускає гармонічні складові сигналу на тих частотах, де зосереджена більша частина енергії корисного сигналу. Більш слабкі складові спектру сигналу послаблюються сильніше, що також дозволяє зменшити вплив спектральних складових широкопasmового шуму. Наприклад, якщо сигнал є періодичним, його спектр має дискретну структуру, тоді УФ матиме гребінчасту форму АЧХ. Такі фільтри широко застосовуються в радіотехніці.

Застосувавши обернене перетворення Фур'є до його частотної характеристики (коефіцієнта передачі) (9.9), отримаємо імпульсну характеристику УФ

$$\begin{aligned} h(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{H}(\omega) e^{j\omega t} d\omega = \frac{c}{2\pi} \int_{-\infty}^{\infty} \dot{S}_0(-\omega) e^{j\omega(t-t_0)} d\omega = \\ &= \frac{c}{2\pi} \int_{-\infty}^{\infty} \dot{S}_0(\omega) e^{-j\omega(t_0-t)} d\omega = c s_0(t_0 - t). \end{aligned} \quad (9.11)$$

Таким чином, імпульсна характеристика УФ є масштабною копією вхідного сигналу, однак, розташованою в дзеркальному порядку вздовж осі часу (на це вказує від'ємний знак при t у формулі (9.11)) [17, 41]. Крім цього, імпульсна характеристика УФ зміщена відносно сигналу $s_0(-t)$ на відрізок часу t_0 (рис. 9.1).

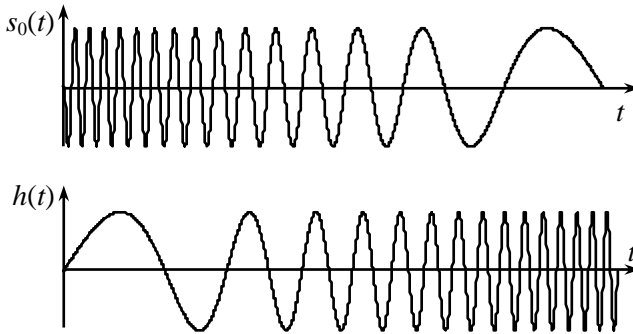


Рисунок 9.1 – Сигнал $s_0(t)$ (а) та імпульсна характеристика узгодженого з ним фільтра $h(t)$ (б)

Необхідна (але не достатня) умова фізичної реалізованості УФ формулюється так: проміжок часу t_0 між початком імпульсу на вході і моментом виникнення максимальної вихідної реакції повинен бути не менше тривалості імпульсу $s_0(t)$. В іншому випадку імпульсна характеристика УФ буде відмінною від нуля при $t < 0$, тобто до моменту надходження дельта-імпульсу на вхід фільтра. Сенс цієї умови такий: для створення максимально можливого миттєвого значення сигналу на виході узгоджений фільтр повинен для цього провести обробку всього вхідного сигналу.

Застосуємо співвідношення (9.11) для $h(t)$ в інтегралі Дюамеля і отримаємо відгук УФ на даний вхідний вплив $s(t)$

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} s(\tau)h(t-\tau)d\tau = c \int_{-\infty}^{\infty} s(\tau)s_0[t_0-(t-\tau)]d\tau = \\ &= c \int_{-\infty}^{\infty} s(\tau)s_0[\tau-(t-t_0)]d\tau. \end{aligned}$$

Видно (див. гл. 3), що останній інтеграл є ВКФ сигналів $s(t)$ і $s_0(t)$ із зсувом на час t_0 , тобто

$$y(t) = cB_{s s_0}(t-t_0). \quad (9.12)$$

В момент часу t_0 миттєве значення сигналу на виході УФ з точністю до коефіцієнта пропорційності дорівнює скалярному добутку цих сигналів

$$y(t_0) = c \int_{-\infty}^{\infty} s(\tau)s_0(\tau)d\tau. \quad (9.13)$$

В ідеальному випадку, коли у вимірювальному каналі шуми відсутні, $s(t) = s_0(t)$ і формула (9.12) приймає вигляд

$$y(t) = cB_{s_0}(t - t_0), \quad (9.14)$$

де $B_{s_0}(\tau)$ – автокореляційна функція очікуваного сигналу. Тобто сигнал на виході УФ пропорційний АКФ вхідного сигналу, із затримкою t_0 в часі. При цьому максимальне значення сигналу при $t = t_0$ визначається його енергією і не залежить від форми сигналу.

Співвідношення (9.13) і (9.14) вказують на можливість реалізації оптимальної обробки сигналу шляхом обчислення кореляції вимірювального сигналу з опорним зразковим сигналом. Однак в цьому випадку потрібна багатоканальна схема, в якій кожний з каналів налаштований на певну часову затримку сигналу.

Принципово надійне виявлення і вимірювання моменту прийому імпульсного сигналу малої амплітуди може бути забезпечене збільшенням його тривалості. Однак це може погіршити якість розділення близько розташованих в часі сигналів особливо при значній різниці їх амплітуд. Для покращення розділення відбитих сигналів при вимірюванні часу їх затримки локаційним методом форма сигналу повинна обиратись з урахуванням пелюсткової форми АКФ сигналу $B_{s_0}(\tau)$. Приклад отримання сигналу на виході фільтра узгодженого з лінійно-частотно-модульованим (ЛЧМ) імпульсом представлено на рис. 9.2. Такий сигнал широко застосовується в локації для забезпечення високого розділення по дальності та її вимірювання, завдяки скороченню тривалості головної пелюстки порівняно з

тривалістю імпульсу на вході УФ обернено пропорційно до ширини смуги частот ЛЧМ сигналу, яка при великих значеннях бази сигналу B близька до його дев'ятиї. В той же час сигнал на виході УФ має значну кількість бічних пелюсток, рівень першої з яких при прямокутній обвідній сигналу на 13 дБ менше величини головної пелюстки (рис. 9.2, а).

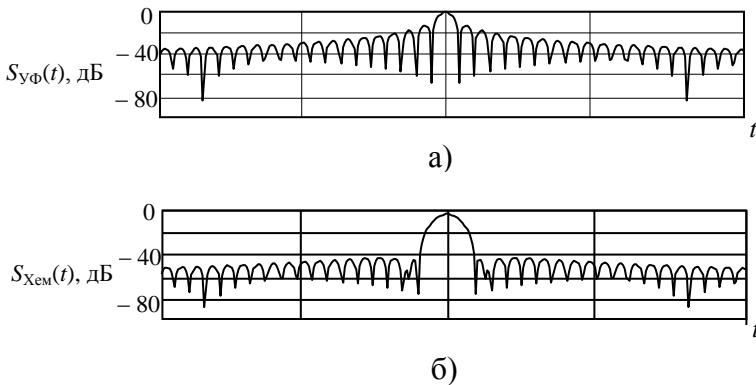


Рисунок 9.2 – Епюри обвідних стиснутого ЛЧМ

Для зниження рівня бічних пелюсток (в першу чергу найбільших), як і при проектуванні цифрових фільтрів, може застосовуватись додаткова віконна обробка сигналу. Це, як у випадку цифрової фільтрації, призводить до деякого зниження рівня головної пелюстки та її розширення (рис. 9.2, б – застосовано вікно Хемінга), оскільки порушується умова узгодженості фільтра.

Типову структуру УФ можна представити у вигляді трансверсального фільтра. У разі обробки радіоімпульсів додатковими елементами є коливальні контури налаштовані на відповідні частоти (рис. 9.3) [1]. Представлені на рис. 9.3 реалізації суміші корисного сигналу і перешкоди на вході і сигналу виході УФ ілюструють його ефективність для виявлення і визначення моменту отримання сигналу.

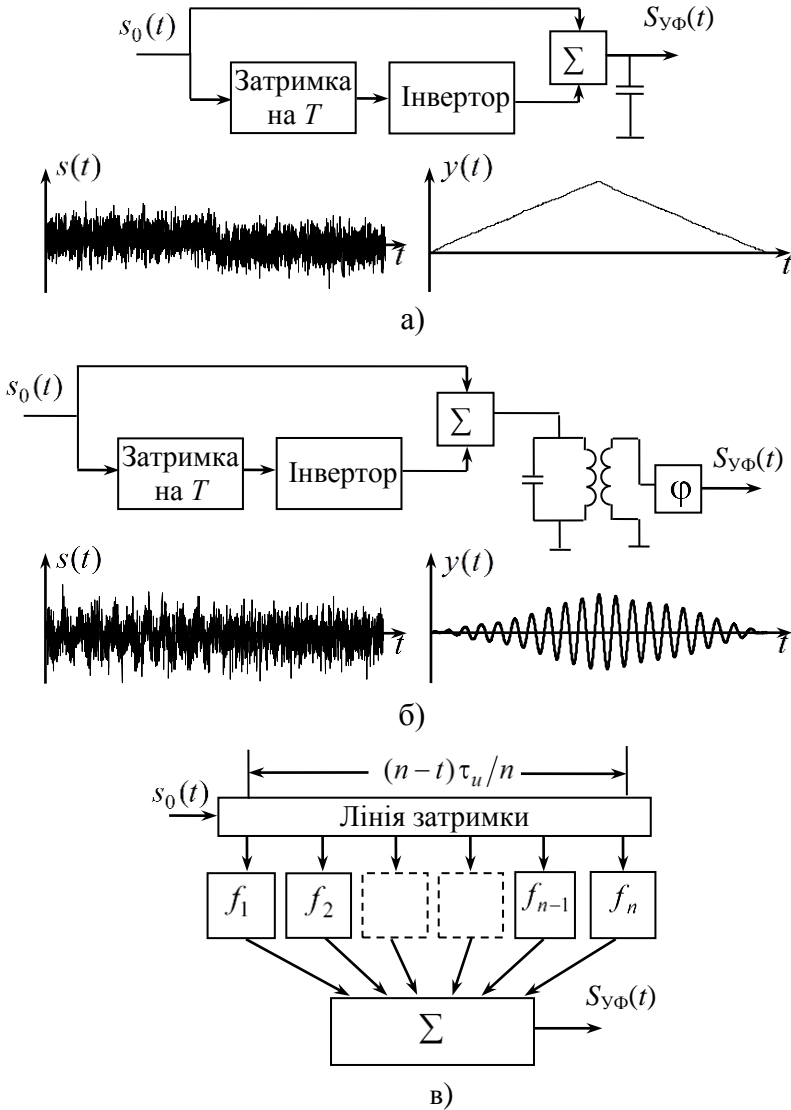


Рисунок 9.3 – Узгоджені фільтри: а) для прямокутного відеоімпульсу; б) для прямокутного радіоімпульсу; в) для частотно-маніпульованого радіоімпульсу

У граничному випадку частотно-маніпульований радіоімпульс відповідатиме неперервному частотно-модульованому радіоімпульсу. У випадку лінійної зміни частоти внутрішньоімпульсної модуляції УФ може бути реалізований з використанням дисперсійної лінії затримки.

Далі розглянемо задачу оптимальної фільтрації неперервного сигналу $s(t) = s_0(t) + n(t)$ в іншій постановці: в результаті фільтрації потрібно отримати на виході ЗВ сигнал $y(t)$ найменш відмінний від сигналу $s_0(t)$.

Рішення цієї задачі було отримане незалежно А. Н. Колмогоровим (1939 р.) і Н. Вінером (1942 р.). Воно базується на трьох основних припущеннях [42, 43]: 1) сигнал $s_0(t)$ і перешкода $n(t)$ є стаціонарними випадковими процесами; 2) операція фільтрації – лінійна; 3) критерієм оптимальності вважається мінімум середньоквадратичної похибки. Суттєвою відмінністю такої постановки задачі є те, що потрібно якнайкраще відтворити сигнал невідомої форми, тоді як у попередньому випадку узгодженої фільтрації форма сигналу була відомою, але на виході УФ могла суттєво спотворюватись задля формування максимально можливого піку в певний момент відліку.

Нехай сигнал $s_0(t)$ і перешкода $n(t)$ є незалежними випадковими процесами з відомими кореляційними функціями $B_{s_0}(\tau)$ і $B_n(\tau)$, а $s(t)$ – сигнал на вході лінійного фільтра з імпульсною характеристикою $h(t)$. Потрібно знайти таку функцію $h(t)$, щоб сигнал на виході фільтра $y(t)$ мінімізував середньоквадратичну похибку

$$\overline{\varepsilon^2(t)} = \overline{[y(t) - s_0(t)]^2}. \quad (9.15)$$

Для фізичної реалізованості фільтра вимагається, щоб імпульсна характеристика фільтра задовольняла умові $h(t) = 0$, якщо $t < 0$.

Вказане обмеження враховується у запису

$$y(t) = \int_{\gamma} s(t-\tau)h(\tau)d\tau,$$

де область інтегрування γ для фільтра, що фізично реалізується, є інтервал від 0 до ∞ , а для фільтра, який фізично не реалізується, – від $-\infty$ до ∞ .

Доведемо спочатку, що характеристика фільтра $h(t)$ оптимальна тоді й тільки тоді, коли

$$\varepsilon(t)s(t-\tau) = 0 \text{ для } \tau \in \gamma, \quad (9.16)$$

тобто коли похибка $\varepsilon(t) = y(t) - s_0(t)$ некорельована з вхідним сигналом $s(t)$ в усі моменти часу в області γ .

Нехай $h_1(t)$ – імпульсна характеристика оптимального фільтра, для якого справедлива умова (9.16), а для $h_2(t)$ – імпульсна характеристика будь-якого іншого лінійного фільтра. Відгуки фільтрів відповідно позначимо через $y_1(t)$ і $y_2(t)$. Тоді

$$\begin{aligned} \overline{[s_0(t) - y_2(t)]^2} &= \overline{[s_0(t) - y_1(t) + y_1(t) - y_2(t)]^2} = \\ &= \overline{[s_0(t) - y_1(t)]^2} + \overline{[y_1(t) - y_2(t)]^2} + 2\overline{[s_0(t) - y_1(t)][y_1(t) - y_2(t)]}. \end{aligned}$$

Оскільки функція $s_0(t) - y_1(t) = \varepsilon(t)$ задовольняє умові (9.16), тоді

$$\begin{aligned} \overline{[s_0(t) - y_1(t)][y_1(t) - y_2(t)]} &= \overline{\varepsilon(t) \int_{\gamma} s(t-\tau)[h_1(\tau) - h_2(\tau)]d\tau} = \\ &= \int_{\gamma} \overline{\varepsilon(t)s(t-\tau)[h_1(\tau) - h_2(\tau)]}d\tau = 0. \end{aligned}$$

Отже

$$\overline{[s_0(t) - y_2(t)]^2} = \overline{\varepsilon^2(t)} + \overline{[y_1(t) - y_2(t)]^2}.$$

Очевидно, що цей вираз буде мати мінімальне значення, коли $y_2(t) = y_1(t)$, що й доводить справедливість умови (9.16).

Представимо співвідношення (9.16) у вигляді $\overline{[s_0(t) - y(t)]s(t-\tau)} = 0$, звідки отримаємо

$$\overline{s_0(t)s(t-\tau)} = \overline{y(t)s(t-\tau)} = \int_{\gamma} \overline{s(t-z)s(t-\tau)}h(z)dz$$

чи

$$B_{s_0s}(\tau) = \int_{\gamma} B_s(\tau-z)h(z)dz. \quad (9.17)$$

У випадку, коли сигнал $s_0(t)$ і перешкода $n(t)$ некорельовані, рівняння (9.17) приймає вигляд

$$B_{s_0}(\tau) = \int_{\gamma} [B_{s_0}(\tau-z) + B_n(\tau-z)]h(z)dz. \quad (9.18)$$

Це інтегральне рівняння є основним в теорії лінійної фільтрації і називається рівнянням Вінера – Хопфа. Його рішенням є шукана функція $h(t)$, що мінімізує середньоквадратичну похибку $\overline{\varepsilon^2(t)}$.

Рівняння (9.18) вирішується порівняно просто, якщо $\gamma \in (-\infty, \infty)$, тобто для фільтрів, що не реалізуються.

Перетворення Фур'є від правої та лівої частин рівняння (9.18) дадуть рівняння в частотній області

$$G_{s_0}(\omega) = [G_{s_0}(\omega) + G_n(\omega)]\dot{K}(\omega), \quad (9.19)$$

де $G_{s_0}(\omega)$ – енергетичний спектр очікуваного сигналу, $G_n(\omega)$ – енергетичний спектр перешкоди, $\dot{K}(\omega)$ – коефіцієнт передачі оптимального фільтра.

З (9.19) отримаємо вираз для коефіцієнта передачі оптимального лінійного фільтра

$$\dot{K}(\omega) = \frac{G_{s_0}(\omega)}{G_{s_0}(\omega) + G_n(\omega)}. \quad (9.20)$$

Права частина виразу (9.20) є дійсною функцією. У разі урахування фізичної реалізованості фільтра, тобто коли $\gamma \in (0, \infty)$, матиме місце додаткова часова затримка t_0 сигналу в фільтрі. Тому коефіцієнт передачі оптимального фільтра прийме вигляд

$$\dot{K}(\omega) = \frac{G_{s_0}(\omega)}{G_{s_0}(\omega) + G_n(\omega)} e^{-i\omega t_0}. \quad (9.21)$$

При цьому середній квадрат похибки у випадку (9.20) дорівнює:

$$\overline{\varepsilon_{\min}^2(t)} = \frac{1}{\pi} \int_0^{\infty} \frac{G_{s_0}(\omega)G_n(\omega)}{G_{s_0}(\omega) + G_n(\omega)} d\omega. \quad (9.22).$$

Зауважимо, що сигнал може бути повністю відділений від перешкоди, тобто похибка буде дорівнювати нулю лише за умови $G_{s_0}(\omega)G_n(\omega) = 0$. Це означає, що спектри сигналу і перешкоди не перекриваються. Якщо перешкода відсутня, тоді $G_n(\omega) = 0$, $\overline{\varepsilon^2} = 0$, $K(\omega) = 1$, тобто фільтр не потрібен.

Фільтр з оптимальною характеристикою (9.20) пропускає коливання різних частот з тим меншою вагою, чим більше відношення $\frac{G_n(\omega)}{G_{s_0}(\omega)}$ на даній частоті. Якщо відношення

сигналу до перешкоди мале: $G_{s_0} \ll G_n$ тоді

$$\overline{\varepsilon^2} \approx \frac{1}{\pi} \int_0^{\infty} G_{s_0}(\omega) d\omega = P_c, \quad \text{тобто похибка дуже велика і}$$

відновлення сигналу за допомогою фільтрації стає неможливим.

9.2. Оптимальна лінійна обробка цифрових сигналів

В основі подальшого розгляду адаптивних фільтрів, який заснований на монографії [43], також лежить концепція оптимального лінійного оцінювання.

Процес адаптивної фільтрації включає два етапи проведення оцінювання: 1) оцінювання шуканого сигналу на виході фільтра і 2) оцінювання вагових коефіцієнтів фільтра, необхідних для досягнення мети оптимізації. Другий з цих

етапів необхідний лише тоді, коли характеристики вхідного сигналу апріорно невідомі.

Задача ж синтезу оптимального цифрового фільтра послідовності відліків вхідного сигналу є аналогічною до розглянутої вище для фільтрів аналогового типу, які повинні найкращим чином відтворювати сигнал невідомої форми.

Нехай спостерігається послідовність відліків сигналу $y(n)$, яка є сптвореною адитивним білим шумом послідовністю $x(n)$:

$$y(n) = x(n) + v(n), \quad (9.23)$$

де $v(n)$ – білий шум з дисперсією σ_v^2 .

Якість оцінювання характеризується різницею між послідовністю на виході фільтра й істинним сигналом [43]:

$$e(n) = x(n) - \hat{x}(n), \quad (9.24)$$

де $e(n)$ – похибка оцінювання; $x(n)$ – дискретна випадкова (або детермінована) величина, яку необхідно оцінити, а $\hat{x}(n)$ – оцінка $x(n)$, виконана за допомогою системи оцінювання, причому

$$\hat{x}(n) = f \{y(n), h(n)\}, \quad (9.25)$$

тобто $\hat{x}(n)$ – лінійна функція послідовності вхідних сигналів $y(n)$ і набору вагових коефіцієнтів фільтра $h(n)$.

Для проведення оптимального оцінювання найчастіше застосовується метод найменших квадратів (МНК). При цьому середньоквадратична похибка визначається як

$$E \{e^2(n)\} = \overline{[x(n) - \hat{x}(n)]^2}. \quad (9.26)$$

Вона мінімізується відносно вагових коефіцієнтів пристрою оцінки для отримання оптимального оцінювання за критерієм МНК [43]. Слід зазначити, що можуть застосовуватись й інші функції вартості. Наприклад, це такі функції як абсолютна величина похибки і нелінійна порогова функція. При використанні критерію найменшого середньоквадратичного малі похибки вносять менший внесок,

ніж великі похибки (на противагу критерію абсолютної величини похибки, який дає однакову вагу для всіх похибок). Нелінійна порогова функція похибки використовується в тому випадку, коли є прийнятний інтервал похибок (тобто, існує задана допустима похибка).

В нерекурсивному фільтрі оцінка $x(n)$ визначається у вигляді кінцевого лінійного полінома $y(n)$

$$\hat{x}(n) = \sum_{k=0}^{N-1} y(n-k)h_k, \quad (9.27)$$

де h_k – вагові коефіцієнти в структурі КІХ-фільтра (рис. 6.3). Вираз (9.27) можна представити у матрично-векторній формі

$$\hat{x}(n) = \mathbf{Y}^T(n)\mathbf{H} = \mathbf{H}^T\mathbf{Y}(n), \quad (9.28)$$

де $\mathbf{Y}^T(n) = [y(n)y(n-1)\dots y(n-N+1)]$ і $\mathbf{H}^T = [h_0h_1\dots h_{n-1}]$, верхній індекс Т позначає транспонування матриці. Тоді функція середньоквадратичної похибки (СКП) приймає вигляд

$$E\{e^2(n)\} = E\{x(n) - \mathbf{H}^T\mathbf{Y}(n)\}^2 \quad (9.28)$$

Вираз (9.28) описує стандартну поверхню квадратичної похибки з одним єдиним мінімумом. Його диференціювання по \mathbf{H}^T дає

$$\frac{\partial e^2(n)}{\partial \mathbf{H}^T} = 2E\left[\{x(n)\mathbf{H}^T\mathbf{Y}(n)\}\mathbf{Y}^T(n)\right]. \quad (9.29)$$

Припускаючи, що (9.29) дорівнює нулю, маємо

$$E\left[\{x(n) - \mathbf{H}^T\mathbf{Y}(n)\}\mathbf{Y}^T(n)\right] = 0, \quad (9.30)$$

$$E\{x(n)\mathbf{Y}^T(n)\} = E\{\mathbf{H}^T\mathbf{Y}(n)\mathbf{Y}^T(n)\}.$$

Якщо ваговий вектор \mathbf{H}^T і вектор сигналу $\mathbf{Y}(n)$ некорельовані, отримаємо

$$E\{(x(n)\mathbf{Y}^T(n))\} = \mathbf{H}_{\text{opt}}^T E\{\mathbf{Y}(n)\mathbf{Y}^T(n)\}. \quad (9.31)$$

Члени математичного сподівання, що входять до (9.31), можна визначити наступним чином:

$\mathbf{P} = E\{(x(n)\mathbf{Y}(n))\}$ – взаємна кореляція між вхідним сигналом і оцінюваним параметром;

$\mathbf{R} = E\{\mathbf{Y}(n)\mathbf{Y}^T(n)\}$ – автокореляційна матриця вхідної сигнальної послідовності.

Тоді (9.31) можна переписати у вигляді

$$\mathbf{P}^T = \mathbf{H}_{\text{opt}}^T \mathbf{R}. \quad (9.32)$$

Співвідношення (9.32) є рівнянням Вінера – Хопфа, яке дає оптимальне за методом найменших квадратів вінерівське рішення для \mathbf{H}

$$\mathbf{H}_{\text{opt}} = \mathbf{R}^{-1}\mathbf{P}. \quad (9.33)$$

Розглянемо практичний *приклад* пристрою вінерівської оцінки для випадку, коли спостерігається сигнал у вигляді суми синусоїди і білого шуму [43]

$$y(n) = \sin \frac{n\pi}{4} + \eta(n). \quad (9.34)$$

Отже, вхідний сигнал $x(n)$ є синусоїдою з частотою, що дорівнює одній восьмій частоти дискретизації фільтра. Для спрощення подальших викладок обмежимо кількість вагових коефіцієнтів отриманого фільтра чотирма. З рішення (9.33) рівняння Вінера – Хопфа видно, що необхідно отримати матрицю \mathbf{R} , яка визначається як

$$\mathbf{R} = E \left\{ \begin{bmatrix} y(n) \\ y(n-1) \\ y(n-2) \\ y(n-3) \end{bmatrix} \begin{bmatrix} y(n)y(n-1)y(n-2)y(n-3) \end{bmatrix} \right\} = \quad (9.35)$$

$$= E \begin{bmatrix} y^2(n) & y(n)y(n-1) & y(n)y(n-2) & y(n)y(n-3) \\ y(n)y(n-1) & y^2(n-1) & y(n-1)y(n-2) & y(n-1)y(n-3) \\ y(n)y(n-2) & \dots\dots\dots & y^2(n-2) & \dots\dots\dots \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots & y^2(n-3) \end{bmatrix}. \quad (9.36)$$

Слід зазначити що, оскільки часовий ряд є стаціонарним, всі члени (2.15) можна отримати з першого рядка матриці:

$$E \{ y^2(n) \} = E \{ y^2(n-1) \} = E \{ y^2(n-2) \} = E \{ y^2(n-3) \};$$

$$E \{ y(n)y(n-1) \} = E \{ y(n-1)y(n-2) \} = E \{ y(n-2)y(n-3) \};$$

$$E \{ y(n)y(n-2) \} = E \{ y(n-1)y(n-3) \},$$

тобто суттєвою є лише часова різниця між двома операндами в операторі математичного сподівання.

Тоді (9.36) перетвориться до вигляду

$$\mathbf{R} = E \begin{bmatrix} y^2(n) & y(n)y(n-1) & y(n)y(n-2) & y(n)y(n-3) \\ y(n)y(n-1) & y^2(n) & y(n)y(n-1) & y(n)y(n-2) \\ y(n)y(n-2) & y(n)y(n-1) & y^2(n) & y(n)y(n-1) \\ y(n)y(n-3) & y(n)y(n-2) & y(n)y(n-1) & y^2(n) \end{bmatrix}. \quad (9.37)$$

Така симетрична матриця вважається матрицею Теплиця, тобто матрицею, в якій члени на всіх діагоналях є рівними між собою. Отримати значення елементів матриці (9.37) для прикладу, що розглядається, особливо просто, оскільки шумова складова впливає лише на члени її головної діагоналі, а всі інші члени можна знайти з детермінованої компоненти сигналу $x(n)$, яка має тільки вісім рівнів квантування з певним інтервалом дискретизації. Значення елементів матриці наступні:

$$\mathbf{R} = \begin{bmatrix} \frac{1}{2} + \sigma_n^2 & \frac{1}{2\sqrt{2}} & 0 & \frac{-1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & \frac{1}{2} + \sigma_n^2 & \frac{1}{2\sqrt{2}} & 0 \\ 0 & \frac{1}{2\sqrt{2}} & \frac{1}{2} + \sigma_n^2 & \frac{1}{2\sqrt{2}} \\ \frac{-1}{2\sqrt{2}} & 0 & \frac{1}{2\sqrt{2}} & \frac{1}{2} + \sigma_n^2 \end{bmatrix}, \quad (9.38)$$

де $\sigma_n^2 = E\{\eta^2(n)\}$.

Подальші перетворення спрощуються, якщо застосувати підстановку

$$\mathbf{R} = \begin{bmatrix} v & \rho & 0 & -\rho \\ \rho & v & \rho & 0 \\ 0 & \rho & v & \rho \\ -\rho & 0 & \rho & v \end{bmatrix}. \quad (9.39)$$

Тоді обернена матриця, визначається як:

$$\mathbf{R}^{-1} = \begin{bmatrix} \frac{-v}{2\rho^2 - v^2} & \frac{\rho}{2\rho^2 - v^2} & 0 & \frac{-\rho}{2\rho^2 - v^2} \\ \frac{\rho}{2\rho^2 - v^2} & \frac{-v}{2\rho^2 - v^2} & \frac{\rho}{2\rho^2 - v^2} & 0 \\ 0 & \frac{\rho}{2\rho^2 - v^2} & \frac{-v}{2\rho^2 - v^2} & \frac{\rho}{2\rho^2 - v^2} \\ \frac{-\rho}{2\rho^2 - v^2} & 0 & \frac{\rho}{2\rho^2 - v^2} & \frac{-v}{2\rho^2 - v^2} \end{bmatrix}. \quad (9.40)$$

Щоб отримати ваговий вектор Вінера, треба також ввести матрицю взаємної кореляції \mathbf{P} , яка визначається за формулою

$$\mathbf{P} = E \left\{ x(n) \begin{bmatrix} y(n) \\ y(n-1) \\ y(n-2) \\ y(n-3) \end{bmatrix} \right\}. \quad (9.41)$$

Підставивши знову істинні значення з (9.34), отримаємо

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2\sqrt{2}} \\ -\frac{1}{2\sqrt{2}} \end{bmatrix} = \begin{bmatrix} v - \sigma_n^2 \\ \rho \\ 0 \\ -\rho \end{bmatrix}. \quad (9.42)$$

Помноживши \mathbf{P} на матрицю \mathbf{R}^{-1} з (9.40), знайдемо

$$\mathbf{H}_{\text{opt}} = \mathbf{R}^{-1}\mathbf{P} = \frac{1}{2(1 + \sigma_n^2)} \begin{bmatrix} 1 \\ \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix}. \quad (9.43)$$

Визначаючи значення \mathbf{H}_{opt} з (9.43), приходимо до узгодженого фільтру для випадку синусоїдального вхідного сигналу. Залишкова середньоквадратична похибка легко обчислюється з (9.43) і (9.28) підстановкою $\mathbf{H} = \mathbf{H}_{\text{opt}}$ в (9.28).

З (9.30) можна отримати такий вираз

$$E \{ e(n)\mathbf{Y}(n) \} = 0 \quad (9.44)$$

і середньоквадратичну похибку записати у вигляді:

$$E \{ e(n) \} = E \left\{ e(n) \left[x(n) - \mathbf{H}_{\text{opt}}^T \mathbf{Y}(n) \right] \right\}. \quad (9.45)$$

Перетворимо тепер формулу для середньоквадратичної похибки:

$$\begin{aligned} E\{e(n)\} &= E\{e(n)x(n)\} = E\{x^2(n)\} - \mathbf{H}_{\text{opt}}^T E\{x(n)\mathbf{Y}(n)\} = \\ &= E\{x^2(n)\} - \mathbf{H}_{\text{opt}}^T \mathbf{P}. \end{aligned} \quad (9.46)$$

Підставляючи значення \mathbf{P} з (9.42) і \mathbf{H}_{opt} з (9.43), отримаємо формулу для залишкової середньоквадратичної похибки

$$E\{e^2(n)\} = \frac{1}{2} - \frac{1}{2(1+\sigma_n^2)} = \frac{\sigma_n^2}{2(1+\sigma_n^2)}. \quad (9.47)$$

Отже, при $\sigma_n^2 = 0$ СКП дорівнює нулю, а при відношенні сигнал/шум рівному 0 дБ (тобто $\sigma_n^2 = 1/2$), кінцева СКП дорівнює 1/6 (величина СКП до процесу оцінювання дорівнювала 1/2).

Збільшення порядку \mathbf{H}_{opt} (тобто використання для оцінювання фільтра з більшим часовим інтервалом обробки сигналу) призводить до відповідного зменшення оптимальної залишкової СКП. Наведений приклад з [43] показує як необхідно на практиці виконувати обробку даних для реалізації вінерівського пристрою оцінювання. При більш високих порядках \mathbf{H}_{opt} можна отримати кращі оцінки, однак при цьому можливе значне збільшення обсягу обчислень.

9.3. Адаптивна обробка сигналів

Розглянуті вище оптимальні (узгоджені) фільтри є прикладом систем обробки сигналів, які не змінюють алгоритм обробки в часі. Тобто вони ґрунтуються на припущенні, що спотворення сигналу є відомим і незмінним в часі. Таким чином, в умовах змін у часі заводової обстановки така обробка може виявитись не оптимальною.

Для покращення обробки сигналів в таких умовах можуть застосовуватись адаптивні системи обробки сигналів [43]. Такі системи діють за принципом замкнутого контуру

(зворотного зв'язку). Вхідний сигнал $s(n)$ зважується у програмованому фільтрі для отримання вихідного сигналу $\hat{y}(n)$, який порівнюється з опорним стандартним сигналом $y(n)$ для знаходження сигналу похибки $e(n)$. Потім сигнал похибки використовується для коригування вагових параметрів процесора (зазвичай ітеративним методом) з метою поступової мінімізації похибки (тобто сигнал на виході пристрою обробки – процесора – впритул наближається до величини опорного сигналу). Такі процесори поділяються на два великі класи: адаптивні фільтри (рис. 9.4) і адаптивні антени [43].

Адаптивний фільтр – це програмований фільтр, частотна характеристика або передавальна функція якого змінюється, тобто адаптується, таким чином, щоб пропустити без спотворення корисні складові сигналу і послабити небажані сигнали або перешкоди, тобто зменшити будь-які спотворення вхідного сигналу. В адаптивних антенах просторова обробка сигналу здійснюється за допомогою антенної решітки, в діаграмі спрямованості якої утворюється головний максимум в напрямку приходу сигналу і формуються нульові провали в напрямку джерел перешкод задля їх придушення.

Адаптивній системі необхідний мінімальний обсяг апріорної інформації про вхідний сигнал. Адаптивний фільтр діє за принципом оцінювання статистичних параметрів сигналу, що поступає, і підстроювання власної імпульсної (перехідної) характеристики таким чином, щоб мінімізувати деяку функцію вартості. Цю функцію можна отримати кількома способами залежно від передбачуваної області застосування, але зазвичай вона отримується за допомогою джерела опорного сигналу $y(n)$ на другому вході, як показано на рис. 9.4.

Сигнал $\hat{y}(n)$ може розглядатись як корисний сигнал на виході фільтра; задача адаптивного алгоритму – підстроювання вагових коефіцієнтів в програмованому фільтрі таким чином, щоб звести до мінімуму $e(n)$ –

відхилення між сигналами на виході $\hat{y}(n)$ і вході $y(n)$ фільтра. Такі адаптивні фільтри часто використовуються для відновлення на виходах вимірювальних каналів сигналів зі змінними в часі характеристиками.

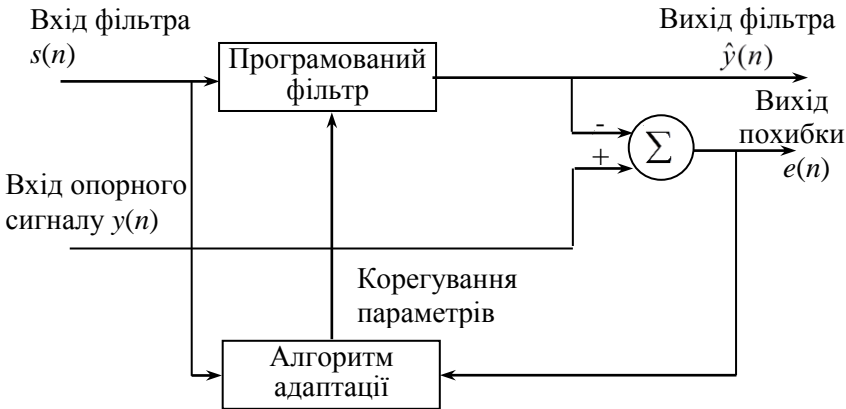


Рисунок 9.4 – Блок-схема адаптивного фільтра

Важливою функцією адаптивного фільтра є моделювання системи. Це ілюструється на рис. 9.5, де первинний сигнал з рівномірною спектральною густиною подається безпосередньо або на вхід s , або на вхід y адаптивного фільтра. Первинний сигнал також надходить на вхід системи з імпульсною характеристикою $\mathbf{H}(n)$, вихід якої з'єднаний з другим входом адаптивного фільтра. Для отримання оптимальних вагових векторів \mathbf{H}_{opt} адаптивного фільтра можуть застосуватись два підходи, які приведуть до абсолютно різних результатів:

1. Невідома система $\mathbf{H}(n)$ підключена до входу y адаптивного фільтра (рис. 9.5, а). В цьому випадку оптимальна імпульсна характеристика адаптивного фільтра є точною моделлю відповідної характеристики системи $\mathbf{H}(n)$.

2. Невідома система $\mathbf{H}(n)$ підключена до входу s адаптивного фільтра (рис. 9.5, б). В цьому випадку

оптимальна імпульсна характеристика адаптивного фільтра є функцією оберненою до відповідної характеристики невідомої системи.

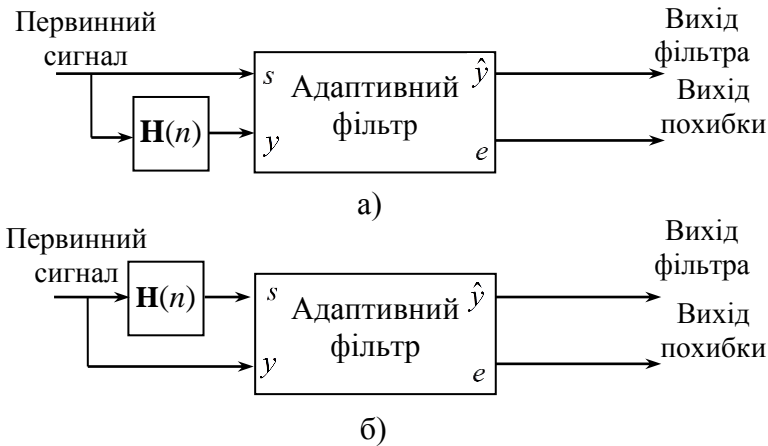


Рисунок 9.5 – Використання адаптивного фільтра для моделювання системи: а) прямого – $\mathbf{H}_{opt} = \mathbf{H}(n)$;

б) оберненого – $\mathbf{H}_{opt} = \mathbf{H}^{-1}(n)$

Практичним прикладом, який ілюструє роботу адаптивного фільтра першого типу (тобто пряме моделювання системи), є придушення відбитого сигналу в гібридній телефонній лінії.

Прикладом, яким можна скористатись для ілюстрації принципу дії адаптивного фільтра, що моделює обернену характеристику системи, є корекція спотворень при передачі даних по телефонних лініях. В цьому випадку вхід телефонної лінії збуджується відомим сигналом, а спотворений сигнал з виходу лінії надходить на вхід $s(n)$ адаптивного фільтра. Потім фільтр перебудовується за допомогою подачі на вхід $y(n)$ послідовної серії відомих (неспотворених) первинних сигналів. Адаптивний фільтр моделює імпульсну характеристику, обернену характеристиці лінії, для отримання на виході відфільтрованих (вільних від спотворень) даних.

Наступна область застосування адаптивних фільтрів – придушення шумів. У цій схемі первинний сигнал, що містить шукану інформацію разом з перешкодовим сигналом, прикладений до входу $y(n)$. Потім від іншого джерела, що не містить ніяких складових шуканого сигналу, надходить незалежний корельований сигнал – зразок сигналу перешкоди. Якщо цей сигнал надходить безпосередньо на вхід $s(n)$ адаптивного фільтра, фільтр формує імпульсну характеристику, що забезпечує отримання на виході сигналу $\hat{y}(n)$, який відтворює небажану складову. Вона когерентно віднімається з $y(n)$, і на виході $e(n)$ залишається тільки шуканий сигнал.

Одним із прикладів використання цього методу є реєстрація серцебиття плода [43]. Первинний сигнал надходить від перетворювача, розташованого на поверхні живота матері. Цей перетворювач виробляє сигнал, що містить імпульси серцевих скорочень плода, які, однак, суттєво маскуються серцебиттям матері. Потім від другого перетворювача, розташованого на грудині матері, отримують вторинний сигнал, що реєструє тільки серцебиття матері. Далі адаптивний фільтр моделює тракт спотворень від перетворювача, розташованого на грудині, до перетворювача, розташованого на животі, для отримання сигналу, який когерентно віднімається від сигналу з поверхні живота.

Ще одне застосування адаптивних фільтрів – це реалізація фільтра, що самоналаштовується, для виділення синусоїди, маскованої широкосмуговим шумом. Таке застосування в адаптивному лінійному підсилювачі здійснюється шляхом подачі сигналу безпосередньо на вхід фільтра $y(n)$ і подачі модифікованого сигналу з часовою затримкою на вхід $s(n)$. У тому випадку, якщо затримка перевищує величину зворотну ширині смуги пропускання фільтра, шумові складові на двох входах будуть некорельовані, а синусоїдальні складові – корельованими. Адаптивний фільтр дає на виході синусоїду зі

збільшеним відношенням сигнал/шум, тоді як на виході сигналу похибки синусоїдальні складові зменшуються.

Для реалізації адаптивного фільтра в схемах, зображених на рис. 9.5, можуть бути застосовані різні типи програмованих фільтрів: НІХ-фільтри, КІХ-фільтри, в тому числі й з перетворенням сигналу в частотну область (рис. 9.6).

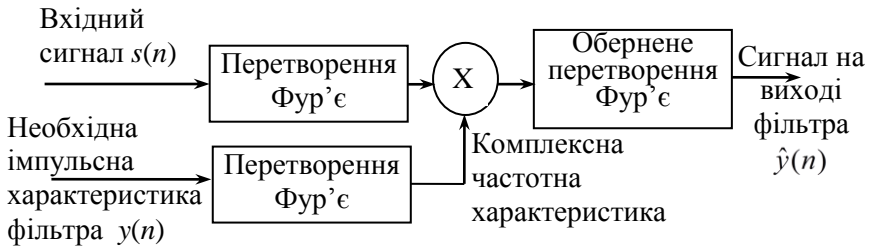


Рисунок 9.6 – Виконання фільтрації методом обробки сигналу в частотній області

9.3.1. Оптимальне рекурсивне калманівське оцінювання

Отримана в розділі 9.2 вінерівська оцінка (9.33) по суті є результатом блокового процесу оцінювання. Він найкраще підходить для випадку, коли в розпорядженні є лише кінцева вибірка (блок) даних, і оцінку можна реалізувати на «автономному» комп'ютері. Однак, якщо ряд є нескінченним за часом, вінерівська оцінка для кожної нової вибірки вимагала б повного перерахування всіх членів АКФ і ВКФ. В оптимальному рекурсивному (або калманівському) фільтрі інформація, що надходить, використовується для коригування рекурсивної оцінки [43].

У даному розділі отримуються співвідношення для скалярного та векторного фільтрів Калмана, а також наведено приклад застосування векторного фільтра Калмана для корекції каналу зв'язку.

Пристрій калманівської оцінки реалізує процес параметричного оцінювання, заснований на авторегресивній (АР) моделі процесу генерації сигналу. АР-модель процесу першого порядку даного типу показана на рис. 9.7, а, а відповідна модель вимірювань – на рис. 9.7, б. (рис. 2.3)

Модель вимірювань являє собою простий підсилювальний ланцюг і джерело адитивного білого шуму $v(n)$. Згідно цієї моделі генерації сигналу на n -му такті сигнал має вигляд

$$y(n) = cx(n) + v(n). \quad (9.48)$$

Рекурсивна формула оцінки першого порядку має вигляд

$$\hat{x}(n) = b(n)\hat{x}(n-1) + k(n)y(n). \quad (9.49)$$

Відзначимо, що в (9.49) коефіцієнти передачі підсилювальних ланок фільтра залежать від часу (див. рис. 9.8).

Для отримання оптимального (з точки зору методу найменших квадратів) пристрою оцінки середньоквадратична похибка $p(n)$ диференціюється за $b(n)$ і $k(n)$, а результати порівнюються до нуля:

$$p(n) = E[\hat{x}(n) - x(n)]^2 = E[b(n)\hat{x}(n-1) + k(n)y(n) - x(n)]^2, \quad (9.50)$$

$$\frac{dp(n)}{db(n)} = 2E\{[b(n)\hat{x}(n-1) + k(n)y(n) - x(n)]\hat{x}(n-1)\} = 0, \quad (9.51)$$

$$\frac{dp(n)}{dk(n)} = 2E\{[b(n)\hat{x}(n-1) + k(n)y(n) - x(n)]y(n)\} = 0. \quad (9.52)$$

Співвідношення між $b(n)$ і $k(n)$ можна вивести, скориставшись (9.51):

$$2E\{[b(n)\hat{x}(n-1) + k(n)y(n) - x(n)]\hat{x}(n-1)\} = 0 \Rightarrow \\ \Rightarrow E\{[b(n)\hat{x}(n-1)]\hat{x}(n-1)\} = E\{[x(n) - k(n)y(n)]\hat{x}(n-1)\}, \quad (9.53)$$

$$\Rightarrow E\{[b(n)[\hat{x}(n-1) - x(n-1)] + b(n)x(n-1)]\hat{x}(n-1)\} = \\ = E\{[x(n) - k(n)y(n)]\hat{x}(n-1)\}. \quad (9.54)$$

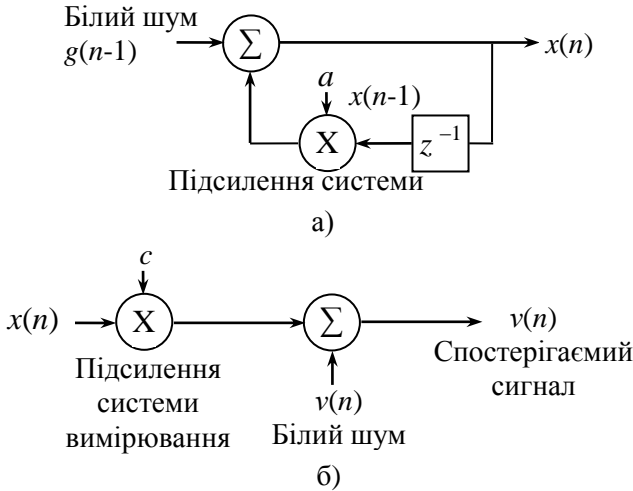


Рисунок 9.7 – Моделі: а) рекурсивної генерації сигналу першого порядку; б) схеми вимірювання даних

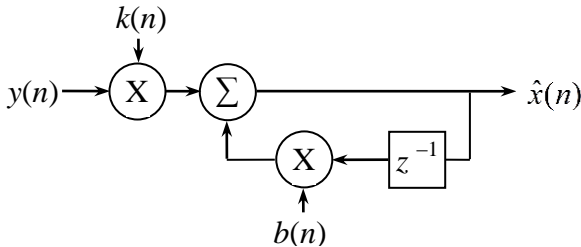


Рисунок 9.8 – Узагальнена структурна схема рекурсивного пристрою оцінювання першого порядку

Підставивши значення $y(n)$ з (9.48), знаходимо

$$\begin{aligned} b(n)E\{e(n-1)\hat{x}(n-1) + x(n-1)\hat{x}(n-1)\} = \\ = E\{[x(n)[1 - ck(n)] - k(n)v(n)]\hat{x}(n-1)\}. \end{aligned} \quad (9.55)$$

Для оптимального пристрою оцінювання повинен виконуватись принцип ортогональності, згідно якого:

$$E\{e(n)\hat{x}(n-1)\} = 0; \quad E\{v(n)\hat{x}(n-1)\} = 0.$$

Тоді рівняння (9.55) набуде вигляду

$$b(n)E[x(n-1)\hat{x}(n-1)] = [1 - ck(n)]E[x(n)\hat{x}(n-1)]. \quad (9.56)$$

З моделі генерації сигналу на рис. 9.7, а маємо

$$x(n) = ax(n-1) + g(n-1). \quad (9.57)$$

Підстановка (9.57) в (9.56), дає

$$\begin{aligned} b(n)E[x(n-1)\hat{x}(n-1)] &= \\ &= [1 - ck(n)]E[ax(n-1)\hat{x}(n-1) + g(n-1)\hat{x}(n-1)]. \end{aligned} \quad (9.58)$$

З рівнянь (9.48) і (9.49) знаходимо

$$\hat{x}(n) = b(n)\hat{x}(n-1) + k(n)cx(n) + k(n)v(n), \quad (9.59)$$

а підстановка $x(n)$ з (9.57) дає:

$$\begin{aligned} \hat{x}(n) &= b(n)\hat{x}(n-1) + k(n)acx(n-1) + k(n)cg(n-1) + k(n)v(n), \\ \hat{x}(n-1) &= b(n-1)\hat{x}(n-2) + ack(n-1)x(n-2) + \\ &+ ck(n-1)g(n-2) + k(n-1)v(n-1), \end{aligned} \quad (9.60)$$

і оскільки середнє всіх добутоків членів (9.60) на $g(n-1)$ дорівнює нулю, можемо записати

$$E\{\hat{x}(n-1)g(n-1)\} = 0.$$

Скориставшись цим співвідношенням, перетворимо (9.58):

$$b(n)E[x(n-1)\hat{x}(n-1)] = a[1 - ck(n)]E[x(n-1)\hat{x}(n-1)]. \quad (9.61)$$

Звідки остаточне співвідношення між $b(n)$ і $k(n)$:

$$b(n) = a[1 - ck(n)]. \quad (9.62)$$

Підставляючи (9.62) в (9.49), знаходимо

$$x(n) = a\hat{x}(n-1) + k(n)[y(n) - ac\hat{x}(n-1)] \quad (9.63)$$

Рівняння (9.63) є визначенням оптимального рекурсивного пристрою оцінки першого порядку або скалярного фільтра Калмана. Перший член $a\hat{x}(n-1)$ прогнозує поточну вибірку, а другий – коригує її на підставі оцінки похибки з урахуванням калманівського коефіцієнта $k(n)$. Структура такого фільтра ілюструється на рис. 9.9.

Визначивши структуру фільтра Калмана, необхідно отримати вираз для змінного у часі коефіцієнта підсилення

Калмана $k(n)$. Спочатку, підставляючи (9.49) в (9.50), визначимо середньоквадратичну похибку у вигляді

$$p(n) = E\{e(n)[b(n)\hat{x}(n-1) + k(n)y(n) - x(n)]\} = \quad (9.64)$$

$$= -E\{e(n)x(n)\}. \quad (9.65)$$

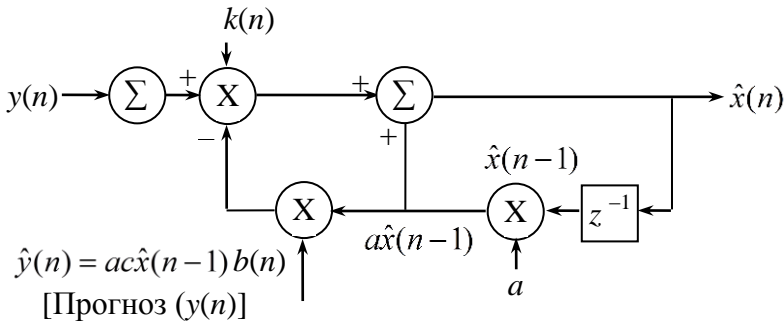


Рисунок 9.9 – Блок-схема скалярного фільтра Калмана першого порядку

Використаємо (9.48) замість $x(n)$ для підстановки в (9.65):

$$cE[e(n)x(n)] = -E[e(n)v(n)]. \quad (9.66)$$

Підставляючи (9.66) в (9.65), знаходимо

$$p(n) = -\frac{1}{c}E[e(n)v(n)]. \quad (9.67)$$

Розкриваючи $e(n)$ і, використовуючи вираз (9.49) для $x(n)$, отримуємо:

$$p(n) = \frac{1}{c}k(n)E[y(n)v(n)] = \frac{1}{c}k(n)\sigma_v^2, \quad (9.68)$$

$$k(n) = \frac{cp(n)}{\sigma_v^2}, \quad (9.69)$$

де $\sigma_v^2 = E[v(n)]^2$. Тепер, підставляючи (9.63) у вираз для середньоквадратичної похибки (9.50), маємо

$$p(n) = E \left\{ a\hat{x}(n-1) + k(n)[y(n) - acx(n-1)] - x(n) \right\}^2, \quad (9.70)$$

а використовуючи (9.48) і (9.57), знаходимо:

$$p(n) = E \left\{ a[1 - ck(n)]e(n-1) - [1 - ck(n)]g(n-1) + k(n)v(n) \right\}^2 = \quad (9.71)$$

$$= a^2 [1 - ck(n)]^2 p(n-1) + [1 - ck(n)]^2 \sigma_g^2 + k^2(n) \sigma_v^2. \quad (9.72)$$

Підстановка (9.68) в (9.72) дає

$$k(n) = \frac{c \left[a^2 p(n-1) + \sigma_g^2 \right]}{\sigma_v^2 + c^2 \sigma_g^2 + c^2 a^2 p(n-1)}. \quad (9.73)$$

Відзначимо, що спочатку, знаючи $p(n-1)$, треба розрахувати, $k(n)$, а потім вже за формулою:

$$p(n) = \frac{1}{c} \sigma_v^2 k(n). \quad (9.74)$$

Три рівняння: (9.63), (9.73) і (9.74) – є рекурсивними рівняннями, які необхідні для реалізації фільтра Калмана першого порядку. На відміну від фільтра Вінера, підсилення фільтра Калмана має бути виражене ітераційним співвідношенням, тобто його не можна представити у вигляді універсального стаціонарного рішення.

9.3.2. Векторний фільтр Калмана

У практичних ситуаціях авторегресійна модель сигналу першого порядку є недостатньою для адекватної характеристики фізичного процесу. Більш адекватною моделлю в цих випадках може виявитись АР-процес N -го порядку. Таким, наприклад, є випадок моделювання трактів передачі даних.

Для розгляду фільтрів вищого порядку, рівняння фільтра Калмана першого порядку, наведені в розд. 9.3.1, можна модифікувати шляхом заміни скалярів векторами N -го порядку. Цей процес найкраще проілюструвати простим прикладом. Найпростішим буде приклад, в якому розглянутий

авторегресійний процес зображується виразом другого порядку, що має вигляд:

$$x(n) = ax(n-1) + bx(n-2) + g(n-1). \quad (9.75)$$

Задамо два змінні стани $x_1(n)$ і $x_2(n)$:

$$x_1(n) = x(n) \text{ і } x_2(n) = x(n-1),$$

та подамо рівняння (9.75) у вигляді двох рівнянь стану:

$$\begin{aligned} x_1(n) &= ax_1(n-1) + bx_2(n-1) + g(n-1), \\ x_2(n) &= x_1(n-1). \end{aligned} \quad (9.76)$$

Записуючи (9.76) у вигляді матричного рівняння, маємо:

$$\begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix} = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(n-1) \\ x_2(n-1) \end{bmatrix} + \begin{bmatrix} g(n-1) \\ 0 \end{bmatrix},$$

або

$$\mathbf{X}(n) = \mathbf{A}\mathbf{X}(n-1) + \mathbf{G}(n-1). \quad (9.77)$$

Рівняння фільтра Калмана для задачі оцінювання представлені тепер у векторній формі. Вони мають вигляд подібний до скалярного фільтра:

$$\hat{\mathbf{X}}(n) = \mathbf{A}\hat{\mathbf{X}}(n-1) + \mathbf{K}(n) \left[\mathbf{Y}(n) - \mathbf{C}\mathbf{A}\hat{\mathbf{X}}(n-1) \right], \quad (9.78)$$

$$\mathbf{K}(n) = \mathbf{P}_1(n) \mathbf{C}^T \left[\mathbf{C}\mathbf{P}_1(n) \mathbf{C}^T + \mathbf{Z}(n) \right]^{-1}, \quad (9.79)$$

$$\mathbf{P}_1(n) = \mathbf{A}\mathbf{P}(n-1)\mathbf{A}^T + \mathbf{Q}(n-1), \quad (9.80)$$

$$\mathbf{P}(n) = \mathbf{P}_1(n) - \mathbf{K}(n)\mathbf{C}\mathbf{P}_1(n), \quad (9.81)$$

де дисперсія скалярного спостережуваного шуму σ_v^2 і дисперсія шуму системи σ_g^2 замінені матрицями $\mathbf{Z}(n)$ і $\mathbf{Q}(n)$ відповідно:

$$\mathbf{Z}(n) = E \left\{ \mathbf{V}(n) \mathbf{V}^T(n) \right\},$$

$$\mathbf{Q}(n) = E \left\{ \mathbf{G}(n) \mathbf{G}^T(n) \right\}.$$

Аналогічно параметр фільтра, або підсилення Калмана, замінений матрицею $\mathbf{K}(n)$ порядку $N \times N$.

Розглянемо *приклад використання векторного фільтра Калмана* для корекції каналу при передачі даних [43].

У випадку, коли довільна послідовність даних передається по каналу зв'язку із спотворенням, для відновлення початкової послідовності даних необхідна фільтрація або корекція сигналу на виході каналу.

Спотворення в каналі можна моделювати за допомогою трансверсального фільтра з лінійною імпульсною характеристикою кінцевої тривалості з доданим шумом на виході. Це модель спостережуваного процесу, показаного на рис. 2.3. Дані на вході каналу мають випадковий розподіл, і, отже, модель генерації сигналу має вигляд:

$$\mathbf{X}(n+1) = \mathbf{A}\mathbf{X}(n) + \mathbf{F}s(n). \quad (9.82)$$

де

$$\mathbf{A} = \begin{bmatrix} 0 & . & . & . & . & 0 \\ 1 & 0 & . & . & . & 0 \\ 0 & 1 & 0 & . & . & 0 \\ \vdots & & & & & \vdots \\ 0 & . & . & 0 & 1 & 0 \end{bmatrix} \text{ і } \mathbf{F} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}. \quad (9.83)$$

Тоді, спостережуваний сигнал каналу під час вибірки з номером n , визначається як:

$$y(n) = \mathbf{C}^T \mathbf{X}(n) + v(n), \quad (9.84)$$

де \mathbf{C} – вектор $N \times 1$ коефіцієнтів каналу, а $v(n)$ – адитивний шум на виході каналу. Якщо \mathbf{A} унітарна матриця, отримане прогнозоване значення $\mathbf{A}\hat{\mathbf{X}}(n-1)$ стає рівним просто $\hat{\mathbf{X}}(n-1)$ і вираз для коригуючого фільтра Калмана має вигляд:

$$\hat{\mathbf{X}}(n) = \hat{\mathbf{X}}(n-1) + \mathbf{K}(n) \left[y(n) - \mathbf{C}^T \hat{\mathbf{X}}(n-1) \right], \quad (9.85)$$

а ітерації для $\mathbf{K}(n)$ визначаються рівняннями (9.79) – (9.81). Тут слід зазначити, що коефіцієнти спотворення каналу \mathbf{C} вважаються апріорно відомими.

Коригуючий фільтр, або вирівнювач, який описується рівнянням (9.85), схематично зображено на рис. 9.10, а, що показує, як добутки різних матриць фізично забезпечують формування прогнозу і корекції. Це можна порівняти з точною багатоканальною матричною реалізацією, показаною на рис. 9.10, б.

У розділі 9.3 на прикладі двох типів лінійних процесів були стисло розглянуті методи оцінювання стохастичних сигналів за допомогою тільки лінійних пристроїв оцінки:

1. Кінцевого або паралельного пристрою оцінювання, що використовує рішення Вінера – Хопфа, яке дає оцінку у вигляді зваженої суми кінцевого числа спостережуваних вибірок.

2. Послідовного або рекурсивного пристрою оцінювання Калмана, що використовує прийнятну модель генерації авторегресивного сигналу, для отримання оцінки, яка може бути суттєво скоригована в результаті аналізу кожної нової вибірки в часовій послідовності.

Розглянуті пристрої оцінки знаходять безпосереднє застосування в ряді областей. Фільтри Вінера найбільш доцільно використовувати у випадку, коли в наявності є лише обмежений обсяг даних. Вони, як правило, застосовуються в таких областях, як сейсмозв'язка і обробка зображень.

Фільтри Калмана найбільш придатні для аналізу неперервного часового ряду i , отже, використовуються автоматичного супроводження цілей в радіолокаційних станціях і обробки сигналів в інерційних навігаційних системах. Поширеним застосуванням методів оцінювання Калмана є розрахунок значень ваг гілок для коригувальних фільтрів з кінцевою імпульсною характеристикою (див. розд. 9.4).

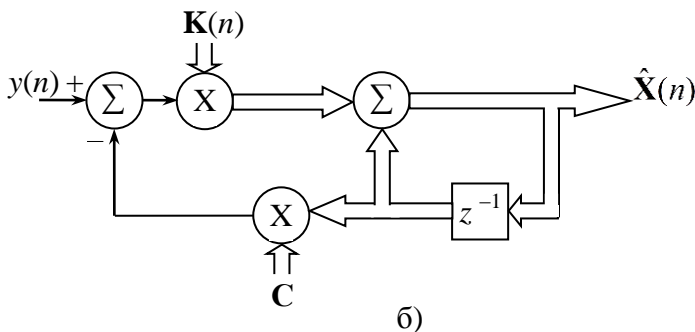
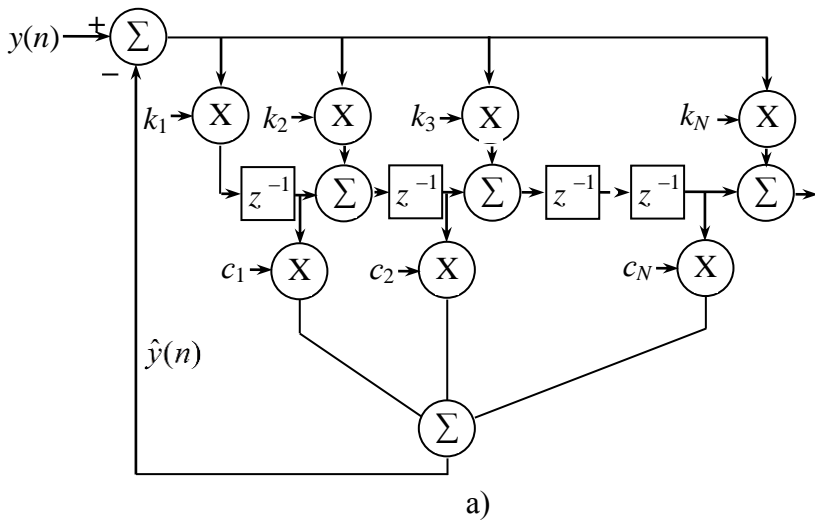


Рисунок 9.10 – Блок-схема коригуючого фільтра Калмана N -го порядку (а); точна реалізація векторного фільтра Калмана (б)

Слід зазначити, що пристрій оцінювання $\hat{x}(n)$ визначений співвідношенням (9.25) неов'язково є лінійною системою. Для вирішення багатьох практичних завдань необхідно застосовувати методи нелінійного оцінювання (наприклад, гомоморфні фільтри в обробці зображень). Однак аналіз методик оцінювання виконаний лише для кількох спеціальних

випадків нелінійних функцій, оскільки загальний нелінійний аналіз поки не розроблений.

9.4. Адаптивні алгоритми для фільтрів з кінцевою імпульсною характеристикою

Адаптивні фільтри зазвичай складаються з двох окремих частин: фільтра, структура якого розрахована на виконання шуканої функції обробки інформації, і адаптивного алгоритму для підстроювання параметрів (коефіцієнтів) цього фільтра. Велика кількість можливих комбінацій структур фільтра і законів адаптивного управління ними призводять іноді до вражаючої різноманітності адаптивних фільтрів.

В цьому навчальному посібнику ми обмежимося розглядом одного з найбільш простих класів структур фільтрів: лінійних фільтрів з кінцевою імпульсною характеристикою. Типовий фільтр КІХ-типу (трансверсальний), реалізований у прямій формі, зображено на рис. 6.3. Кожна вибірка сигналу на виході фільтра є лінійною комбінацією кінцевого числа попередніх вибірок вхідного сигналу. Фільтр не є рекурсивним (тобто не містить зворотного зв'язку). Зазначена властивість призводить до винятково простих адаптивних алгоритмів, як буде показано в наступних розділах.

Після визначення структури фільтра виникає необхідність розробки адаптивного алгоритму для підстроювання його коефіцієнтів. У даній главі ми розглядаємо адаптивні закони, мета яких полягає в тому, щоб мінімізувати енергію на виході фільтра (тобто дисперсію сигналу або вихідну суму квадратів на виході). Необхідність у мінімізації подібної цій функції вартості виникає в багатьох областях застосування, що включають оцінювання методом найменших квадратів, наприклад при адаптивному придушенні шуму, адаптивному лінійному підсиленні та адаптивному спектральному оцінюванні.

У двох наступних розділах розглядаються два адаптивних алгоритми для КІХ-фільтрів типу: рекурсивний алгоритм найменших квадратів (РНК) і алгоритм методу найменших квадратів (МНК) Уідроу-Хопфа. Алгоритм МНК завдяки його простоті є привабливим для багатьох застосувань, в яких необхідно мінімізувати обчислювальні процедури. Алгоритм РНК широко застосовується для ідентифікації систем та аналізу часових рядів. Незважаючи на потенційні переваги, його застосування для обробки сигналів обмежується через більшу обчислювальну складність. В останні роки можливості застосування алгоритму РНК зросли, особливо завдяки його «швидким» версіям, які дозволяють більш ефективно проводити розрахунки.

Алгоритм РНК використовується для адаптивної корекції каналів зв'язку, адаптивної обробки сигналів в антенних решітках та інших завдань [43].

Виведення алгоритму РНК і короткі відомості про його основні властивості представлені в розд. 9.4.1. В розд. 9.4.2 виводиться алгоритм МНК, як апроксимація алгоритму РНК, і обговорюються його властивості і виконувани перетворення. І, нарешті, в розд. 9.4.3 розглядається розрахунок адаптивного фільтра із заданою структурою. Структура фільтра вибирається такою, щоб гарантувати лінійні фазові характеристики.

9.4.1. Рекурсивний алгоритм найменших квадратів

В [43] розглядається наступна вихідна задача, яка є основою для багатьох методів адаптивної фільтрації. На відміну від позначень на рис. 6.3, нехай $y(n)$ – вхідний сигнал фільтра КІХ-типу N -го порядку, а сигнал на виході фільтра

$$e(n) = y(n) - h_1 y(n-1) - \dots - h_N y(n-N), \quad n \geq 0, \quad (9.86)$$

вектор коефіцієнтів фільтра позначимо \mathbf{H} :

$$\mathbf{H}^T = [h_1 \ h_2 \ \dots \ h_N], \quad (9.87)$$

тобто $b_0 = 1$, $b_i = -h_i$ для $i \geq 1$.

Сигнал на виході фільтра, зрозуміло, залежить від коефіцієнтів фільтра.

Далі розглянемо алгоритм, розроблений для мінімізації суми квадратів вихідних вибірок

$$V(n) = \sum_{s=0}^n e^2(s). \quad (9.88)$$

Вектор коефіцієнтів фільтра, для якого ця сума мінімізована, позначимо $\hat{\mathbf{H}}(n)$. Цей тип фільтра можна інтерпретувати як прогнозуючий пристрій (предиктор) за критерієм найменших квадратів для вхідної послідовності $y(n)$. А вихід $e(n)$ – як похибку передбачення. Такий тип задач часто виникає при аналізі часових рядів і обробці сигналів. Відзначимо, що не зробилось ніяких припущень про статистичні властивості послідовності $y(n)$. Проведемо спочатку розгляд детермінованої задачі мінімізації. Щоб краще усвідомити цю проблему, перепишемо (9.86) в матричній формі

$$\begin{bmatrix} e(0) \\ e(1) \\ \vdots \\ \vdots \\ e(n) \end{bmatrix} = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ \vdots \\ y(n) \end{bmatrix} - \begin{bmatrix} 0 & \dots & \dots & \dots & \dots \\ y(0) & & & & 0 \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & y(0) \\ & & & & \vdots \\ y(n-1) & \dots & & & y(n-N) \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ \vdots \\ \vdots \\ h_N \end{bmatrix}. \quad (9.89)$$

$\mathbf{e}(n)$ $\mathbf{y}(n)$ $\mathbf{Y}(n)$ \mathbf{H}

При складанні рівняння (9.89) вважалось, що $y(n) = 0$ для $n < 0$. Іншими словами, дані є «пропущеними через вікно», тобто помноженими на ступінчасту функцію $u_1(n)$, де $u_1(n) = 0$ для $n < 0$ і $u_1(n) = 1$ для $n \geq 0$. Можна вибрати й інші ступінчасті функції, що призведе до різних адаптивних алгоритмів.

Наприклад, коли дані не множаться на ступінчасту функцію, отримаємо «не пропущену через вікно» або «коваріантну» форму цих рівнянь:

$$\begin{bmatrix} e(N) \\ \dots \\ e(n) \end{bmatrix} = \begin{bmatrix} y(N) \\ \dots \\ y(n) \end{bmatrix} - \begin{bmatrix} y(N-1) & \dots & y(0) \\ \dots & \dots & \dots \\ y(n-1) & \dots & y(n-N) \end{bmatrix} \begin{bmatrix} h_1 \\ \dots \\ h_N \end{bmatrix}. \quad (9.90)$$

Спочатку розглянемо випадок, коли дані пропущені через вікно; це призведе до найбільш простої структури рівняння.

Відзначимо, що мінімізація функції вартості $V(n)$ еквівалентна мінімізації квадрата норми в лівій частині рівняння (9.89), тому що

$$V(n) = \|\mathbf{e}(n)\|^2 \triangleq \mathbf{e}^T(n)\mathbf{e}(n). \quad (9.91)$$

Вектор, що мінімізує цю норму, визначається рівнянням:

$$\hat{\mathbf{H}}(n) = \left[\mathbf{Y}^T(n)\mathbf{Y}(n) \right]^{-1} \mathbf{Y}^T(n)y(n). \quad (9.92)$$

Це рівняння випливає з стандартного положення лінійної алгебри, що розглядає рішення перевизначених систем лінійних рівнянь. В принципі, рівняння (9.92) дає розв'язок для задачі адаптивної фільтрації. На кожному відрізку часу можна розрахувати значення (9.92) і визначити величину коефіцієнтів фільтра. Однак це призводить до значного обсягу обчислень, оскільки коефіцієнти щоразу перераховуються з самого початку. Більш прийнятною є рекурсивна форма алгоритму обчислення $\hat{\mathbf{H}}(n)$. Описаний, далі, рекурсивний алгоритм найменших квадратів (РНК) може коригувати коефіцієнти фільтра шляхом повного використання інформації, що міститься в поточному наборі коефіцієнтів. На кожному відрізку часу потрібне лише збільшення обсягу обчислень. Це по суті виконується за допомогою процесу оцінювання калманівського типу, описаних в розділі 9.3.1.

Для отримання рекурсивного алгоритму найменших квадратів розглянемо випадок, коли виконується тільки одне додаткове спостереження $y(n+1)$. Визначаючи

$$\mathbf{P}(n) = \left[\mathbf{Y}^T(n) \mathbf{Y}(n) \right]^{-1}, \quad (9.93)$$

маємо

$$\begin{aligned} \mathbf{P}(n+1) &= \left[\mathbf{Y}^T(n) \quad \boldsymbol{\varphi}(n+1) \right] \left[\begin{array}{c} \mathbf{Y}(n) \\ \boldsymbol{\varphi}^T(n+1) \end{array} \right]^{-1} = \\ &= \left[\mathbf{Y}^T(n) \mathbf{Y}(n) + \boldsymbol{\varphi}(n+1) \boldsymbol{\varphi}^T(n+1) \right]^{-1}, \end{aligned} \quad (9.94)$$

де

$$\boldsymbol{\varphi}^T(n+1) = [y(n) \dots y(n-N+1)]. \quad (9.95)$$

Інверсію матриць в (9.94) можна переписати у вигляді

$$\begin{aligned} \mathbf{P}(n+1) &= \mathbf{P}(n) - \mathbf{P}(n) \boldsymbol{\varphi}(n+1) \times \\ &\times \left[1 + \boldsymbol{\varphi}^T(n+1) \mathbf{P}(n) \boldsymbol{\varphi}(n+1) \right]^{-1} \boldsymbol{\varphi}^T(n+1) \mathbf{P}(n). \end{aligned} \quad (9.96)$$

Це рівняння випливає з матричної тотожності

$$\left[\mathbf{A} + \mathbf{B} \mathbf{C} \mathbf{D} \right]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{C} + \mathbf{D} \mathbf{A}^{-1} \mathbf{B})^{-1} \mathbf{D} \mathbf{A}^{-1}, \quad (9.97)$$

яка виконується для всіх матриць \mathbf{A} , \mathbf{B} , \mathbf{C} і \mathbf{D} сумісних розмірностей (для невивродженої матриці \mathbf{A}).

Потім відзначимо, що

$$\begin{aligned} \mathbf{H}(n+1) &= \mathbf{P}(n+1) \mathbf{Y}^T(n+1) \mathbf{Y}(n+1) = \\ &= \mathbf{P}(n+1) \left[\mathbf{Y}^T(n) \mathbf{y}(n) + \boldsymbol{\varphi}(n+1) y(n+1) \right] = \\ &= \left\{ \mathbf{P}(n) - \mathbf{P}(n) \boldsymbol{\varphi}(n+1) \left[1 + \boldsymbol{\varphi}^T(n+1) \mathbf{P}(n) \boldsymbol{\varphi}(n+1) \right]^{-1} \times \right. \\ &\times \boldsymbol{\varphi}(n+1) \mathbf{P}(n) \left. \right\} \mathbf{Y}^T(n) \mathbf{y}(n) + \mathbf{P}(n+1) \boldsymbol{\varphi}(n+1) y(n+1) = \\ &= \mathbf{H}(n) - \mathbf{P}(n) \boldsymbol{\varphi}(n+1) \left[1 + \boldsymbol{\varphi}^T(n+1) \mathbf{P}(n) \boldsymbol{\varphi}(n+1) \right]^{-1} \times \\ &\times \boldsymbol{\varphi}^T(n+1) \hat{\mathbf{H}}(n) + \mathbf{P}(n+1) \boldsymbol{\varphi}(n+1) y(n+1). \end{aligned} \quad (9.98)$$

Визначимо

$$\begin{aligned}
 \mathbf{K}(n+1) &= \mathbf{P}(n+1)\boldsymbol{\varphi}(n+1) = \\
 &= \mathbf{P}(n)\boldsymbol{\varphi}(n+1) - \mathbf{P}(n)\boldsymbol{\varphi}(n+1)\left[1 + \boldsymbol{\varphi}^T(n+1)\mathbf{P}(n)\boldsymbol{\varphi}(n+1)\right]^{-1} \times \\
 &\times \boldsymbol{\varphi}^T(n+1)\mathbf{P}(n)\boldsymbol{\varphi}(n+1) = \mathbf{P}(n)\boldsymbol{\varphi}(n+1) \times \\
 &\left[1 - \left[1 + \boldsymbol{\varphi}^T(n+1)\mathbf{P}(n)\boldsymbol{\varphi}(n+1)\right]^{-1} \boldsymbol{\varphi}^T(n+1)\mathbf{P}(n)\boldsymbol{\varphi}(n+1)\right] = \\
 &= \mathbf{P}(n)\boldsymbol{\varphi}(n+1)\left[1 + \boldsymbol{\varphi}^T(n+1)\mathbf{P}(n)\boldsymbol{\varphi}(n+1)\right]^{-1}.
 \end{aligned} \tag{9.99}$$

Використовуючи визначення (9.99), з (9.98) маємо

$$\hat{\mathbf{H}}(n+1) = \hat{\mathbf{H}}(n) + \mathbf{K}(n+1)\left[y(n+1) - \boldsymbol{\varphi}^T(n+1)\hat{\mathbf{H}}(n)\right]. \tag{9.100}$$

Рівняння (9.96), (9.99) і (9.100) представляють алгоритм РНК. Початкове значення \mathbf{P} можна отримати або оцінюючи $\left[\mathbf{Y}^T(0)\mathbf{Y}(0)\right]^{-1}$ з початкового набору даних, або просто допускаючи, що $\mathbf{P}(0) = \sigma\mathbf{I}$, де σ – велике (додатне) число; як правило, $\sigma = 100/\text{Var}\{y(0)\}$. Аналогічно, початкове значення \mathbf{H} можна отримати, вважаючи, що $\hat{\mathbf{H}}(0) = -\mathbf{P}(0)\mathbf{Y}^T(0)\mathbf{y}(0)$, або просто вважаючи $\hat{\mathbf{H}}(0) = 0$.

9.4.2. Алгоритм РНК з експоненціальним зважуванням

Адаптивний алгоритм, виведений вище, має нескінченну пам'ять. Значення коефіцієнтів фільтра є функціями всіх попередніх значень вхідного сигналу. Як буде розглянуто далі, часто корисно вводити в алгоритм «множник забування», щоб останні дані мали більшу значущість, ніж більш ранні. Одним із способів реалізації цього є заміна функції вартості у вигляді суми квадратів на експоненціально зважену суму квадратів значень вихідного сигналу:

$$\bar{V}(n) = \sum_{s=0}^n \lambda^{n-s} e^2(s). \tag{9.101}$$

де константа $0 \leq \lambda \leq 1$ визначає ефективну пам'ять алгоритму. Коли $\lambda = 1$, алгоритм, як і раніше, має нескінченну пам'ять. Коли $\lambda < 1$, алгоритм буде мати ефективну пам'ять протягом $\tau = -1/\lg \lambda \approx 1/(1-\lambda)$ вибірок сигналу. Співвідношення для τ можна пояснити, якщо розглянути випадок, коли $e^2(s) = 1$. Тоді

$$\bar{V}(n) = \sum_{s=0}^n \lambda^{n-s} = \frac{1-\lambda^{n+1}}{1-\lambda}. \quad (9.102)$$

Якщо визначити тривалість ефективної пам'яті як величину $n+1 = \tau$, при якій $\bar{V}(\tau)/\bar{V}(\infty) \approx 0,9$, тоді $1-\lambda^\tau = 0,9$ або $\tau = \lg 0,1/\lg \lambda$. Коли константа λ дуже близька до одиниці, $\lg \lambda \approx \lambda - 1$.

Щоб зрозуміти вплив множника забування на рекурсивний алгоритм, запишемо новий критерій похибки у вигляді

$$\bar{V}(n) = \mathbf{e}^T(n) \Lambda(n) \mathbf{e}(n), \quad (9.103)$$

де

$$\Lambda(n) = \text{diag} \{ \lambda^n, \lambda^{n-1}, \dots, \lambda, 1 \}. \quad (9.104)$$

Множення (9.89) на $\Lambda^{1/2}(n)$ і рішення відносно $\mathbf{H}(n)$, як і раніше (див. (9.92)), дасть

$$\mathbf{H}(n) = \left[\mathbf{Y}^T(n) \Lambda(n) \mathbf{Y}(n) \right]^{-1} \mathbf{Y}^T(n) \Lambda(n) \mathbf{y}(n). \quad (9.105)$$

Тепер, знову визначимо $\mathbf{P}(n)$ у вигляді

$$\mathbf{P}(n) = \left[\mathbf{Y}^T(n) \Lambda(n) \mathbf{Y}(n) \right]^{-1}. \quad (9.106)$$

Звідки випливає, що

$$\mathbf{P}^{-1}(n+1) = \lambda \mathbf{Y}^T(n) \Lambda(n) \mathbf{Y}(n) + \boldsymbol{\varphi}(n+1) \boldsymbol{\varphi}^T(n+1) = \quad (9.107)$$

$$\lambda \mathbf{P}^{-1}(n) + \boldsymbol{\varphi}(n+1) \boldsymbol{\varphi}^T(n+1).$$

Обернення (9.107) дає

$$\mathbf{P}(n+1) = \frac{\mathbf{P}(n) - \mathbf{P}(n)\boldsymbol{\varphi}(n+1)\left[\lambda + \boldsymbol{\varphi}^T(n+1)\mathbf{P}(n)\boldsymbol{\varphi}(n+1)\right]^{-1}\boldsymbol{\varphi}^T(n+1)\mathbf{P}(n)}{\lambda} \quad (9.108)$$

Далі відзначимо, що

$$\hat{\mathbf{H}}(n+1) = \mathbf{P}(n+1)\mathbf{Y}^T(n+1)\boldsymbol{\Lambda}(n+1)y(n+1) = \mathbf{P}(n+1)\left[\lambda\mathbf{Y}^T(n)\boldsymbol{\Lambda}(n)y(n) + \boldsymbol{\varphi}(n+1)y(n+1)\right]. \quad (9.109)$$

Подібно до попереднього випадку, легко показати, що (9.100) все ще виконується, хоча при розрахунку $\mathbf{K}(n+1)$ використовується нове визначення $\mathbf{P}(n+1)$:

$$\begin{aligned} \mathbf{K}(n+1) &= \mathbf{P}(n+1)\boldsymbol{\varphi}(n+1) = \\ &= \mathbf{P}(n)\boldsymbol{\varphi}(n+1)\left[\lambda + \boldsymbol{\varphi}^T(n+1)\mathbf{P}(n)\boldsymbol{\varphi}(n+1)\right]^{-1}. \end{aligned} \quad (9.110)$$

Повний алгоритм РНК, таким чином, має наступний вигляд:

- Початок

$$\hat{\mathbf{H}}(0) = 0, \mathbf{P}(0) = \sigma\mathbf{I}, \boldsymbol{\varphi}(0) = 0.$$

- На кожному часовому етапі зробити наступне:

$$e(n) = y(n) - \boldsymbol{\varphi}^T(n)\hat{\mathbf{H}}(n-1) \left[= e_{\hat{\mathbf{H}}(n-1)}(n) \right];$$

Розрахувати сигнал на виході фільтра

$$\mathbf{P}(n) = \frac{\mathbf{P}(n-1) - \mathbf{P}(n-1)\boldsymbol{\varphi}(n)\left[\lambda + \boldsymbol{\varphi}^T(n)\mathbf{P}(n-1)\boldsymbol{\varphi}(n)\right]^{-1}\boldsymbol{\varphi}^T(n)\mathbf{P}(n-1)}{\lambda};$$

Корегувати підсилення

$$\mathbf{K}(n) = \mathbf{P}(n-1)\boldsymbol{\varphi}(n)\left[\lambda + \boldsymbol{\varphi}^T(n)\mathbf{P}(n-1)\boldsymbol{\varphi}(n)\right]^{-1};$$

Корегувати коефіцієнти фільтра

$$\hat{\mathbf{H}}(n) = \hat{\mathbf{H}}(n-1) + \mathbf{K}(n)e(n);$$

Корегувати стани фільтра

$$\boldsymbol{\varphi}(n+1) = \mathbf{Z}\boldsymbol{\varphi}(n) + [y(n+1), 0, \dots, 0]^T,$$

де $\mathbf{Z} = \begin{bmatrix} 0 & & & 0 \\ 1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & 0 & 1 & 0 \end{bmatrix}$ – оператор зсуву вниз.

Структура РНК-адаптивного фільтра зображена на рис. 9.11.

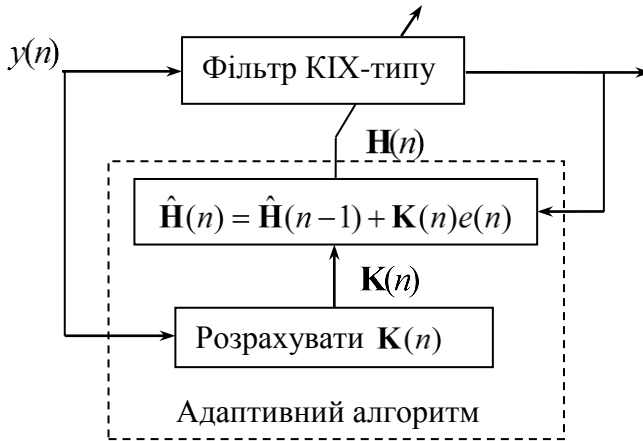


Рисунок 9.11 – Адаптивний фільтр на основі рекурсивного метода найменших квадратів. На виході фільтра отримуємо

$$e(n) = y(n) - \boldsymbol{\varphi}^T(n)\hat{\mathbf{H}}(n-1)$$

9.4.3. Обчислювальна складність

Головною причиною, що може обмежити придатність алгоритму РНК для ряду застосувань, є його обчислювальна складність. Мірою його обчислювальної складності може слугувати кількість операцій, які необхідно виконати на одному кроці алгоритму.

Використовуючи той факт, що $\mathbf{P}(n)$ – симетрична матриця, можна реалізувати алгоритм, який потребуватиме

$2,5N^2 + 4N$ множень і додавань на кожному часовому кроці (N -порядок фільтра). Для фільтрів вищого порядку часто виявляється неприйнятною обчислювальна складність, яка зростає пропорційно квадрату порядку фільтра.

Відзначимо, що для коригування вектора підсилення $\mathbf{K}(n)$ потрібна найбільша кількість операцій. При заданому векторі підсилення на іншу частину алгоритму припадає лише $2N$ множень і додавань. Отже, для зменшення обчислювальної складності алгоритму, необхідно спростити коригування підсилення. Значно простішу формулу для коригування можна отримати, апроксимуючи алгоритм МНК за допомогою градієнтних методів, як буде показано в розд. 3.3. У цьому алгоритмі вектор підсилення пропорційний вектору даних $\boldsymbol{\varphi}(n)$ і, отже, додаткові обчислення не потрібні.

9.4.4. Стохастична інтерпретація

Алгоритм РНК був виведений без прийняття будь-яких припущень про вхідні данні. Його визначає властивість – мінімізації суми квадратів значень сигналу на виході після кожного часового кроку. Цей факт сам по собі дуже корисний. Однак часто буває, що про різні особливості характеристик фільтра бажано знати більше. Щоб детальніше дослідити властивості такого фільтра, необхідно прийняти деякі припущення щодо статистичних характеристик даних. Однак, важливо пам'ятати, що алгоритм буде виконувати задану функцію [мінімізацію $V(n)$ або $\bar{V}(n)$] для довільних вхідних послідовностей. Розглянемо випадок стаціонарного ергодичного стохастичного процесу з даними $y(n)$, нульовим середнім і функцією кореляції

$$r(\tau) = E\{y(n)y(n-\tau)\}. \quad (9.111)$$

Завдяки властивості ергодичності математичні сподівання можна замінити середніми за часом значеннями випадкових величин $y(n)$. Відзначимо також, що кореляційна функція

$$r(\tau) = E \{ y(n)y(n-\tau) \} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{s=\tau}^n y(n)y(n-\tau). \quad (9.112)$$

Звідси випливає

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbf{Y}^T(n) \mathbf{Y}(n) = \mathbf{R} \triangleq \begin{bmatrix} r(0) & \dots & r(N-1) \\ \dots & \dots & \dots \\ r(N-1) & \dots & r(0) \end{bmatrix} \quad (9.113)$$

і

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbf{Y}^T(n) \mathbf{y}(n) = \mathbf{r} \triangleq \begin{bmatrix} r(1) \\ \dots \\ r(N) \end{bmatrix}, \quad (9.114)$$

а, отже,

$$\lim_{n \rightarrow \infty} n \mathbf{P}(n) = \lim_{n \rightarrow \infty} \left[\frac{1}{n} \mathbf{Y}^T(n) \mathbf{Y}(n) \right]^{-1} = \mathbf{R}^{-1} \quad (9.115)$$

і

$$\mathbf{H}_{\text{opt}} \triangleq \lim_{n \rightarrow \infty} \hat{\mathbf{H}}(n) = \lim_{n \rightarrow \infty} \left[\mathbf{Y}^T(n) \mathbf{Y}(n) \right]^{-1} \mathbf{Y}^T(n) \mathbf{y}(n) = \mathbf{R}^{-1} \mathbf{r}. \quad (9.116)$$

Інакше кажучи, вектор коефіцієнтів фільтра сходиться до рішення рівнянь Юла – Уокера

$$\mathbf{R} \mathbf{H}_{\text{opt}} = \mathbf{r}, \quad (9.117)$$

і відповідним чином нормалізована матриця $\mathbf{P}(n)$ сходиться до матриці, яка є зворотною до ковариаційної. Наведені вище міркування є лише правдоподібною аргументацією, але не формальним доказом. Однак, ці факти можна строго довести при деяких не жорстких умовах регулярності.

Залишається розглянути обмежувальний фільтр, тобто фільтр з вектором коефіцієнтів \mathbf{H}_{opt} . Оскільки алгоритм РНК мінімізує суму квадратів вибірок сигналу на виході, тому не дивно, що обмежувальний фільтр, як можна показати, мінімізує дисперсію вихідного процесу. Для розуміння цього розглянемо наступну проблему лінійного оцінювання методом найменших квадратів. Позначимо $\hat{y}(n|n-1)$

предиктор N -го порядку процесу $y(n)$, що має випередження на один такт

$$\hat{y}(n | n-1) = \sum_{i=1}^N a(i)y(n-i). \quad (9.118)$$

Коефіцієнти предиктора $a(i)$ вибираються так, щоб мінімізувати дисперсію похибки передбачення $s(n)$ за методом найменших квадратів

$$\varepsilon(n) = y(n) - \hat{y}(n | n-1) = y(n) - \sum_{i=1}^N a(i)y(n-i). \quad (9.119)$$

Добре відома властивість оцінки за методом найменших квадратів – властивість ортогональності, згідно з якою похибка передбачення $\varepsilon(n)$ не корельована з усіма попередніми даними:

$$E\{\varepsilon(n)y(s)\} = 0 \text{ для } s < n. \quad (9.120)$$

Помноживши рівняння (9.119) на $y(s)$ і, визначивши математичне сподівання, отримаємо:

$$0 = r(n-s) - \sum_{i=1}^N a(i)r(n-s-i). \quad (9.121)$$

Вважаючи $s = n-1, \dots, n-N$ можна переписати рівняння (9.121) в матричній формі

$$\begin{bmatrix} r(0) & \dots & r(N-1) \\ \dots & \dots & \dots \\ r(N-1) & \dots & r(0) \end{bmatrix} \begin{bmatrix} a(1) \\ \dots \\ a(N) \end{bmatrix} = \begin{bmatrix} r(1) \\ \dots \\ r(N) \end{bmatrix}. \quad (9.122)$$

R **A** **r**

Порівнявши з (9.117), дійдемо висновку, що $\mathbf{H}_{\text{opt}} = \mathbf{A}$. Тобто, для стаціонарних вхідних даних РНК-адаптивний фільтр буде сходиться до лінійного фільтра-предиктора на основі МНК з випередженням на один такт (цей фільтр іноді називають фільтром Вінера: він був більш докладно розглянутий в розд. 9.2).

Відзначимо, що вище обговорювався випадок, коли $\lambda = 1$. По суті, для збіжності коефіцієнтів фільтра необхідно, щоб вони залежали від нескінченної кількості вибірок даних. У наступному розділі будуть зроблені деякі висновки, що стосуються випадку кінцевої пам'яті: $\lambda \neq 1$.

Розглянемо отримання оцінки асимптотичної точності алгоритму МНК. Для цього в [43] застосовано теорему, яка формулюється наступним чином: при не жорстких умовах стаціонарності і регулярності для пристрою оцінки за МНК $\hat{\mathbf{H}}(n)$ [див. формулу (9.122)] характерним є асимптотичне нормальний розподіл, тому має місце співвідношення

$$\sqrt{n}(\hat{\mathbf{H}}(n) - \mathbf{H}_{opt}) \rightarrow x, \quad (9.123)$$

де

$$x \sim N(0, \sigma_\varepsilon^2 \mathbf{R}^{-1}), \quad (9.124)$$

$$\sigma_\varepsilon^2 \triangleq E\{e^2(n)\} \text{ (дисперсія похибки передбачення),} \quad (9.125)$$

\mathbf{H}_{opt} визначається з (9.117), а \mathbf{R} – ковариаційна матриця даних (9.113). З цієї теореми витікає, що для достатньо великих значень n ковариаційна матриця вектора похибки оцінювання параметрів фільтра

$$\tilde{\mathbf{H}}(n) \triangleq \hat{\mathbf{H}}(n) - \mathbf{H}_{opt} \quad (9.126)$$

приблизно визначається виразом

$$E\{\tilde{\mathbf{H}}(n)\tilde{\mathbf{H}}^T(n)\} \triangleq \frac{\sigma_\varepsilon^2}{n} \mathbf{R}^{-1}. \quad (9.127)$$

Таким чином, згідно цієї теореми (9.127) вказує нижню границю дисперсії похибок оцінювання параметрів адаптивного фільтра, оскільки при коротких послідовностях даних ковариація $\tilde{\mathbf{H}}(n)$ може бути значно більшою за $(\sigma_\varepsilon^2/n)\mathbf{R}^{-1}$.

Представлена теорема також є корисною для оцінювання «шумових властивостей» коефіцієнтів адаптивного фільтра як

функції числа вибірок даних і статистичного розподілу даних. Якщо $\lambda = 1$, дисперсія коефіцієнтів фільтра буде зменшуватись обернено пропорційно кількості вибірок даних. В усталеному режимі коефіцієнти можна вважати постійними. У випадку, коли $\lambda < 1$, коефіцієнти завжди мають шумову складову з дисперсією $(\sigma_\varepsilon^2/\tau)\mathbf{R}^{-1}$, тобто залежать від ефективної тривалості вікна даних τ . Коли вікно є достатньо тривалим, цілком задовільною оцінкою дисперсії коефіцієнтів фільтра є асимптотична границя (9.127). Якою саме повинна бути тривалість τ для конкретної реалізації адаптивного фільтра може бути визначено методом імітаційного моделювання.

Корисним критерієм робочих характеристик адаптивного фільтра є дисперсія сигналу на його виході. Сигнал на виході фільтра з довільним вектором вагових коефіцієнтів \mathbf{H} визначається формулою [див. (9.86)]

$$e_{\mathbf{H}}(n) = y(n) - \boldsymbol{\varphi}^T(n)\mathbf{H}. \quad (9.128)$$

Сигнал на виході можна представити у вигляді суми «оптимальної» похибки прогнозу і ще одного члена:

$$e_{\mathbf{H}}(n) = \underbrace{y(n) - \boldsymbol{\varphi}^T(n)\mathbf{H}_{opt}}_{\varepsilon(n)} + \boldsymbol{\varphi}^T(n)[\mathbf{H}_{opt} - \mathbf{H}]. \quad (9.129)$$

Звідки маємо

$$E\{e_{\mathbf{H}}^2(n)\} = E\{\varepsilon^2(n)\} + E\left\{\left(\mathbf{H}_{opt} - \mathbf{H}\right)^T \boldsymbol{\varphi}(n)\boldsymbol{\varphi}^T(n)\left(\mathbf{H}_{opt} - \mathbf{H}\right)\right\} + 2E\left\{\varepsilon(n)\boldsymbol{\varphi}^T(n)\left(\mathbf{H}_{opt} - \mathbf{H}\right)\right\}. \quad (9.130)$$

Останній член дорівнює нулю, оскільки похибка прогнозу некорельована з попередніми вибірками даних. Таким чином,

$$E\{e_{\mathbf{H}}^2(n)\} = \sigma_\varepsilon^2 + \left(\mathbf{H}_{opt} - \mathbf{H}\right)^T \mathbf{R}\left(\mathbf{H}_{opt} - \mathbf{H}\right). \quad (9.131)$$

Рівняння (9.129) може бути застосоване для оцінювання дисперсії сигналу на виході фільтра при довільних значеннях коефіцієнтів \mathbf{H} . Як і слід було очікувати, дисперсія на виході

буде мінімізованою тоді й тільки тоді, коли $\mathbf{H} = \mathbf{H}_{opt}$. Її значення σ_ε^2 можна оцінити наступним чином

$$\begin{aligned} E\{\varepsilon^2(n)\} &= E\left\{\left[y(n) - \boldsymbol{\Phi}^T(n)\mathbf{H}_{opt}\right]^2\right\} = \\ &= E\{y^2(n)\} + E\{\mathbf{H}_{opt}^T \boldsymbol{\Phi}(n) \boldsymbol{\Phi}^T(n) \mathbf{H}_{opt}\} - \\ &- 2E\{\mathbf{H}_{opt}^T \boldsymbol{\Phi}(n) y(n)\} = r(0) + \mathbf{H}_{opt}^T \mathbf{R} \mathbf{H}_{opt} + 2\mathbf{H}_{opt} \mathbf{r}. \end{aligned} \quad (9.132)$$

Скориставшись визначенням $\mathbf{H} = \mathbf{H}_{opt}$ з (9.117), отримаємо

$$\sigma_\varepsilon^2 = r(0) - \mathbf{H}_{opt} \mathbf{r} = r(0) - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r}. \quad (9.133)$$

Слід зазначити, що в процесі виконання адаптивного алгоритму коефіцієнти фільтра входять у вирази як випадкові величини. Тому для оцінювання дисперсії потрібно скористатись (9.130), замінивши \mathbf{H} на $\hat{\mathbf{H}}(n-1)$:

$$\begin{aligned} E\{e^2_{\mathbf{H}(n-1)}(n)\} &= \sigma_\varepsilon^2 + E\left\{\left[\mathbf{H}^T(n-1)\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)\tilde{\mathbf{H}}(n-1)\right]^2\right\} = \\ &= \sigma_\varepsilon^2 + \text{tr}E\left\{\left[\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)\right]\left[\tilde{\mathbf{H}}(n-1)\tilde{\mathbf{H}}^T(n-1)\right]\right\} = \\ &= \sigma_\varepsilon^2 + \text{tr}\left\{E\left[\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)\right]E\left[\tilde{\mathbf{H}}(n-1)\tilde{\mathbf{H}}^T(n-1)\right]\right\} = \\ &\sigma_\varepsilon^2 + \text{tr}\left\{\mathbf{R} \frac{\sigma_\varepsilon^2}{n} \mathbf{R}^{-1}\right\} = \sigma_\varepsilon^2 + \frac{N\sigma_\varepsilon^2}{n}; \end{aligned} \quad (9.134)$$

тут припускається, що \mathbf{H} і $\boldsymbol{\Phi}$ некорельовані. Іншими словами, збільшення дисперсії вихідного сигналу внаслідок «шумових властивостей» коефіцієнтів визначається формулою

$$M = \frac{E\{e^2_{\hat{\mathbf{H}}(n-1)}(n)\} - \sigma_\varepsilon^2}{\sigma_\varepsilon^2} = \frac{N}{n}. \quad (9.135)$$

У літературі з обробки сигналів це відношення іноді називають «коефіцієнтом розузгодження» адаптивного

фільтра. Вище наведений результат був введений для випадку нескінченної пам'яті ($\lambda = 1$). У разі кінцевої пам'яті можливо використовувати (9.135), замінивши n на ефективну довжину вікна $\tau \approx 1/(1-\lambda)$, і розглядаючи її як нижню межу істинного коефіцієнта розузгодженості.

9.4.5. Адаптивний алгоритм методу найменших квадратів

Алгоритм методу найменших квадратів, розроблений Уїдроу і Хопфом раніше за адаптивний алгоритм РНК. Алгоритм МНК використовувався для обробки сигналів в багатьох областях застосування. Простота і легкість реалізації даного алгоритму дають йому перевагу при вирішенні багатьох практичних задач. Основні недоліки алгоритму МНК пов'язані з його властивостями збіжності. Цей алгоритм заснований на градієнтному методі мінімізації квадратичної функції, тоді як алгоритм РНК являє собою процедуру типу Ньютона – Рафсона, яка зазвичай сходиться швидше, ніж засновані на градієнтному методі [43].

Спираючись на [43], отримаємо алгоритм МНК і вивчимо деякі його властивості.

Перш ніж ввести алгоритм МНК, перепишемо вираз для дисперсії на виході фільтра в ще одній формі [див. (9.131)]:

$$\begin{aligned} J(\mathbf{H}) &\triangleq E\{e_{\mathbf{H}}^2(n)\} = E\left\{\left[y(n) - \boldsymbol{\phi}^T(n)\mathbf{H}\right]^2\right\} = \\ &= E\{y^2(n)\} - 2E\{y(n)\boldsymbol{\phi}^T(n)\mathbf{H}\} + \mathbf{H}^T E\{\boldsymbol{\phi}(n)\boldsymbol{\phi}^T(n)\}\mathbf{H} = \quad (9.136) \\ &= r(0) - 2\mathbf{r}^T\mathbf{H} + \mathbf{H}^T\mathbf{R}\mathbf{H}, \end{aligned}$$

де $J(\mathbf{H})$ – квадратична функція елементів вектора коефіцієнтів \mathbf{H} .

Метод прискореного спуску є способом оптимізації, в якому використовуються градієнти поверхні для знаходження її мінімуму. Градієнт в будь-якій точці поверхні

розраховується шляхом диференціювання $J(\mathbf{H})$ по вектору параметрів:

$$\frac{\partial J(\mathbf{H})}{\partial \mathbf{H}} \triangleq \begin{bmatrix} \frac{\partial J(\mathbf{H})}{\partial H(1)} \\ \vdots \\ \frac{\partial J(\mathbf{H})}{\partial H(N)} \end{bmatrix} = 2(\mathbf{R}\mathbf{H} - \mathbf{r}) . \quad (9.137)$$

З іншого боку [див. (9.131)],

$$\frac{\partial J(\mathbf{H})}{\partial \mathbf{H}} = 2\mathbf{R}(\mathbf{H} - \mathbf{H}_{\text{opt}}) . \quad (9.138)$$

Прирівнюючи до нуля градієнт вектора, отримуємо оптимальний вектор \mathbf{H}_{opt} , як і в (9.117). Для обчислення цього оптимального вектора параметрів використовується метод прискореного спуску, тобто ітеративна процедура. Почнемо з довільного вектора $\mathbf{H}(0)$, і на кожному кроці будемо змінювати поточний вектор на величину, пропорційну градієнту вектора з протилежним знаком:

$$\mathbf{H}(k+1) = \mathbf{H}(k) + \mu \left[-\frac{\partial J(\mathbf{H})}{\partial \mathbf{H}} \right]_{\mathbf{H}=\mathbf{H}(k)} , \quad (9.139)$$

де скалярний параметр μ – коефіцієнт збіжності, що визначає стійкість і швидкість адаптації.

Для вивчення властивостей цього різницевого рівняння розглянемо вектор похибки

$$\tilde{\mathbf{H}}(k) = \mathbf{H}(k) - \mathbf{H}_{\text{opt}} . \quad (9.140)$$

З (9.138) – (9.140) отримаємо наступне різницеве рівняння для цього вектора похибки:

$$\tilde{\mathbf{H}}(k+1) = (\mathbf{I} - 2\mu\mathbf{R})\tilde{\mathbf{H}}(k) . \quad (9.141)$$

Якщо обрати μ досить малою величиною, тоді завжди можна забезпечити стійкість матриці $(\mathbf{I} - 2\mu\mathbf{R})$. Це гарантує, що похибка буде асимптотично прямувати до нуля.

Щоб ясніше оцінити характер збіжності вектора похибки, перетворимо (9.141) в систему N непарних скалярних рівнянь. Для цього спочатку відзначимо, що коваріаційну матрицю \mathbf{R} , яка є симетричною і додатно визначеною, можна представити у вигляді:

$$\mathbf{R} = \mathbf{U}\mathbf{D}\mathbf{U}^T, \quad (9.142)$$

де \mathbf{U} – ортонормована матриця ($\mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{I}$), а \mathbf{D} – діагональна матриця, що містить характеристичні числа матриці \mathbf{R} :

$$\mathbf{D} = \text{diag}\{\lambda_1, \dots, \lambda_N\}. \quad (9.143)$$

Підставляючи (9.142) в (9.141) і множачи зліва на \mathbf{U}^T , отримуємо:

$$\mathbf{U}^T \tilde{\mathbf{H}}(k+1) = (\mathbf{I} - 2\mu\mathbf{D})(\mathbf{U}^T \tilde{\mathbf{H}}(k)). \quad (9.144)$$

Позначаючи перетворений вектор похибки як

$$\mathbf{X}(k) = \begin{bmatrix} x_1(k) \\ \vdots \\ x_N(k) \end{bmatrix} = \mathbf{U}^T \tilde{\mathbf{H}}(k), \quad (9.145)$$

знаходимо

$$x_i(k+1) = (1 - 2\mu\lambda_i) x_i(k). \quad (9.146)$$

Збіжність забезпечується, якщо виконується нерівність $|1 - 2\mu\lambda_i| < 1$ для всіх характеристичних чисел λ_i . Ця умова буде обов'язково задоволеною при

$$0 < \mu < \frac{1}{\lambda_{\max}}, \quad (9.147)$$

де λ_{\max} – найбільше з характеристичних чисел матриці \mathbf{R} . Відзначимо також, що до тих пір, поки $\mu < \frac{1}{2}\lambda_{\max}$, швидкості збіжності для всіх рішень будуть зростати зі збільшенням μ . Однак, як тільки $\mu = \frac{1}{2}\lambda_{\max}$, мода, яка

відповідає λ_{\max} , почне сповільнюватись. Отже, вибір μ в околиці $\frac{1}{2}\lambda_{\max}$ зазвичай забезпечує найкращу швидкість збіжності для різницевого рівняння (9.141).

Рівняння (9.146) також дозволяє знайти швидкість, з якою різні рішення рівняння для вектора похибки наближаються до нуля. Позначивши часову постійну i -ої моди, як τ_i , отримаємо (вважаючи $\mu\lambda_i \ll 1$)

$$1 - 2\mu\lambda_i = \exp(-1/\tau_i) = 1 - \frac{1}{\tau_i}, \quad (9.148)$$

або

$$\tau_i \approx \frac{1}{2\mu\lambda_i}. \quad (9.149)$$

Таким чином, найбільша постійна часу, що входить в похибку системи, дорівнює

$$\tau_{\max} = \frac{1}{2\mu\lambda_{\min}}, \quad (9.150)$$

де λ_{\min} – найменше характеристичне число коваріаційної матриці \mathbf{R} . З (9.147) і (9.150) знаходимо

$$\tau_{\max} > \frac{\lambda_{\max}}{2\lambda_{\min}}. \quad (9.151)$$

Іншими словами, чим більше розкид характеристичних чисел коваріаційної матриці, тим більше часу потрібно для збіжності при використанні методу прискореного спуску.

Для вищеприписаного ітераційного методу необхідно знати точний градієнт функції вартості $J(\mathbf{H})$. На практиці градієнт апріорно невідомий, і його необхідно оцінити на основі сигналу, що поступає. В алгоритмі МНК для цієї мети використовується градієнт квадратичної похибки єдиної вибірки:

$$\frac{\partial e_{\mathbf{H}}^2(n)}{\partial \mathbf{H}} = 2e_{\mathbf{H}}(n) \frac{\partial e_{\mathbf{H}}(n)}{\partial \mathbf{H}} = -2e_{\mathbf{H}}(n)\boldsymbol{\varphi}(n). \quad (9.152)$$

Отже, алгоритм МНК можна записати як

$$\mathbf{H}(n) = \mathbf{H}(n-1) + 2\mu e_{\mathbf{H}}(n)\boldsymbol{\varphi}(n). \quad (9.153)$$

Підсумовування багатьох членів виду $e(n)\boldsymbol{\varphi}(n)$ дає грубе наближення градієнта єдиної вибірки до істинного градієнту, а очікуване значення вектора коефіцієнтів наблизатиметься до оптимального вектору коефіцієнтів. Відзначимо, що

$$\begin{aligned} E\{-2e_{\mathbf{H}}(n)\boldsymbol{\varphi}(n)\} &= 2\left[E\{\boldsymbol{\varphi}(n)\boldsymbol{\varphi}^T(n)\}\mathbf{H} - E\{y(n)\boldsymbol{\varphi}(n)\}\right] = \\ &= 2[\mathbf{R}\mathbf{H} - \mathbf{r}] = \frac{\partial J(\mathbf{H})}{\partial(\mathbf{H})}. \end{aligned} \quad (9.154)$$

Таким чином, математичне сподівання наближеного значення градієнта дорівнює істинному значенню градієнта. Коротко обговоримо більш точні докази збіжності. Відзначимо простоту МНК-рівняння коректування. Кожен коефіцієнт фільтра коригується шляхом додавання доданка, що містить зважену похибку:

$$h_i(n) = h_i(n-1) + 2\mu y(n-i)e(n). \quad (9.155)$$

Похибка $e(n)$ є спільною для всіх коефіцієнтів $2\mu y(n-i)$, тоді як ваговий коефіцієнт $[2\mu y(n-i)]$ пропорційний поточному значенню даних, накопичених в i -й секції фільтра (див. рис. 6.3). Для обчислення $[2\mu e(n)]$ необхідно виконати $N+1$ операцій (множення і додавання). Таким чином, для кожного кроку коригування фільтра буде потрібно, в цілому, $(2N+1)$ операцій (в порівнянні з $2,5N^2 + 4N$ операцій в алгоритмі РНК).

В [43] при аналізі збіжності алгоритму МНК вказано, що при виконанні умови збіжності алгоритму $\mu\mathbf{D} \ll \mathbf{I}$, мають місце наближені рівняння

$$E\left\{\left[\mathbf{U}^T \tilde{\mathbf{H}}(n)\right]\left[\mathbf{U}^T \tilde{\mathbf{H}}(n)\right]^T\right\} \approx \mu\sigma_{\varepsilon}^2 \mathbf{I}, \quad (9.156)$$

або

$$E\left\{\tilde{\mathbf{H}}(n)\tilde{\mathbf{H}}(n)^T\right\} \approx \mu\sigma_{\varepsilon}^2 \mathbf{I}. \quad (9.157)$$

Як згадувалося в розділі 3.1. дисперсія сигналу на виході фільтра часто використовується в якості міри його робочої характеристики. Розглянемо дисперсію вихідного сигналу та її вплив на ітеративне обчислення вектора коефіцієнтів. З виразів (3.46), (3.58) і (3.63) виходить

$$\begin{aligned} J[\mathbf{H}(k)] &= \sigma_\varepsilon^2 + \tilde{\mathbf{H}}^T(k) \mathbf{R} \tilde{\mathbf{H}}(k) = \sigma_\varepsilon^2 + \mathbf{X}^T(k) \mathbf{D} \mathbf{X}(k) = \\ &= \sigma_\varepsilon^2 + \sum_{i=1}^N \lambda_i [x_i(k)]^2. \end{aligned} \quad (9.157)$$

Було показано, що елементи $\mathbf{X}(k)$ експоненціально прагнуть до нуля з постійними часу, що задаються формулою (9.150). Оскільки функція $J[\mathbf{H}(k)]$ включає суму квадратів елементів $\mathbf{X}(k)$, вона буде експоненціально зменшуватися зі швидкістю, в 2 рази перевищує швидкість загасання $\mathbf{X}(k)$. Інакше кажучи, постійні часу, пов'язані з, дорівнюватимуть половині постійних часу, пов'язаних з. Отже, постійна часу, пов'язана з i -ою модою, дорівнює:

$$\tau_i^J = \frac{1}{2} \tau_i = \frac{1}{4\mu\lambda_i}. \quad (9.158)$$

Криву залежності $J[\mathbf{H}(k)]$ від числа ітерацій в літературі називають «навчальною кривою». Крива починається з

$$J[\mathbf{H}(0)] = \sigma_\varepsilon^2 + \tilde{\mathbf{H}}^T(0) \mathbf{R} \tilde{\mathbf{H}}(0) = \sigma_\varepsilon^2 + \mathbf{X}^T(0) \mathbf{D} \mathbf{X}(0) \quad (9.159)$$

і спадає до кінцевого значення

$$J[\mathbf{H}(\infty)] = \sigma_\varepsilon^2. \quad (9.160)$$

В алгоритмі МНК істинна дисперсія на виході буде більша, ніж передвіщена формулою (9.157), через «шумових властивостей» вектора похибки коефіцієнтів $\tilde{\mathbf{H}}(n)$. З (9.134) і того факту, що $\tilde{\mathbf{H}}(n-1)$ і $\boldsymbol{\varphi}(n)$ не корельовані, слід:

$$E\{e^2_{\mathbf{H}(n-1)}(n)\} = \sigma_\varepsilon^2 + \text{tr}\left[\mathbf{R} E\{\tilde{\mathbf{H}}(n-1) \tilde{\mathbf{H}}^T(n-1)\}\right]. \quad (9.161)$$

Підставляючи їх асимптотичну коваріацію похибки з (3.84) в (3.89), одержуємо

$$E\{e^{2\mathbf{h}(n-1)}(n)\} = \sigma_{\varepsilon}^2 + \mu\sigma_{\varepsilon}^2 \text{tr}\{\mathbf{R}\} = \mu\sigma_{\varepsilon}^2 \sum_{i=1}^N \lambda_i. \quad (9.162)$$

Використовуючи термінологію, введену в літературі з адаптивної фільтрації, визначимо коефіцієнт розузгодження M :

$$M = \frac{E\{e^{2\mathbf{h}(n-1)}(n)\} - \sigma_{\varepsilon}^2}{\sigma_{\varepsilon}^2} = \mu \sum_{i=1}^N \lambda_i = \mu \text{tr}\{\mathbf{R}\}. \quad (9.163)$$

Ця формула добре працює для малих величин розузгодження ($M \leq 0,25$), оскільки для її справедливості необхідно, щоб вектор був досить близький до оптимального значення. Для стаціонарного вхідного сигналу завжди можна досягти малого розузгодження, вибравши досить малі величини μ .

Нагадаємо, що коефіцієнт розузгодження алгоритму РНК визначається як $N/\tau \approx (1-\lambda)N$ [див. (9.135)]. Коефіцієнт розузгодження для алгоритму МНК задається величиною $\mu N \text{Var}\{y(n)\}$ [див. (9.163)]. Можна скласти два алгоритми для роботи з однаковими коефіцієнтами розузгодження, обравши такі параметри μ і λ , щоб $(1-\lambda) = \mu N \text{Var}\{y(n)\}$. Якщо вхідний сигнал нормований так, що його дисперсія дорівнює одиниці, отримаємо $(1-\lambda) = \mu$. Ця формула є корисною для порівняння характеристик двох алгоритмів.

Для розрахунку коефіцієнт розузгодження корисно представити через порядок фільтра і швидкість адаптації. Відзначимо, наприклад, що

$$M = \mu \sum_{i=1}^N \lambda_i = \mu N \lambda_{\text{cp}}, \quad (9.164)$$

де λ_{cp} – середнє значення характеристичних чисел коваріаційної матриці. З (9.158) слід

$$\lambda_i = \frac{1}{4\mu\tau_i^J}. \quad (9.165)$$

Визначивши середню постійну часу

$$\frac{1}{\tau_{\text{cp}}^J} = \frac{1}{N} \sum_{i=1}^N \frac{1}{\tau_i^J}, \quad (9.166)$$

отримуємо,

$$\lambda_{\text{cp}} = \frac{1}{4\mu\tau_{\text{cp}}^J}. \quad (9.167)$$

Таким чином,

$$M = \frac{N}{4\tau_{\text{cp}}^J}. \quad (9.168)$$

Ця формула пов'язує коефіцієнт розузгодження із середнім часом встановлення навчальної кривої і числом коефіцієнтів фільтра.

Ще одне розрахункове обмеження, яке необхідно враховувати, пов'язано зі стійкістю різницевого рівняння

$$\left[\mathbf{U}^T \tilde{\mathbf{H}}(n) \right] = \left[\mathbf{I} - 2\mu\mathbf{D} \right] \left[\mathbf{U}^T \tilde{\mathbf{H}}(n-1) \right] + 2\mu\epsilon(n) \left[\mathbf{U}^T \varphi(n) \right] \quad (9.169)$$

або еквівалентного детермінованого рівняння (9.144). Для забезпечення стійкості матриці $\mathbf{I} - 2\mu\mathbf{D}$ необхідно, щоб $0 < \mu < 1/\lambda_{\text{max}}$. Щоб позбутися обчислення характеристичних чисел ковариаційної матриці вибірок, на практиці часто користуються наступним емпіричним правилом:

$$\frac{1}{\text{tr}\mathbf{R}} > \mu > 0. \quad (9.170)$$

Ця умова впливає з того, що для додатно визначених матриць:

$$\text{tr}\mathbf{R} = \sum_{i=1}^N \lambda_i > \lambda_{\text{max}}, \quad (9.171)$$

і, отже, величина $\mu < \frac{1}{\text{tr}\mathbf{R}}$ гарантує, що $\mu < \frac{1}{\lambda_{\text{max}}}$. Якщо використовувати (9.170), тоді $\text{tr}\mathbf{R} = Nr(0) = N\text{Var}\{y(n)\}$ можна

легко вирахувати (при величині $\text{Var}\{y(n)\}$, оціненої по вхідному сигналу).

Як і в детермінованому випадку [див. (9.151)], швидкість збіжності лежить в межах розкиду характеристичних чисел ковариационної матриці даних.

Адаптивний алгоритм МНК можна розглядати з іншої точки зору - в якості окремого випадку методу стохастичною апроксимації. Стохастична апроксимація – це термін, присвоєний загального класу рекурсивних алгоритмів, що використовуються для вирішення рівнянь виду:

$$g(\mathbf{H}) = E\{G(y, \mathbf{H})\} = 0, \quad (9.174)$$

де G – відома функція вхідного сигналу y , а \mathbf{H} – невідомий вектор коефіцієнтів. У вихідному алгоритмі стохастичною апроксимації, розвиненому в роботі [43], використовується наступна формула коригування для отримання \mathbf{H} на підставі спостережуваного часового ряду:

$$\tilde{\mathbf{H}}(n-1) = \tilde{\mathbf{H}}(n) + \gamma(n)G[y(n+1), \tilde{\mathbf{H}}(n)], \quad (9.175)$$

де $\{\gamma(n)\}$ – відповідним чином обрана скалярна послідовність. Можна показати, що вищенаведений алгоритм сходиться до істинного значення параметрів \mathbf{H} , якщо послідовність $\{\gamma(n)\}$, функція G і сигнал $y(n)$ задовольняють певним умовам:

$$\sum_{n=1}^{\infty} \gamma(n) = \infty, \quad (9.176)$$

$$\sum_{n=1}^{\infty} \gamma^p(n) < \infty \text{ для деяких } p > 1. \quad (9.177)$$

Як правило, вибирають послідовність $\{\gamma(n)\}$ такою, щоб $\gamma(n) = 1/n$. Такий вибір майже напевно призведе до збіжності $\tilde{\mathbf{H}}(n)$ до істинного значення \mathbf{H} .

Для розглянутої нами завдання адаптивної фільтрації маємо:

$$g(\mathbf{H}) = \frac{\partial J(\mathbf{H})}{\partial \mathbf{H}} = E\{2\varphi(n)e(n)\}, \quad (9.172)$$

і в цьому випадку з (9.171) отримуємо

$$\tilde{\mathbf{H}}(n+1) = \tilde{\mathbf{H}}(n) + 2\gamma(n)\varphi(n)e(n). \quad (9.173)$$

Для багатьох застосувань бажано не допускати того, щоб величина $\gamma(n)$ стала занадто малою і не змогла забезпечити реакцію адаптивного алгоритму на зміни статистичних характеристик вхідних даних $y(n)$. Простий спосіб досягти цього полягає в тому, щоб встановити $\gamma(n) = \mu$, де μ – деяке мале додатне число. У такому випадку коефіцієнти фільтра не будуть прямувати до сталого значення, а будуть коливатись поблизу деякого значення. Дана ситуація відповідає випадку кінцевої пам'яті алгоритму РНК, тоді як вибір $\gamma(n) = 1/n$ відповідає випадку нескінченної пам'яті.

Як приклади застосування алгоритмів адаптації на рисунках 9.12 – 9.15 представлені результати імітаційного моделювання, які ілюструють настройку адаптивної антенної решітки (ААР) з виділення корисного сигналу на фоні шумових перешкод, надані професором Скачковим В.В. На рис. 9.12 маємо залежність відношення миттєвої потужності обвідних трьох корисних радіоімпульсів у суміші з власними шумами каналів на виході антенної решітки з $N=11$ приймальними каналами. Відстань між елементами одновимірної лінійної решітки складала $0,75\lambda$, де λ – довжина хвилі, на якій здійснюється приймання корисного сигналу.

Імпульсні сигнали спостерігаються на 30, 100 та 190 кроках спостереження. Відношення сигнал-шум для кожного імпульсу складало 10 дБ. Далі імітувався вплив зовнішніх шумових перешкод від двох джерел розташованих під кутами 15° та -15° по відношенню до напрямку прийому корисного сигналу 0° . Потужність завад від джерел перешкод перевищувала потужність власних шумів в кожному з приймальних каналів відповідно на 30 і 17 дБ. Як видно з

рис. 9.13 – 9.15 (верхня крива) вплив зовнішніх перешкод практично не дозволяє спостерігати корисні сигнали. Нижні криві на цих рисунках ілюструють функціонування алгоритму адаптивного налаштування коефіцієнтів ААР (просторового фільтра): рис. 9.13 відповідає алгоритму РНК, рис. 9.14 – МНК, рис. 9.15 – ілюструє випадок, коли настроювання коефіцієнтів виконується без використання зворотного зв'язку за всією вибіркою отриманих даних.

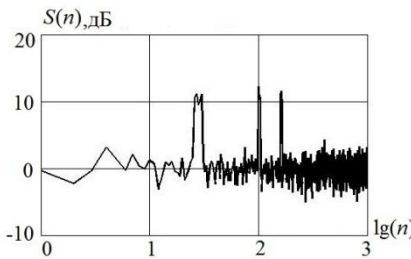


Рисунок 9.12 – Відношення сигнал-шум

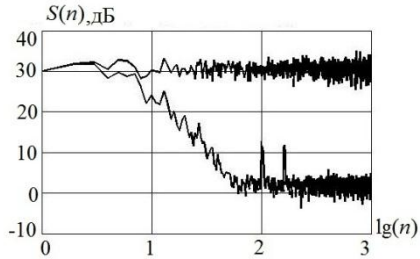


Рисунок 9.13 – Алгоритм адаптації Ньютона

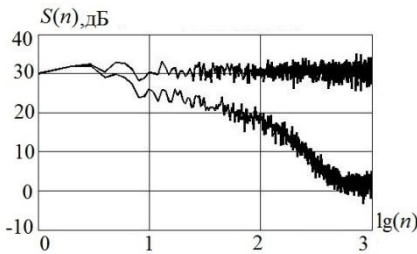


Рисунок 9.14 – Градієнтний алгоритм адаптації

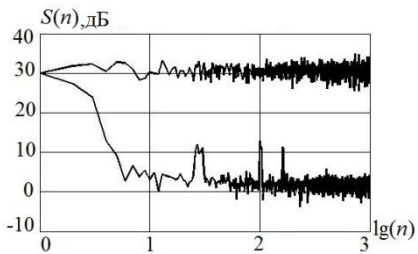


Рисунок 9.15 – Прямий алгоритм адаптації

Контрольні запитання

1. Які лінійні фільтри називають оптимальними? Назвіть критерії, що застосовуються для оптимізації лінійних стаціонарних систем.

2. Запишіть співвідношення для імпульсної характеристики та комплексного коефіцієнта передачі

узгодженого фільтра. Для якого виду перешкод такий фільтр є оптимальним?

3. Сформулюйте необхідну умову фізичної реалізованості узгодженого фільтра.

4. Яку форму та максимальне значення має сигнал на виході узгодженого фільтра?

5. Яким чином можливо знизити рівень бічних пелюсток сигналу на виході фільтра при обробці ЛЧМ сигналу?

6. Імпульсна характеристика якого фільтра є рішенням рівняння Вінера – Хопфа?

7. Запишіть співвідношення для коефіцієнта передачі оптимального фільтра з відтворення сигналу невідомої форми. За яких умов середньоквадратична похибка його відтворення буде дорівнювати нулю?

8. З якою вагою пропускають коливання різних частот оптимальний фільтр з відтворення сигналу невідомої форми та узгоджений з цим сигналом фільтр?

9. Який вигляд для цифрового нерекурсивного фільтра мають рівняння Вінера – Хопфа та його рішення?

10. Як впливає підвищення порядку вінерівського КІХ-фільтра на залишкову середньоквадратичну похибку?

11. Які етапи проведення оцінювання включає в себе процес адаптивної фільтрації.

12. Дайте визначення адаптивного фільтра. Назвіть приклади застосування таких фільтрів.

13. Побудуйте блок-схему та поясніть принцип дії адаптивного фільтра.

14. Які два підходи можуть застосуватись для моделювання невідомої лінійної системи? В чому відмінність результатів їх застосування?

15. Які різновиди цифрових фільтрів можуть застосовуватись для реалізації адаптивного фільтра?

16. Запишіть рівняння та побудуйте узагальнену структурну схему рекурсивного пристрою оцінювання першого порядку.

17. Побудуйте блок-схему та запишіть рівняння

необхідні для реалізації фільтра Калмана першого порядку скалярного фільтра Калмана.

18. Запишіть рівняння фільтра Калмана для задачі оцінювання у векторній формі.

19. Побудуйте блок-схему коригуючого фільтра Калмана N -го порядку.

20. На яких методах мінімізації квадратичної функції засновані алгоритми МНК та РНК?

21. З яких двох основних блоків складається адаптивний фільтр?

22. Що є головною причиною, яка може обмежити придатність алгоритму РНК?

23. Які особливості має алгоритм РНК з експоненціальним зважуванням?

24. Запишіть співвідношення для «коефіцієнту розузгодження» адаптивного фільтра.

25. Які переваги і недоліки має алгоритм МНК порівняно з алгоритмом РНК?

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Теория передачи сигналов: Учебник для вузов / А. Г. Зюко, Д. Д. Кловский, М. В. Назаров, Л. Ф. Финк. – 2-е изд., перераб. и доп. – М.: Радио и связь, 1986. – 304 с.
2. ДСТУ 2681-94. Метрологія. Терміни та визначення.
3. Величко О. М. Оцінювання результатів вимірювань: основи і нормативне забезпечення / О. М. Величко, Л. В. Коломієць, Т. Б. Гордієнко. – Одеса: ВМВ, 2010. – 380 с.
4. Сергеев А. Г. Метрология: Учеб. пособие для вузов / А. Г. Сергеев, В. В. Крохин. – М.: Логос, 2001. – 408 с.
5. Бендат Дж. Прикладной анализ случайных данных : Пер. с англ. / Дж. Бендат, А. Пирсол. – М.: Мир, 1989. – 540 с.
6. Сато Ю. Обработка сигналов. Первое знакомство / Ю. Сато. – ОДЕКА, 1999. – 175 с.
7. Компьютер как источник помех. – Режим доступу до статті: <http://www.electrosad.ru/Jornal/VidPom.htm>.
8. Защита от радиопомех / М. В. Максимов, М. П. Бобнев, Б. Х. Кривицкий [и др.] ; ред. Максимов М. В. – М.: Сов. радио, 1976. – 497 с.
9. Волощук Ю. Л. Сигнали та процеси в радіотехніці. Том. 1-4 / Ю. Л. Волощук . – Харків: «СМІТ», 2003.
10. Баскаков С. И. Радиотехнические цепи и сигналы. – М.: Высшая школа, 1988. – 448 с.
11. Сергиенко А. Б. Цифровая обработка сигналов / А. Б. Сергиенко. – СПб.: Питер, 2002. – 608 с.
12. Цифровой сигнальный процессор. Материал из Википедии – свободной энциклопедии. – Электронний ресурс. – Режим доступу: https://ru.wikipedia.org/wiki/Цифровой_сигнальный_процессор.

13. Экстраполяция. Материал из Википедии – свободной энциклопедии. Электронный ресурс. – Режим доступа: <http://ru.wikipedia.org/wiki/Экстраполяция>.

14. Аліасинг. Матеріал з Вікіпедії – вільної енциклопедії. Електронний ресурс. – Режим доступу: <http://uk.wikipedia.org/wiki/Аліасинг>.

15. Гетманов В. Г. Цифровая обработка сигналов: Учебное пособие. Изд. 2-е, расш. и перераб. – М.: НИЯУ МИФИ, 2010. – 232 с.

16. Вострокнутов Н. Н. Цифровые измерительные устройства. Теория погрешностей, испытания, поверка / Н. Н. Вострокнутов. – М.: Энергоатомиздат, 1990. – 208 с.

17. Баскаков С. И. Радиотехнические цепи и сигналы / С. И. Баскаков. – М.: «Высшая школа», 2003. – 446 с.

18. Гоноровский И. С. Радиотехнические цепи и сигналы : Учебник для вузов. – 4-е изд. перераб. и доп. / И. С. Гоноровский. – М.: Радио и связь, 1986. – 512 с.

19. Тихонов В. И. Статистический анализ и синтез радиотехнических устройств и систем: Учеб. пособие для вузов / В. И. Тихонов, В. Н. Харисов. – М.: Радио и связь, 1991. – 608 с.

20. Алмазов В. Б. Основы теории радиолокации / В. Б. Алмазов. – Харьков: ВИРТА им. Говорова Л. А., 1992. – 204 с.

21. Поддубный Г. В. Математический анализ для радиоинженеров / Г. В. Поддубный, Р. К. Романовский. – М.: Воениздат, 1976. – 344 с.

22. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров: Пер. с англ. / Под ред. П. Г. Арамановича. – М.: Наука, 1978. – 832 с.

23. Электронные учебники Мордовского университета. Электронный ресурс. – Режим доступа: http://do1.fet.mrsu.ru/text/ucheba/asoju/metodichki/OTY_ELESTR/resources/resource_0/id_16/1.htm#p1.

24. Рабинер Л. Теория и применение цифровой обработки сигналов / Л. Рабинер, Б. Гоулд; пер. с англ. под ред. Ю. Н. Александрова. – М.: Мир, 1978. – 848 с.

25. Марпл.-мл. С. Л. Цифровой спектральный анализ и его приложения: Пер. с англ. – М.: Мир, 1990. – 584 с.

26. Спектральный анализ на ограниченном интервале времени. Оконные функции. [Электронный ресурс]. – Режим доступа <http://www.dsplib.ru/content/win/win.html>.

27. Моделирование в радиолокации / А. И. Леонов, В. Н. Васенев, Ю. И. Гайдуков и др. / Под ред. А. И. Леонова. – М.: Советское радио, 1979. – 264 с.

28. Радіотехніка: Енциклопедичний навчальний довідник: Навч. посібник / За ред. Ю. Л. Мазора, Є. А. Мачуського, В. Ш. Правди. – К.: Вища школа, 1999. – 838 с.

29. Коваль Ю. О., Гринченко Л. В., Милютченко І. О., Рибін О. І. Основи теорії кіл. Підручник для студентів вищих навчальних закладів. Ч. 2 / За заг. редакцією В. М. Шокала та В. І. Правди. – Х.: Компанія СМІТ, 2008. – 560 с.

30. Калашніков А. Ю. Перетворення сигналів: навч. посіб. для бакалаврів / Калашніков А. Ю., Шкуліпа А. В., Горелік С. М. – Одеса: ОНАЗ ім. О. С. Попова, 2012. – Ч. 1 і 2. – 100 с.

31. Двайт Г. Б. Таблицы интегралов и другие математические формулы / Г. Б. Двайт. – М.: Наука, 1977.

32. Павленко Ю. Ф. Измерение параметров частотно-модулированных колебаний / Ю. Ф. Павленко, П. А. Шпаньон. – М.: Радио и связь, 1986. – 208 с.

33. Заездный А. М. Основы расчетов по статистической радиотехнике. – М.: Издательство «Связь», 1969. – 447 с.

34. Приемно-передающие устройства радиотехнических систем. Часть 2. / Ю. Н. Седышев,

В. Н. Рогачев, В. И. Гапон и др.; Под общей ред. Ю. Н. Седышева. – Харьков: ВИРТА им. Говорова Л. А., 1992. – 321 с.

35. Ельчанинов А. М. Радиопередающие устройства радиоэлектронной техники / А. М. Ельчанинов, Д. А. Шаров, А. П. Омельчук. – М.: Воениздат, 1991. – 320 с.

36. Новиков Л. В. Основы вейвлет-анализа сигналов. Учебное пособие, 1999. – 152 с.

37. Дьяконов В. П. Вейвлеты. От теории к практике. Изд. 2-е, перераб. и доп. – М.: Солон-Пресс, 2004. – 400 с.

38. Залманзон Л. А. Преобразование Фурье, Уолша, Хаара и их применение в управлении, связи и других областях. – М.: Наука. Гл. ред. физ.-мат. лит., 1989. – 496 с.

39. Смоленцев Н. К. Основы теории вейвлетов. Вейвлеты в MATLAB. – М.: ДМК Пресс, 2005. – 304 с.

40. Астафьева Н. М. Вейвлет-анализ: основы теории и примеры применения // Успехи физических наук. – 1996. – том 166, № 11. – С. 1145–1170.

41. Вакман Д. Е. Сложные сигналы и принцип неопределенности в радиолокации / Д. Е. Вакман. – М.: Сов. радио, 1965. – 304 с.

42. Зюко А. Г. Теория передачи сигналов. Учебник для вузов / А. Г. Зюко, Ю. Ф. Коробов. – М.: «Связь», 1972. – 282 с.

43. Адаптивные фильтры: Пер. с англ. / Под. ред. К. Ф. Н. Коуэна и П. М. Гранта. – М.: Мир, 1988. – 392 с.

Теорема Котельникова

Доведена радянським вченим В. А. Котельниковим у 1933 році.

Теорема. Довільний сигнал, неперервний спектр якого не містить частот вище $f_g = \omega_g / 2\pi$, може бути повністю відновлений, якщо відомі відлікові значення цього сигналу узяті через рівні інтервали часу $\Delta t = 1/2f_g = \pi/\omega_g$.

Теорема Котельникова (в англійській літературі теорема Шенона-Найквіста. Шенон довів теорему незалежно у 1949 р.) встановлює принципову можливість повного відновлення детермінованого сигналу з обмеженим спектром і вказує граничне значення кроку (інтервалу) дискретизації, при якому таке відновлення можливе.

Доведення теореми

Нехай функція $x(t)$, що описує сигнал, і дискретизується, має обмежену спектральну густину $S(j\omega)$, причому

$$S(j\omega) = 0 \text{ при } |\omega| > \omega_g, \quad (\text{A.1})$$

де ω_g – найбільша частота в спектрі сигналу $x(t)$. Використовуючи обернене перетворення Фур'є з урахуванням співвідношення (A.1), запишемо: 10

$$x(t) = \frac{1}{2\pi} \int_{-\omega_g}^{\omega_g} S(j\omega) e^{j\omega t} dt. \quad (\text{A.2})$$

Для будь-яких моментів часу, наприклад $t = k\Delta t = k \frac{\pi}{\omega_g}$

, де k – будь-яке ціле число, функція $x(t)$ набуває значень

$$x(k\Delta t) = \frac{1}{2\pi} \int_{-\omega_g}^{\omega_g} S(j\omega) e^{j\frac{k\pi}{\omega_g}\omega} d\omega. \quad (\text{A.3})$$

Розглядаючи спектральну щільність $S(j\omega)$ як функцію частоти з періодом $2\omega_g$ і періодично продовжуючи її з цим періодом, розкладемо $S(j\omega)$ в ряд Фур'є на інтервалі частот $[-\omega_g, \omega_g]$

$$S(j\omega) = \sum_{k=-\infty}^{\infty} C_k e^{j\frac{k\pi}{\omega_g}\omega},$$

де коефіцієнти розкладання дорівнюють

$$C_k = \frac{1}{2\omega_g} \int_{-\omega_g}^{\omega_g} S(j\omega) e^{j\frac{k\pi}{\omega_g}\omega} d\omega. \quad (\text{A.4})$$

Порівнюючи (A.3) і (A.4), бачимо, що $2C_k\omega_g = 2\pi x(-k\Delta t)$, звідки визначаємо

$$C_k = \frac{\pi}{\omega_g} x(-k\Delta t).$$

Виразимо $S(j\omega)$ через відліки початкової функції

$$S(j\omega) = \frac{\pi}{\omega_g} \sum_{k=-\infty}^{\infty} x(-k\Delta t) e^{j\frac{k\pi}{\omega_g}\omega} = \frac{\pi}{\omega_g} \sum_{k=-\infty}^{\infty} x(-k\Delta t) e^{j\omega k\Delta t}. \quad (\text{A.5})$$

Оскільки підсумовування ведеться як по додатних, так і по від'ємних числах k , знак перед k у виразі (A.5) можна змінити на зворотний, тому

$$S(j\omega) = \frac{\pi}{\omega_g} \sum_{k=-\infty}^{\infty} x(k\Delta t) e^{-j\omega k\Delta t}. \quad (\text{A.6})$$

Підставивши (A.6) у вираз (A.2), визначимо значення початкової функції у будь-який момент часу

$$S(j\omega) = \frac{1}{2\omega_g} \int_{-\omega_g}^{\omega_g} \left\{ \sum_{k=-\infty}^{\infty} x(k\Delta t) e^{-j\omega k\Delta t} \right\} e^{j\omega t} d\omega.$$

Враховуючи збіжність ряду Фур'є, змінимо порядок підсумовування й інтегрування

$$x(t) = \frac{1}{2\omega_g} \sum_{k=-\infty}^{\infty} x(k\Delta t) \int_{-\omega_g}^{\omega_g} e^{j\omega(t-k\Delta t)} d\omega. \quad (\text{A.6})$$

У отриманому виразі обчислимо інтеграл

$$\int_{-\omega_g}^{\omega_g} e^{j\omega(t-k\Delta t)} d\omega = \frac{e^{j\omega(t-k\Delta t)} \Big|_{-\omega_g}^{\omega_g}}{j(t-k\Delta t)} = \frac{2 \sin \omega_g(t-k\Delta t)}{(t-k\Delta t)}.$$

Підставивши результат обчислення інтеграла в (A.6), остаточно отримаємо

$$x(t) = \sum_{k=-\infty}^{\infty} x(k\Delta t) \frac{\sin \omega_g(t-k\Delta t)}{\omega_g(t-k\Delta t)}. \quad (\text{A.6})$$

Отже, неперервна функція з обмеженим спектром може бути представлена безліччю своїх значень (відліків), узятими в моменти часу $t = k\Delta t = k \frac{\pi}{\omega_g}$. Вираз (A.6) є рядом Котельникова, в якому роль коефіцієнтів виконують відліки функції $x(k\Delta t)$, а базисними є функції виду

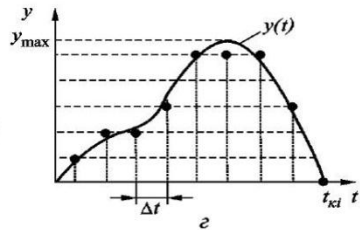
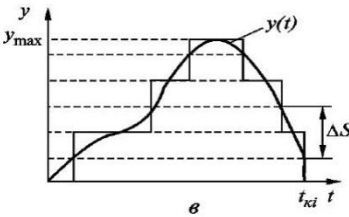
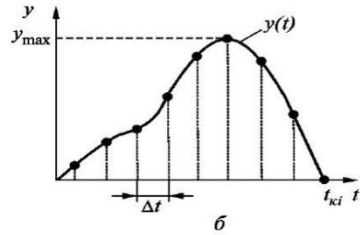
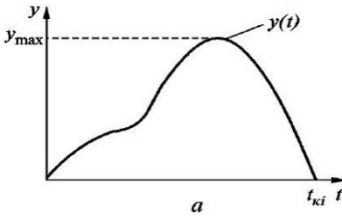
$$\varphi_k(t) = \frac{\sin \omega_g(t-k\Delta t)}{\omega_g(t-k\Delta t)},$$

які називаються функціями відліків.

**Рекомендований перелік питань
для тестування**

1. На зображеннях сигналів поруч з *a*, *б*, *в*, *г* записати відповідні їм цифри: 1, 2, 3, 4, де сигнали:

- 1 – Дискретний; 2 – Аналоговий;
3 – Цифровий; 4 – Квантований.



2. Відмітити вірні ознаки для сигналу *a* на рисунку в питанні 1.

- | | |
|----------------|------------------|
| а. імпульсний | д. неперервний |
| б. випадковий | е. не випадковий |
| г. періодичний | ж. неперіодичний |

3. Середня потужність сигналу $s(t) = A \cdot \sin(\omega t + \varphi)$, де $A=1$ В, період сигналу $T=2$ мс, на резисторі 1 Ом, складає:

- а. 0,5 Вт; б. 0,25 Вт; в. 2 Вт.

4. Миттєва потужність сигналу $s(t)$: $p(t) =$
 а. $s^2(t)/2$; б. $s^2(t)$; в. $2s^2(t)$.

5. Прямокутний імпульс $s(t) = A[\sigma(t - \tau) - \sigma(t - \tau - T)]$
 має параметри: амплітуда 5 В; тривалість 1 мс; затримка за
 часом 2 мс ($\sigma(t)$ – функція Хевісайда). Вказати значення

$$\tau = \quad ; \quad A = \quad ; \quad T = \quad .$$

6. Енергія прямокутного імпульса з вказаними у 5
 питанні параметрами складає:

а. $0,005 \text{ В}^2 \cdot \text{с}$; б. $0,025 \text{ В}^2 \cdot \text{с}$; в. $0,01 \text{ В}^2 \cdot \text{с}$.

7. $\{\varphi_j(t)\}_{j=1}^N$ система базисних функцій, для якої
 виконується співвідношення

$$(\varphi_i, \varphi_j) = \int_{-\infty}^{\infty} \varphi_i(t) \varphi_j(t) dt = \begin{cases} c, & \text{при } i = j, \\ 0, & \text{при } i \neq j. \end{cases}$$

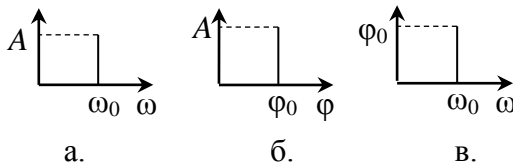
Система таких базисних функцій при $c = \text{const} \in$:

- а. Ортонормованою;
- б. Ортогональною;
- в. Загального вигляду.

8. $\delta(t - t_0) = \begin{cases} 0, & \text{при } t \neq t_0, \\ \infty, & \text{при } t = t_0. \end{cases}$ це функція

- а. Уолша;
- б. Хевісайда;
- в. Хемінга;
- г. Дірака.

9. Амплітудний спектр сигналу $s(t) = A \cdot \cos(\omega_0 t + \varphi_0)$ має
 вигляд:



10. На якому з рисунків питання 9 представлено фазовий спектр сигналу $s(t)$?

а. б. в.

11. Оберіть вираз, що є рядом Котельникова:

а. $V(t) = \sum_{j=0}^N a_j (t - t_0)^j,$

б. $x(t) = \sum_{k=-\infty}^{\infty} x(k\Delta t) \frac{\sin \omega_c (t - k\Delta t)}{\omega_c (t - k\Delta t)},$

в. $Y(t) = \sum_{k=0}^{\infty} A_m [\sigma(t - kT) - \sigma(t - kT - \tau)].$

12. Автокореляційна функція детермінованого сигналу $s(t)$ з кінцевою енергією обчислюється за формулою

а. $B_S(\tau) = \int_{-\infty}^{\infty} s(t)s(t - \tau) dt;$

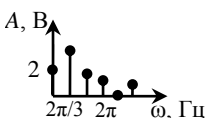
б. $B_{12}(\tau) = \int_{-\infty}^{\infty} s_1(t)s_2(t - \tau) dt;$

в. $B_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}_{12}(\omega) e^{j\omega\tau} d\omega.$

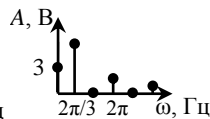
13. Дискретизований неперіодичний сигнал на вході дискретного фільтра має такі відліки: 1, 2, 4. Імпульсна характеристика фільтра має вигляд: 1, 2, 3, 1. Яким буде максимальне значення сигналу на виході фільтра:

а. 12; б. 17; в. 15; г. 14.

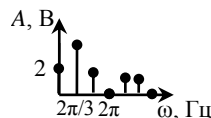
14. Амплітудний спектр послідовності імпульсів за умов: $T=3$ с; $\tau=1$ с і $A = 6$ В має вигляд:



а.



б.



в.

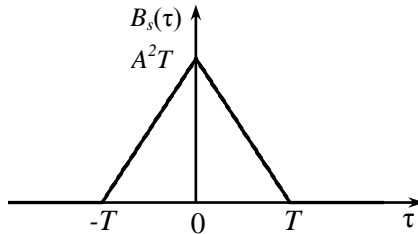
15. На якому з рисунків а-в питання 14 представлено спектр сигналу типу «меандр»?

а. б. в.

16. Автокореляційна функція (АКФ) детермінованого сигналу $s(t)$ з кінцевою енергією є _____ функцією свого аргументу τ . Обрати пропущене слово:

а. парною; б. непарною; в. періодичною.

17. АКФ імпульса приведено на рисунку.



Яку форму він має?

а. Трикутний;
 б. Прямокутний;
 в. Гаусівський.

18. Рівність $E = \int_{-\infty}^{\infty} s^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\dot{S}(\omega)|^2 d\omega$ має назву

а. Фур'є;
 б. Парсеваля;
 в. Рэлея.

19. Теорема Вінера-Хинчина визначається співвідношенням

а. $D_x = \frac{1}{2\pi} \int_{-\infty}^{\infty} W(\omega) d\omega;$

б. $K_x(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2, t_1, t_2) dx_1 dx_2;$

в. $R(\tau) = \frac{1}{\pi} \int_0^{\infty} W(\omega) \cos(\omega\tau) d\omega.$

20. Пряме дискретне перетворення Фур'є (ДПФ) обчислюється за формулою

$$\text{а. } \dot{S}(\omega) = \int_{-\infty}^{\infty} s(t) e^{-j\omega t} dt;$$

$$\text{б. } \dot{C}_i = \frac{1}{T} \int_{-T/2}^{T/2} s(t) e^{-ji\omega_1 t} dt;$$

$$\text{в. } \dot{c}(n) = \frac{1}{N} \sum_{k=0}^{N-1} x(k) e^{-j\frac{2\pi nk}{N}} .;$$

$$\text{г. } x(k) = \sum_{n=0}^{N-1} c(n) e^{j\frac{2\pi}{N} kn};$$

21. Перехідна характеристика RL -ланки

$$g_R(t) = 1 - \exp(-Rt/L),$$

тоді імпульсна характеристика $h_R(t)$ дорівнює (використати інтеграл Дюамеля у формі

$$s_{\text{вих}}(t) = s_{\text{вх}}(t)g(0) + \int_0^t \frac{g(t')}{dt'} s_{\text{вх}}(t-t') dt');$$

$$\text{а. } (R/L)\exp(-Rt/L);$$

$$\text{б. } \delta(t) - (R/L)\exp(-Rt/L);$$

$$\text{в. } (R/L)\exp(Rt/L);$$

$$\text{г. } \sigma(t) - (R/L)\exp(-Rt/L).$$

22. Як, за умов питання 21, обчислюється $h_L(t)$, якщо

$$g_L(t) = \exp(-Rt/L).$$

а. б. в. г.

23. Якщо дискретна послідовність $\{ax(k)+by(k)\}$, тоді n -на складова її спектру (тут спектри $x(k)$ і $y(k) - c_1(n)$ та $c_2(n)$)

$$\text{а. } ac_1(n) \times b c_2(n);$$

$$\text{б. } ac_1(n) - b c_2(n);$$

$$\text{в. } ac_1(n)/[b c_2(n)];$$

$$\text{г. } ac_1(n) + b c_2(n).$$

24. Спектральна густина довільного сигналу $s(t)$ обчислюється за формулою:

$$\text{а. } \dot{S}(\omega) = \int_{-\infty}^{\infty} s(t)e^{-j\omega t} dt$$

$$\text{б. } \dot{C}_i = \frac{1}{T} \int_{-T/2}^{T/2} s(t)e^{-ji\omega_1 t} dt$$

$$\text{в. } A(\omega) = \int_{-\infty}^{\infty} s(t) \cos \omega t dt$$

$$\text{г. } B(\omega) = \int_{-\infty}^{\infty} s(t) \sin \omega t dt .$$

25. За якими формулами з пит. 24 можна обчислити спектральну густина сигналу, якщо $s(t)$ – парна функція

а. б. в. г.

26. Білий шум має спектральну густина потужності

$$\text{а. } W(\omega) = W_0 = \text{const}$$

$$\text{б. } W_x(\omega) = W_0 \exp(-\beta \omega^2)$$

$$\text{в. } N_x(f) = \frac{2\sigma_x^2 \beta}{\beta^2 + (2\pi f)^2}$$

27. Нехай виконується рівність

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

При якому значенні n процес є стаціонарним у широкому сенсі?

а. 3 б. 1 в. 2

28. Дисперсія стаціонарного випадкового процесу на виході лінійної системи не обчислюється за формулою

а. $\frac{1}{2\pi} \int_{-\infty}^{\infty} W_{\text{вх}}(\omega) |\dot{K}(\omega)|^2 d\omega;$

б. $\int_{-\infty}^{\infty} N_{\text{вх}}(f) \cdot |\dot{K}(f)|^2 df;$

в. $\int_{-\infty}^{\infty} h(t)h(t-\tau) dt.$

29. ДПФ кругової згортки послідовностей $\{x_1(k)\}$ і $\{x_2(k)\}$ однакової довжини N , ДПФ яких відповідно рівні $c_1(n)$ і $c_2(n)$, дорівнює (тут * позначає операцію згортки):

а. $c_1(n) \times c_2(n);$

б. $c_1(n) + c_2(n);$

в. $c_1(n) * c_2(n).$

30. Різницеве рівняння дискретного фільтра має вигляд $u(n) = x(n) + \frac{1}{2}u(n-1) - \frac{1}{4}u(n-2);$ Тоді його імпульсна характеристика є:

а. кінцевою;

б. нескінченною;

в. не можна вказати.

31. Сигнал $S(t) = A(t)\cos(\omega_0 t + \varphi_0)$ є :

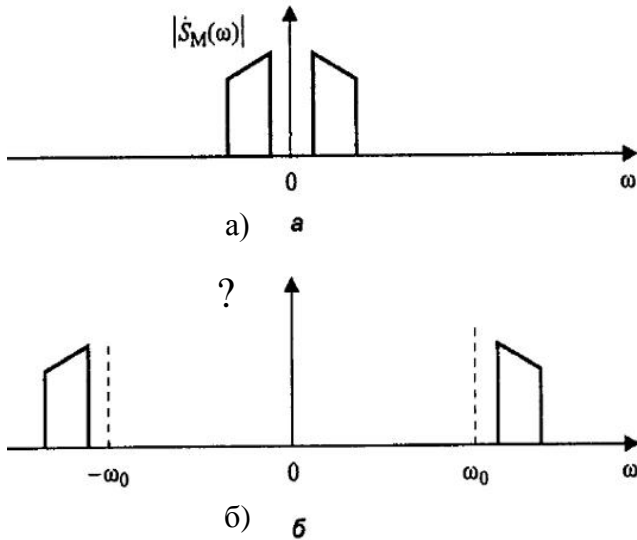
а. Амплітудно-модульованим (АМ).

б. Частотно-модульованим (ЧМ).

в. Квадратурно-модульованим (КАМ).

г. Фазово-модульованим (ФМ).

32. На рисунку а представлено спектр модулюючого сигналу. Спектр якого сигналу представлено на рисунку б?



- а. Частотно-модульованого;
- б. Фазово-модульованого;
- в. Однополосного АМ сигналу з верхньою бічною смугою;
- г. Однополосного АМ сигналу з нижньою бічною смугою.

33. Вкажіть сигнал з фазовою модуляцією:

- а. $s(t) = A \cos(\omega_0 t + k s_M(t))$;
- б. $s(t) = A \cos[\omega_0 t + k \int s_M(t') dt' + \varphi_0]$;
- в. $s(t) = A(t) \cos(\omega_0 t + \varphi(t))$.

34. Спосіб модуляції, коли параметри коливання-носія змінюються стрибкоподібно, називається

- а. маніпуляція;
- б. квадратурна модуляція;
- в. кутова модуляція.

35. Спектральна густина потужності випадкового процесу на виході лінійної системи $W_{\text{вих}}(\omega)$ визначається формулою

а. $W_{\text{вих}}(\omega) | \dot{K}(\omega) |^2$;

б. $W_{\text{вих}}(\omega) | \dot{K}(\omega) |$;

в. $W_{\text{вих}}(\omega) \dot{K}(\omega)$.

36. Сигнал – це _____, властивості якого визначаються взаємодією між матеріальним об'єктом та засобом його дослідження.

а. умовний знак;

б. фізичний процес;

в. функціонал.

37. У відгуку нелінійної системи присутні спектральні складові з частотами, на яких

а. відсутні спектральні складові у вхідній дії;

б. присутні спектральні складові у вхідній дії;

в. працює модулятор.

38. Нелінійні системи мають дві такі основні властивості (вказати які):

а. не виконується принцип суперпозиції;

б. виконується принцип суперпозиції;

в. відбувається перетворення спектрів сигналів;

г. спектри сигналів залишаються не змінними.

39. Періодичними називають сигнали,

а. миттєві значення яких мають частково відомий характер зміни за часом;

б. миттєві значення яких повторюються через постійний інтервал часу;

в. які мають дискретний амплітудний спектр з некрatними частотами.

40. Модуль спектральної густини визначають за формулою:

а. $\int_{-\infty}^{\infty} s(t) e^{-j\omega t} dt$;

б. $A(\omega) - jB(\omega)$;

в. $\sqrt{A^2(\omega) + B^2(\omega)}$;

г. $\text{Re}[S(\omega)]$.

Навчальне видання

Братченко Геннадій Дмитрович
Перелигін Борис Вікторович
Банзак Оксана Вікторівна
Казакова Надія Феліксівна
Григор'єв Денис Вікторович

МЕТОДИ ТА ЗАСОБИ ОБРОБКИ СИГНАЛІВ

Навчальний посібник

За загальною редакцією доктора технічних наук
Г. Д. Братченка

Макет обкладинки: А.І. Розмариця

Підписано до друку 10.12.2014 р.
Формат 60x90/16. Папір офсетний.
Гарнітура Times New Roman. Друк офсетний.
Ум. друк. арк. 26,27. Наклад 100 прим.
Зам. № 1012/2

Надруковано з готового оригінал-макету у друкарні
«Плутон»
м. Одеса, вул. Ніжинська, 56
тел.: +38 048 700 42 42