

МІНІСТЕРСТВО ТРАНСПОРТУ ТА ЗВ'ЯЗКУ УКРАЇНИ
УКРАЇНСЬКА ДЕРЖАВНА АКАДЕМІЯ ЗАЛІЗНИЧНОГО
ТРАНСПОРТУ

О.П. Батаєв, І.В. Ковтун, Н.А. Корольова

ТЕОРІЯ ЕЛЕКТРИЧНОГО ЗВ'ЯЗКУ

Навчальний посібник

*Рекомендовано Міністерством освіти і науки України
як навчальний посібник для студентів
вищих навчальних закладів*

Харків 2010

ББК 32.811
УДК 621.396
Б-25

Батаєв О.П., Ковтун І.В., Корольова Н.А. Теорія електричного зв'язку: Навч. посібник. – Харків: УкрДАЗТ, 2010. - 630 с.

У навчальному посібнику викладено детальні відомості про повідомлення, сигнали, завади, спектральний і кореляційний аналіз детермінованих сигналів та випадкових процесів, різні перетворення сигналів.

Наведені основні характеристики каналів зв'язку з урахуванням підвищення їхньої ефективності. Викладені основи оптимального прийому дискретних і безперервних сигналів, основи теорії інформації та завадостійкого кодування. Наведено принципи розподілу сигналів у багатоканальних системах передачі. Особливу увагу приділено цифровим методам передачі інформації.

Посібник може використовуватися на практичних заняттях та при виконанні розрахункових робіт.

Для студентів і слухачів Української державної академії залізничного транспорту.

Іл. 162, табл. 10, бібліогр.: 51 назв.

Рекомендовано Міністерством освіти і науки України як навчальний посібник для студентів вищих навчальних закладів
(№ 14/18 – Г – 21 від 9.01.09)

Рецензенти:
професори В.А. Краснобаєв (ХНУСГ),
В.В. Поповський (ХНУРЕ)

ISBN 978-966-2033-26-7

© О.П. Батаєв, І.В. Ковтун,
Н.А. Корольова
© Українська державна академія
залізничного транспорту, 2010

О.П. Батаєв, І.В. Ковтун, Н.А. Корольова

ТЕОРІЯ ЕЛЕКТРИЧНОГО ЗВ'ЯЗКУ

Навчальний посібник

Відповідальний за випуск О.П. Батаєв
Редактор Ібрагімова Н.В.

КВ № 8617 від 06.04.2004р. Підписано до друку 7.12.2009 р.
Формат паперу 60x84 1/16. Папір писальний.
Умовн.-друк.арк. 19,0. Обл.-вид.арк. 19,25.
Замовлення № . Тираж 300. Ціна договірна

Видавництво УкрДАЗТу, свідоцтво ДК № 2874 від 12.06.2007 р.
Друк. ТОВ «Енергозберігаючі технології»
61050, Харків, Харківська набережна,8.
Свідоцтво про реєстрацію ДК № 1360 від 19.05.03р.

ЗМІСТ

Передмова	10
Список скорочень	14
Основні позначення.....	16
Глава 1. Загальні відомості про системи та мережі електрозв'язку	22
1.1. Інформація, повідомлення, сигнали, як форма уявлення матерії.....	22
1.2. Системи, канали та мережі зв'язку	27
1.3. Завади та спотворення у каналах зв'язку	39
1.4. Кодування та модуляція	42
1.5. Демодуляція та декодування.....	54
1.6. Загальні характеристики систем зв'язку	57
Висновки до глави 1	61
Контрольні питання та задачі.....	62
Глава 2. Математичні моделі повідомлень, сигналів і завод	64
2.1. Функціональні простори та їх базиси	64
2.2. Спектральний аналіз сигналів на основі рядів Фур'є.....	72
2.2.1. Спектральне подання періодичних сигналів	76
2.2.2. Спектральне подання неперіодичних сигналів.....	79
2.3. Ортогональні функції Радемахера і Уолша	87
2.4. Дискретизація в часі безперервних сигналів і їх відновлення	89
2.4.1. Дискретизація безперервних сигналів	89
2.4.2. Спектральне подання дискретизованих сигналів ...	95
2.4.3. Особливості дискретизації сигналів.....	97
2.4.4. Відновлення безперервного сигналу	104
2.5. Випадкові процеси та їх загальні характеристики	109
2.5.1. Функції розподілу випадкових процесів.....	110
2.5.2. Моментні (числові) характеристики випадкових процесів.....	113
2.5.3. Приклади деяких випадкових процесів.....	118

2.5.4.	Кореляційні функції детермінованих і випадкових процесів	134
2.6.	Аналітичний та вузькосмуговий сигнал	149
2.6.1.	Математичні моделі і характеристики аналітичного сигналу	149
2.6.2.	Імовірнісні характеристики обвідної і фази вузькосмугового випадкового гаусівського процесу	159
2.7.	Марковські процеси	165
	Висновки до глави 2	172
	Контрольні питання та задачі	175
Глава 3. Модульовані сигнали в системах передачі інформації		
	3.1. Загальні положення	177
	3.2. Основні види аналогової амплітудної модуляції ...	179
	3.3. Кутова модуляція	188
	3.4. Дискретна модуляція гармонічного переносника ..	199
	3.5. Імпульсна модуляція	204
	3.6. Модуляція шумової несучої	210
	3.7. Енергетичний спектр модульованих сигналів	217
	3.8. Цифрові методи модуляції	220
	3.9. Класифікація видів модуляції	233
	Висновки до глави 3	237
	Контрольні питання та задачі	238
Глава 4. Математичні моделі каналів зв'язку		
	4.1. Загальні відомості про канали зв'язку	240
	4.2. Перетворення випадкових сигналів у лінійних каналах з постійними параметрами	243
	4.3. Перетворення випадкових сигналів у нелінійних каналах з постійними параметрами	245
	4.4. Проходження сигналів через канали зв'язку з випадковими параметрами	246
	4.5. Адитивні завади в каналах зв'язку	252
	4.6. Квантовий шум в оптичному каналі зв'язку	256
	4.7. Моделі безперервних каналів зв'язку	259
	4.7.1. Ідеальний канал без завад	260

4.7.2. Канал з адитивним гаусівським шумом	260
4.7.3. Канал з невизначеною фазою сигналу й адитивним шумом	261
4.7.4. Канал з міжсимвольною інтерференцією й адитивним шумом	263
4.8. Моделі дискретних каналів зв'язку.....	264
4.8.1. Деякі моделі дискретних каналів з пам'яттю.....	268
4.8.2. Модель дискретно-безперервного каналу	272
Висновки до глави 4	272
Контрольні питання та задачі	274
Глава 5. Інформаційні характеристики джерел повідомлень та каналів зв'язку	276
5.1. Загальні відомості про джерела повідомлень і канали зв'язку	276
5.2. Основні поняття теорії інформації	280
5.2.1. Часткова кількість інформації.....	280
5.2.2. Ентропія джерела повідомлень	281
5.3. Інформаційні характеристики дискретних каналів зв'язку	285
5.3.1. Кількість інформації, що передається по каналу зв'язку (взаємна інформація)	285
5.3.2. Пропускна спроможність дискретного каналу зв'язку	291
5.4. Теорема кодування Шеннона для дискретного каналу зв'язку.....	298
5.4.1. Теорема про властивість асимптотичної рівноімовірності (ВАР)	299
5.4.2. Теорема 1. Про кодування джерела повідомлень	301
5.4.3. Теорема 2. Про кодування в каналі без завад.....	302
5.4.4. Теорема кодування в дискретному каналі з завадами	304
5.4.5. Інтерпретація теореми кодування.....	309
5.4.6. Енергетичний вииграш кодування.....	315
5.5. Потенційні можливості безперервних каналів зв'язку при передачі дискретних повідомлень.....	318

5.5.1. Кодування й декодування в безперервному каналі зв'язку	318
5.5.2. Пропускна спроможність безперервного каналу зв'язку	319
5.6. Теорема про кодування в безперервному каналі з завадами	330
5.7. Потенційні можливості каналів з багатьма користувачами	334
Висновки до глави 5	336
Контрольні питання та задачі	337
Глава 6. Основи теорії завадостійкості систем передачі дискретних повідомлень	339
6.1. Завдання синтезу оптимальних демодуляторів	339
6.2. Критерії якості й правила прийому дискретних повідомлень	344
6.3. Оптимальні алгоритми прийому при повністю відомих сигналах (когерентний прийом)	352
6.4. Прийом сигналів на узгоджений фільтр	361
6.5. Завадостійкість оптимального когерентного прийому	370
6.6. Прийом сигналів з невизначеною фазою (некогерентний прийом)	380
6.7. Прийом дискретних повідомлень в умовах флуктуації фаз й амплітуд сигналів	395
6.8. Прийом дискретних повідомлень у каналах із зосередженими за спектром й імпульсними завадами	401
6.9. Завадостійкість прийому дискретних повідомлень в оптичному діапазоні хвиль	406
6.10. Порівняння завадостійкості систем передачі дискретних повідомлень	410
Висновки до глави 6	412
Контрольні питання та задачі	414
Глава 7. Завадостійке кодування	417
7.1. Основні поняття і визначення	417

7.2.	Принцип виявлення і виправлення помилок у завадостійких кодах.....	421
7.3.	Класифікація й основні характеристики завадостійких кодів	426
7.3.1.	Класифікація завадостійких кодів	426
7.3.2.	Основні характеристики завадостійких кодів	429
7.4.	Лінійні, блокові коди з виявленням і виправленням помилок	431
7.4.1.	Особливості кодування лінійних блокових кодів	432
7.4.2.	Кодування в коді Хеммінга	435
7.4.3.	Особливості декодування лінійних блокових кодів	436
7.4.4.	Декодування в коді Хеммінга	438
7.5.	Циклічні коди.....	439
7.5.1.	Способи завдання і кодування	439
7.5.2.	Декодування циклічних кодів	445
7.5.3.	Мажоритарне декодування циклічних кодів	449
7.6.	Згорткові коди.....	451
7.6.1.	Визначення, способи завдання і кодування.....	451
7.6.2.	Декодування згорткових кодів.....	454
7.7.	Застосування завадостійких кодів у системах залізничної автоматики, телемеханіки і зв'язку	455
7.8.	Способи підвищення правильності передачі інформації в каналах зв'язку з завадами.....	460
	Висновки до глави 7	462
	Контрольні питання та задачі.....	464

Глава 8. Завадостійкість передачі безперервних повідомлень.....	465
8.1. Критерії завадостійкості прийому безперервних повідомлень	465
8.2. Оптимальна оцінка окремих параметрів сигналу	468
8.3. Оптимальна демодуляція безперервних сигналів.....	474
8.4. Завадостійкість систем передачі безперервних повідомлень при слабких завадах.....	486

8.4.1. Системи з піднесучими частотами	490
8.4.2. Системи з імпульсною модуляцією	491
8.5. Поріг завадостійкості широкосмугового демодулятора.....	494
8.6. Оптимальна лінійна фільтрація безперервних сигналів	500
8.7. Оптимальна лінійна фільтрація безперервних повідомлень	505
8.8. Загальні відомості про нелінійну фільтрацію	510
8.9. Загальні відомості про цифрову передачу безперервних повідомлень	518
8.10. Завадостійкість імпульсно-кодової модуляції.....	523
8.10.1. Потужність шуму квантування.....	524
8.10.2. Шум помилкових імпульсів при декодуванні.....	528
8.11. Кодування з передбаченням	531
Висновки до глави 8	539
Контрольні питання та задачі	540
Глава 9. Принципи багатоканального зв'язку	542
9.1. Багатоканальна система передачі повідомлень.....	542
9.2. Основні положення лінійної теорії розподілу сигналів	545
9.3. Умова лінійного розподілу сигналів	547
9.4. Основні способи розподілу сигналів.....	549
9.4.1. Частотний розподіл каналних сигналів.....	549
9.4.2. Часовий розподіл каналів	553
9.4.3. Розподіл каналних сигналів за фазою	557
9.5. Розподіл сигналів за формою	559
9.6. Системи передачі із шумоподібними сигналами.....	562
9.7. Комбінаційний розподіл сигналів.....	572
9.7.1. Принцип комбінаційного ущільнення.....	572
9.7.2. Принцип комбінаційного розподілу сигналів	573
9.8. Кодовий розподіл сигналів	577
9.9. Використання багатопозиційних сигналів.....	580
9.10. Коригувальні коди	588
9.11. Використання сигнально-кодових конструкцій.....	591
9.12. Оптимізація систем зв'язку	594
9.12.1. Загальна методологія оптимізації системи.....	594

9.12.2. Оптимізація системи передачі інформації	597
Висновки до глави 9	602
Контрольні питання та задачі	604
Загальні висновки	606
Бібліографічний список	609
Додаток 1. Прості розривні функції	613
Додаток 2. Спектри та спектральна щільність сигналів різної форми.....	614
Додаток 3. Значення інтегралів імовірності	623
Додаток 4. Функції розподілу безперервних випадкових величин	624
Додаток 5. Функції розподілу дискретних випадкових величин	627
Додаток 6. Значення логарифмів	629
Додаток 7. Поліноми, що не приводяться.....	630

Передмова

Дисципліна «Теорія електричного зв'язку» є базовою дисципліною в системі підготовки бакалаврів, фахівців і магістрів за напрямом 7.092507 «Автоматика і автоматизація на транспорті», зокрема за спеціалізацією 7.092507.03 «Автоматичні системи технологічного зв'язку на залізничному транспорті», а також за напрямом 6.092401 «Телекомунікаційні системи і мережі».

Даний курс базується на знаннях, отриманих студентами при вивченні фізики, вищої математики, інформатики, електроніки і мікросхемотехніки, теорії електричних ланцюгів і сигналів.

Сучасна теорія зв'язку використовує як детерміновані моделі сигналів, так і імовірнісні (стохастичні) моделі для повідомлень, що передаються, а також для відповідних сигналів і завад (шумів) у каналі зв'язку. Імовірнісний підхід враховує випадковий (невідомий для одержувача) характер передачі повідомлень і завад у каналі зв'язку і дозволяє синтезувати оптимальні приймальні пристрої, що забезпечують максимально можливу якість прийому і граничні показники систем передачі повідомлень (систем зв'язку).

Основи сучасної (статистичної) теорії зв'язку беруть свій початок у фундаментальних роботах В.О. Котельникова з теорії потенційної завадостійкості (1947 р.) і К. Шеннона з теорії інформації і кодування (1948 р.). Сюди слід віднести і ранішні роботи з теорії зв'язку учених Х. Найквіста (1928 р.) і В.О. Котельникова (1933 р.), в яких була незалежно сформульована і доведена теорема відліків, Р. Хартлі (1928 р.), в якій введена логарифмічна міра кількості інформації, Д.В. Агєєва (1935 р.), присвячена теорії лінійного розподілу сигналів і каналів.

У подальшому розвитку статистичної теорії зв'язку велику роль відіграли роботи вчених О.Я. Хінчина (1938 р.) з кореляційної теорії стаціонарних випадкових процесів, А.М. Колмогорова (1941 р.) і Н. Вінера (1943 р.) з інтерполяції і екстраполяції стаціонарних випадкових процесів, А. Вальда (1950 р.) з теорії статистичних рішень.

Істотний внесок у теорію зв'язку зробили також відомі учені Р. Райс, О.О. Харкевич, В.І. Сифоров, Р. Галлагер, А.І. Берг, Б.Р. Левін, В.І. Тихонов, Х. Хелстром, Р. Фано, Л.М. Фінк, Д. Віттербі, Д.Д. Кловський, П.Ф. Поляков, В.В. Поповський і багато інших вітчизняних і зарубіжних діячів науки.

Курс «Теорія електричного зв'язку» належить до фундаментальних дисциплін, необхідних для підготовки висококваліфікованих фахівців, що володіють сучасними методами аналізу і синтезу телекомунікаційних систем і пристроїв різного призначення.

Метою даного курсу є вивчення основних закономірностей і методів передачі різних повідомлень по каналах зв'язку, для чого в курсі ставляться і вирішуються завдання аналізу і синтезу основних видів систем зв'язку і окремих їх функціональних елементів.

У курсі розглядаються основні способи математичного подання повідомлень, сигналів і завад, методи функціонування і перетворення сигналів у різних електричних ланцюгах, питання аналізу завадостійкості і пропускнуї спроможності систем електрозв'язку, методи ефективного та завадостійкого кодування, способи оптимального прийому повідомлень у каналах різної структури, принципи багатоканальної передачі сигналів по одній лінії зв'язку, способи розподілу інформації в мережах зв'язку, основи цифрової обробки сигналів, а також питання оптимізації систем зв'язку.

Навчальний посібник містить 9 глав. У кінці кожної глави наводяться висновки, що містять підсумковий аналіз основних положень і результатів, а також контрольні питання і задачі, направлені на закріплення даного матеріалу і глибше, усвідомлене його осмислення та запам'ятовування.

За змістом і методичним викладом навчальний посібник розрахований на студентів денного і заочного навчання. У його основу покладені лекції, що читаються авторами впродовж ряду років.

У першій главі висловлюються загальні відомості про системи і мережі електрозв'язку. Тут також розглядаються такі поняття, як інформація, повідомлення, сигнал, завада, лінія зв'язку, канал зв'язку, наводяться їх основні характеристики.

Даються загальні відомості про модуляції та демодуляції, кодування та декодування, а також про особливості багатоканального зв'язку.

У другій главі розглядаються математичні моделі повідомлень, сигналів і завад. Основна увага при цьому приділяється спектральному аналізу сигналів з використанням рядів Фур'є, функцій Уолша, дискретизації безперервних сигналів і їх відновленню, аналізу випадкових процесів і їх характеристик, аналітичному і вузькосмуговому сигналам, а також марковським процесам.

У третій главі дається опис модульованих сигналів, які використовуються в системах передачі повідомлень. З обліком класифікації різних видів модуляції розглядаються характеристики сигналів (часові, спектральні, кореляційні) безперервної, дискретної та імпульсної модуляції, наводяться схемні реалізації модуляторів і демодуляторів.

У четвертій главі розглядаються математичні моделі каналів зв'язку, перетворення випадкових сигналів у детермінованих лінійних і нелінійних каналах, проходження сигналів через канали зв'язку з випадковими параметрами.

У п'ятій главі проводиться аналіз інформаційних характеристик джерел повідомлень і каналів зв'язку. Тут розглядаються інформаційні характеристики джерел дискретних і безперервних повідомлень, дискретних і безперервних каналів зв'язку, питання узгодження джерел повідомлень з каналами зв'язку, потенційні можливості передачі повідомлень по каналах зв'язку.

У шостій главі розглядаються питання завадостійкості передачі дискретних повідомлень. На основі аналізу критеріїв правильності передачі повідомлень описуються оптимальні алгоритми прийому дискретних повідомлень при повністю відомих сигналах, прийом сигналів на узгоджений фільтр, кореляційний прийом сигналів, прийом сигналів з невизначеною фазою, наводиться оцінка завадостійкості систем передачі дискретних повідомлень.

У сьомій главі наводяться основи теорії завадостійкого кодування. Тут дається класифікація кодів і їх основні характеристики, детально розглядаються коди Хеммінга,

циклічні, згорткові та інші. Аналізуються системи із зворотним зв'язком як способи підвищення правильності передачі повідомлень у каналах з помилками.

У восьмій главі розглядаються питання завадостійкості безперервних повідомлень. З використанням критеріїв завадостійкості безперервних повідомлень описується оптимальна лінійна фільтрація на основі фільтрів Колмогорова – Вінера, Калмана, даються загальні відомості про нелінійну фільтрацію, наводяться цифрові способи передачі безперервних повідомлень, кодування з передбаченням.

У дев'ятій главі проводиться аналіз принципів багатоканального зв'язку. На основі загальної теорії розподілу сигналів розглядаються такі способи розподілу сигналів і каналів: частотне, часове, фазове, комбіноване, кодове. Наводяться загальні відомості про системи передачі з шумоподібними сигналами, розглядаються перспективи розвитку систем і мереж зв'язку.

Глави 1, 6 написані Н.А. Корольовою, глави 2, 4, 7, 8, 9 – О.П. Батаєвим, глави 3,5 – І.В. Ковтун, передмова і загальні висновки - О.П. Батаєвим. Редагування всіх глав навчального посібника виконане О.П. Батаєвим.

Автори виражають особливу вдячність д.т.н., проф. П.Ф. Полякову за допомогу при підготовці матеріалів для розділів 4, 8, а також за ретельний та критичний розгляд всього матеріалу навчального посібника. Автори вельми вдячні співробітникам кафедри «Транспортний зв'язок» УкрДАЗТ І.М. Мещеряковій, О.Ф. Співак, З.В. Осмоловській, які узяли на себе трудомістку роботу з комп'ютерного набору навчального посібника. Автори також вдячні д.т.н., проф., завідувачеві кафедри «Телекомунікаційні системи» ХНУРЕ В.В. Поповському і д.т.н., проф. кафедри «Автоматизація і комп'ютерні технології» ХНТУСГ В.А. Краснобаєву за ряд корисних зауважень, що сприяли поліпшенню даного навчального посібника.

Відгуки за навчальним посібником просимо присилати у видавництво за адресою м. Харків, пл. Фейєрбаха, 7.

Список скорочень

ААСЗ	- асинхронні адресні системи зв'язку
АВ	- алгоритм Вітербі
АКН	- алгоритм Кловського – Ніколаєва
АКФ	- автокореляційні функції
АМ	- амплітудна модуляція
АМн	- амплітудна маніпуляція
АМП	- алгоритм максимальної правдоподібності
АРК	- апаратура розподілу каналів
АЧХ	- амплітудно–частотна характеристика
БГШ	- білий гаусівський шум
БК	- безперервний канал
БСП	- багатоканальні системи передачі
ВАХ	- вольт-амперна характеристика
ВВ	- випадкова величина
ВКФ	- взаємно кореляційні функції
ВСЗ	- відношення сигнал – завада (відношення середніх потужностей сигналу і завади)
ВФМн	- відносна фазова маніпуляція
ДБК	- дискретно–безперервний канал
ДК	- дискретний канал
ДПФ	- дискретне перетворення Фур'є
ДЧМ	- двократна частотна модуляція
ЕВК	- енергетичний виграш коду
ЕВС	- енергетичний виграш системи
ЕОМ	- електронно-обчислювальна машина
ЗДПФ	- зворотне дискретне перетворення Фур'є
ІФР	- інтегральна функція розподілення
ІКМ	- імпульсно-кодова модуляція
ІХ	- імпульсна характеристика
КЗ	- канал зв'язку
КМ	- кутова модуляція
ЛЗ	- лінія зв'язку
ЛРП	- лінійні рекурентні послідовності
ММЗ	- модуляція мінімальним (частотним) зсувом
МО	- математичне очікування
МПО	- максимально правдоподібна оцінка

МСІ	- міжсимвольна інтерференція
ОМ	- односмугова модуляція
ПГ	- перетворення (перетворювач) Гілберта
Пер	- перетворювач
ПрД	- передавач
ПрМ	- приймач
РКЧ	- часовий розподіл каналів
РРЛ	- радіорелейна лінія
РС-коди	- коди Рида - Соломона
СКК	- сигнально-кодові конструкції
СКВ	- середньоквадратичне відхилення
СКП	- середньоквадратична похибка
СПДП	- система передачі дискретних повідомлень
СПІ	- системи передачі інформації
СФ	- смуговий фільтр
СЩП	- спектральна щільність потужності (середньої)
СЩЕ	- спектральна щільність енергії
ТВ	- телевізійний
ТФ	- телефонний
УФ	- узгоджений фільтр
ФК	- функція кореляції, фільтр Калмана
ФКВ	- фільтр Колмогорова – Вінера
ФМ	- фазова модуляція
ФМн	- фазова маніпуляція
ФНЧ	- фільтр нижніх частот
ФЧХ	- фазочастотна характеристика
ЧМ	- частотна модуляція
ЦК	- циклічні коди
ЦОС	- цифрова обробка сигналів
ЦФ	- цифровий фільтр
ЧМБФ	- частотна модуляція з безперервною фазою
ЧМн	- частотна маніпуляція
ЧРК	- частотний розподіл каналів
ШПФ	- швидке перетворення Фур'є
ШШС	- широкосмугові шумоподібні сигнали
ЩІ	- щільність імовірності

Основні позначення

$\{A\}$	- ансамбль (множина) повідомлень
\vec{a}	- реалізація вектора (ланцюжка символів) повідомлення
a	- реалізація елемента повідомлення
$\{B\}$	- ансамбль (множина) первинних сигналів
$B(t)$	- випадковий первинний сигнал
$B = 2\Delta FT$	- база сигналу
$B(t_1, t_2); B(\tau)$	- функція кореляції процесу (сигналу)
\vec{b}	- реалізація вектора (ланцюжка) кодівих символів
$b(t)$	- реалізація первинного сигналу
$b_l^{(i)}$	- l -тий елемент первинного цифрового сигналу (кодівий символ) i -ої позиції
C	- максимальна кількість інформації у каналі зв'язку (біт/символ або біт/відлік)
C'	- пропускна спроможність каналу зв'язку (біт/с)
$D(\bullet)$	- дисперсія випадкової величини або процесу
d	- відстань між сигнальними точками, відстань між двійковими послідовностями
d_{min}	- мінімальна відстань між комбінаціями лінійного блокового коду
E	- енергія сигналу
$F(\bullet)$	- інтегральна функція розподілу імовірності
ΔF	- смуга частот сигналу (каналу)
$f_\Delta = \frac{1}{\Delta t_\Delta}$	- частота дискретизації безперервного сигналу
f	- циклічна частота
$G(f)$	- спектральна щільність потужності

$G_0(f)$	- одностороння спектральна щільність потужності (на позитивних частотах)
$G(t), G(t, \tau)$	- випадкова імпульсна характеристика лінійного ланцюга
$g(t), g(t, \tau)$	- детермінована імпульсна характеристика лінійного ланцюга
g, g'	- виграш, узагальнений виграш системи модуляції
$H(X), H(X Y)$	- ентропія, умовна ентропія дискретної випадкової величини (дискретного джерела повідомлень)
$H'(X)$	- продуктивність дискретного джерела повідомлень
$H_\varepsilon(A), H'_\varepsilon(A)$	- епсилон ентропія, епсилон продуктивність дискретного джерела повідомлень
$h(x), h(x y)$	- диференціальна ентропія, умовна диференціальна ентропія безперервної випадкової величини
$h^2 = \frac{E}{N_0}$	- відношення енергії елемента сигналу на вході демодулятора до односторонньої спектральної щільності потужності білого шуму
$h^2 = \frac{P}{N_0 R_i} = \frac{h^2}{R \log_2 m}$	- нормована енергія $\left(h^2 = \frac{E}{N_0} \right)$ сигналу на 1 біт інформації (бітова енергія)
$I(X, Y), I(Y, X)$	- взаємна інформація між X і Y (інформація, що міститься в Y відносно X або в X відносно Y)
$I'(X, Y)$	- швидкість передачі інформації від X до Y
j	- знак уявної одиниці, $j = \sqrt{-1}$
$\{K\}$	- об'єм алфавіту дискретного джерела
$\dot{K}(f), K(j\omega)$	- передавальна функція або комплексний коефіцієнт передачі

$K(f) = \dot{K}(f) $	- амплітудно-частотна характеристика (АЧХ)
k	- число інформаційних символів у кодовій комбінації
$M(X)$	- математичне очікування випадкової величини (процесу)
$m_{км}$	- індекс кутової модуляції
m, m_{AM}	- основа коду, коефіцієнт глибини амплітудної модуляції
m_1	- математичне очікування випадкової величини (або процесу)
$N = m^n$	- об'єм алфавіту коду
$N(t)$	- випадкова адитивна завада в безперервному каналі, випадковий процес
N_0	- одностороння (на позитивних частотах) спектральна щільність потужності квазібілого і білого шуму
$n(t)$	- реалізація випадкового процесу
n	- довжина (загальне число символів) кодової комбінації
P_{cp}	- середня потужність сигналу
$P\{\bullet\}, P_x$	- імовірність події, вказаної в дужках або позначеної індексом
p_b, p_e	- імовірність помилки на один інформаційний біт, еквівалентна імовірність помилки
$p_{нп}$	- імовірність невиявлення помилки
$p_0 = p_c$	- імовірність помилкового прийому символу
$p_{од}$	- імовірність помилкового декодування блока символів
p_d	- допустима (задана) імовірність помилки

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt$$

- додаткова функція помилок

$R = \frac{k}{n}$	- швидкість коду
$R_i = \frac{R \log_2 m}{T}$	- максимальна продуктивність при $R=1$ (інформаційна швидкість) дискретного джерела (біт/с), швидкість передачі інформації
$r = n - k$	- число перевірочних символів у кодових комбінаціях блокового коду
$\frac{R(t_1, t_2)}{R_{max}}, \frac{R(\tau)}{R(0)}$	- нормована функція кореляції, коефіцієнт кореляції
$S(t)$	- випадковий сигнал на вході приймача (детектора) без урахування адитивних завад
$\dot{S}_x(f), S_x(j\omega)$	- спектральна щільність перетворення Фур'є сигналу $x(t)$
$S_x(f) = \left \dot{S}_x(f) \right $	- амплітудний спектр сигналу $x(t)$
$s(t)$	- реалізація випадкового сигналу на вході приймача без адитивних завад
T	- тривалість тактового інтервалу, тривалість фінітного сигналу, знак транспонування матриці
T_a	- інтервал аналізу коливання, що приймається (сигнал плюс завада)
t	- поточний час
τ_i	- тривалість імпульсу
$U(t)$	- випадковий сигнал на виході модулятора
$u(t)$	- реалізація випадкового сигналу на виході модулятора
$V(t), W(t), X(t), Y(t)$	- випадкові процеси з реалізаціями (вибірковими функціями), відповідно $v(t), w(t), x(t), y(t)$
ν	- швидкість передачі (число символів у секунду) дискретного джерела (каналу), число відліків в одну секунду безперервного сигналу

$W(f)$	- спектральна щільність енергії випадкового процесу
$W_0(f)$	- одностороння спектральна щільність енергії
$w_1(x, t)$	- одновимірна щільність розподілу імовірності випадкової величини (випадкового процесу)
$w_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$	- n-мірна щільність імовірності сукупності випадкових величин (випадкового процесу)
$\{X\}, \{Y\}$	- алфавіт повідомлень на вході і виході дискретного каналу
\vec{x}, \vec{y}	- вектор (ланцюжок символів) повідомлень на вході і виході дискретного каналу
$Z(t)$	- сума сигналу і адитивної завади на вході приймача (детектора)
$z(t)$	- реалізація суми сигналу і адитивної завади на вході приймача (детектора)
$\beta = \frac{1}{h^2_e}$	- енергетична ефективність системи
$\delta(t)$	- дельта-функція
$E(t)$	- помилка в оцінюванні випадкового параметра або випадковий процес, шум спостереження або квантування
$\varepsilon(t), \varepsilon$	- реалізація помилки в оцінюванні випадкового параметра або випадковий процес
$\gamma = \frac{R_i}{\Delta F}$	- частотна ефективність системи, коефіцієнт передачі каналу
Δt_δ	- крок дискретизації безперервного сигналу в часі
$\eta = \frac{R_i}{C'}$	- коефіцієнт інформаційної ефективності системи
θ	- фазове зрушення, поточний безрозмірний час
$\theta_{\text{кан}}$	- фазове зрушення у каналі

A	- відношення правдоподібності
λ	- інформаційний параметр
ν	- кодове обмеження згорткового коду
Π	- пік – чинник повідомлення або сигналу (відношення максимального сигналу до середньоквадратичного)
ρ	- відношення середніх потужностей сигналу і шуму
$\sigma^2 = D$	- дисперсія випадкової величини (процесу)
$\sigma = \sqrt{D}$	- середньоквадратичне відхилення
τ	- інтервал між двома перетинами процесу ($t_2 - t_1$), затримка
$\varphi(t)$	- фаза сигналу при кутовій модуляції
$\psi(t) = \omega_0 t + \varphi(t)$	- повна фаза сигналу
$\varphi(f) = \arg \dot{K}(f)$	- фазочастотна характеристика (ФЧХ)
ψ, φ	- фазове зрушення, фаза сигналу
$\psi_i(t), \varphi_i(t)$	- функції ортогонального (або ортонормованого) базису
ω	- кутова (кругова) частота, вага кодової комбінації блокового коду
$\omega(t) = \frac{d\psi(t)}{dt}$	- миттєва частота сигналу

Точка зверху означає комплексний вираз

Знак * зверху виразу означає комплексне сполучення

Пряма позначка (риска) над символом або формулою означає статистичне усереднювання (за ансамблем), хвиляста за часом

Знак ^ над символом означає оцінку, що видається демодулятором, детектором або фільтром

Знак \cap означає перетворення Гілберта

Знак \otimes означає згортку двох функцій

Знак \oplus означає складання по модулю 2

(n, k) - позначення лінійного блокового коду довжини n і k інформаційними символами

$L_2(T)$ - простір Гілберта

ГЛАВА 1

ЗАГАЛЬНІ ВІДОМОСТІ ПРО СИСТЕМИ ТА МЕРЕЖІ ЕЛЕКТРОЗВ'ЯЗКУ

1.1. Інформація, повідомлення, сигнали як форма уявлення матерії

Поняття *інформації* й *повідомлення* вживаються досить часто. Ці близькі за змістом поняття складні й дати їхнє точне визначення неможливо. Поняття *інформація* походить від латинського *informatio* — роз'яснення, ознайомлення, поінформованість. Під інформацією розуміють сукупність відомостей, даних про які-небудь події, явища або предмети. Ми живемо в інформаційному світі. Усе, що ми бачимо, чуємо, пам'ятаємо, знаємо, переживаємо, - все це різні форми інформації. Сукупність відомостей, даних стає знанням лише після їх інтерпретації з урахуванням цінності й змісту цих відомостей. Отже, інформацію в широкому розумінні можна визначити як сукупність знань про навколишнє середовище. У такому розумінні інформація є найважливішим ресурсом науково-технічного й соціально-економічного розвитку суспільства. На відміну від матеріального й енергетичного ресурсів, інформаційний ресурс не зменшується при споживанні, накопичується згодом, порівняно легко й просто за допомогою технічних засобів обробляється, зберігається й передається на значні відстані.

Для *передачі* або *зберігання* інформації використовують різні знаки (символи), що дозволяють подати її в деякій формі. Цими знаками можуть бути слова й фрази в людській мові, жести й малюнки, форми коливань, математичні знаки й т. п. Сукупність знаків, що відображають ту або іншу інформацію, називають *повідомленням*. Так, при телеграфній передачі повідомленням є текст телеграми, що являє собою послідовність окремих знаків - букв і цифр. При розмові по телефону повідомленням є безперервна зміна в часі звукового тиску, що відображає не тільки зміст, але й інтонацію, тембр, ритм й інші властивості мови. При передачі зображень, що рухаються, у телевізійних системах повідомлення являє собою зміну в часі яскравості та кольору елементів зображення.

Передача повідомлень (а отже, і інформації) на відстань здійснюється за допомогою якого-небудь матеріального носія (CD та DVD дисків, паперу, магнітної стрічки й т. д.) або фізичного процесу (звукових або електромагнітних хвиль, струму й т. д.). Фізичний процес, що відображає (несе) передане повідомлення, називається *сигналом*.

Сигнал можна використовувати для опису будь-якого фізичного процесу, що змінюється відповідно до повідомлення. У сучасних системах управління й зв'язку найчастіше використовують електричні сигнали. Фізичною величиною, що визначає такий сигнал, є струм або напруга. Сигнали формуються шляхом зміни тих або інших параметрів фізичного носія відповідно до переданого повідомлення. Цей процес (зміни параметрів носія) прийнято називати *модуляцією*.

Повідомлення можуть бути функціями часу, наприклад мова при передачі телефонних розмов, температура або тиск при передачі телеметричних даних, вистава в театрі при передачі по телебаченню й т. п. В інших випадках повідомлення не є функцією часу (наприклад, текст телеграми, нерухоме зображення й т. д.).

Сигнал передає (розгортає) повідомлення в часі. Отже, він завжди є функцією часу, навіть якщо повідомлення (наприклад, нерухоме зображення) таким не є. Якщо сигнал являє собою функцію $x(t)$, що приймає тільки певні дискретні значення x (наприклад, 1 й 0), то його називають *дискретним* або *дискретним за рівнем* (амплітудою). Так само й повідомлення, що приймає тільки деякі певні рівні, називають дискретним. Якщо ж сигнал (або повідомлення) може приймати безліч множин значень у деякому інтервалі, то він називається *безперервним* або *аналоговим*.

У деяких випадках повідомлення або сигнал задають не на всій осі часу, а в певні моменти t . Такі повідомлення (сигнали) називають *дискретними за часом* на відміну від безперервних за часом, заданих на всій осі t . Наприклад, мова є повідомленням безперервним як за рівнем, так і за часом, а датчик температури, що видає її значення через кожні 5 хв, служить джерелом повідомлень, безперервних за величиною, але дискретних за часом. На рис. 1.1 наочно проілюстровані різні види сигналів.

Сигнал з кінцевим числом дискретних рівнів часто називають цифровим, оскільки рівні можна пронумерувати числами з кінцевим числом розрядів.

Дискретні повідомлення не обов'язково перетворюються на дискретні сигнали, а безперервні повідомлення — на безперервні сигнали. Найчастіше саме безперервні сигнали використовують для передачі дискретних повідомлень (у якості сигналів-переносників). Дискретні ж сигнали можуть використовуватися для передачі безперервних повідомлень (після їх дискретизації).

Повідомлення за допомогою пристроїв (датчиків) звичайно перетворюються на електричну величину $b(t)$ — *первинний сигнал*. При передачі мови таке перетворення виконує мікрофон, при передачі зображення — відеокамера. У більшості випадків первинний сигнал є низькочастотним коливанням, що відображає передане повідомлення.

У деяких випадках первинний сигнал безпосередньо передають по фізичним ланцюгам. Так надходять сигнали, наприклад, при звичайному міському телефонному зв'язку. Для передачі на більші відстані (по кабелю або радіоканалу) первинний сигнал перетворюють на *високочастотний*.

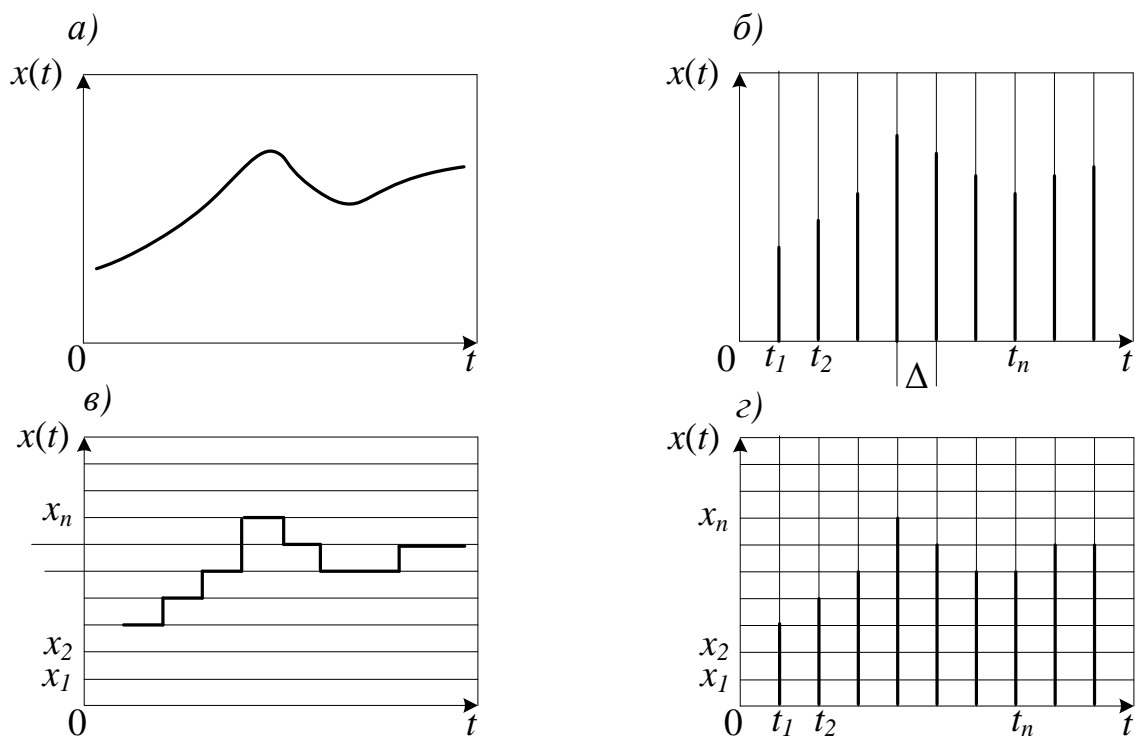


Рис. 1.1. Види сигналів: а) безперервний; б) дискретизований; в) квантований; г) дискретний

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

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

Описом конкретного сигналу може бути деяка функція часу $x(i)$. Визначивши так чи інакше цю функцію, визначаємо й сигнал. Однак такий повний опис сигналу не завжди потрібний. Для вирішення ряду питань досить більш загального опису у вигляді декількох параметрів, що характеризують основні властивості сигналу, подібно до того, як це робиться в системах транспортування. Зазначаючи габаритні розміри й масу, характеризуємо основні властивості предмета з погляду умов його перевезення; інші властивості (наприклад, кольори) із цього погляду є несуттєвими.

Сигнал також є об'єктом транспортування, а *техніка зв'язку* - технікою транспортування (передачі) сигналів *каналами зв'язку*. Тому доцільно визначити параметри сигналу, які є основними з погляду його передачі. Такими параметрами є *тривалість* сигналу T_c , його *динамічний діапазон* D_c і *ширина спектра* ΔF_c .

Тривалість сигналу показує, що будь-який сигнал, розглянутий як часовий процес, має початок і кінець. Тому тривалість сигналу T_c є природним його параметром, що визначає інтервал часу, у межах якого сигнал існує (с, мс, мкс, нс).

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

Ширина спектра сигналу ΔF_c є параметром, який дає уявлення про швидкості зміни сигналу усередині інтервалу його існування. Спектр сигналу в принципі може бути необмеженим. Однак для будь-якого сигналу можна вказати діапазон частот, у межах якого зосереджена його основна енергія. Цим діапазоном і визначається ширина спектра сигналу (Гц, кГц, МГц, ГГц).

У техніці зв'язку спектр сигналу часто свідомо скорочують. Це обумовлено тим, що апаратура й лінія зв'язку мають обмежену *смугу частот*.

Скорочення спектра здійснюється виходячи із припустимих перекручувань сигналу. Наприклад, при телефонному зв'язку потрібно, щоб мова була розбірлива й щоб кореспонденти могли розпізнати один одного за голосом. Для виконання цих умов досить передати мовний сигнал у смузі від 300 до 3400 Гц. Передача більш широкого спектра мови в цьому випадку недоцільна, тому що веде до технічних ускладнень і збільшення витрат.

Аналогічно необхідна ширина спектра телевізійного сигналу визначається необхідною чіткістю зображення. При стандарті в 625 рядків верхня частота сигналу досягає 6 МГц. Спектр сигналу зображення багато ширше за спектр сигналу звукового супроводу. Це істотно ускладнює побудову систем телевізійного віщання в порівнянні із системами звукового віщання. Ширина спектра телеграфного сигналу залежить від швидкості передачі й звичайно приймається рівною $\Delta F_c = 1,5\nu$, де ν - швидкість передачі (телеграфування) у Бодах (Бод), тобто число символів, переданих у секунду. Так, при телетайпній передачі $\nu = 50\text{ Бод}$ й тоді $\Delta F_c = 75\text{ Гц}$.

Спектр модульованого сигналу звичайно ширше за спектр вхідного повідомлення (первинного сигналу) і залежить від виду модуляції.

Добуток основних параметрів сигналу називають об'ємом сигналу.

$$V_c = T_c \Delta F_c D_c. \quad (1.1)$$

Об'єм сигналу є узагальненою характеристикою сигналу як носія інформації. На практиці іноді використовується геометричне зображення об'єму сигналу у вигляді прямокутного паралелепіпеда в тривимірному просторі з координатами T , ΔF і D . Чим більше обсяг сигналу, тим більше інформації можна "вкласти" у цей обсяг і тим складніше такий сигнал передавати по каналу зв'язку з необхідною якістю.

1.2. Системи, канали та мережі зв'язку

Мережа зв'язку складається в загальному випадку із джерела та одержувача повідомлень, перетворювачів повідомлення на сигнал та сигналу на повідомлення, каналу зв'язку. На рис. 1.2 зображена структурна схема найпростішої *одноканальної системи зв'язку*. Джерелом повідомлень й одержувачем в одних системах зв'язку може бути людина, в інших — різного роду пристрої (автомат, обчислювальна машина й т. д.).

Пристрій, що перетворює повідомлення на сигнал, називають *передавальним*, а пристрій, що перетворює прийнятий сигнал на повідомлення, — *приймальним*. За допомогою первинного перетворювача в передавальному пристрої повідомлення a , що може мати будь-яку фізичну природу (зображення, звукове коливання й т. п.), перетвориться на первинний електричний сигнал $b(t)$. У телефонії, наприклад, ця операція зводиться до перетворення акустичних коливань на електричну напругу, що пропорційно змінюється, на виході мікрофона. У телеграфії за допомогою телеграфного апарата послідовність елементів повідомлення (букв) замінюється

послідовністю кодових символів (0, 1 або крапка, тире), що одночасно перетворюється на послідовність електричних імпульсів постійного струму, у передавачі первинний сигнал $b(i)$ (низькочастотний) перетворюється на вторинний (високочастотний) сигнал $u(t)$, придатний для передачі по використовуваному каналу. Це здійснюється за допомогою модуляції.

Перетворення повідомлення на сигнал повинно бути оборотним. У цьому випадку за вихідним сигналом можна відновити вхідний первинний сигнал, тобто одержати всю інформацію, що міститься в переданому повідомленні. В іншому випадку частина інформації буде загублена при передачі.

Лінією зв'язку називається фізичне середовище й сукупність апаратних засобів, що використовуються для передачі сигналів від передавача до приймача. У системах електричного зв'язку — це насамперед кабель або хвилевід, у системах радіозв'язку — область простору, у якому поширюються електромагнітні хвилі від передавача до приймача.

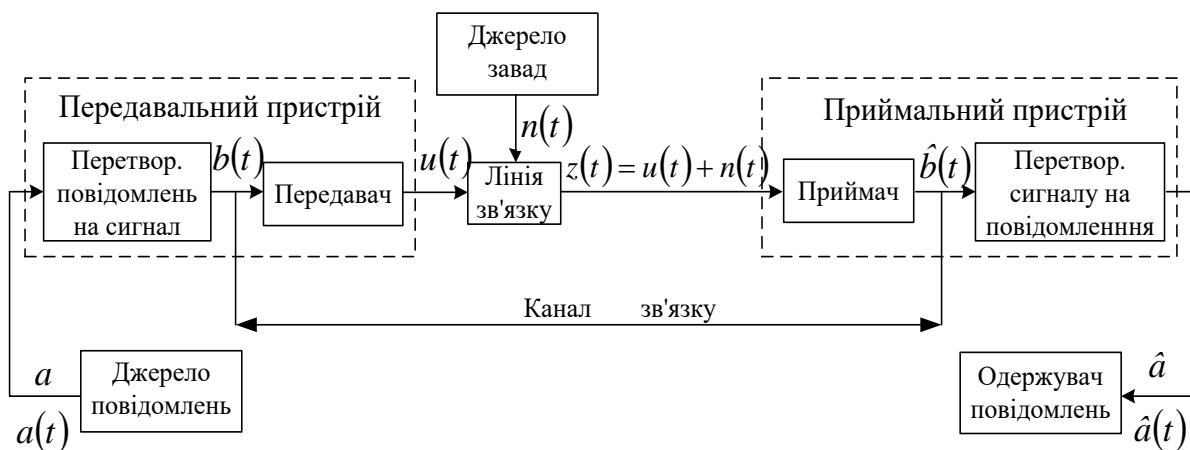


Рис. 1.2. Узагальнена структурна схема системи передачі інформації

При передачі каналний сигнал $u(t)$ може спотворюватися й на нього можуть накладатися завади $n(t)$. Приймальний пристрій обробляє прийняте коливання $z(t) = x(t) + n(t)$, яке є сумою перекрученого сигналу $x(t)$ і завади $n(t)$, і відновлює по ньому

повідомлення \hat{a} , що з деякою похибкою відображає передане повідомлення a . Інакше кажучи, приймач повинен на основі аналізу коливання $z(t)$ визначити, яке з можливих повідомлень передавалось. Тому прийомний пристрій є одним з найбільш відповідальних і складних елементів системи зв'язку.

Сукупність технічних засобів для передачі повідомлень від джерела до споживача та лінії зв'язку називається *системою зв'язку*. Цими засобами є передавальний пристрій, лінія зв'язку й прийомний пристрій. Іноді в поняття система зв'язку включається джерело й споживач повідомлень.

За видом переданих повідомлень розрізняють такі системи зв'язку: передача мови (телефонія); передача тексту (телеграфія); передача нерухомих зображень (фототелеграфія); передача рухомих зображень (телебачення), телевимірювання, телекерування й передача даних.

За призначенням телефонні й телевізійні системи поділяють на віщальні, що відрізняються високим ступенем художності відтворення повідомлень, і професійні, що мають спеціальне застосування (службовий зв'язок, промислове телебачення й т. п.). У системі телевимірювання фізична величина, що підлягає вимірюванню (температура, тиск, швидкість і т. п.), за допомогою датчиків перетворюється на первинний електричний сигнал, що надходить на передавач. На прийомному кінці передану фізичну величину або її зміни виділяють із сигналу й спостерігають або реєструють за допомогою записуючих приладів. У системі телекерування здійснюється передача команд для автоматичного виконання певних дій. Нерідко ці команди формують автоматично на підставі результатів вимірювання, переданих телеметричною системою.

Впровадження вискоефективних електронно-обчислювальних машин (ЕОМ) призвело до необхідності швидкого розвитку *систем передачі даних*, що забезпечують обмін інформацією між обчислювальними засобами й об'єктами автоматизованих систем керування. Цей вид електрозв'язку в порівнянні з телеграфним відрізняється більше високими вимогами до швидкості й правильності передачі інформації.

Каналом зв'язку називається сукупність засобів, що забезпечують передачу сигналу від деякої точки *A* системи до точки *B* (рис. 1.3). На цьому рисунку позначено Пер1 - перетворювач повідомлення на сигнал; Пер2 - перетворювач сигналу на повідомлення; ПрД – передавач; ПрМ – приймач; ЛЗ – лінія зв'язку. Точки *A* та *B* можуть бути обрані довільно, аби тільки між ними проходив сигнал. Частина системи зв'язку, розташована до точки *A*, є джерелом сигналу для цього каналу. Якщо сигнали, що надходять на вхід каналу і з його виходу, є дискретними (за рівнями), то *канал* називається *дискретним*. Якщо вхідні й вихідні сигнали є безперервними (за рівнем), то й *канал* називається *безперервним*.

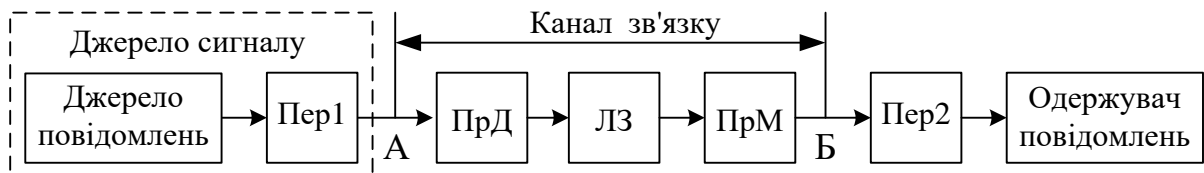


Рис. 1.3. Структурна схема каналу зв'язку

Зустрічаються також *дискретно-безперервні* й *безперервно-дискретні* канали, на вхід яких поступають дискретні сигнали, а з виходу знімаються безперервні, або навпаки. Канал може бути дискретним або безперервним незалежно від характеру переданих повідомлень. Більш того, в одній і тій самій системі зв'язку можна виділити як дискретний, так і безперервний канал. Все залежить від того, обрані точки *A* і *B* входу й виходу.

За аналогією з сигналами канал зв'язку, застосовуваний для передавання сигналів, характеризують такими основними параметрами: часом роботи каналу зв'язку T_k , смугою частот пропускання каналів зв'язку ΔF_k і динамічним діапазоном D_k (перевищенням середньої потужності сигналу над середньою потужністю завад на виході каналу зв'язку):

$$D_k = 10 \log \frac{P_c}{P_3}. \quad (1.2)$$

Добуток основних параметрів каналу зв'язку називають об'ємом (ємністю) каналу зв'язку:

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

Основна умова узгодження каналу зв'язку з сигналом, виконання якої забезпечує можливість неспотвореного сигналу, така:

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

Повідомлення має бути перетворене на сигнал так, щоб об'єм сигналу в тривимірному просторі $T, \Delta F, D$ відповідав об'єму каналу в тому самому тривимірному просторі, тобто час передавання сигналу має збігатися з часом роботи каналу зв'язку; смуга частот сигналу має збігатися зі смугою частот каналу зв'язку; динамічні діапазони на вході і виході каналу зв'язку мають дорівнювати один одному. Тільки в цьому випадку канал зв'язку може забезпечити неспотворене передавання сигналу.

Якщо об'єм сигналу менший за об'єм каналу зв'язку або дорівнює йому, тоді завжди можна здійснити таке перетворення сигналу, при якому виконуватимуться умови (1.4). Отже, умова можливості узгодження сигналу і каналу зв'язку

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

Відношення називається резервом ємності каналу зв'язку та визначається як

$$r_k = \frac{V_k}{V_c}. \quad (1.6)$$

Резерв ємності каналу зв'язку характеризує можливість підвищення надійності каналу зв'язку.

Загальними ознаками безперервних каналів є наступні. По-перше, більшість каналів можна вважати *лінійними*.

Типи каналів, по яких передаються сигнали, численні й різноманітні. Розрізняють канали провідного зв'язку (повітряні, кабельні, світловоди й ін.) і канали радіозв'язку. Кабельні лінії зв'язку є основою магістральних мереж далекого зв'язку, по них здійснюється передача сигналів у діапазоні частот від десятків кГц до сотень МГц.

Досить перспективними є волоконно-оптичні лінії зв'язку. Вони дозволяють у діапазоні 600...900 ТГц (0,5...0,3 мкм) забезпечити дуже більшу пропускну спроможність (сотні ТВ або сотні тисяч ТФ каналів). Поряд із провідними лініями зв'язку широко використовуються радіолінії різних діапазонів (від сотень кГц до десятків ГГц). Ці лінії більш економічні й незамінні для зв'язку з рухомими об'єктами. Найбільше поширення для багатоканального радіозв'язку одержали радіорелейні лінії (РРЛ) метрового, дециметрового й сантиметрового діапазонів на частотах від 60 МГц до 15 ГГц.

Різновидом РРЛ є тропосферні лінії з використанням відбиттів від неоднорідностей тропосфери. Все більше застосування знаходять супутникові лінії зв'язку — РРЛ із ретранслятором на штучних супутниках Землі (ШСЗ). Для цих ліній (систем) зв'язку відведені діапазони частот 4...6 і 11...275 ГГц. Дальність при одному ретрансляторі на супутнику, гнучкість і можливість організації глобального зв'язку - важливі переваги супутникових систем.

Однак нерівність (1.5) може виконуватися й тоді, коли одне або два з нерівностей (1.4) не виконані. Це означає, що можна робити "обмін" тривалості на ширину спектра. Записаний на плівку телефонний сигнал, який має ширину спектра 3 кГц, необхідно передати через канал, смуга пропускання якого 300 Гц. Це можливо здійснити, відтворюючи сигнал зі швидкістю, до 10 разів меншою за ту, з якою він був записаний. При цьому всі частоти вихідного сигналу зменшаться в 10 разів і в стільки ж раз збільшиться час передачі. Прийнятий сигнал також записується на плівку, а потім, відтворивши його зі швидкістю, в 10 разів більшою, можна відновити вихідний сигнал. Аналогічно можна передати сигнал швидше, якщо смуга пропускання каналу ширше за спектр сигналу.

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

У системі зв'язку, наведеній на рис. 1.2, передача повідомлень здійснюється в одному напрямку від джерела до одержувача. Такий режим зв'язку називається *симплексним*. Режим, при якому забезпечується можливість одночасної передачі повідомлень у прямому й зворотному напрямку, називається *дуплексним*. Можливий і *напівдуплексний* режим, коли обмін повідомлень здійснюється по черзі.

Система зв'язку називається *багатоканальною*, якщо вона забезпечує передачу декількох повідомлень по одній загальній лінії зв'язку. Структурна схема найпростішої *багатоканальної* системи зв'язку зображена на рис. 1.4.

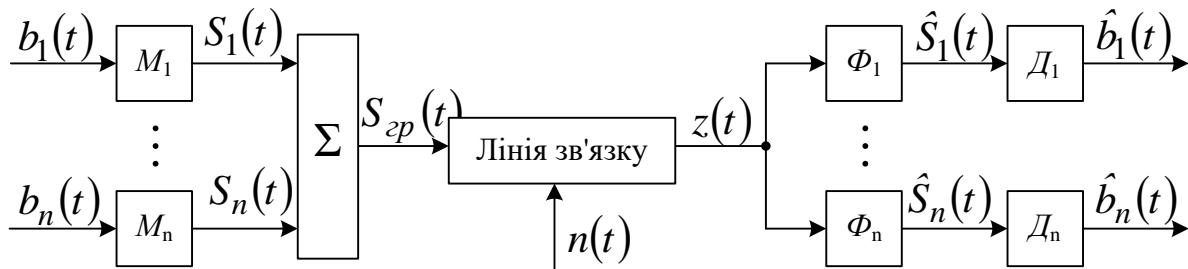


Рис. 1.4. Структурна схема багатоканальної системи передачі

Для обміну повідомленнями між багатьма територіально рознесеними користувачами (абонентами) створюються мережі зв'язку, що забезпечують передачу й розподіл повідомлень за заданими адресами (у заданий час і з установленою якістю). Розподіл потоків повідомлень за заданими адресами здійснюється на вузлах зв'язку за допомогою комутаційних пристроїв. За способом розподілу повідомлень мережі поділяються на ті, що не комутують, і ті, що комутуються. У першому випадку зв'язок між абонентами здійснюється по постійно закріплених каналах за

принципом "кожний з кожним". У другому випадку абоненти зв'язуються між собою не безпосередньо, а через вузли комутації. Мережа зв'язку являє собою сукупність кінцевих (абонентських) пристроїв, каналів зв'язку (сполучних ліній) і вузлів комутації. Залежно від числа абонентів і розмірів обслуговуваної території мережі можуть мати різну структуру: лінійну, радіальну, кільцеву, радіально-вузлову й т. п. Завдання оптимальної побудови мереж зв'язку є одним з найважливіших завдань теорії й техніки зв'язку. Вирішується це завдання за допомогою теорії графів і теорії масового обслуговування.

Під мережею зв'язку розуміється сукупність вузлів комутації і ліній зв'язку, об'єднаних через фізичне середовище передачі [42]. Мережа зв'язку дозволяє якнайповніше і оперативніше забезпечити потреби абонентів різних категорій у широкому асортименті послуг зв'язку. Залежно від своєї протяжності мережі поділяються на глобальні, регіональні, міські, кампусні і локальні тощо.

Так, на залізничному транспорті України з урахуванням специфіки організації його структури телефонна мережа зв'язку будується переважно за радіально-вузловим принципом і поділяється на магістральну, дорожню, відділкову і станційну (рис. 1.5) [14].

Магістральний зв'язок (МС) з'єднує Центральну станцію зв'язку (ЦСЗ) Укрзалізниці з дорожніми вузлами (ДВ) і останні між собою. Дорожня мережа (ДМ) організовується в межах однієї залізниці. Відділкова мережа (ВМ) з'єднує відділковий вузол (ВВ) з крупними залізничними станціями (ЗС) у відділеннях дороги (дирекціях). Лінійні споруди зв'язку спільно з каналоутворюючою апаратурою, що забезпечує з'єднання між адміністративними пунктами, утворюють первинну мережу зв'язку. Вторинні мережі організовуються на базі первинної мережі і об'єднують місцеві телефонні мережі, мережі телеграфного зв'язку і передачі даних тощо.

З'єднання абонентів і абонентських пристроїв вторинної мережі для передачі і прийому інформації здійснюється за допомогою комутаційних станцій, розташованих у вузлах комутації (вузлах зв'язку).

У сучасному електровз'язку розрізняють такі способи комутації: комутація каналів, комутація пакетів, комутація повідомлень.

Під комутацією каналів розуміється процедура встановлення фізичного з'єднання між двома кінцевими пристроями за допомогою одного або декількох комутаторів на весь час встановлення з'єднання.

Під комутацією пакетів розуміється технологія передачі даних, що характеризується тим, що потік інформації який передається, розбивається на пакети фіксованої або змінної довжини, які обробляються і комутуються в мережі як незалежні блоки. Встановлення з'єднання здійснюється лише на час передачі пакету.

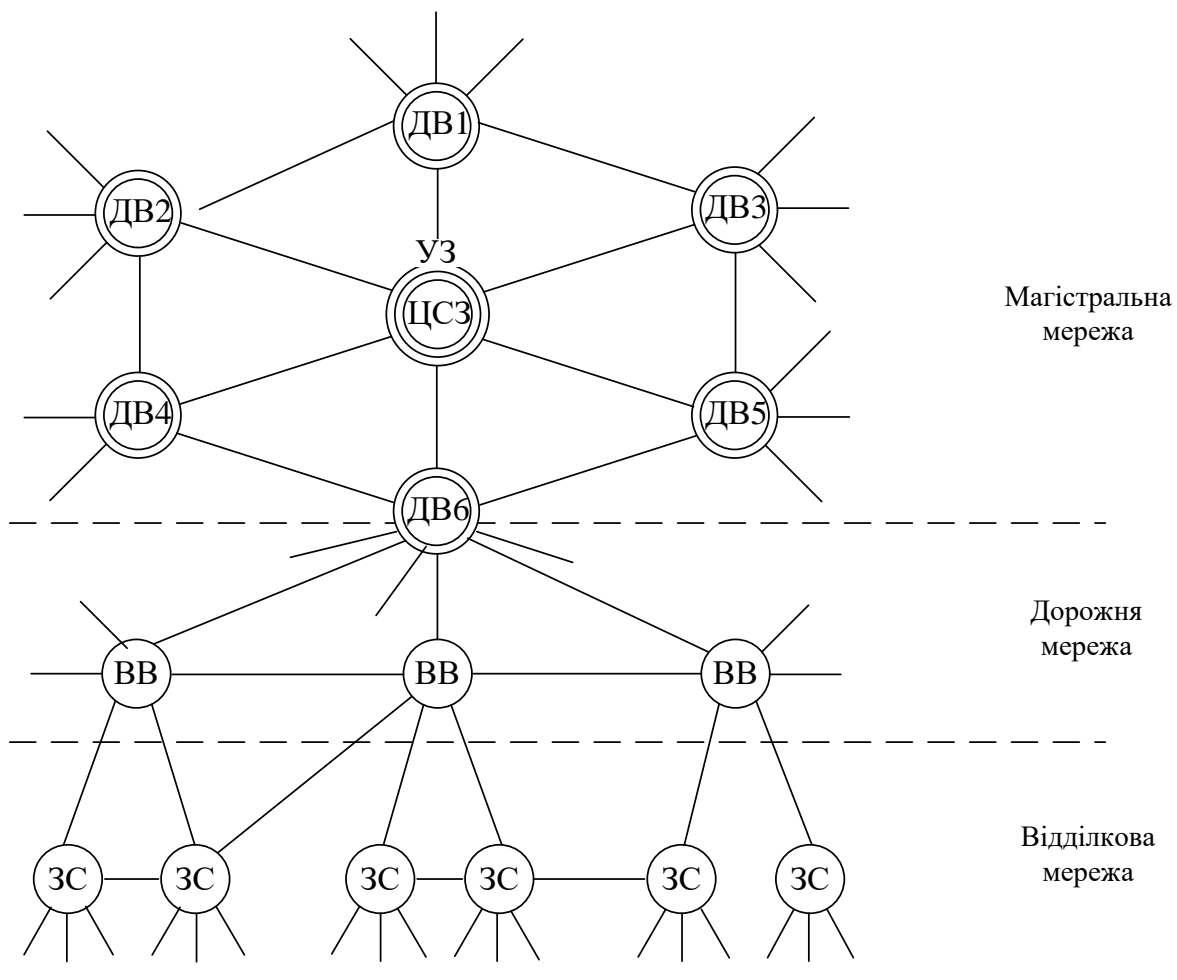


Рис. 1.5. Принцип побудови мереж магістрального, дорожнього і відділкового видів зв'язку на залізничному транспорті

Під комутацією повідомлень розуміється транзитна передача через комутатор блоку даних, що належать до даного повідомлення, без розділення його на частини (або пакети при комутації пакетів). Повідомлення, на відмінну від пакету, має довжину, яка визначається інформаційним змістом.

Оскільки сучасні телекомунікаційні системи і мережі зв'язку є достатньо складними структурами, що включають багато компонентів, то для зображення взаємодії між собою різних частин мережі Міжнародна організація стандартизації (ISO, International Organization for Standardization) в 1977 р. розробила модель відкритої системної взаємодії (OSI, Open Systems Interconnection). Відтоді вона стала широко використовуватися для пояснення мережних комунікацій. У сучасному електрозв'язку розрізняють такі способи комутації: комутація каналів, комутація пакетів, комутація повідомлень [43].

У процесі побудови будь-якої багаторівневої структури виникає завдання визначення оптимального числа її рівнів. Так, при розробленні еталонної моделі число її рівнів визначалося з таких міркувань:

1. Розбиття на рівні має максимально відображати логічну структуру проектованої мережі.

2. Міжрівневі межі повинні бути визначені так, щоб забезпечувалося мінімальне число і простота міжрівневих зв'язків.

3. Необхідно враховувати, що велике число рівнів, з одного боку, спрощує внесення змін до системи, а з іншого – збільшує кількість міжрівневих протоколів і ускладнює опис моделі в цілому.

З урахуванням цього Міжнародною організацією стандартів для мереж зв'язку різного застосування була запропонована семирівнева модель взаємодії відкритих систем (рис. 1.6).

У цій системі основним, з погляду користувача, є прикладний рівень. Сьомий рівень забезпечує виконання прикладних процесів користувачів і визначає семантику, тобто смисловий зміст інформації, якою обмінюються відкриті системи в процесі їх взаємодії. З цією метою даний рівень, окрім протоколів взаємодії прикладних процесів, підтримує протоколи

передачі файлів, віртуального терміналу, електронної пошти і подібні до них.

Шостий рівень називається представницьким (рівень подання даних); він визначає єдиний для всіх відкритих систем синтаксис інформації, що передається. Необхідність даного рівня обумовлена різною формою подання інформації в мережі передачі даних і кінцевих обладнаннях. Даний рівень грає важливу роль в забезпеченні "відкритості" систем, дозволяючи їм взаємодіяти між собою незалежно від їх внутрішньої мови.

П'ятий рівень називається сеансовим, оскільки основним його призначенням є організація сеансів зв'язку між прикладними процесами, розташованими в різних абонентських системах. На даному рівні створюються порти для передачі і прийому повідомлень і організуються з'єднання – логічні канали між собою. Необхідність протоколів даного рівня визначається відносною складністю мережі передачі даних і прагненням забезпечити достатньо високу надійність передачі інформації.

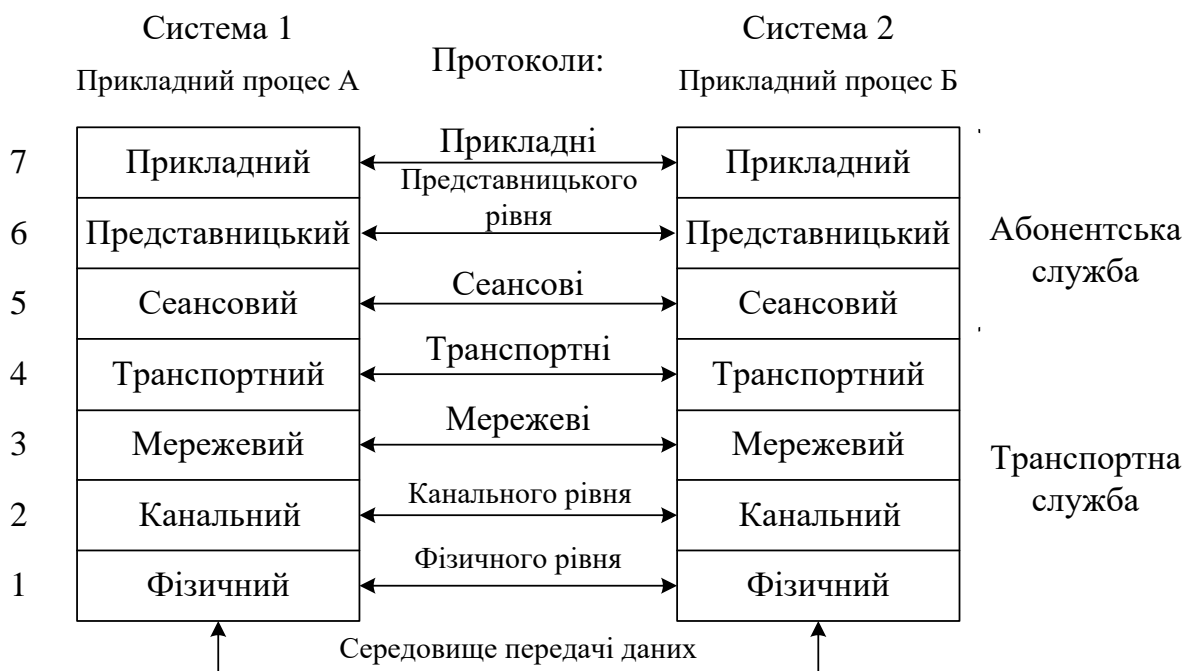


Рис. 1.6. Еталонна модель взаємодії відкритих систем

Четвертий рівень називається транспортним (рівень наскрізної передачі) і служить для забезпечення передачі даних між двома взаємодіючими відкритими системами і організації

процедури сполучення абонентів мережі з системою передачі даних. На цьому рівні визначається взаємодія абонентських систем і підтримується логічний канал (транспортне з'єднання) між абонентами.

Третій рівень називається мережевим і призначений для забезпечення процесів маршрутизації інформації (повідомлень, що передаються) й управління мережею передачі даних. На відміну від попередніх, даний рівень більшою мірою орієнтовано на мережу передачі даних. На цьому рівні вирішуються питання управління мережею передачі даних, зокрема маршрутизації і управління інформаційними потоками.

Другий рівень називається каналним і забезпечує функціональні і процедурні засоби для встановлення, підтримки і розмикання з'єднань на рівні каналів передачі даних. Процедури каналного рівня забезпечують виявлення і, можливо, виправлення помилок, що виникають на фізичному рівні.

Перший рівень називається фізичним і служить для забезпечення механічних, електричних, функціональних і процедурних засобів організації фізичних з'єднань при передачі даних між фізичними пристроями (об'єктами).

Інтенсивний розвиток інформаційних технологій у 70-ті р. минулого сторіччя викликав такий розвиток мікропроцесорної техніки, яка стимулювала подальше вдосконалення цифрових методів передачі мови і даних, що призвело до розроблення і впровадження не тільки технологій локальних мереж: ARCnet, Ethernet, Token Ring і FDDI, але і нових високошвидкісних технологій глобальних мереж: PDH, SONET, SDH, ISDN, FramRelay, ATM.

Найбільш поширеними цифровими технологіями в глобальних мережах зв'язку є PDH і SDH [43].

Так, в ієрархії PDH (плезіохронна цифрова технологія) широко застосовується цифровий канал (його часто називають основним), він має швидкість передачі 64 кбіт/с. На його базі будують цифрові групові тракти:

а) первинний – 30 каналів інформаційних і 2 канали додаткових (один для синхронізації, другий, – для управління комутаційним устаткуванням на АТС). Він має швидкість 2048 кбіт/с;

б) вторинний – 120 каналів інформаційних і додаткові канали для сигналізації і синхронізації. Швидкість передачі сигналів у цьому тракті 8,448 Мбіт/с;

в) третинний – 480 каналів інформаційних і додаткові канали для сигналізації і синхронізації. Швидкість передачі сигналів у цьому тракті 34,468 Мбіт/с;

г) четвертинний – 1920 каналів інформаційних і додаткові канали для сигналізації і синхронізації. Швидкість передачі сигналів дорівнює 139,264 Мбіт/с.

Ієрархія SDH (синхронна цифрова технологія) в порівнянні з ієрархією PDH має такі переваги: спрощення побудови мережі, вища надійність, самовідновлюваність мережі, гнучкість управління мережею, виділення смуги пропускання на вимогу, прозорість для передачі будь-якого графіка, універсальність застосування, простота нарощування потужності.

Основною функціональною одиницею в ієрархії SDH є синхронний транспортний модуль (STM), швидкість передачі якого складає $V_{STM-1} = 155,52$ Мбіт/с (для STM - 1). Для вищих рівнів передачі швидкість визначається за формулою

$$V_{II} = V_{STM-1} \cdot N,$$

де рівень ієрархії SDH, причому $N = 1, 4, 16$ і т. д.

Так, для другого рівня ієрархії SDH швидкість передачі

$$V_{I2} = 4 \times 155,52 = 622,08 \text{ Мбіт/с.}$$

Аналогічно можна отримати швидкості передачі для третього рівня $V_{I3} = 16 \times 155,52 = 2488,32$ Гбіт/с і для четвертого $V_{I4} = 64 \times 155,52 = 9953,28$ Гбіт/с.

1.3. Завади та спотворення у каналах зв'язку

У реальному каналі сигнал при передачі спотворюється й повідомлення відтворюється з деякою помилкою. Причиною таких помилок є як перекручування, внесені самим каналом, так і завади, що впливають на сигнал. Частотні й часові характеристики каналу визначають так названі лінійні

спотворення. Крім того, канал може вносити й нелінійні спотворення, обумовлені нелінійністю тих або інших ланок каналу.

Якщо лінійні й нелінійні спотворення обумовлені відомими характеристиками каналу, то вони, принаймні в принципі, можуть бути усунуті належною корекцією. Варто відрізнити спотворення від завад, що мають випадковий характер. Завади заздалегідь не відомі й тому не можуть бути повністю усунуті.

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

У провідних каналах зв'язку основним видом завад є імпульсні шуми й спотворення зв'язку. Поява імпульсних завад часто пов'язане з автоматичною комутацією й перехресними наведеннями. Спотворення зв'язку є явищем, при якому сигнал у лінії різко загасає або зникає.

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

вплив завади $n(t)$ на корисний сигнал $u(t)$ можна виразити оператором

$$z(t) = L[u(t), n(t)]. \quad (1.7)$$

В окремому випадку, коли оператор вироджується в суму, завада називається *адитивною*.

$$z(t) = u(t) + n(t). \quad (1.8)$$

Якщо ж оператор може бути поданий у вигляді добутку, то заваду називають *мультиплікативною*.

$$z(t) = \sum_{n=1}^N k_n(t) u(t - \tau_n), \quad (1.9)$$

де $k(t)$ - випадковий процес;

τ_n – час затримки n -го каналу.

У реальних каналах звичайно мають місце й адитивні, і мультиплікативні завади, і тому

$$z(t) = \sum_{n=1}^N k_n(t) u(t - \tau_n) + n(t). \quad (1.10)$$

Серед адитивних завад різного походження виділяють зосереджені за спектром (вузькосмугові) завади, зосереджені в часі (імпульсні) завади й так названу флуктуаційну заваду, не обмежену в часі й спектрі. Флуктуаційна завада (флуктуаційний шум) являє собою випадковий процес із нормальним розподілом (гаусівський процес). Така завада найбільш вивчена й становить найбільший інтерес як у теоретичному, так й у практичному відношенні. Цей вид завад практично має місце у всіх реальних каналах (рис. 1.7). У діапазоні оптичних частот істотне значення має квантовий шум, викликаний дискретною природою сигналу. Мультиплікативні завади обумовлені випадковими змінами параметрів каналу зв'язку. Зокрема ці завади проявляються в зміні рівня сигналу.

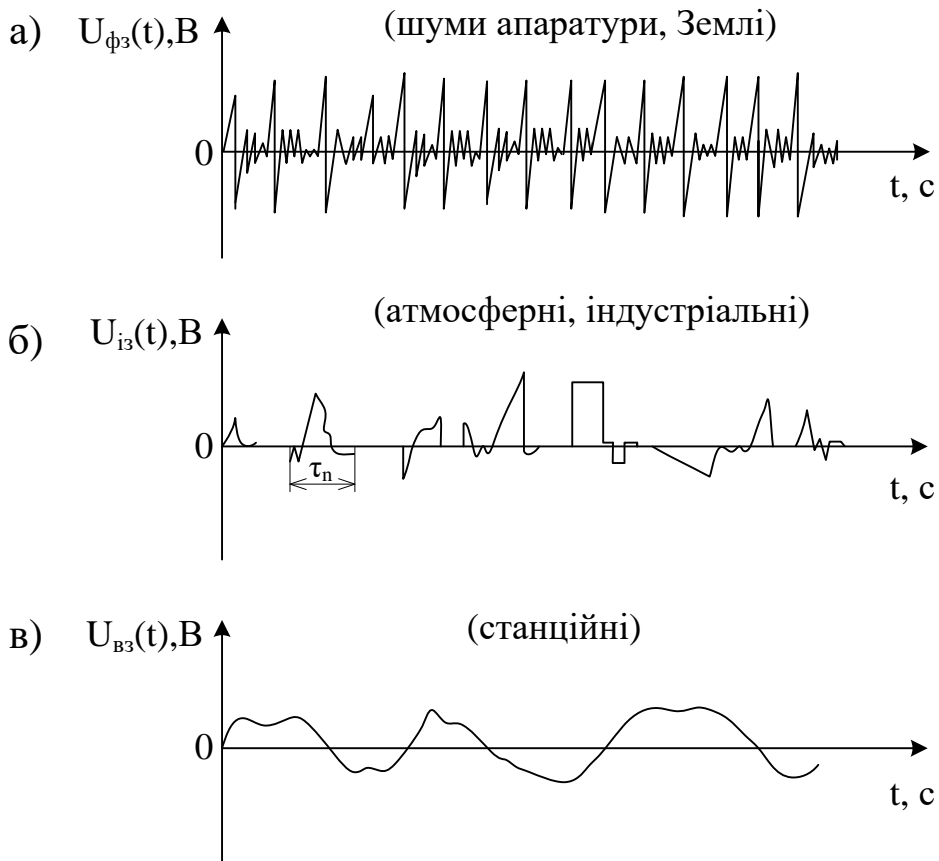


Рис. 1.7. Графіки завад:
 а) флуктуаційна; б) імпульсна; в) вузькосмугова

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

1.4. Кодування та модуляція

Перетворення дискретного повідомлення на сигнал звичайно здійснюється у вигляді двох операцій — *кодування* й *модуляції*. Кодування являє собою перетворення повідомлення на послідовність кодових символів, а модуляція - перетворення цих

символів на сигнали, придатні для передачі по каналу. За допомогою кодування й модуляції джерело повідомлень узгоджується з каналом.

Найпростішим прикладом дискретного повідомлення є текст. Будь-який текст складається з кінцевого числа елементів: букв, цифр, розділових знаків. Їхня сукупність називається алфавітом джерела повідомлення. Оскільки число елементів в алфавіті звичайне, то їх можна пронумерувати й тим самим звести передачу повідомлення до передачі послідовності чисел.

Так, для передачі великих букв російського алфавіту (32 букви) необхідно передати числа від 0 до 31. Для передачі будь-якого числа, записаного в десятковій формі, потрібна передача.

Процес перетворення повідомлень на сигнал у теорії інформації називається кодуванням відображенням або кодуванням, а алгоритм, що визначає правила, за якими здійснюється це перетворення, - кодом.

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

За табличним способом задання алфавітного оператора складається таблиця, в ліву частину якої вписуються всі слова з області визначення даного алфавітного оператора, а в праву частину – слова, що утворюються внаслідок застосування даного алфавітного оператора до кожного слова лівої частини таблиці.

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

Найпростішим видом кодування є кодування за літерами, коли кожній літері алфавіту повідомлення $A(x_1, \dots, x_n)$ відповідає деяка скінченна послідовність літер (слово) в алфавіті сигналу $B(y_1, \dots, y_m)$, яка називається кодом відповідної літери. Різним літерам алфавіту мають відповідати різні коди.

При передаванні повідомлень системи зв'язку обов'язковою вимогою є вимога зворотності кодування, котра полягає в тім,

щоб за кодом будь-якого слова можна було б відновити спочатку закодоване слово. Тільки у разі виконання цієї вимоги можна за сигналом відновити передане повідомлення.

Для того щоб визначити умови зворотності кодування слів за літерами, введемо поняття про початковий відрізок слова, заданого в певному алфавіті. Домовимося називати слово p початковим відрізком слова q , якщо слово q можна подати у вигляді $q=pl$, де l – будь-яке слово, в тому числі й слово нульової довжини. Наприклад, слово $p=xu$ є початковим відрізком слова $q=xuux$, тому що слово q можна подати у вигляді $q=pl$, де $l=ux$.

Для того щоб кодування слів за літерами в алфавіті $A(x_1, \dots, x_n)$ було зворотним, необхідними і достатніми мають виконуватися такі умови:

1. Коди різних літер вихідного алфавіту повинні відрізнятися один від одного.

2. Код кожної з літер вихідного алфавіту не повинен збігатися з жодним із початкових відрізків кодів інших літер цього алфавіту.

Друга умова зворотності кодування завжди виконується, якщо коди всіх літер вихідного алфавіту мають однакову довжину. У цьому випадку в коді слова q досить виділити початковий відрізок s_1 , що має довжину, рівну довжині коду літери, і за ним визначити першу літеру слова p .

Аналогічно можуть бути знайдені й всі інші літери слова p . Кодування за літерами, при якому коди всіх літер вихідного алфавіту мають рівну довжину, називається нормальним або рівномірним кодуванням.

Кодування дозволяє перетворювати слова, подані в будь-якому довільному алфавіті, на слова в обраному стандартному алфавіті. Найчастіше за такий стандартний алфавіт береться двійковий алфавіт, що складається з двох літер, які звичайно ототожнюються з цифрами 0 і 1.

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

$$m^k > n. \quad (1.11)$$

Оскільки число різних слів довжиною k в алфавіті, що складається з m літер, дорівнює m^k , то нерівність (1.11) показує, що можна закодувати всі літери алфавіту $A(x_1, \dots, x_n)$ словами довжиною k і алфавіту $B(y_1, \dots, y_m)$ так, що коди різних літер відрізняються один від одного. Будь-яке таке кодування буде нормальним, а отже, зворотнім.

Найпростішим прикладом кодування за літерами повідомлень є коди Бодо і Морзе, що застосовуються при телеграфуванні.

Алфавіт сигналів коду Бодо складається з двох літер (0 і 1). Код кожної літери повідомлення складається з п'яти літер алфавіту сигналу, що дозволяє передати 32 різні літери ($2^5 = 32$). Коди деяких літер українського алфавіту для цього випадку наведені нижче.

Літери українського алфавіту	А	Б	В	Г	Д
Код Бодо	10000	01100	01101	01010	11110

Код Бодо є рівномірним кодом. При побудові сигналів у кодї Бодо для подання одиниці використовується посилення струму в лінію зв'язку, а для подання нуля – відсутність посилення струму.

У кодї Морзе алфавіт сигналу складається з чотирьох літер: точки, тире, паузи між літерами і паузи між словами. Цей код є нерівномірним, тобто коди букв вихідного алфавіту мають у ньому різну довжину. Наприклад, слово «КІЇВ» у цьому кодї має такий вигляд:

К	И	Ї	В
– • –	• •	•	• – –

При побудові сигналів у кодї Морзе точці відповідає коротка послідовність струму і рівна їй за тривалістю пауза, тире – послідовність струму потроєної тривалості і коротка пауза, інтервалу

між літерами – пауза подвоєної тривалості, інтервалу між словами – пауза збільшеної в чотири рази тривалості.

Щоб з'ясувати зміст двійкової системи, звернемося до десяткової системи, основою числення якої є число 10. Будь-яке ціле число K можна подати у вигляді

$$K = a_n \cdot 10^n + \dots + a_2 \cdot 10^2 + a_1 \cdot 10^1 + a_0 \cdot 10^0, \quad (1.12)$$

де a_0, a_1, \dots, a_n – коефіцієнти, що набувають значень від 0 до 9.

Так, число 265 можна записати як $2 \cdot 10^2 + 6 \cdot 10^1 + 5 \cdot 10^0$. Вочевидь, замість числа 10 можна взяти будь-яке інше ціле число m і подати число K так [9]:

$$K = a_n m^n + \dots + a_2 m^2 + a_1 m^1 + a_0 m^0, \quad (1.13)$$

де a_0, a_1, \dots, a_n – коефіцієнти, що набувають значень від 0 до $m-1$.

Задаючи величину m , можна побудувати будь-яку систему числення.

При $m=2$ одержимо двійкову систему, в якій числа записуються за допомогою двох цифр – 0 і 1. Наприклад, число 13 у двійковій системі записується 1101, що відповідає виразу $1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$. Арифметичні дії в двійковій системі дуже прості. Так, додавання здійснюється за такими правилами: $0+0=0$; $0+1=1$; $1+0=1$; $1+1=10$. Розрізняють ще порозрядне додавання без перенесення в старший розряд, або «додавання за модулем 2». Правила цього додавання такі: $0 \oplus 0=0$; $0 \oplus 1=1$; $1 \oplus 0=1$; $1 \oplus 1=0$.

Якщо перетворити послідовність елементів повідомлення на послідовність двійкових чисел, то для передавання останніх по каналу зв'язку достатньо лише двох різних сигналів. Наприклад, символи 0 і 1 можуть передаватися коливаннями з різними частотами чи імпульсами струму різної полярності. Завдяки своїй простоті двійкова система числення широко застосовується для кодування дискретних повідомлень.

Кодування – це процес перетворення елементів повідомлення на відповідні їм числа (кодові символи). Кожному

елементу повідомлення присвоюється певна сукупність кодових символів, яка називається кодовою комбінацією. Сукупність кодових комбінацій, що відображають дискретні повідомлення, утворює код. Правило кодування може бути виражене кодовою таблицею, в якій наводяться алфавіт повідомлень, що кодуються, і відповідні їм кодові комбінації.

Множина можливих кодових символів називається кодовим алфавітом, а їх кількість m – основою коду. У загальному випадку при основі коду m правила кодування K елементів повідомлення зводиться до правил запису K різних чисел у m -ковій системі числення. Число розрядів n , що утворюють кодову комбінацію, називається розрядністю коду чи довжиною кодової комбінації. У залежності від системи числення, використовуваної при кодуванні, розрізняють двійкові і m -кові (недвійкові коди).

Коди, в яких усі комбінації мають однакову довжину, називають рівномірними. Для рівномірного коду число можливих комбінацій дорівнює m^n . Прикладом такого коду є п'ятизначний код Бодо, що містить п'ять двійкових елементів ($m=2$, $n=5$). Число можливих кодових комбінацій ($2^5=32$) є достатнім для кодування всіх літер українського алфавіту, проте недостатнім для передавання повідомлення, яке містить букви інших алфавітів, цифри, різні умовні знаки (крапка, кома, додавання, множення).

Тому наданий час використовується «Міжнародний телеграфний код 2» (МТК-2), побудований за реєстровим принципом, відповідно до якого та сама п'ятиелементна кодова комбінація може застосовуватися до трьох разів у залежності від положення регістра: український, латинський, цифровий. Загальне число різних знаків при цьому дорівнює 84, що є достатнім для кодування телеграми.

Для передавання даних рекомендовано семиелементний код МТК-5. Коди МТК-2 і МТК-5 є первинними (простими). Основними параметрами кодів є основа коду m , довжина кодової комбінації n , відстань між кодовими комбінаціями d_{ij} і вага кодової комбінації w .

Відстань d_{ij} характеризує розбіжність між двома кодовими комбінаціями і визначається, за Хеммінгом, числом незбіжних

розрядів у них, тобто числом одиниць у сумі двох комбінацій за модулем 2. Число ненульових елементів у кодовій комбінації визначає її вага w . Застосування рівномірних кодів спрощує побудову автоматичних літеродрукувальних пристроїв і не потребує передавання розділових символів між кодовими комбінаціями.

Нерівномірні коди характерні тим, що в них кодові комбінації відрізняються одна від одної не тільки взаємним розташуванням символів, але й їхньою кількістю, тому різні комбінації мають різну тривалість. Такі коди потребують або спеціальних розділових знаків, що вказують кінець однієї й початок іншої кодової комбінації, або ж такої побудови, щоб жодна кодова комбінація не була початком іншої. Коди, які задовольняють цю умову, називаються незвідними (префіксними). Зазначимо, що рівномірний код також є незвідним. Будувати код зручно у вигляді графа (кодового дерева), з кожного вузла якого виходить кількість гілок, яка дорівнює основі коду (для двійкового коду, наприклад, крок наверх означає 0, крок униз – 1).

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

За ознакою завадозахищеності коди поділяють на дві групи. Коди, в яких усі можливі кодові комбінації використовуються для передавання інформації, називаються простими або кодами без надлишковості (примітивними). У простих рівномірних кодах перетворення одного символу комбінації на інший, наприклад 1 на 0 чи 0 на 1, призводить до появи нової дозволеної комбінації, тобто до помилки.

Коректувальні коди будуються так, що для передавання повідомлення використовуються не всі кодові комбінації, а лише деяка їх частина (дозволені кодові комбінації). Тим самим створюється можливість виявлення і виправлення помилки у разі

неправильного відтворення деякого числа символів. Коректувальні властивості кодів досягаються введенням у кодові комбінації додаткових (надлишкових) символів.

Декодування полягає у відновленні повідомлення за прийнятими кодовими символами. Пристрої, що здійснюють кодування і декодування, називають відповідно кодером і декодером. Як правило, це логічні пристрої. На рис. 1.8 зображена структурна схема системи передавання дискретних повідомлень, а на рис. 1.9 пояснюється процес перетворення дискретного повідомлення на сигнал.

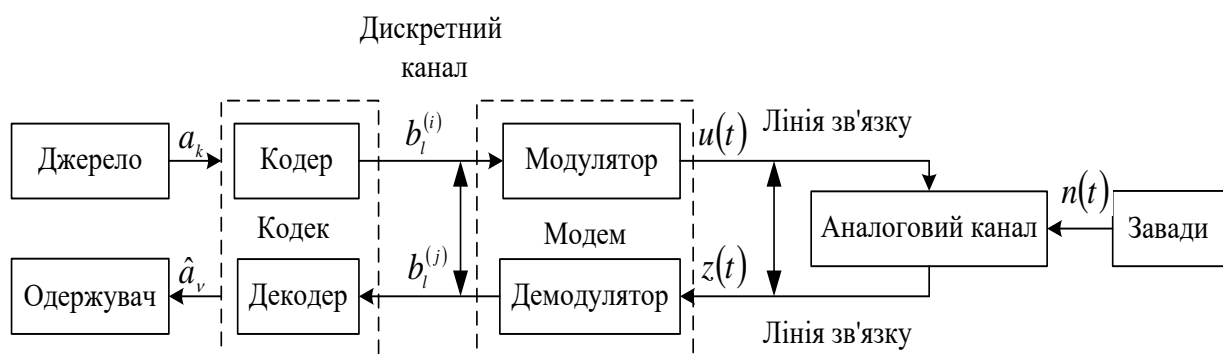


Рис. 1.8. Структурна схема системи передавання дискретних повідомлень

Передане повідомлення позначено a_k , кодоване повідомлення (чи первинний цифровий сигнал) – $b_u(t)$, його компоненти – $b_l^{(i)}$ (l – номер послідовно переданого символу, i – номер позиції коду, $i = \overline{0, m-1}$).

Сигнал, що надходить у лінію зв'язку, позначено $u(t)$, прийняте коливання – $z(t)$, відновлена послідовність кодових символів – $\hat{b}_u(t)$ (її компоненти – b_i) і декодоване (відновлене) повідомлення – \hat{a}_v .

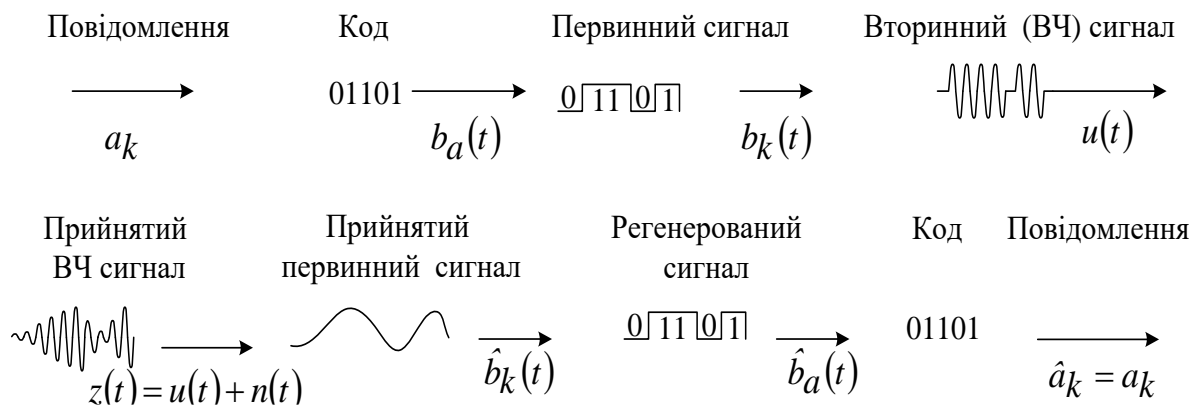


Рис. 1.9. Процес перетворення дискретного повідомлення на сигнал та зворотнє перетворення сигналу на повідомлення

Позначення прийнятих сигналів, кодових символів і відновленого повідомлення інші, ніж позначення переданих. Цим підкреслюється те, що через дію завад прийнятий сигнал відрізняється від переданого, а відновлене повідомлення може не збігатися з первинним.

У сучасних системах передавання дискретних повідомлень розрізняють дві групи самостійних пристроїв: кодеки і модем. Кодек – це пристрій, що перетворює повідомлення на коди (кодер) і код на повідомлення (декодер), а модем – пристрій, що перетворює код на сигнал (модулятор) і сигнал на код (демодулятор). Канальні пристрої (смугові підсилювачі передавача і приймача, коректори та ін.) разом з лінією зв’язку утворюють неперервний канал, а останній разом з модемом – дискретний канал.

Слід мати на увазі, що в системах радіозв’язку після передавача за допомогою передавальних антен утворюється просторово-часовий сигнал $u(t, r)$ (електромагнітна хвиля), що залежить не тільки від часу t , але й від просторових координат точки спостереження $r(x, y, z)$. Крім декартових, можна ввести й інші просторові координати, наприклад полярні.

Сигнал, який залежить від багатьох координат, називають електромагнітним полем. У місце прийому (на вихід антени) надходить поле (просторово-часовий сигнал) $z(t, r) = s(t, r) + n(t, r)$. Найчастіше воно спочатку за допомогою приймальної антени

перетворюється лише на часовий сигнал $z(t)$, який надалі підлягає лише часовій обробці.

Питання формування й обробки просторово-часових сигналів розглядаються в спеціальних курсах, а в цьому посібнику вважатимемо, що пристрої перетворення часовий сигнал - поле на передаванні і поле - часовий сигнал на прийманні знаходиться всередині заданої лінії зв'язку.

Для передавання дискретного повідомлення a (рис. 1.10) його спочатку перетворюють на дискретний первинний електричний сигнал $b_a(t)$, а потім, як правило, за допомогою модулятора формують канальний сигнал $u(t)$, який і посиляють у лінію зв'язку.

При передачі дискретних (цифрових) сигналів використовують амплітудну (АМн), частотну (ЧМн), фазову (ФМн) та відносно фазову маніпуляцію (ВФМн).

Прийняте коливання $z(t)$ піддається зворотним перетворенням, внаслідок яких виділяється первинний сигнал \hat{b} , за яким і відновлюється з тією чи іншою точністю повідомлення \hat{a} .

Загальний принцип модуляції полягає в зміні одного чи кількох параметрів несучого коливання (переносника) $f(t, \alpha, \beta, \dots)$ відповідно до переданого повідомлення. Так, якщо переносником є гармонічне коливання $f(t) = \cos(\omega_0 t + \varphi)$, то можна створити три види модуляції: амплітудну, частотну і фазову.

Якщо переносником є періодична послідовність імпульсів $f(t) = U_0 \sum_{l=-\infty}^{\infty} v(t - lT - t_0)$, то при заданій формі імпульсів $v(t)$ можна створити чотири основних види імпульсної модуляції: амплітудно-імпульсну, широтно-імпульсну, часово-імпульсну і частотно-імпульсну. Застосування радіоімпульсів дозволяє одержати ще два види модуляції за частотою і за фазою високочастотного заповнення.

При дискретній (цифровій) модуляції закодоване повідомлення у вигляді послідовності кодових символів $\{b_i\}$ перетворюється на послідовність елементів (посилок) сигналу

$\{u_i(t)\}$ шляхом впливу кодових символів на переносник $f(t)$. За допомогою модуляції один з параметрів переносника змінюється за знаком, який визначається кодом.

У разі безпосереднього передавання переносником може бути постійний струм, сила якого і напрямок змінюються. Зазвичай за переносник, як і в неперервній модуляції, використовується змінний струм (гармонічне коливання). У цьому випадку можна одержати амплітудну, частотну і фазову модуляції.

На рис. 1.10 наведені форми сигналу при двійковому коді для різних видів дискретної (або цифрової) модуляції (або маніпуляції). Тут проілюстрований вид модуляції, коли закон передаваного сигналу цифровий (первинний), а сигнал-переносник (несуче коливання) аналоговий (неперервний). Таке перетворення і називається дискретною модуляцією (або маніпуляцією). При амплітудній маніпуляції (АМн) символу 1 відповідає передавання несучого коливання протягом часу T (посилка), символу 0 – відсутність коливання (пауза).

При часовій маніпуляції (ЧМн) передавання несучого коливання з частотою f_1 відповідає символу 1, а передавання коливання з частотою f_0 - символу 0. При двійковій фазовій модуляції змінюється фаза несучого коливання на π у разі кожного переходу від 0 до 1 і від 1 до 0.

На практиці застосовують систему відносної фазової модуляції (ВФМн). На відміну від фазової, при відносній фазовій маніпуляції фазу сигналів відраховують не від деякого еталона, а від фази попереднього елемента сигналу. Наприклад, символ 0 передається відрізком синусоїди з початковою фазою попереднього елемента сигналу, а символ 1 – таким самим відрізком з початковою фазою, що відрізняється від початкової фази попереднього елемента сигналу на π .

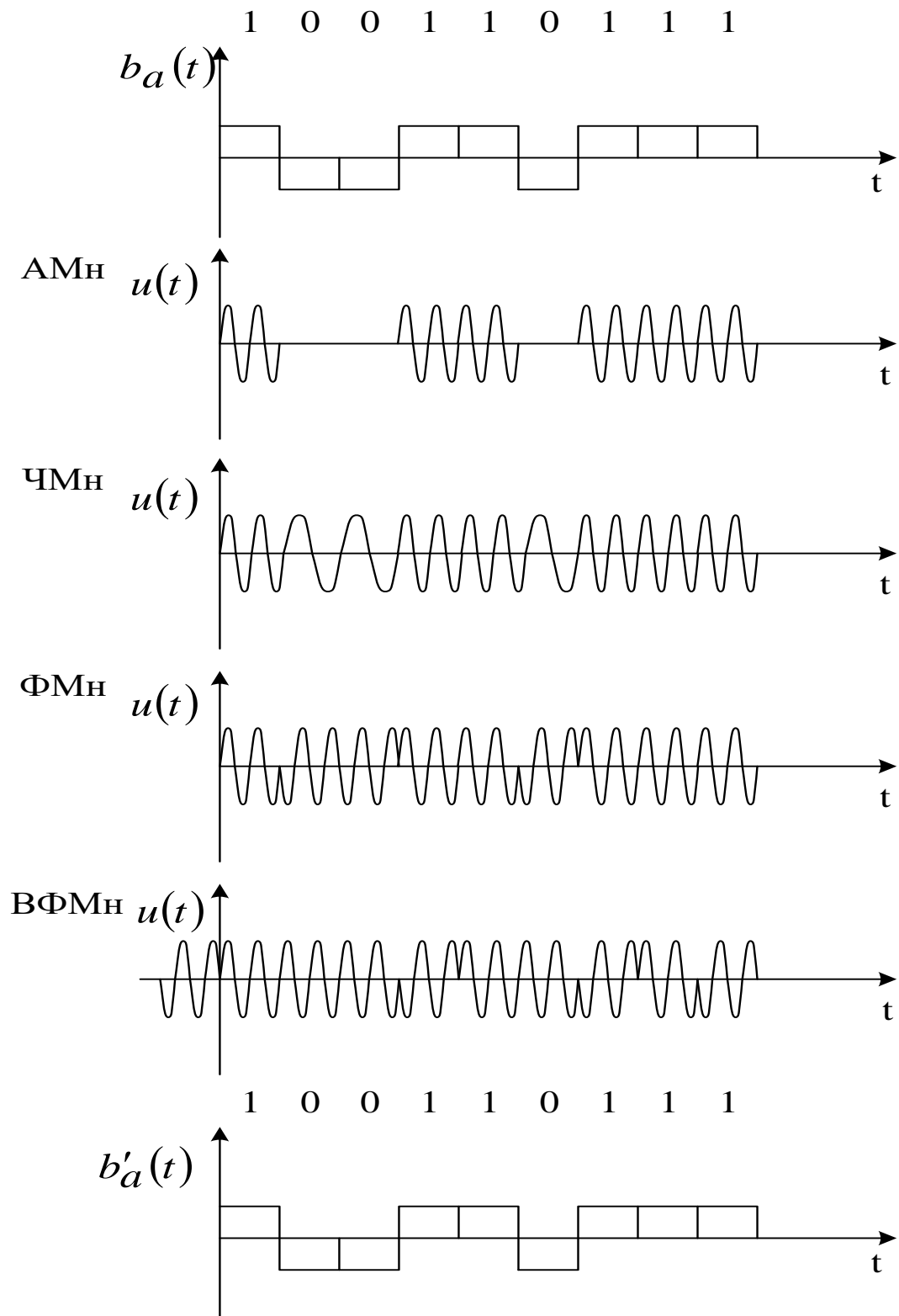


Рис. 1.10. Форми сигналів при двійковому коді для різних видів дискретної модуляції

У більш загальному випадку дискретну модуляцію доцільно розглядати як перетворення кодових символів $0, 1, \dots, m-1$ на відрізки сигналу $u_i(t)$, де $i=0, 1, \dots, m-1$ – переданий символ. При цьому вид сигналу $u_i(t)$, у принципі, може бути довільний. На практиці його вибирають так, щоб задовольнити вимоги, висунуті до системи зв'язку (зокрема за швидкістю передавання і за займаною смугою частот), і щоб сигнали добре розрізнялися з урахуванням впливу завад.

Тривалість послілки первинного сигналу $b_u(t)$ при дискретному передаванні визначає швидкість передавання посилок (технічну швидкість чи швидкість модуляції).

Ця швидкість v виражається числом посилок, переданих за одиницю часу, і вимірюється в Бодах. 1 Бод – це швидкість, з якою протягом 1 с передається одна послілка. Якщо тривалість послілки T виражена в секундах, то швидкість модуляції в Бодах $v=1/T$. Якщо смугу частот обмежити третьою гармонікою, то ширина спектра первинного сигналу $F=1,5v$ Гц.

1.5. Демодуляція та декодування

Передане повідомлення в приймачі звичайно відновлюється в такій послідовності. Спочатку сигнал демодулюється. У системах передачі безперервних повідомлень у результаті демодуляції відновлюється первинний сигнал, що відображає передане повідомлення. Цей сигнал потім надходить на відтворюючий або записуючий пристрій.

У радіосистемах таким пристроєм може бути гучномовець або магнітофон. У системах передачі дискретних повідомлень звичайно в результаті демодуляції послідовність елементів сигналу перетворюється на послідовність кодових символів. Потім за ними відновлюються повідомлення, які доводяться до одержувача. Останнє перетворення називається *декодуванням*.

Не слід вважати, що демодуляція і декодування – це прості операції, які є зворотними до модуляції і кодування. Через різні спотворення і вплив завад отриманий сигнал може істотно відрізнитися від переданого. Тому завжди можна висловити ряд припущень (гіпотез) щодо того, яке повідомлення передавалося.

Задачею приймального пристрою є прийняття рішення про те, яке з можливих повідомлень дійсно передавалося джерелом. Для цього прийнятий сигнал підлягає аналізу з урахуванням усіх відомостей про джерело (наприклад, про ймовірність, з якою джерело посилає те чи інше повідомлення), про застосований код і метод модуляції, а також про властивості каналу.

За результатами аналізу, як правило, можна визначити умовні (апостеріорні) ймовірності можливих гіпотез і на цій підставі прийняти рішення, яке й надходить до одержувача. Та частина приймального пристрою, що здійснює аналіз отриманого сигналу і приймає рішення про передане повідомлення, називається розв'язувальною схемою.

У системах передавання неперервних повідомлень при аналоговій модуляції розв'язувальна схема визначає за спотвореним канальним отриманим (вторинним) сигналом найбільш імовірний переданий первинний сигнал і відновлює його. Тут розв'язувальною схемою є демодулятор. У системах передавання дискретних повідомлень розв'язувальна схема найчастіше складається з двох частин: демодулятора і декодера.

Іноді в процесі передавання дискретних повідомлень операції демодуляції і декодування виконує один пристрій, який перетворює послідовність елементів отриманого сигналу відразу на послідовність символів (літер) повідомлення. Такий метод приймання називають спільною демодуляцією-декодуванням чи прийманням у цілому на відміну від поелементного прийому з двома частинами розв'язувальної схеми.

У першому випадку аналізується цілком відрізок сигналу, що відповідає кодовій комбінації, і на підставі того чи іншого критерію відновлюється переданий елемент повідомлення (літера).

У другому випадку спочатку аналізуються окремі елементи сигналу, що відповідають кодовим символам, а потім відновлена кодова комбінація декодується, тобто перетворюється на елемент (літеру) повідомлення.

У деяких випадках роль розв'язувальної схеми виконує повністю чи частково людина. Так, приймаючи телеграфні сигнали на слух, оператор вирішує, який сигнал (точка чи тире) був переданий. Він же здійснює й операцію декодування.

У приймачах дискретних повідомлень, призначених для запису інформації, всі зазначені операції виконуються автоматично.

У найпростішому випадку перша частина розв'язувальної схеми являє собою граничний пристрій - реле або тригер, що працюють за принципом «Так» чи «Ні». Якщо прийнятий елемент сигналу вищий за поріг (наприклад, за допустиму напругу), видається один символ коду (наприклад, 1), якщо нижчий - інший (наприклад, 0).

У деяких випадках застосовують розв'язувальні схеми з двома порогами. Якщо рівень сигналу знаходиться між двома порогами, рішення не приймається – замість сумнівного елемента сигналу видається спеціальний символ стирання, котрий полегшує можливість правильного декодування прийнятої кодової комбінації.

Для прийняття рішення про те, яке повідомлення передавалося, необхідно проаналізувати отриманий сигнал, піддавши його різним перетворенням, які називають обробкою сигналу. Однією з задач теорії зв'язку є пошук правил отриманої обробки сигналу, застосування яких дає найбільш імовірне рішення про передане повідомлення.

Ці правила залежать від властивостей каналу і методів передавання (кодування і модуляції). Іноді оптимальні правила обробки виявляються складними, і для спрощення апаратури використовують іншу, не оптимальну, обробку.

Якість прийому й обробки сигналів істотно залежить від точності синхронізації переданих і прийнятих сигналів. Розрізняють синхронізацію тактову (визначення границь одиничних елементів сигналу), циклову (правильний поділ кодових комбінацій), несучих частот тощо. Похибки синхронізації призводять до зниження вірогідності прийому повідомлень, а в ряді випадків – до неправильного прийому всього повідомлення чи його частини.

Найпростішим методом, що дозволяє на прийомі відокремити одну кодову комбінацію від іншої, є стартозастопний режим передавання, коли на початку і в кінці кожної комбінації передається спеціальний сигнал («старт» і «стоп»). Такий метод передавання належить до асинхронних, тому що передавання

будь-якої кодової комбінації може починатися в будь-який момент часу після закінчення переданої комбінації.

При синхронних методах передавання елементи сигналу передаються безперервно через однакові проміжки часу. Поділ кодових комбінацій здійснюється в цьому випадку за допомогою циклової синхронізації.

1.6. Загальні характеристики системи зв'язку

При оцінці роботи системи зв'язку необхідно насамперед урахувати, яку точність передачі повідомлення забезпечує система й з якою швидкістю передається інформація. Перше визначає якість передачі, друге - кількість. У реальній системі зв'язку якість передачі залежить від ступеня спотворення прийнятого повідомлення. Ці спотворення залежать від властивостей і технічного стану системи, а також від інтенсивності й характеру завад.

У правильно спроектованій і технічно справній системі зв'язку необоротні спотворення повідомлень обумовлені лише впливом завад, у цьому випадку якість передачі повністю визначається *завадостійкістю* системи. Під завадостійкістю звичайно розуміють здатність системи протистояти шкідливому впливу завад на передачу повідомлень. Оскільки дія завад проявляється в тім, що прийняте повідомлення відрізняється від переданого, то кількісно завадостійкість при заданій заваді можна характеризувати *ступенем відповідності прийнятого повідомлення переданому*. Назвемо цю величину загальним терміном — *правильність*.

Кількісну міру правильності доводиться вибирати по-різному, залежно від характеру повідомлення й вимог одержувача. Нехай повідомлення являє собою дискретну послідовність елементів з деякої кінцевої безлічі. Вплив завади на передачу такого повідомлення проявляється в тім, що замість фактично переданого елемента може бути прийнятий який-небудь інший. Така подія називається помилкою. Як кількісну міру правильності можна взяти ймовірність помилки p або будь-яку монотонну функцію цієї ймовірності.

При передачі безперервних повідомлень ступенем відповідності прийнятого повідомлення $B'(t)$ (з реалізацією $b'(t)$) переданому $B(t)$ (з реалізацією $b(t)$) може служити деяка величина, що являє собою "відстань" між $b(t)$ і $b'(t)$. Часто приймають критерій квадратичного відхилення поданий як [21]

$$\varepsilon^2(t) = \frac{1}{T} \int_0^T [\hat{b}(t) - b(t)]^2 dt. \quad (1.14)$$

Кількісну міру правильності можна також визначити як ймовірність того, що відхилення ε не перевершить деякої заданої величини ε_0 :

$$Q = P[|\varepsilon| \leq \varepsilon_0]. \quad (1.15)$$

Правильність передачі залежить від відношення середніх потужностей сигналу й завади (ВСЗ). При даній інтенсивності завади ймовірність помилки тим менше, чим сильніше відрізняються один від одного сигнали, що відповідають різним повідомленням. Тоді вираз (1.15) полягає в тім, щоб вибрати для передачі сигнали з більшим розходженням. Правильність передачі залежить і від способу прийому. Необхідно вибрати такий спосіб прийому, який щонайкраще реалізує розходження між сигналами при даному відношенні сигналу до завади. Істотним є розходження між прийомом в аналогових й дискретних системах передачі повідомлень.

В аналогових системах будь-який, навіть якнайменший, вплив на сигнал викликає спотворення модульованого параметра, завжди спричиняє внесення відповідної похибки в повідомлення. Тому абсолютно повне відновлення переданого повідомлення неможливе.

У дискретних системах помилка при передачі повідомлень виникає тільки тоді, коли сигнал зорієнтується неправильно, а це відбувається лише при спотвореннях, що перевищують деякий поріг.

У теорії завадостійкості, розробленої В.О. Котельниковим [29], показано, що при обраному критерії й заданій безлічі сигналів, прийнятих при адитивному білому гаусівському шумі (БГШ), існує гранична (потенційна) завадостійкість, що за жодного способу прийому не може бути перевершена. Прийомний пристрій, що реалізує потенційну завадостійкість, називається оптимальним (найкращим) за даним критерієм.

Поряд з правильністю найважливішим показником роботи системи зв'язку є швидкість передачі. У системах передачі дискретних повідомлень швидкість вимірюється числом переданих символів в одиницю часу ν (Бод). Кількість переданої інформації прийнято вимірювати в бітах (двійкових одиницях). Максимальна кількість інформації, яку можна передати двійковим символом, дорівнює 1 біт. При використанні не двійкових, а m -них символів максимальна кількість інформації, яку можна передати, дорівнює $\log m$ біт. Тому дискретне джерело може забезпечити максимальну швидкість видачі інформації (максимальну продуктивність), біт/с

$$R_i = \frac{\log_2 m}{T}, \quad (1.16)$$

де T - тривалість посилки;
 m — основа коду.

При $m = 2$, $R_i = 1/T$ біт/с швидкість передачі інформації R_i чисельно дорівнює технічній швидкості ν . При $m > 2$ можливо, що швидкість передачі інформації $R_i > \nu$.

Однак нерідко в дискретних системах зв'язку швидкість передачі інформації $R_i < \nu$. Це буває, коли не всі посилки використовуються для передачі інформації, наприклад якщо частина з них служить для синхронізації або виявлення й виправлення помилок (при використанні коригувального коду).

При аналізі систем зв'язку вводиться також характеристика середньої швидкості передачі інформації для безперервного каналу в одиницю часу із заданою правильністю. Існує максимально можлива (гранична) швидкість передачі, що

називається *пропускною спроможністю каналу* - C . Це поняття визначає потенційні можливості системи зв'язку, що використовує даний канал, а вираз має вигляд, біт/с

$$C' = \Delta F_k \log_2 \left(1 + \frac{P_c}{P_3} \right),$$

де P_c - потужність сигналу;

P_3 - потужність завади.

У реальній системі середня швидкість передачі інформації завжди менше пропускної спроможності каналу. У теорії інформації К. Шенноном [50] доведена теорема, відповідно до якої для джерела без надмірності при $R_i < C'$ можна знайти такий спосіб кодування-декодування, при якому можлива передача повідомлень по каналі з завадами з якнайменшою помилкою.

Універсальним показником ефективності системи зв'язку є коефіцієнт η , що характеризує використання системою пропускної спроможності каналу $\eta = R_i/C'$ (інформаційна ефективність).

Своєчасність передачі повідомлень визначається припустимою затримкою, обумовленою перетворенням повідомлень і сигналів, а також кінцевим часом поширення сигналу по каналу зв'язку. Вона залежить, по-перше, від характеру й довжини каналу, по-друге, від тривалості обробки сигналу в передавальних і прийомних пристроях.

Швидкість передачі й затримка є незалежними характеристиками, практично не зв'язаними один з одним.

Існують і багато інших параметрів, що характеризують із різних точок зору якість системи зв'язку. До них, зокрема, належать надійність системи, габаритні розміри й маса апаратури, вартість устаткування, експлуатаційні витрати й т. п. Ці характеристики в курсі "Теорія електричного зв'язку" не розглядаються.

Висновки до глави 1

1. Передача повідомлень по каналах зв'язку здійснюється за допомогою сигналів, які є матеріальними носіями повідомлень, що відображають ту або іншу інформацію. Характерною рисою повідомлень (сигналів) є їхня непередбачуваність. Про будь-яке повідомлення можна говорити лише як про можливий з деякою ймовірністю події. Повідомлення про відому подію інформації не несе. Процес передачі повідомлень завжди є імовірнісним (стохастичним).

2. Повідомлення й відповідні їм сигнали можуть бути дискретними й безперервними. Безперервний каналний сигнал формується за допомогою модуляції, а дискретний - за допомогою кодування й модуляції. Кодування визначає закон побудови сигналу, а модуляція - вид сигналу, що передається по каналу.

3. Безперервне повідомлення (сигнал) у цифрових системах зв'язку перетворюється на цифровий сигнал за допомогою трьох операцій; дискретизації за часом, квантування за рівнем і кодування.

4. Основними пристроями системи передачі дискретних повідомлень є кодек і модем. Канальні пристрої разом з лінією зв'язку утворюють безперервний канал, а разом з модемом - дискретний канал.

5. Причинами появи помилок при передачі повідомлень по каналу зв'язку є спотворення, внесені самим каналом, і завади, що впливають на сигнал. Завади можуть бути адитивними й мультиплікативними. Серед адитивних завад найпоширенішими є флуктуаційні, зосереджені за спектром та імпульсні.

6. Найважливішими характеристиками системи зв'язку є правильність і швидкість передачі повідомлень. Перша визначає якість передачі, а друга - кількість. При передачі дискретних повідомлень правильність (завадостійкість) визначається ймовірністю помилки, а при передачі безперервних повідомлень - середньоквадратичною помилкою. Швидкість вимірюється найчастіше числом переданих двійкових одиниць інформації в одиницю часу (біт/с).

7. Сучасна теорія зв'язку інтенсивно розвивається як статистична теорія, основи якої становлять теорія сигналів, теорія завадостійкості й теорія кодування.

Контрольні питання та задачі

1.1. Канал зв'язку зі смугою пропускання ΔF_k використовується для передачі сигналів протягом часу T_k . У каналі діє шумова гаусівська завада з рівномірною спектральною щільністю середньої потужності $G_z(\omega) = N_0$. Визначити граничну потужність сигналу, що може бути переданий по даному каналу, якщо фізичний об'єм каналу V_k відомий. Вихідні дані наведені в табл. 1.1.

Таблиця 1.1

Вихідні дані для обчислювання граничної потужності сигналу

Параметр	Варіант									
	1	2	3	4	5	6	7	8	9	10
ΔF_k , кГц	10	10	15	10	2	5	7	10	3	5
T_k , с	10	5	10	1	10	2	4	10	8	4
$G_z(\omega)$	10^{-4}	10^{-4}	10^{-5}	10^{-5}	10^{-5}	10^{-4}	10^{-5}	10^{-3}	10^{-4}	10^{-4}
V_k , дБ	10^5	10^4	10^6	10^6	10^5	10^4	10^6	10^4	10^5	10^5

1.2. Що називається інформацією, повідомленням, сигналом?

1.3. Як класифікуються повідомлення і сигнали?

1.4. Якими параметрами характеризуються повідомлення і сигнали?

1.5. Що таке лінія зв'язку? Що таке канал зв'язку?

1.6. Якими параметрами характеризується канал зв'язку?

1.7. Як класифікуються канали зв'язку?

1.8. За якими ознаками класифікуються завади в каналах зв'язку?

1.9. Які основні параметри мають адитивні завади?

1.10. Що таке модуляція і демодуляція?

1.11. Що таке кодування, декодування?

1.12. В чому відмінність простих кодів від завадостійких?

1.13. Задана таблиця трьох кодів (табл. 1.2).

Таблиця 1.2

Вихідні данні для визначення різних кодів

Код	Повідомлення							
	а	б	в	г	д	е	ж	з
1	000	001	010	011	100	101	110	111
2	0	1	00	01	10	11	110	111
3	00	01	100	101	1100	1101	1110	1111

Необхідно: 1) визначити вид коду; 2) визначити, які із цих кодів не вимагають розділових знаків і чому; 3) закодувати слова: зірка, вітер, база; 4) розшифрувати послідовності кодових комбінацій: а) коду 1: 111000001000010000; 001101100000; б) коду 3: 0100101001110; 01110110100; 10000111100.

1.14. Стартостопний телеграфний апарат передає одну букву сімома послілками; однією стартовою (20 мс), п'ятьма кодовими (20 мс кожна) і однією стоповою (30 мс). Визначити швидкість модуляції v , Бод, і максимальну швидкість передачі інформації R_i , біт/с.

1.15. Що розуміють під вирішальною схемою при прийомі дискретних повідомлень?

1.16. Яке число розрядів повинен мати рівномірний код, призначений для первинного кодування 32-буквеного алфавіту, при основі коду $m = 2, 8, 16, 32$?

1.17. Покажіть, як за рахунок кодування можна здійснити обмін смуги частот і відношення сигнал-завада в каналі при незмінному обсязі сигналу.

ГЛАВА 2

МАТЕМАТИЧНІ МОДЕЛІ ПОВІДОМЛЕНЬ, СИГНАЛІВ І ЗАВАД

2.1. Функціональні простори та їх базиси

Електричні сигнали складної форми за своєю фізичною природою не завжди подібні до звичних для нас представленням про вектори як про направлені відрізки. Часто на практиці використовують векторне подання сигналів (функцій). Річ у тому, що серед різних математичних прийомів, застосовуваних при дослідженні електричних кіл і сигналів, найчастіше використовується подання довільної функції у вигляді суми більш простих (“елементарних”) функцій, які можна подати векторами.

Такий підхід лежить в основі принципу незалежності дії (суперпозиції) при вивченні перетворень сигналів у лінійних електричних колах. Наочні геометричні подання, пов'язані з відображенням функцій як векторів простору сигналів, допомагають часто усвідомити фізичну сутність процесів формування, передачі у лінійних електричних колах і поділу сигналів, синтезу оптимальних сигналів і пристроїв обробки сигналів за наявності завад.

Розкладання сигналу складної форми на найпростіші складові подібне до розкладання звичайного вектора x тривимірного простору на його складові за координатним базисом одиничних ортогональних векторів i, j, k (рис. 2.1). Таке подання можна записати як

$$x = x_1i + x_2j + x_3k. \quad (2.1)$$

Складовими вектора x за базисом (i, j, k) будуть вектори x_1i, x_2j, x_3k . Коефіцієнти x_1, x_2, x_3 є проєкціями вектора x на координатні осі i, j, k і називаються координатами вектора x . Інакше кажучи, вектор x у тривимірному просторі повністю визначається сукупністю його координат $x = x_1, x_2, x_3$. Безперервні функції можуть бути відображені за допомогою деяких простих розривних функцій (дод. 1).

Щоб перейти до узагальнення поняття вектора тривимірного простору для випадку n -мірного простору, розглянемо приклад. Деяке наближене уявлення про функцію, яка обумовлює сигнал $x(t)$, можна скласти з її відображення послідовністю прямокутних імпульсів, які мають на інтервалах $i\Delta$ значення $x(i\Delta)$ (рис. 2.2). Якщо тепер умовно представити функцію $x(t)$ на інтервалі $(0; T)$ “вектором”, то для його визначення буде потрібно $n = T/\Delta$ координат $x_i = x(i\Delta)$. Це означає, що функцію $x(t)$ за аналогією з виразом (2.1) можна подати у вигляді суми [17]

$$x(t) = \sum_{i=0}^{n-1} x_i \psi_i(t), \quad (2.2)$$

де $\psi_i(t) = \psi(t - i\Delta)$ - елементарні базисні функції;

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

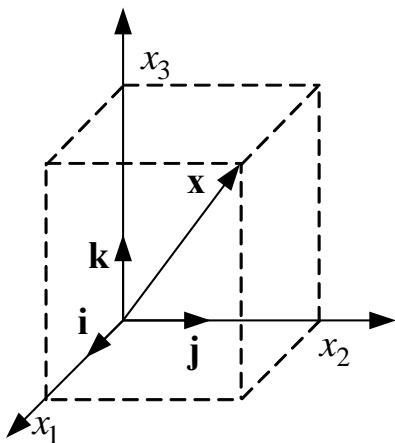


Рис. 2.1. Зображення вектора x в тривимірній ортогональній системі координат

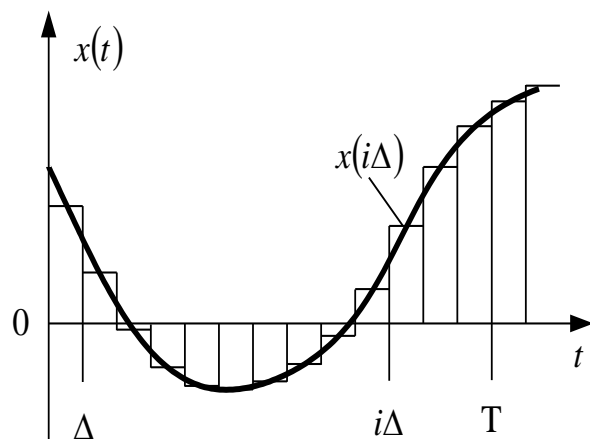


Рис. 2.2. Відображення неперервної функції послідовністю прямокутних імпульсів

Вектор x , що відповідає функції $x(t)$, у n -вимірному просторі одиничних ортів ψ_i буде повністю визначатися його координатами $x = x_0, x_1, \dots, x_{n-1}$. Таким чином, сигнал $x(t)$ довільної форми набуде вигляду суми n найпростіших елементарних сигналів, у даному випадку у вигляді імпульсів прямокутної форми. Слово простір використовується тут, щоб надати множині сигналів геометричного змісту і тим самим наочності. Найпростішим і в той же час фізично досить змістовним є трактування сигналів як елементів нормованого лінійного метричного простору.

Лінійним або векторним називається простір, для елементів якого виконуються правила додавання і множення на будь-яке число з деякої множини $\{\lambda\}$, яка називається множиною скалярів. Додавання векторів відбувається покоординатно, тобто сумою векторів x (функції $x(t)$) і y (функції $y(t)$) називається вектор $x+y = (x_0 + y_0, x_1 + y_1, \dots, x_n + y_n)$, що належить даному простору, а добуток λx вектора x на число λ дає вектор $\lambda x = (\lambda x_0, \lambda x_1, \dots, \lambda x_{n-1})$, що також належить даному простору. У лінійному просторі існує нульовий елемент 0 , такий, що $x + 0 = x$ і кожному елементові x відповідає протилежний елемент $-x$, отже, $x + (-x) = 0$. Вектор, утворений додаванням n лінійно незалежних (базисних) векторів ψ_i зі скалярними коефіцієнтами x_i , називається їхньою лінійною комбінацією $x = \sum_{i=0}^{n-1} x_i \psi_i$.

Множина векторів $\{\psi_i\}$ називається лінійно незалежним (базисом), якщо умова

$$\sum_{i=0}^{n-1} x_i \psi_i = 0 \quad (2.3)$$

виконується лише тоді, коли усі $x_i = 0$. Інакше кажучи, лінійно незалежною називається множина $\{\psi_i\}$, для якої жодна з її компонент не може бути утворена лінійною комбінацією інших. Розмірність лінійного простору визначається кількістю

будь-яких лінійно незалежних базисних векторів $\{\psi_i\}$, які утворюють цей простір. Лінійно незалежні вектори $\{\psi_i\}$ можна розглядати як координатні осі простору.

Метричним називається простір, у якому визначена відстань між елементами (векторами) простору (метрика), тобто кожній парі елементів, скажімо x і y , може бути поставлене деяке дійсне невід'ємне число $d(x, y)$ і спосіб, за яким знаходиться це число.

Відстань задовольняє такі правила:

1). $d(x, y) = 0$, якщо $x = y$;

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

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

де x, y, z - елементи (точки) простору.

Смисл перших двох умов очевидний. Третю умову називають нерівністю трикутника: довжина сторони трикутника менша (або дорівнює) за суму двох інших сторін.

Нормовані простори. Серед лінійних метричних просторів важливе місце займають нормовані простори. Цей вид простору визначається заданням норми $\|x\|$, яка задовольняє такі аксіоми:

1). $\|x\| \geq 0$;

2). $\|\lambda x\| = |\lambda| \cdot \|x\|$;

3). $\|x + y\| \leq \|x\| + \|y\|$.

Спочатку наведемо терміни і визначення, які стосуються n -мірного дійсного евклідового простору R_n . Будь-який вектор x у цьому просторі визначається сукупністю його координат $x = x_0, x_1, \dots, x_{n-1}$. Сукупність n лінійно незалежних векторів утворює n -мірний евклідовий простір, який позначається R_n . Простір R_n можна визначити як безліч точок, представлених кінцями векторів, для яких норма

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

Як бачимо, норма є узагальненням довжини вектора в двомірному і тривимірному просторі. Відстань між двома векторами x і y визначається як норма різниці векторів

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

Для простору Евкліда R_n можна ввести поняття скалярного добутку двох векторів x і y [30]

$$(x, y) = \sum_{i=0}^{n-1} x_i y_i = \|x\| \cdot \|y\| \cos \varphi, \quad (2.4)$$

де φ - кут між двома векторами.

Для проєкцій x на y і навпаки, y на x , маємо

$$\|x\| \cos \varphi = \frac{(x, y)}{\|y\|}; \quad \|y\| \cos \varphi = \frac{(x, y)}{\|x\|}.$$

Координати вектора являють собою проєкції вектора на координатні осі аналогічно до виразу (2.1). Зі співвідношення (2.4) випливає очевидна нерівність

$$\|(x, y)\| \leq \|x\| \cdot \|y\|, \quad (2.5)$$

відома в літературі як нерівність Буняковського-Шварца. Знак рівності має місце лише тоді, коли $y = kx$, де k – скаляр, тобто коли вектори x і y колінеарні. Для відповідних сигналів $x(t)$ і $y(t)$ це означає, що вони збігаються за формою $y(t) = kx(t)$. Квадрат вище визначеної норми вектора x можна знайти як скалярний добуток вектора самого на себе

$$\|x\|^2 = (x, x). \quad (2.6)$$

При $n \rightarrow \infty$ простір R_n переходить у нескінченновимірний простір Гілберта, що позначається L_2 . Гілбертовим простором є, зокрема, простір усіх безперервних комплексних функцій аргументу t , заданих в інтервалі $\left(-\frac{T}{2}; \frac{T}{2}\right)$, у якому скалярний добуток визначається співвідношенням [19]

$$(x, y) = \int_{-\frac{T}{2}}^{\frac{T}{2}} \dot{x}(t)x^*(t)dt, \quad (2.7)$$

а квадрат норми запишемо у вигляді

$$\|x\|^2 = \int_{-\frac{T}{2}}^{\frac{T}{2}} \dot{x}(t)x^*(t) = \int_{-\frac{T}{2}}^{\frac{T}{2}} |x(t)dt|^2 dt. \quad (2.8)$$

Норма (2.4) має не тільки геометричний, але і виразний фізичний зміст. Так, якщо сигнал $x(t)$ – речовинний електричний струм в одиничному опорі 1 Ом, то квадрат норми

$$\|x\|^2 = \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2 dt = E_x$$

визначає енергію сигналу. Елементи гілбертового простору L_2 характеризуються квадратом, що інтегрується, тобто якщо елементи цього простору – речовинні сигнали $x(t)$, визначені на інтервалі $\left(-\frac{T}{2}; \frac{T}{2}\right)$, то виконується умова

$$E_x = \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt < \infty. \quad (2.9)$$

Гілбертовий простір позначається при цьому $L_2(t)$. При $T \rightarrow \infty$ одержуємо простір $L_2(\infty)$. Для деяких сигналів (функцій) простору $L_2(\infty)$ умова (2.5) при $T \rightarrow \infty$ може не виконуватися, але виконується умова

$$P_x = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt < \infty. \quad (2.10)$$

У цьому випадку можна замість виразу (2.10) увести скалярний добуток з розмірністю потужності (для струмів і напруг на одиничному опорі)

$$(x + y)_P = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) y^*(t) dt. \quad (2.11)$$

Квадрат норми вектора x у цьому випадку запишеться як

$$\|x\|_P^2 = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \|\dot{x}(t)\|^2 dt = P_x. \quad (2.12)$$

При виконанні умови (2.10) у просторі $L_2(\infty)$ визначені співвідношення (2.11) та (2.12) при $T \rightarrow \infty$.

Надалі, говорячи про функції з квадратом, що інтегрується, у просторі $L_2(\infty)$, маємо на увазі виконання умови (2.9) або умови (2.10) при $T \rightarrow \infty$. Квадрат відстані між двома векторами в речовинному просторі $L_2(T)$ визначається співвідношенням

$$d^2(x, y) = \|x - y\|^2 = \int_{-\frac{T}{2}}^{\frac{T}{2}} (x(t) - y(t))^2 dt \quad (2.13)$$

або

$$d^2(x, y) = \|x - y\|_p^2 = \int_{-\frac{T}{2}}^{\frac{T}{2}} (x(t) - y(t))^2 dt. \quad (2.14)$$

Формулу (2.14) можна використовувати і при $T \rightarrow \infty$ для сигналів з кінцевою середньою потужністю.

Простір L_2 являє собою природне узагальнення простору R_n , одержуваного шляхом переходу від дискретизованої функції до функції безперервного аргументу. У курсі ТЕЗ простір L_2 має особливе значення, тому що він дозволяє застосувати загальні геометричні уявлення до повідомлень, сигналів і завад, визначених як функції безперервного аргументу. Враховуючи у виразі (2.2), що $n \rightarrow \infty$, одержуємо вираз безперервної функції $x(t)$ в просторі Гілберта

$$x(t) = \sum_{i=0}^{\infty} x_i \psi_i(t). \quad (2.15)$$

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

Більш загальним поняттям є поняття функціонала. Функціонал установлює відповідність між множиною чисел, з

одного боку, і деякою множиною функцій - з іншого. Можна сказати, що функціонал $\Phi [\bullet]$ встановлює залежність числа від функції $y = \Phi[f(x)]$. Прикладом функціонала є визначений інтеграл, величина якого (при незмінних межах) залежить від виду підінтегральної функції.

Дуже корисним є поняття функціонального оператора, який встановлює відповідність між двома множинами функцій, тобто за допомогою оператора L встановлюється така залежність: $y(t) = L[x(t)]$. Через те що функції можуть бути подані векторами і множина функцій визначається як векторний простір, дію оператора можна описати в геометричних термінах як перетворення простору X векторів x на простір Y векторів y . Зворотне перетворення Y на X позначають L^{-1} . У задачах перетворення повідомлень і сигналів використовуються поряд з лінійними операторами також нелінійні і параметричні оператори.

Якщо функція $x(t)$ на кожному інтервалі $i\Delta$ може приймати одне з m можливих значень $x_i^{(k)}$, де $k = \overline{0, m-1}$, то на відрізку тривалістю T вона буде цілком визначена $n = \frac{T}{\Delta}$ значеннями $x_i^{(k)}$ або, що те ж саме, сукупністю коефіцієнтів $\{x_0^{(k)}, x_1^{(l)}, \dots, x_{n-1}^{(v)}\}$ ($k, l, v \in \overline{0, m-1}$), яка називається n -набором. Зокрема, при $m = 2$ коефіцієнт $x_i^{(k)}$ приймає значення 0 або 1, n -набір це просто кодова комбінація n -значного двійкового ($m = 2$) коду, яка відображає символ (букву, цифру) переданого повідомлення. Двійкові n -набори відображаються векторами (точками) у просторі Хеммінга 2_n [24].

Скалярний добуток у цьому просторі зручно задати функцією

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

де знак \sum – сума в звичайному розумінні.

Звідси норма двійкового вектора

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

Можна бачити, що норма двійкового вектора визначається кількістю одиниць, що містяться в ньому. Цю норму називають також вагою вектора (кової комбінації) і позначають ω .

Відстань у просторі Хеммінга запишеться у вигляді

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

де знак \oplus означає операцію додавання по модулю 2 ($\text{mod } 2$): $0 \oplus 0 = 0$, $0 \oplus 1 = 1$, $1 \oplus 0 = 1$, $1 \oplus 1 = 0$.

Наведемо приклад додавання по $\text{mod } 2$ двох векторів:

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

Додавання і віднімання по $\text{mod } 2$ еквівалентні. У просторі Хеммінга відстань між двійковими векторами визначається за кількістю позицій у кодовій комбінації, у яких вектори x і y мають символи, що розрізняються [24]. У розглянутому прикладі $d(x, y) = 4$. У більш загальному випадку, якщо кількість розрізняваних значень дорівнює m , використовується різниця по модулю m .

2.2. Спектральний аналіз сигналів на основі рядів Фур'є

Важливим поняттям у просторах Евкліда, Гілберта і Хеммінга є ортогональність векторів. Два вектори x та y ортогональні, якщо $x \times y = 0$ [44].

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

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

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} \psi_i(t)\psi_j(t)dt = \begin{cases} 0, & j \neq i, \\ 1, & j = i. \end{cases} \quad (2.16)$$

Замість вираз (2.15) маємо

$$x(t) = \sum_{i=0}^{\infty} C_i \psi_i(t), \quad (2.17)$$

де C_i – коефіцієнти (координати) розкладання в ортонормованому базисі $\{\psi_i(t)\}$.

Представлення (2.17) називають узагальненим рядом Фур'є [19]. Для визначення коефіцієнтів C_i знайдемо скалярний добуток

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

Урахувавши вираз (2.16), отримаємо

$$C_i = (x, \psi_j) = \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)\psi_j(t)dt. \quad (2.18)$$

Таким чином, коефіцієнти C_i узагальненого ряду Фур'є є проєкціями вектора x на ортогональні осі (одиничні орти) ψ_i . Урахувавши вирази (2.8) і (2.17), можна одержати вираз

$$E_x = \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt = \int_{-\frac{T}{2}}^{\frac{T}{2}} \left[\sum_{i=0}^{\infty} C_i \psi_i(t) \right]^2 dt = \sum_{i=0}^{\infty} C_i^2, \quad (2.19)$$

який є окремим випадком рівності Парсеваля [42]. З урахуванням ортонормованого базису $\{\psi_i(t)\}$ легко бачити, що скалярний добуток і норму в просторі R_n можна знаходити за формулами (2.11) і (2.12).

Представимо тепер приблизно функцію $x(t)$ розкладанням у скорочений ряд за ортонормованими базисними функціями $\{\psi_i(t)\}$ у вигляді

$$x_\varepsilon(t) = \sum_{i=0}^{n-1} \gamma_i \psi_i(t) \quad (2.20)$$

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

$$\begin{aligned} \varepsilon^2(t) &= \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} [x(t) - x_\varepsilon(t)]^2 dt = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \left[x(t) - \sum_{i=0}^{n-1} \gamma_i \psi_i(t) \right]^2 dt = \\ &= \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt - \frac{2}{T} \sum_{i=0}^{n-1} \gamma_i \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \psi_i(t) dt + \sum_{i=0}^{n-1} \gamma_i^2 \frac{1}{T}. \end{aligned} \quad (2.21)$$

Урахувавши вираз (2.18), можна записати

$$\overline{\varepsilon^2(t)} = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt + \frac{1}{T} \sum_{i=0}^{n-1} (\gamma_i - C_i)^2 - \frac{1}{T} \sum_{i=0}^{n-1} C_i^2. \quad (2.22)$$

Похибка $\overline{\varepsilon^2(t)}$ має мінімальне значення, коли $\gamma_i = C_i$, тобто коли коефіцієнти розкладання в укороченому поданні (2.20) є коефіцієнтами узагальненого ряду Фур'є. Позначивши $\overline{\varepsilon^2(t)} = \Delta_n$, можна написати виходячи з виразу (2.22) умову

$$\Delta_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt - \frac{1}{T} \sum_{i=0}^{n-1} C_i^2 \geq 0, \quad \text{або} \quad \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt \geq \frac{1}{T} \sum_{i=0}^{n-1} C_i^2. \quad (2.23)$$

Нерівність (2.23) називають нерівністю Бесселя [30]. Коли зростає n , величина Δ_n зменшується. Якщо при $n \rightarrow \infty$ СКП прагне до нуля, то систему базисних функцій називають $\{\psi_i(t)\}$ повною. Маючи на увазі, що при $n \rightarrow \infty$ справедливий вираз (2.19), можна стверджувати, що в просторі Гілберта система базисних функцій $\{\psi_i(t)\}$ є повною. Ця система функцій є також замкнутою, тому що для будь-якої функції $x(t)$ з $L_2(T)$ нерівність (2.23) переходить при $n \rightarrow \infty$ в рівність.

2.2.1. Спектральне подання періодичних сигналів

При формуванні та обробці сигналів часто доводиться мати справу з періодичними коливаннями складної форми. Періодичну функцію $x(t) = x(t - nT)$, де T період повторення, можна представити розкладанням в узагальнений ряд Фур'є (2.17) за базисними функціями основної тригонометричної системи

$$\left\{ \begin{array}{l} 1, \quad \cos \omega_1 t, \quad \cos 2\omega_1 t, \quad \dots, \quad \cos k\omega_1 t, \quad \dots \\ \sin \omega_1 t, \quad \sin 2\omega_1 t, \quad \dots, \quad \sin k\omega_1 t, \quad \dots \end{array} \right\}. \quad (2.24)$$

Усі функції системи (2.24) попарно ортогональні на інтервалі $\left(-\frac{T}{2}; \frac{T}{2}\right)$. Узагальнений ряд Фур'є за базисними функціями (2.24) можна записати як

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\omega_1 t + b_k \sin k\omega_1 t); \quad (2.25)$$

$$a_k = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \cos(k\omega_1 t) dt; \quad b_k = \frac{2}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \sin(k\omega_1 t) dt. \quad (2.26)$$

Представлення (2.25) називають рядом Фур'є. Ряд (2.25) можна записати у вигляді

$$x(t)_k = \sum_{k=0}^{\infty} A_k \cos k\omega_1 t - \varphi_k, \quad (2.27)$$

де

$$A_0 = \frac{a_0}{2}; \quad A_k = \sqrt{a_k^2 + b_k^2}; \quad \varphi_k = \arctg \frac{b_k}{a_k}. \quad (2.28)$$

Відповідно до формули (2.27) періодичну функцію $x(t)$ можна представити сумою гармонічних коливань з частотами, кратними основній частоті $f_1 = 1/T$ з амплітудами A_k і початковими фазами φ_k . Сукупність амплітуд A_k ($k = 0, 1, 2, \dots$) утворює амплітудний спектр сигналу (рис. 2.3), а сукупність фаз φ_k ($k = 0, 1, 2, \dots$) – фазовий спектр сигналу.

Ряд Фур'є (2.27) часто подається в комплексній формі

$$x(t) = \sum_{k=-\infty}^{\infty} C_k e^{jk\omega_1 t}, \quad (2.29)$$

де C_k – комплексна амплітуда

$$C_k = \frac{a_k - jb_k}{2} = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) e^{-jk\omega_1 t} dt. \quad (2.30)$$

Варто звернути увагу на те, що сума у виразі (2.29) включає не тільки додатні значення k , але і від’ємні (з’являються “від’ємні частоти”).

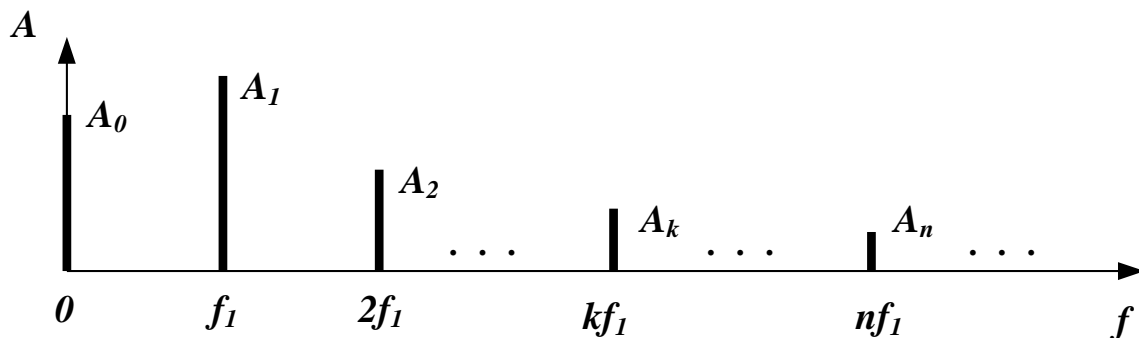


Рис. 2.3. Амплітудний спектр періодичного сигналу з періодом проходження $T = \frac{1}{f_1}$

Для переходу з виразу (2.27) у вираз (2.29) можна скористатися формулою Ейлера [17]

$$\cos(k\omega_1 t - \varphi_k) = \frac{e^{+j(k\omega_1 t - \varphi_k)} + e^{-j(k\omega_1 t - \varphi_k)}}{2}. \quad (2.31)$$

Вираз (2.31) можна інтерпретувати як представлення гармонічного сигналу одиничної амплітуди з позитивною частотою $k\omega_1$ у вигляді суми двох гармонічних коливань (половинної амплітуди) на додатній частоті $k\omega_1$ і від’ємній частоті $-k\omega_1$. Для речовинних функцій $x(t)$, як впливає з виразів (2.26) і (2.28), $a_k = a_{-k}$, $b_k = -b_{-k}$, $A_k = A_{-k}$, $\varphi_k = -\varphi_{-k}$ (амплітудний спектр – парна функція частоти, фазовий – непарна

функція частоти). Як наслідок $C_{-k} = \frac{a_{-k} - jb_{-k}}{2} = \frac{a_k - jb_k}{2} = \dot{C}_k$,

$$\frac{A_k}{2} = |\dot{C}_k| = |\dot{C}_{-k}|, \quad \arg \dot{C}_k = -\varphi_k = -\arg \dot{C}_{-k}. \quad \text{Комплексне}$$

представлення ряду Фур'є виявляється дуже зручним при виконанні різних розрахунків. Спектри деяких періодичних сигналів наведені у дод. 2.

2.2.2. Спектральне подання неперіодичних сигналів

Розкладання в тригонометричний ряд Фур'є (2.29) може бути узагальнене для випадку неперіодичних сигналів $x(t)$ при спрямуванні $T \rightarrow \infty$ або $f_1 = \frac{1}{T} \rightarrow \infty$. Для цього запишемо вираз (2.29) у такому вигляді:

$$x(t) = \sum_{k=-\infty}^{\infty} C_k \text{Te}^{jk\omega_1 t} \Delta f, \quad (2.32)$$

де Δf – частотне рознесення між лініями спектра періодичного сигналу, $\Delta f = f_1 = \frac{1}{T}$.

Розглянемо поточну частоту спектра $\omega = k\omega_1$ і визначимо спектральну щільність (СЩ) за Фур'є для неперіодичного сигналу:

$$\dot{S}(f) = \lim_{\Delta f \rightarrow 0} \frac{C}{\Delta f} = \lim_{T \rightarrow \infty} \dot{C}_k T. \quad (2.33)$$

Тоді з виразу (2.32) при $\Delta f \rightarrow 0$ отримаємо вираз

$$x(t) = \int_{-\infty}^{\infty} \dot{S}(f) e^{j\omega t} df, \quad (2.34)$$

І тоді з виразів (2.30) і (2.33) формула для визначення СЩ буде мати такий вигляд:

$$\dot{S}(f) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt. \quad (2.35)$$

Згідно з виразом (2.34) неперіодична функція $x(t)$ має вигляд суми гармонійних компонентів $e^{j\omega t}$ (на додатних і від'ємних частотах) з нескінченно малими амплітудами $\dot{S}(f)df$. Модуль визначає $|\dot{S}(f)|$ неперервний спектр неперіодичного сигналу, а $\arg \dot{S}(f) = \varphi(f)$ – безперервний фазовий спектр неперіодичного сигналу. Спектр по Фур'є можна записати як

$$\arg \dot{S}(f) = A(f) - jB(f) = \dot{S}(f) \cdot e^{j\varphi(f)},$$

де $A(f)$ – парна функція частоти,

$$A(f) = \int_{-\infty}^{\infty} x(t) \cdot \cos dt; \quad (2.36,а)$$

$B(f)$ – непарна функція частоти,

$$B(f) = \int_{-\infty}^{\infty} x(t) \cdot \sin dt. \quad (2.36,б)$$

З виразів (2.36 а, б) видно, що для дійсних функцій $x(t)$ амплітудний спектр $|\dot{S}(f)| = \sqrt{A^2(f) - B^2(f)}$ є парною функцією частоти, фазовий спектр $\varphi(f) = -\arctg \frac{B(f)}{A(f)}$ – непарна функція частоти.

Дискретний (лінійчатий) спектр амплітуд \dot{S}_k періодичного сигналу $x(t) = x(t + nT)$ з урахуванням виразу (2.30) можна знайти за формулою

$$\dot{C}_k = \frac{1}{T} \dot{S}\left(\frac{k}{T}\right). \quad (2.37)$$

Перетворення Фур'є $x(t) \rightarrow \dot{S}(f)$ (пряме і зворотне) $\dot{S}(f) \rightarrow x(t)$ описуються, як видно з виразів (2.32) і (2.33), лінійним оператором. Тому для цих перетворень справедливий принцип суперпозиції (накладення): СЩ для сигналу $x(t) = \sum_{k=1}^N x_k(t)$ визначається сумою СЩ доданків $x_k(t)$. Варто підкреслити, що, строго кажучи, СЩ (2.33) існує для сигналів $x(t)$, які задовольняють умову абсолютної інтегрованості

$$\int_{-\infty}^{\infty} |x(t)| dt < \infty. \quad (2.38)$$

Проте можна визначити СЩ і для сигналів $x(t)$, що не задовольняють умову (2.38), якщо скористатися введеною вище узагальненою δ -функцією. Наприклад, нехай $x(t) = e^{j\omega_0 t}$, тоді СЩ за Фур'є такого сигналу за визначенням знаходиться з виразу

$$\dot{S}(f) = \int_{-\infty}^{\infty} e^{j\omega_0 t} e^{-j\omega t} dt = \dot{S}(f) \int_{-\infty}^{\infty} e^{-j(\omega - \omega_0)t} dt. \quad (2.39)$$

Скориставшись інтегральним визначенням δ -функції, з виразу (2.39) одержимо результат $\dot{S}(f) = \delta(f - f_0)$. Аналогічно можна показати, що СЩ для сигналу $x(t) = e^{-j\omega_0 t}$ дорівнює $\dot{S}(f) = \delta(f + f_0)$.

Як наслідок цього, СЩ для сигналу $\cos \omega_0 t = \frac{e^{j\omega_0 t} + e^{-j\omega_0 t}}{2}$ набуде вигляду $\dot{S}(f) = 0,5 \cdot \delta(f - f_0) + 0,5\delta(f + f_0)$, спектральна

щільність для сигналу $\sin \omega_0 t = \frac{e^{j\omega_0 t} - e^{-j\omega_0 t}}{2j}$ запишеться

формулою $\dot{S}(f) = 0,5 \cdot e^{-j\frac{\pi}{2}} \delta(f - f_0) + 0,5 \cdot e^{-j\frac{\pi}{2}} \delta(f + f_0)$.

Скалярний добуток функцій $x(t)$ і $y(t)$ (у загальному випадку комплексних) у просторі Гілберта $L_2(T)$ можна виразити і через їхній СЩ за Фур'є:

$$(x, y) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \dot{x}(t) \cdot y^*(t) dt = \int_{-\infty}^{\infty} \dot{S}_x(f) \cdot S_y^*(f) df. \quad (2.40)$$

Співвідношення (2.40) називають узагальненою формулою Релея (або рівністю Парсеваля).

Якщо у формулу (2.40) припустити, що $y(t) = x(t)$, то для дійсного сигналу $x(t)$ маємо

$$(x, x) = E = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x^2(t) dt = \int_{-\infty}^{\infty} \dot{S}_x(f) \cdot S_x^*(f) df = \int_{-\infty}^{\infty} S_x^2(f) df. \quad (2.41)$$

Характеристика $S_x^2(f) = W_x(f)$ має смисл СЩ енергії, і вираз (2.41) можна записати у такому вигляді:

$$(x, x) = E = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x^2(t) dt = \int_{-\infty}^{\infty} W_x(f) \cdot S_x^*(f) df. \quad (2.42)$$

Можна також ввести поняття спектральної щільності потужності (СЩП) сигналу тривалістю T :

$G_x(f) = \frac{1}{T} W_x(f) = \frac{S_x^2(f)}{T}$. Тоді вираз для потужності сигналу $x(t)$

з виразу (2.42) запишеться як

$$(x, x)_P = P = \frac{1}{T} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x^2(t) dt = \int_{-\infty}^{\infty} G_x(f) df. \quad (2.43)$$

Характеристики $W_x(f)$ і $G_x(f)$ відіграють важливу роль у перетворенні сигналів і їхніх спектрів. З їхнього визначення стає зрозумілим, що ці характеристики є парними функціями частоти. Тоді вираз для енергії і потужності сигналу запишемо як

$$E = 2 \int_0^{\infty} W_x(f) df = \int_0^{\infty} W_{0,x}(f) df; \quad (2.44)$$

$$P = 2 \int_0^{\infty} G_x(f) df = \int_0^{\infty} G_{0,x}(f) df, \quad (2.45)$$

де $W_{0,x}(f)$, $G_{0,x}(f)$ – відповідно СЩ енергії і потужності, визначені на частотах, $W_{0,x}(f) = 2W_x(f)$, $G_{0,x}(f) = 2G_x(f)$.

Співвідношення (2.43) корисно узагальнити. Визначимо скалярний добуток і запишемо у такому вигляді:

$$(x, y_\tau) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \dot{x}(t) \cdot y^*(t - \tau) dt. \quad (2.46)$$

Враховуючи, що спектр Фур'є для затриманого на час τ сигналу $y_\tau(t) = y(t - \tau)$ дорівнює $\dot{S}_y(f) \cdot e^{-j\omega\tau}$, а для сигналу $y^*(t - \tau)$ спектр Фур'є дорівнює $S_y^*(f) \cdot e^{j\omega\tau}$, одержуємо з виразу (2.46) співвідношення

$$(x, y_\tau) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \dot{x}(t) \cdot y^*(t - \tau) dt = \int_{-\infty}^{\infty} S_x(t) \cdot S_y^*(t) \cdot e^{j\omega\tau} df. \quad (2.47)$$

Якщо у формулі (2.47) припустити $y(t) = x(t)$ і ввести позначення $B_{E,x}(\tau)$ для функції кореляції (ФК) сигналу $x(t)$ з розмірністю енергії, то можна записати

$$\begin{aligned}
 B_{E,x}(\tau) &= \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \cdot y^*(t - \tau) dt = \int_{-\frac{T}{2}}^{\frac{T}{2}} W_x(f) \cdot e^{j\omega\tau} df = \\
 &= \int_0^{\infty} W_{0,x}(f) \cdot \cos \omega\tau df.
 \end{aligned} \tag{2.48}$$

Ввівши позначення $B_x(\tau) = (x, x_\tau)$ для ФК сигналу $x(t)$ з розмірністю потужності, одержуємо співвідношення

$$\begin{aligned}
 B_x(t) &= \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \cdot x^*(t - \tau) dt = \int_{-\infty}^{\infty} G_x(f) \cdot e^{j\omega\tau} df = \\
 &= \int_0^{\infty} G_{0,x}(f) \cdot \cos \omega\tau df
 \end{aligned} \tag{2.49}$$

і як наслідок перетворень Вінера-Хінчина отримаємо [20]

$$W_x(t) = \int_{-\infty}^{\infty} B_{E,x}(\tau) \cdot e^{-j\omega\tau} d\tau; \tag{2.50, а}$$

$$G_x(f) = \int_{-\infty}^{\infty} B_{E,x}(\tau) \cdot e^{-j\omega\tau} d\tau. \tag{2.50, б}$$

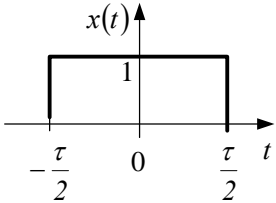
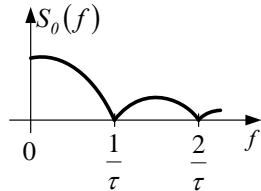
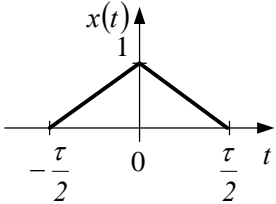
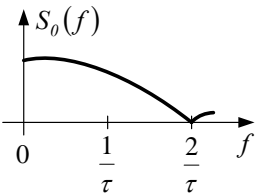
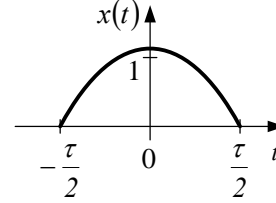
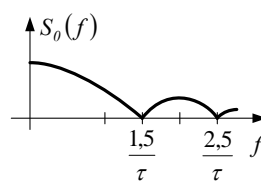
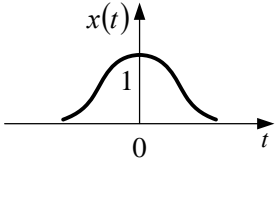
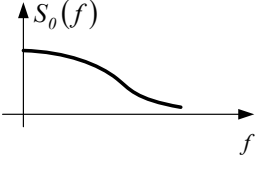
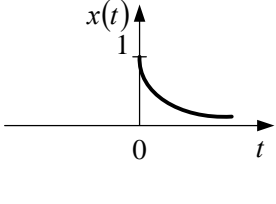
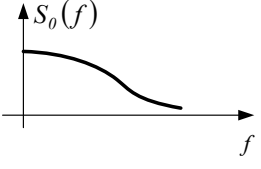
Таким чином, ФК $B_x(\tau)$ сигналу $x(t)$ і його СЩ потужності $G_x(f)$ (аналогічно до ФК сигналу $B_{E,x}(\tau)$ і його СЩ енергії) $W_x(f)$ утворять пари перетворень Фур'є.

У табл. 2.1 для наочності наведені приклади спектрів деяких імпульсів (неперіодичних функцій) і дані графіки їхніх амплітудних спектрів в області додатних частот. З наведених

прикладів видно, що імпульси обмеженої тривалості теоретично мають нескінченний спектр. Практично під шириною спектра будемо розуміти ефективну область частот ΔF , у межах якої сконцентровано 90...99 % енергії.

Таблиця 2.1

Спектри імпульсів різної форми при $\gamma = 0,9$

Сигнал $x(t)$	Спектральна щільність $S(f)$	Амплітудний спектр
$x(t) = \begin{cases} 1, & t \leq \frac{\tau}{2} \\ 0, & t > \frac{\tau}{2} \end{cases}$		$\tau \cdot \frac{\sin\left(\frac{\omega\tau}{2}\right)}{\left(\frac{\omega\tau}{2}\right)}$ 
$x(t) = \begin{cases} 1 - \frac{2 t }{\tau}, & t \leq \frac{\tau}{2} \\ 0, & t > \frac{\tau}{2} \end{cases}$		$\frac{\tau}{2} \cdot \frac{\sin^2\left(\frac{\omega\tau}{4}\right)}{\left(\frac{\omega\tau}{4}\right)^2}$ 
$x(t) = \begin{cases} \cos\omega_0 t, & t \leq \frac{\tau}{2} \\ 0, & t > \frac{\tau}{2} \end{cases}$		$\frac{4\tau}{\pi} \cdot \frac{\cos\left(\frac{\omega\tau}{2}\right)}{1 - \left(\frac{\omega\tau}{2}\right)^2}$ 
$\{x(t) = e^{-\beta^2 t^2}$		$\frac{\sqrt{\pi}}{\beta} e^{-\left(\frac{\omega}{2\beta}\right)^2}$ 
$\{x(t) = \begin{cases} e^{-\beta t}, & 1 \geq 0 \\ 0, & t < 0 \end{cases}$		$\frac{1}{\beta + j\omega}$ 

Для порівняння в табл. 2.2 наведені значення добутків $\Delta F \tau$ при $\gamma = 0,9$ для імпульсів з табл. 2.1.

Для гаусівського і експонентного імпульсів, що мають теоретично нескінченну тривалість, для зручності розрахунків також вводять поняття ефективної тривалості τ , розуміючи під цим інтервал часу, у межах якого зосереджена основна частка енергії сигналу. Якщо прийняти за основну частину всієї енергії сигналу $E_\gamma = \gamma E$ $\gamma = 0,9 \dots 0,99$, то ефективна ширина спектра й ефективна тривалість відповідно знаходяться з виразів

$$E_\gamma = \int_{-\frac{\tau_E}{2}}^{\frac{\tau_E}{2}} x^2(t) dt = \gamma \int_{-\infty}^{\infty} x^2(t) dt \text{ та } E_\gamma = \int_0^{F_E} W_{0,x}(f) df = \gamma \int_{-\infty}^{\infty} W_{0,x}(f) df .$$

Характерною рисою є те (табл. 2.1), що для всіх імпульсів (простих сигналів) виконується співвідношення $\Delta F \tau \approx 1$, тобто добуток $\Delta F \tau$ – величина порядку одиниці. Це співвідношення вказує на явний зв'язок між шириною спектра і тривалістю імпульсу: чим коротший імпульс, тим ширший його спектр.

Таблиця 2.2

Характеристики імпульсних сигналів

№ п/п	Імпульс	τ	ΔF	$\Delta F \tau$
1.	Прямокутний	$0,90^\tau$	$0,81^{1/\tau}$	0,73
2.	Трикутний	$0,541^\tau$	$0,84^{1/\tau}$	0,46
3.	Косинусоїдальний	$0,596^\tau$	$0,73^{1/\tau}$	0,43
4.	Гаусівський	$0,825^{1/\beta}$	$0,26^\beta$	0,22
5.	Експонентний	$1,155^{1/\beta}$	$0,98^\beta$	1,13

Для складних сигналів з базою $B_S = T_S \Delta F_S \gg 1$ виконується нерівність $\Delta F \tau \gg 1$. Спектральна щільність деяких неперіодичних сигналів наведена у дод. 2.

2.3. Ортогональні функції Радемахера і Уолша

В останні роки успішно розвиваються цифрові методи обробки і передачі сигналів на основі дискретних ортогональних послідовностей у вигляді функцій Радемахера, Уолша й ін. Функції Радемахера утворюються із синусоїдальних функцій за допомогою співвідношення [18]

$$r_k(\theta) \operatorname{sign}[\sin(2^k \pi \theta)], \quad 0 \leq \theta \leq 1,$$

де θ – безрозмірний час, $\theta = t/T$;

T – період функції;

$k = 0, 1, 2, \dots$ – порядок функції (ціле число);

$\operatorname{sign}(x)$ – знак дійсного числа, $(\operatorname{sign}(x) = +1$ при $x > 0$, $\operatorname{sign}(x) = -1$ при $x < 0$).

Інакше кажучи, функції Радемахера, які приймають значення ± 1 , можна трактувати як функції «прямокутного синуса». На рис. 2.4 наведені для прикладу графіки перших чотирьох функцій Радемахера $r_k(\theta)$ для $k = 0, 1, 2, 3$, звідки випливає, що функції $r_k(\theta)$ ортонормовані на інтервалі

$$0 \leq \theta \leq 1 \quad \int_0^1 r_i(\theta) \cdot r_k(\theta) d\theta = \begin{cases} 1, & i = k, \\ 0, & i \neq k. \end{cases}$$

Подальшим розвитком систем функцій, які мають форму «прямокутної хвилі», є система функцій Уолша $\{wal(m, \theta)\}$. Вона утворюється в такий спосіб. За визначенням вводиться функція $wal(0, \theta) = 1$ при $m = 0$. Для одержання функції при $wal(m, \theta)$ $m \geq 1$ досить записати число m в двійковій системі числення,

тобто подати у вигляді суми $m = 2^{\mu_1} + 2^{\mu_2} + \dots + 2^{\mu_p}$, де $\mu_1 < \mu_2 < \dots < \mu_p$ - додатні цілі числа. При цьому функція Уолша запишеться у вигляді $wal(m, \theta) = r_{\mu_1+1}(\theta)r_{\mu_2+1}(\theta)\dots r_{\mu_p+1}(\theta)$. На рис. 2.5 наведені графіки перших восьми функцій Уолша $wal(0, \theta)$, $wal(1, \theta)$, ..., $wal(7, \theta)$, побудованих за чотирма функціями Радемахера.

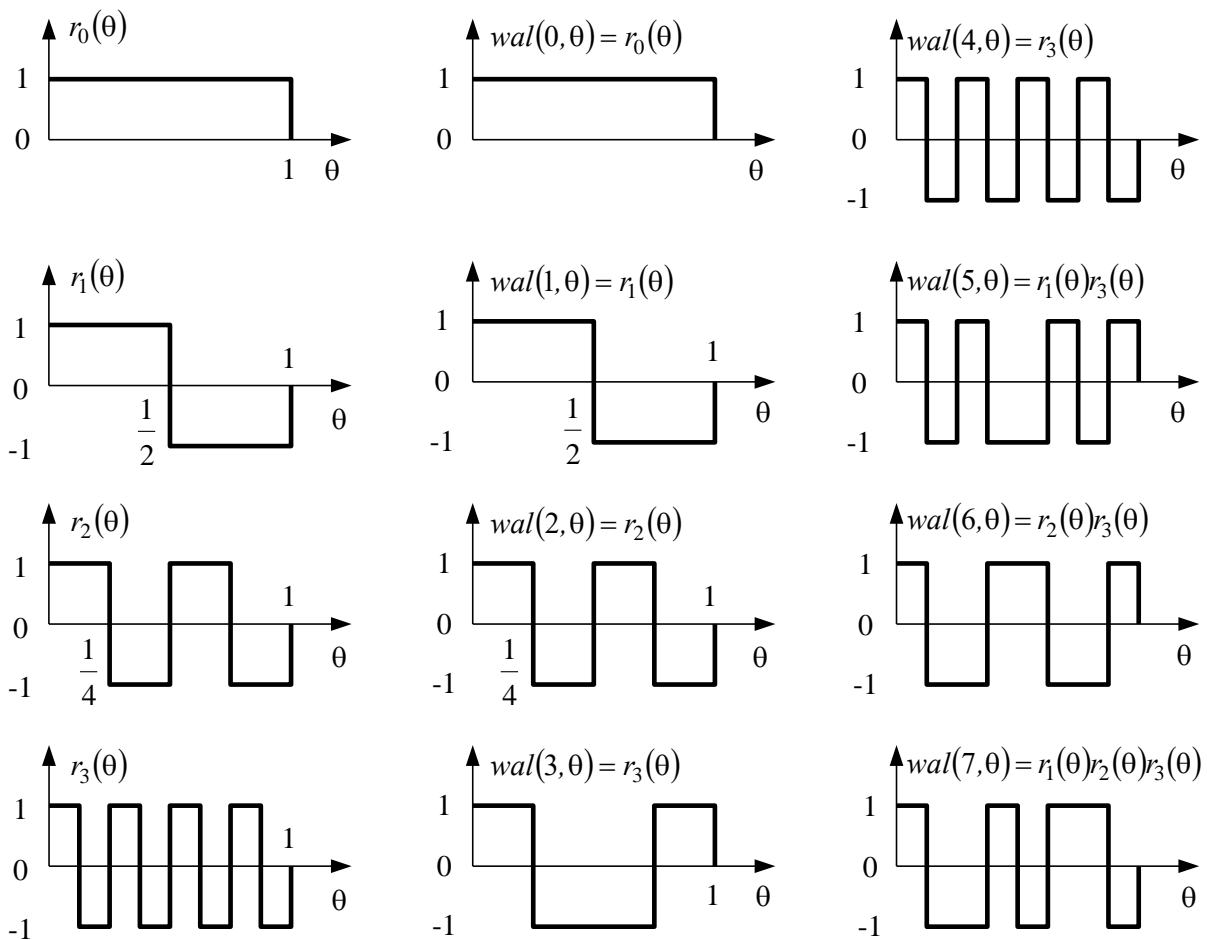


Рис. 2.4. Графіки функцій Радемахера 1, 2 і 3-го порядків

Рис. 2.5. Графіки перших восьми функцій Уолша

Функції Уолша не тільки ортогональні, вони мають ще і властивість мультиплікативності. Це означає, що добуток будь-яких двох функцій Уолша також є функцією Уолша

$wal(i, \theta)wal(k, \theta) = wal(p, \theta)$, де $p = i \oplus k$. У зв'язку з можливістю застосування до функцій Уолша логічних операцій вони знаходять застосування при розробленні пристроїв формування і перетворення сигналів на базі мікропроцесорної техніки. Сигнали на основі функцій Уолша використовуються в цифрових багатоканальних системах передачі інформації.

У теорії зв'язку, особливо в задачах апроксимації, знаходять застосування також й інші ансамблі ортогональних функцій. Серед них функції Лежандра, функції Лагерра, функції Ерміта [18].

2.4. Дискретизація в часі безперервних сигналів і їх відновлення

2.4.1. Дискретизація безперервних сигналів

При дискретному і цифровому представленні безперервних сигналів необхідно попередньо замінити ці сигнали їх дискретними відліками (вибірками). З цією метою в теорії і техніці сигналів широко використовується теорема В.О. Котельникова. Оpubлікована в 1933 р., вона визначає умови, за яких можлива заміна безперервного в часі сигналу його дискретними значеннями. Суть її зводиться до такого: якщо в спектрі досліджуваного сигналу $S(t)$, заданого на нескінченній осі часу, верхня частота в спектрі обмежена деякою максимальною частотою ($f_0 < f_m$), то такий сигнал повністю визначається послідовністю своїх значень (відліків) у моменти часу, що знаходяться один від одного на відстані не більш, ніж на

$$\Delta t_0 \leq \frac{1}{2f_m}, \text{ де } \Delta t_0 - \text{інтервал (період) дискретизації початкового}$$

сигналу. Причому відновити на приймальному боці початковий сигнал за його відліками можна з якою завгодно малою похибкою.

Тоді відповідно до цієї теореми, званої також теоремою відліків у часовій області, заданий сигнал $S(t)$ можна подати у вигляді ряду Котельникова [44]

$$S(t) = \sum_{n=-\infty}^{\infty} S\left(\frac{n}{2f_M}\right) \frac{\sin \omega_M(t - n/2f_M)}{\omega_M(t - n/2f_M)} = \sum_{n=-\infty}^{\infty} S(n\Delta t_\delta) \varphi_n(t), \quad (2.51)$$

де Δt_δ - часовий інтервал між двома сусідніми відліками (вибірками), $\Delta t_\delta = \frac{1}{2f_M} = \frac{\pi}{\omega_M}$;

S - вибірки (відліки) початкового безперервного сигналу $S(t)$ в моменти часу $(n\Delta t_\delta)$, $S(n/2f_M) = S(n\Delta t_\delta)$;

ω_M - кутова частота, $\omega_M = 2\pi f_M$;

$\varphi_n(t)$ - функції відліків початкового сигналу $S(t)$, які повинні бути ортогональними, $\varphi_n(t) = \frac{\sin \omega_M(t - n/2f_M)}{\omega_M(t - n/2f_M)} = \frac{\sin \omega_M(t - n\Delta t_\delta)}{\omega_M(t - n\Delta t_\delta)}$.

Представлення довільної безперервної функції (сигналу) $S(t)$ за допомогою ряду Котельникова графічно показано на рис. 2.6.

Функції відліків $\varphi_n(t) = \text{sinc}[\omega_M(t - n\Delta t_\delta)]$ як функції вигляду $\frac{\sin x}{x}$ має такі властивості:

1. У точках $t = (n\Delta t_\delta)$ на тимчасовій осі, де $n=0, \pm 1, \pm 2, \dots$ функції $\varphi_n(n\Delta t_\delta) = 1$, а в точках $t = k\Delta t_\delta$, де $k = \pm 1, \pm 2, \dots$, відмінне від n , функції $\varphi_n(k\Delta t_\delta) = 0$.

2. Спектральна щільність функцій $\varphi_n(t)$ рівномірно розподілена в смузі частот $|\omega| \leq \omega_M$, а модуль її дорівнює $|\omega(j\omega)| = 1/2f_M = \frac{\pi}{\omega_M} = \Delta t_\delta$. Це витікає з властивості симетрії (властивості взаємної оборотності частоти і часу для парних сигналів) як властивості перетворення Фур'є. На рис. 2.7 зображені часова функція $\varphi_0(t)$ і відповідна їй функція спектральної щільності за модулем $|\varphi_0(j\omega)|$.

Оскільки функція $\varphi_n(t)$ відрізняється від функції $\varphi_0(t)$ тільки зсувом на осі часу на $n\Delta t_\delta$, то спектральна щільність функції $\varphi_n(t)$ може бути записана в такому вигляді:

$$\varphi_n(j\omega) = \begin{cases} \frac{1}{2f_m} e^{-jn\Delta t_\delta \omega} = \Delta t_\delta e^{-jn\Delta t_\delta \omega}, & \text{при } -\omega_m \leq \omega \leq \omega_m, \\ 0, & \text{при } \omega < -\omega_m, \omega > \omega_m. \end{cases} \quad (2.52)$$

Ряд Котельникова (2.51) точно визначає заданий сигнал $S(t)$ тільки в точках відліку, оскільки коефіцієнтами цього ряду є самі вибірки (відліки) даного сигналу, тобто значення $S(n\Delta t_\delta)$. Отже, якщо відповідно до виразу (2.51) виконати всі операції з відліками і функціями відліків, то можна відновити початковий сигнал $S(t)$ із заданою точністю.

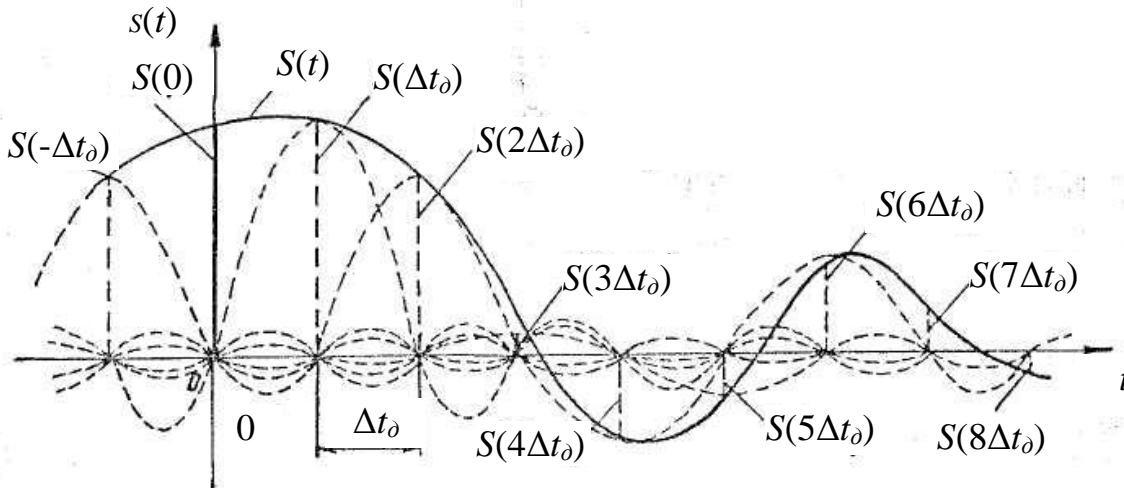


Рис. 2.6. Графічне подання довільного безперервного сигналу рядом Котельникова

Можна також показати, що ряд (2.51) визначає сигнал $S(t)$ у будь-який момент часу t , а не тільки в точках відліків $t = n\Delta t_\delta$.

Скористаємося для цього відомою методикою розкладання функції, що задовольняє умови Діріхле, за ортогональною системою. У даному випадку як ортогональні використовуються функції відліків $\varphi_n(t)$, для яких інтервал ортогональності прагне до нескінченності, а квадрат норми визначається відповідно до виразу

$$\begin{aligned} \|\varphi_n\|^2 &= \int_{-\infty}^{\infty} \varphi_n^2(t) dt = \int_{-\infty}^{\infty} \frac{\sin^2 \omega_M(t - n\Delta t_\delta)}{\omega_M^2(t - n\Delta t_\delta)^2} dt = \\ &= \frac{1}{\omega_M} \int_{-\infty}^{\infty} \frac{\sin^2 \omega_M(t - n\Delta t_\delta)}{\omega_M^2(t - n\Delta t_\delta)^2} d\omega_M t = \frac{\pi}{\omega_M} = \Delta t_\delta. \end{aligned} \quad (2.53)$$

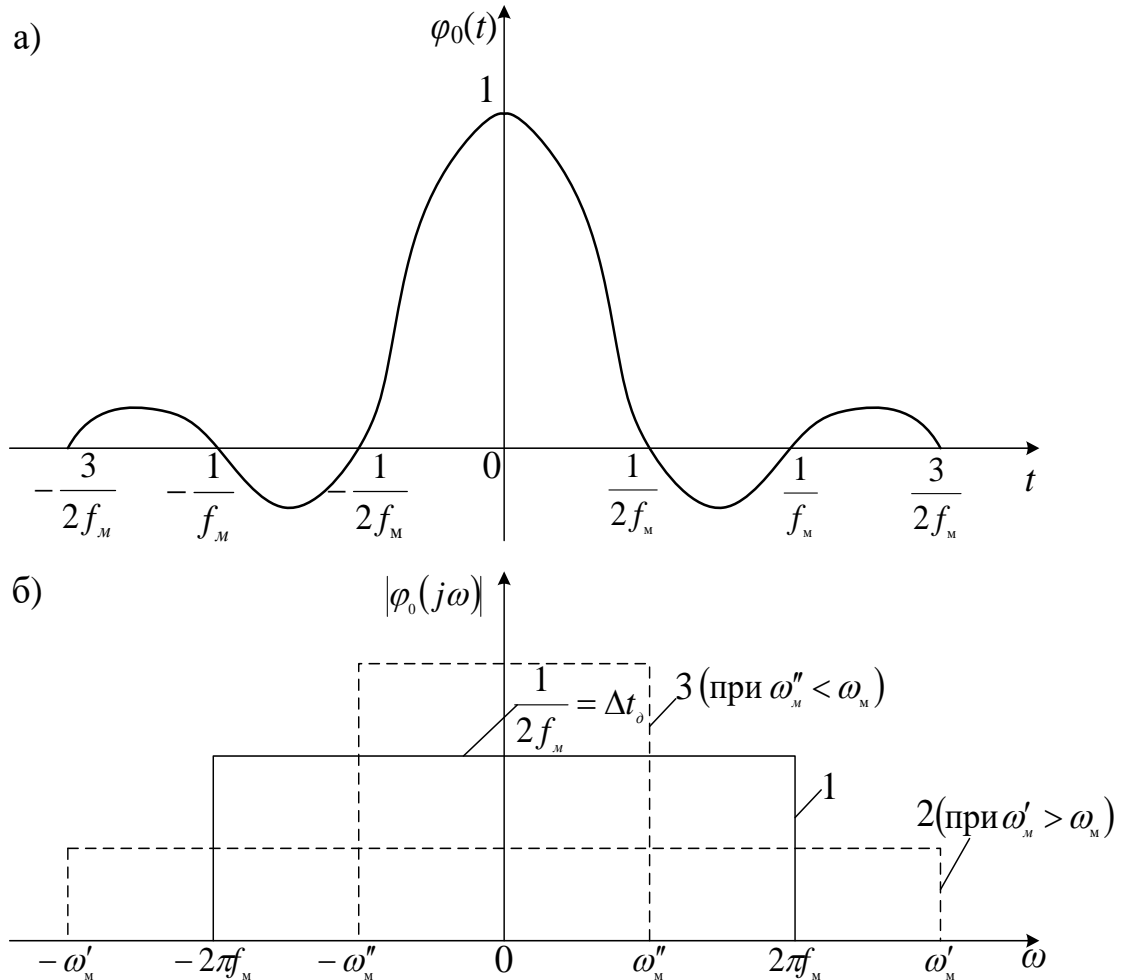


Рис. 2.7. Графічне зображення функцій: а) $\varphi_0(t)$; б) $|\varphi_0(j\omega)|$

Використовуючи відомий вираз для коефіцієнтів ряду Фур'є (2.30), можна записати

$$C_n = \frac{1}{\Delta t_\delta} \int_{-\infty}^{\infty} S(t) \varphi_n^*(t) dt, \quad (2.54)$$

де $\varphi_n^*(t)$ - комплексно зв'язана функція з функцією $\varphi_n(t)$.

Для обчислення виразу (2.54) скористаємося властивістю згортки перетворення Фур'є. Тоді отримаємо

$$\begin{aligned} \int_{-\infty}^{\infty} S(t)\varphi_n^*(t)dt &= \frac{1}{2\pi} \int_{-\omega_M}^{\omega_M} S(j\Omega)\varphi_n^*[j(\omega - \Omega)]d\omega = \\ &= \frac{1}{2\pi} \cdot \frac{1}{2f_M} \int_{-\infty}^{\infty} S(j\omega)e^{jn\omega\Delta t_\delta} d\omega. \end{aligned} \quad (2.55)$$

Межі інтеграції тут обмежені максимальною частотою $\omega_i = 2\pi f_i$ в спектрі сигналу $S(t)$, а також у спектрі функції відліків $\varphi_n(t)$ відповідно до рис. 2.7.

Інтеграл $\int_{-\infty}^{\infty} S(j\omega)e^{jn\omega\Delta t_\delta} d\omega$ у виразі (2.55) відповідно до зворотного перетворення Фур'є є значенням сигналу $S(t)$ у момент часу $t = n\Delta t_\delta$. Тоді маємо

$$\int_{-\infty}^{\infty} S(t)\varphi_n^*(t)dt = \frac{1}{2f_M} S(n\Delta t_\delta) = \Delta t_\delta S(n\Delta t_\delta). \quad (2.56)$$

Після підстановки виразу (2.55) в вираз (2.54) отримаємо

$$C_n = S(n\Delta t_\delta). \quad (2.57)$$

Звідси витікає, що коефіцієнтами ряду (2.51) є вибірки (відліки) функції сигналу $S(t)$ в моменти часу $t = n\Delta t_\delta$.

Зазначимо також, що оскільки обмеження спектра частот сигналу $S(t)$ максимальною частотою ω_M забезпечує безперервність функції $S(t)$, то ряд Котельникова (2.51) збігається до функції (сигналу) $S(t)$ при будь-якому значенні часу $t \rightarrow \infty$.

Якщо зменшити інтервал Δt_δ між сусідніми вибірками сигналу, то ширина спектра функції відліку $|\varphi_n(j\omega)|$ збільшиться

відповідно до властивості розтягування і стиснення сигналу на часовій осі і стає більше, ніж у початкового сигналу $S(t)$.

На рис. 2.7, б це вказано пунктирною лінією 2. Таке розширення смуги частот підвищує точність представлення сигналу $S(t)$, оскільки при цьому враховуються найбільш вищі частоти ($f_c > f_m$), крім того, знижуються вимоги до АЧХ фільтрів нижніх частот, поновлюючих безперервний сигнал з відліків у приймальному пристрої.

При збільшенні величини Δt_δ спектр функції $\varphi_n(t)$ звужується і стає менше за спектр початкового сигналу $S(t)$, що вказане пунктирною лінією 3 на рис. 2.7, б. Тоді при обчисленні інтеграла (2.32) межі інтегрування зменшуються і набуті значення коефіцієнтів у виразі (2.54) будуть вибірками не заданого сигналу, а іншого, спектр якого виявляється обмеженим меншою частотою, ніж ω_m . Тобто сигнал $S(t)$ відтворюється з деякою помилкою.

Якщо тривалість сигналу $S(t)$ кінцева (фінітний сигнал) і дорівнює T_c , то спектр його нескінченний і має пелюстковий характер.

Проте на практиці завжди можна визначити максимальну частоту спектра сигналу f_m так, щоб пелюстки спектра цього сигналу за межами частоти f_m містили нехтувана малу частку енергії в порівнянні з енергією сигналу $S(t)$.

При такому допущенні для сигналу тривалістю T_c і шириною спектра частот $\Delta f_c \approx f_m$ загальне число незалежних (ортогональних) параметрів, тобто відліків $S(n\Delta t_\delta)$, яке необхідне для повного завдання сигналу і відновлення його на приймальному боці, має дорівнювати

$$N = T_c / \Delta t_\delta = 2f_m T_c. \quad (2.58)$$

Тоді вираз (2.59) для ряду Котельникова матиме вигляд

$$S(t) = \sum_{n=0}^{2f_m T_c} S(n\Delta t_\delta) \frac{\sin \omega_m(t - n\Delta t_\delta)}{\omega_m(t - n\Delta t_\delta)}. \quad (2.59)$$

Число N у виразі (2.59) називають числом ступенів свободи сигналу $S(t)$, оскільки навіть при довільному виборі значень $S(n\Delta t_\partial)$ сума (2.59) повністю визначає функцію, що задовольняє умови заданого спектра і заданої тривалості сигналу $S(t)$. Число N іноді називають базою сигналу.

Енергію і середню потужність сигналу $S(t)$, представленого рядом Котельникова, можна виразити через послідовність його часових вибірок (відліків):

$$E_c = \sum_{n=0}^{2f_M T_c} |C_n|^2 \|\varphi_n\|^2 = \sum_{n=0}^{2f_M T_c} |S(n\Delta t_\partial)|^2 \|\varphi_n\|^2 = \Delta t_\partial \sum_{n=0}^{2f_M T_c} |S(n\Delta t_\partial)|^2; \quad (2.60)$$

$$P_{cp} = S^2(t) = \frac{E_c}{T_c} = \frac{\Delta t_\partial}{T_c} \sum_{n=0}^{2f_M T_c} |S(n\Delta t_\partial)|^2 = \frac{1}{2f_M T_c} \sum_{n=0}^{2f_M T_c} |S(n\Delta t_\partial)|^2, \quad (2.61)$$

причому середня потужність у виразі (2.61) визначається на опорі навантаження $R_H=1$ Ом.

2.4.2. Спектральне подання дискретизованих сигналів

На практиці часто потрібно подати сигнал $S(t)$ у вигляді частотних вибірок спектральної щільності $S(j\omega)$ сигналу, а не його часових вибірок $S(n\Delta t_\partial)$.

Тоді для функції спектра $S(j\omega)$ можна скласти ряд аналогічний до ряду (2.51), користуючись властивістю взаємної заміни частоти на час і навпаки для парних функцій. З цією метою базисна функція $\varphi_n(t) = \sin c[\omega_M(t - n\Delta t_\partial)]$ повинна бути замінена базисною функцією вигляду

$$\varphi_n(\omega) = \frac{\sin\left[\frac{T_c}{2}(\omega - n\Delta\omega)\right]}{\frac{T_c}{2}(\omega - n\Delta\omega)} = \text{sinc}\left[\frac{T_c}{2}\left(\omega - n\frac{2\pi}{T_c}\right)\right], \quad (2.62)$$

яка отримана з функції $\varphi_n(t)$ заміною t на ω , ω_m на $T_c/2$ і $\Delta t_\partial = 1/2 f_m$ на $\Delta\omega = 2\pi/T_c$.

Тоді вираз для ряду Котельникова в частотній області запишеться як

$$S(j\omega) = \sum_{n=-f_m T_c}^{f_m T_c} S(n\Delta\omega) \frac{\sin\left[\frac{T_c}{2}(\omega - n\Delta\omega)\right]}{\frac{T_c}{2}(\omega - n\Delta\omega)} =$$

$$= \sum_{n=-f_m T_c}^{f_m T_c} S\left(n\frac{2\pi}{T_c}\right) \operatorname{sinc} \frac{T_c}{2} \left(\omega - n\frac{2\pi}{T_c}\right). \quad (2.63)$$

Відповідно до виразу (2.63) на рис. 2.8 наведено приклад дискретизації спектра довільного сигналу за теоремою Котельникова, де $\Delta\omega$ – ширина частотної вибірки спектра дискретизованого сигналу.

Аналіз виразу (2.63) і рис. 2.8 показує, що якщо здійснюються часові вибірки відповідно до ряду Котельникова (2.51), то часовий інтервал між ними (інтервал дискретизації за часом) не повинен перевищувати значення $\Delta t_\partial \leq 1/2 f_m$, а при реалізації частотних вибірок частотний інтервал (інтервал дискретизації за частотою) не повинен перевищувати значення $\Delta\omega \leq 2\pi/T_c$.

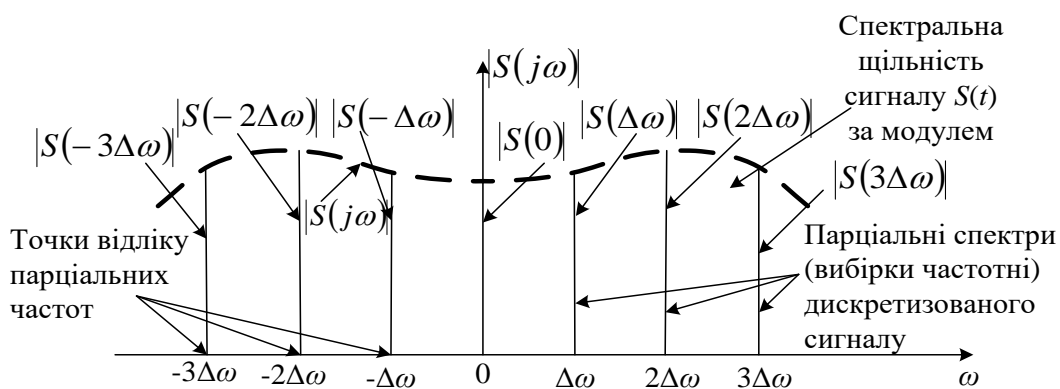


Рис. 2.8. Приклад дискретизованого спектра довільного сигналу за теоремою Котельникова

При ширині спектра сигналу, що дорівнює $\Delta\omega_c = 2\omega_m$, число частотних вибірок складе $N = 2\omega_m/\Delta\omega = 2f_m T_c$, тобто відповідає представленню сигналу $S(t)$ рядом Котельникова (2.51).

У загальному випадку вибірки $S(n2\pi/T_c)$ у виразі (2.63) є комплексними числами і в кожній точці відліку на осі частот повинні бути задані два параметри: дійсна і уявна частини цієї вибірки (модуль і аргумент). Таким чином, при частотному представленні сигналу $S(t)$ загальне число параметрів виявляється удвічі більшим, ніж при часовому представленні, коли вибірки $S(n/2f_m)$ є дійсними числами. Така надмірність у частотному представленні сигналу усувається за умови, що $S(n2\pi/T_c)$ і $S(-n2\pi/T_c)$ є комплексно-зв'язаними функціями, тому завдання однієї з них однозначно визначає іншу. Отже, спектр дискретизованого сигналу повністю характеризується сукупністю комплексних вибірок (відліків), узятих тільки в області позитивних частот, і числом незалежних параметрів (мір свободи) сигналу $N = 2f_m T_c$, як і при представленні сигналу в часовій області.

2.4.3. Особливості дискретизації сигналів

Процес дискретизації сигналів, здійснюваний на боці, що передає, по суті є одним з видів імпульсної модуляції сигналів. Він лежить в основі всіх імпульсних і цифрових видів зв'язку. Тому технічні пристрої, що забезпечують дискретизацію сигналів, можна назвати модуляторами або дискретизаторами, в якості яких часто використовуються різні електронні ключі. Процес дискретизації можна також розглядати як результат множення сигналу $S(t)$ на допоміжну періодичну послідовність $y_T(t)$ коротких тактових імпульсів прямокутної форми з періодом $T = \Delta t_\partial = 1/2f_m$, тривалістю $\tau_i \ll T$ і амплітудою $U_i \gg S_{max}(t)$. Отже, дискретизований сигнал з кроком (тактом) Δt_∂ можна записати за допомогою виразу

$$S_T(n\Delta t_\partial) = S(t)y_T(n\Delta t_\partial). \quad (2.64)$$

Структурна схема дискретизатора зображена на рис. 2.9.

Під час вступу чергового імпульсу з генератора тактових імпульсів (ГТІ) на електронний ключ останній відкривається і на виході дискретизатора з'являється відлік (вибірка) сигналу $S(t)$ дискретизатора, що діє на вході, в даний момент часу. За відсутності імпульсу на вході 2 від ГТІ електронний ключ закритий і на виході дискретизатора напруга дорівнює нулю. Процес часової дискретизації сигналу $S(t)$ зображений на рис. 2.10. Для визначення вимог до величини $\tau_i/T = \tau_i/\Delta t_\partial = 2\tau_i f_m$ розглянемо структуру спектра дискретизованого сигналу.

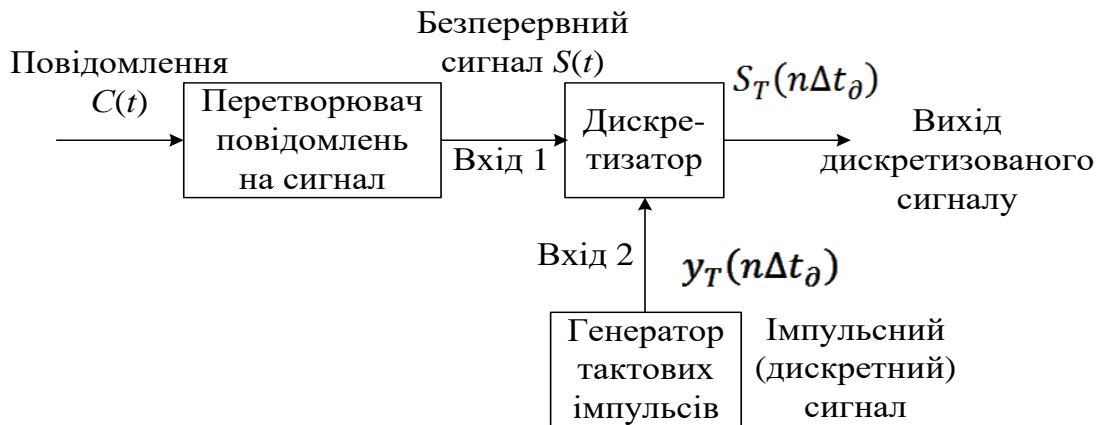


Рис. 2.9. Структурна схема дискретизатора

Вважаємо, що спектральна щільність $S(j\omega)$ початкового континуального сигналу $S(t)$ відома відповідно до прямого перетворення Фур'є

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

Запишемо вираз для спектра періодичної послідовності прямокутних імпульсів зі шпаруватістю $q = T/\tau_i = \Delta t_\partial/\tau_i$ відповідно до ряду Фур'є в такому вигляді:

$$y_T(n\Delta t_\delta) = U_i \left[\frac{\tau_i}{T} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin\left(\frac{n\omega_1\tau_i}{2}\right) \cos n\omega_1\tau_i \right] = \quad (2.65)$$

$$= U_i \frac{\tau_i}{T} \left[1 + 2 \sum_{n=1}^{\infty} \text{sinc}\left(\frac{n\pi\tau_i}{T}\right) \cos n\omega_1\tau_i \right],$$

де ω_1 - частота дискретизації сигналу $S(n\Delta t_\delta)$, $\omega_1 = \frac{2\pi}{T} = \omega_\delta$.

Підставивши вираз (2.65) у вираз (2.64), отримаємо

$$S_T(n\Delta t_\delta) = U_i \frac{\tau_i}{T} \left[S(t) + 2S(t) \sum_{n=1}^{\infty} \text{sinc}\left(\frac{n\pi\tau_i}{T}\right) \cos n\omega_1\tau_i \right]. \quad (2.66)$$

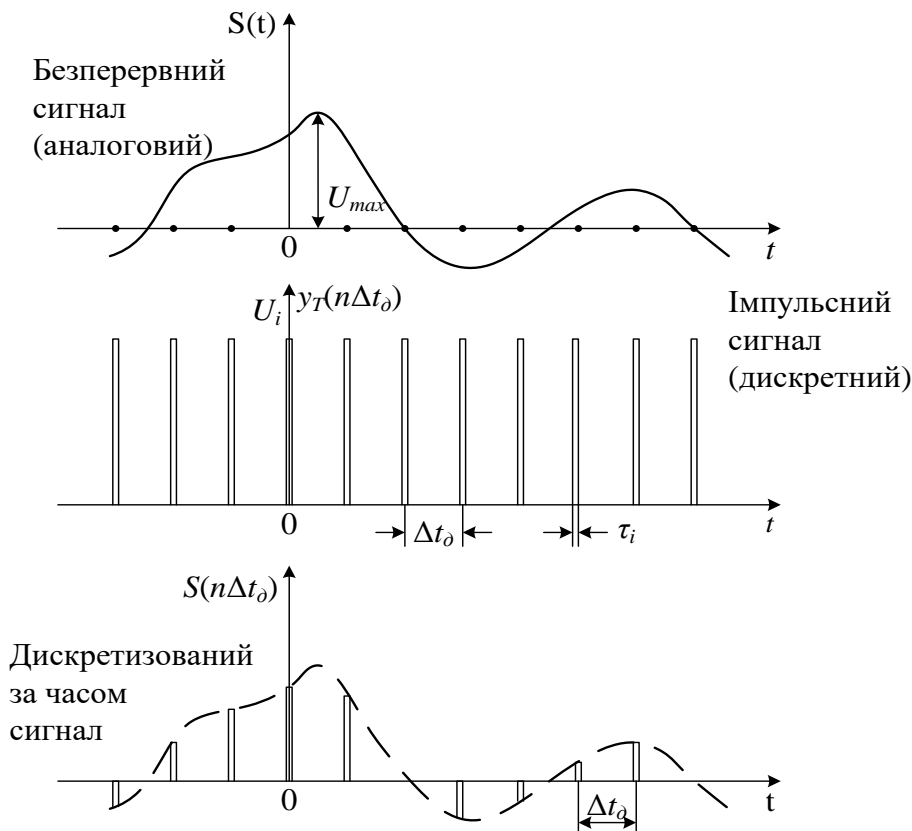


Рис. 2.10. Графіки, що пояснюють процес дискретизації безперервного сигналу $S(t)$

Аналіз виразу (2.66) показує, що першому його доданку відповідає спектральна щільність $S(j\omega)$ початкового сигналу, а кожному з добутоків $S(t)\cos n\omega_1 t$ з урахуванням теореми про зсув спектра сигналу (властивість модуляції) відповідає спектральна щільність з двох парціальних (окремих) складових $0,5\{S[j(\omega - n\omega_1)] + S[j(\omega + n\omega_1)]\}$. Таким чином, спектральна щільність дискретизованого сигналу $S(t)$ запишеться як

$$S_T(nj\Delta\omega) = U_i \frac{\tau_i}{T} \left\{ S(j\omega) + \sum_{n=1}^{\infty} \operatorname{sinc}\left(\frac{n\pi\tau_i}{T}\right) S[j(\omega - n\omega_1)] + \sum_{n=1}^{\infty} \operatorname{sinc}\left(\frac{n\pi\tau_i}{T}\right) S[j(\omega + n\omega_1)] \right\}. \quad (2.67)$$

З урахуванням того, що $\lim_{n \rightarrow 0} \operatorname{sinc}(0) = 1$, вираз (2.67) можна записати в компактній формі:

$$S_T(nj\Delta\omega) = U_i \frac{\tau_i}{T} \sum_{n=-\infty}^{\infty} \operatorname{sinc}\left(\frac{n\pi\tau_i}{T}\right) S[j(\omega - n\omega_1)]. \quad (2.68)$$

На рис. 2.11, а, б, в зображені відповідно спектральна щільність початкового сигналу $S(t)$, спектр періодичної послідовності прямокутної форми зі шпаруватістю $\Delta t_\delta / \tau_i = 4$ і спектральна щільність дискретизованого сигналу.

На підставі вищевикладеного можна зробити такі висновки:

1. Спектр дискретизованого сигналу $S(n\Delta t_\delta)$ складається з нескінченної послідовності парціальних спектральних щільностей (спектрів) $S(j\omega n)$ початкового сигналу $S(t)$, зрушених на частотні інтервали $\omega_1 = \omega_\delta = \frac{2\pi}{T} = \frac{2\pi}{\Delta t_\delta}$, модуль максимальної щільності яких убиває згідно із законом $\operatorname{sinc}\left(\frac{\pi n \tau_u}{T}\right)$.

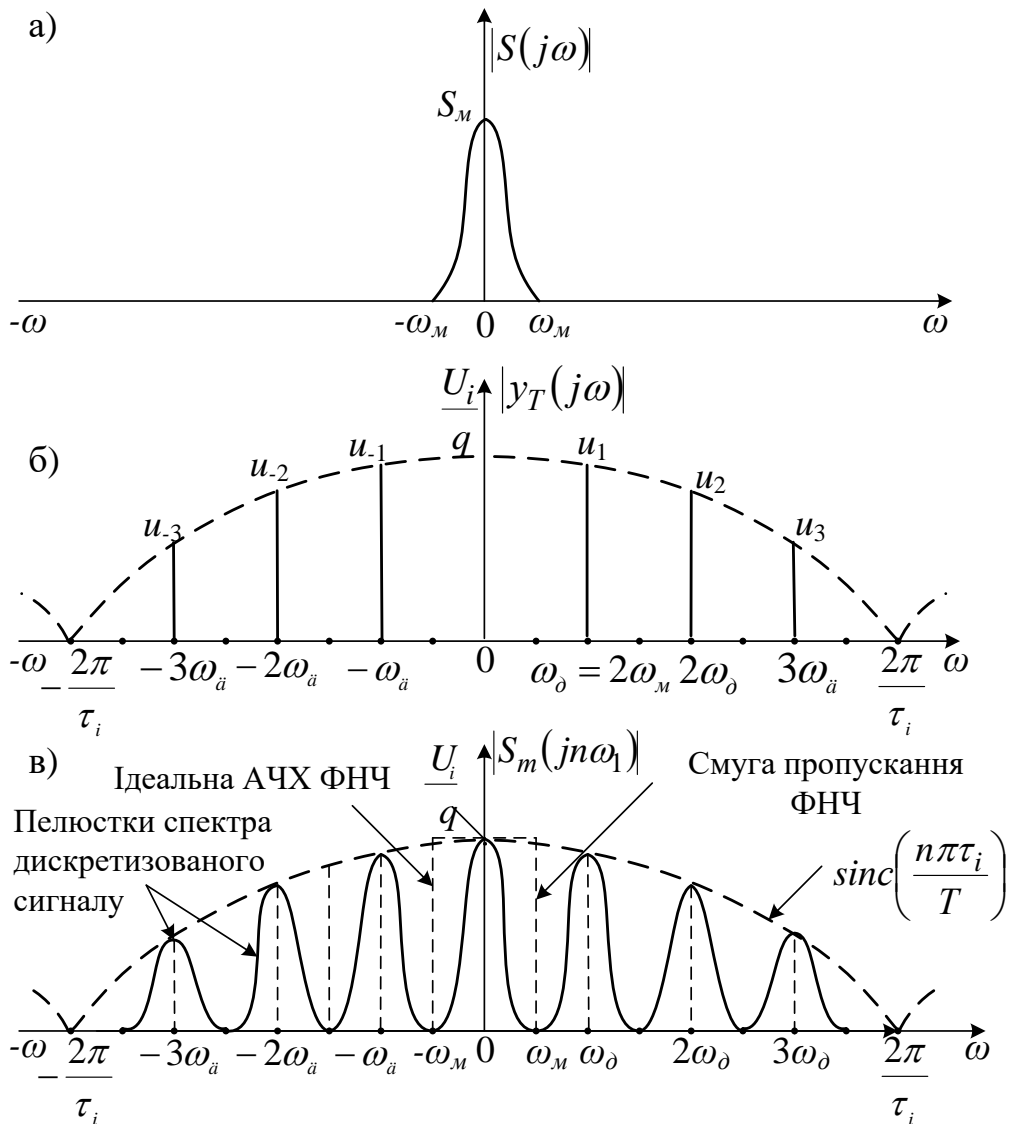


Рис. 2.11. Графіки спектрів: а) початкового сигналу; б) періодичної послідовності прямокутних імпульсів зі шпаруватістю $q=4$; в) дискретизованого сигналу

2. Якщо крок (інтервал) дискретизації відповідно до теореми відліків Котельникова вибрати з умови $T = \Delta t_\delta \leq 1/2f_m = \pi/\omega_m$, то окремі (парціальні) спектри не перекриваються і можуть бути виділені фільтром нижніх частот (ФНЧ). На практиці інтервал дискретизації Δt_δ зазвичай вибирають у декілька разів менше порівняно з $\Delta t_\delta = 1/2f_m$. Це необхідно для підвищення точності відтворення сигналу і полегшення реалізації ФНЧ на приймальному боці.

3. Із зменшенням відношення $\tau_i/\Delta t_\delta$ пелюстки спектру дискретизованого сигналу убувають швидше при $\tau_i/\Delta t_\delta \rightarrow 0$, тобто під час переходу сигналу $y_T(n\Delta t_\delta)$ до δ -функції спектр дискретизованого сигналу набуває строго періодичної структури, а ширина його пелюсток прагне до нуля. Таким чином, послідовність часових відліків набуває вигляд послідовності δ -функції з ваговими коефіцієнтами, що дорівнюють значеннями сигналу $S(t)$ в точках відліку $n\Delta t_\delta$, що ілюструється рис. 2.12.

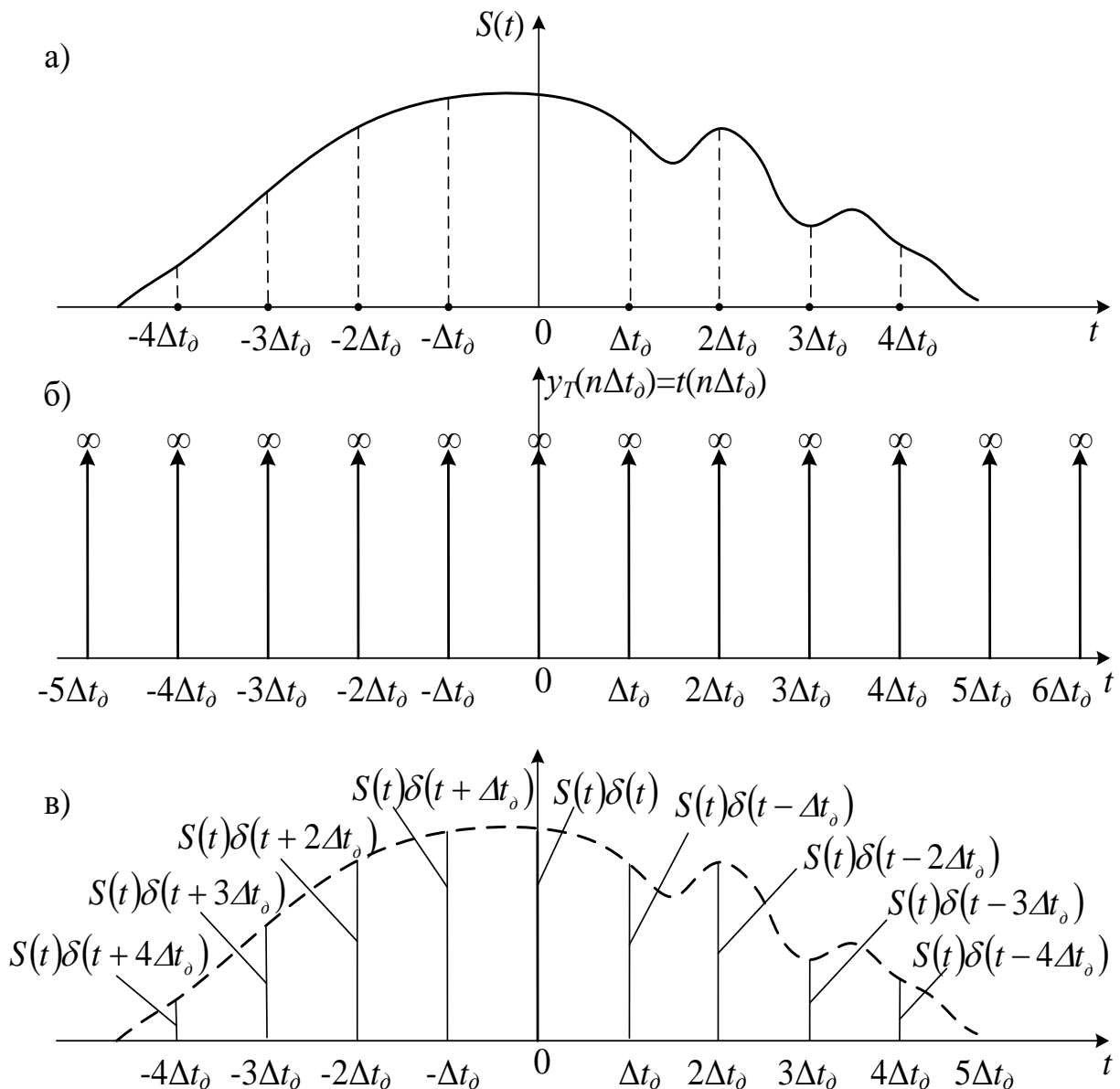


Рис. 2.12. Графіки часових функцій:
 а) початкового сигналу; б) послідовності δ -функції;
 в) відліків дискретизованого сигналу при $\tau_i/\Delta t_\delta \rightarrow 0$

Тоді вираз (2.65) запишеться у вигляді δ -функції

$$y_T(n\Delta t_\delta) = \sum_{n=-\infty}^{\infty} \delta(t - nT). \quad (2.69)$$

Підставивши вираз (2.69) у вираз (2.64), отримаємо

$$S_T(n\Delta t_\delta) = \sum_{n=-\infty}^{\infty} S(n\Delta t_\delta) \delta(t - n\Delta t_\delta). \quad (2.70)$$

Тоді спектральну щільність дискретизованого сигналу з урахуванням виразу (2.69) можна записати як

$$S_T(nj\Delta\omega) = \frac{1}{\Delta t_\delta} \sum_{n=-\infty}^{\infty} S[j(\omega - n\omega_1)]. \quad (2.71)$$

При цьому слід зазначити, що енергія дискретизованого сигналу, виражена через δ -функції, нескінченно велика. Отже, і енергія його спектра нескінченно велика. На практиці використовують тактові імпульси, що виробляються ГТІ, з кінцевою енергією. У цьому випадку спектр $S_T(nj\Delta\omega)$ при $\omega \rightarrow \infty$ убуває відповідно до виразу (2.68). Таким чином, представлення сигналу $S_\delta(n\Delta t_\delta)$ з урахуванням виразу (2.70) істотно спрощує спектральний аналіз дискретизованих сигналів. Наприклад, спектральну щільність $S_T(nj\Delta\omega)$ можна визначити безпосередньо за сукупністю відліків $S(n\Delta t_\delta)$ без звернення до спектра $S(j\omega)$ початкового сигналу $S(t)$. Якщо застосувати пряме перетворення Фур'є до виразу (2.71) при $n=0,1,2,\dots$, то отримаємо

$$\begin{aligned} S_T(j\omega) &= \int_{-\infty}^{\infty} S_T(t) e^{-j\omega t} dt = \int_{-\infty}^{\infty} e^{-j\omega t} \sum_{n=0}^{\infty} S_T(n\Delta t_\delta) \delta(t - n\Delta t_\delta) dt = \\ &= \sum_{n=0}^{\infty} S_T(n\Delta t_\delta) \int_{-\infty}^{\infty} e^{-j\omega t} \delta(t - n\Delta t_\delta) dt = \sum_{n=0}^{\infty} S_T(n\Delta t_\delta) e^{-jn\omega_1 t}. \end{aligned} \quad (2.72)$$

За своєю розмірністю функції $S(j\omega)$ і $S_T(nj\Delta\omega)$ відрізняються одна від одної. Перша з них (спектральна щільність) має розмірність $\frac{B}{Гц} = B \cdot c$, а друга має розмірність B , оскільки спектр дискретизованого сигналу дискретний (лінійчатий).

2.4.4. Відновлення безперервного сигналу

Відновлення безперервного сигналу з дискретизованого здійснюється на приймальному боці системи зв'язку. Відповідно до ряду Котельникова (2.51) для того, щоб отримати початковий сигнал $S(t)$, необхідно на приймальному боці мати формувач ортогональних функцій відліків $\varphi_n(t)$ на відліки, що поступають, перемножувач відповідного відліку на свою ортогональну функцію і виконати підсумок за всіма результатами цих перемножень. Такі операції здійснює фільтр нижніх частот, що має імпульсну характеристику, яка є відгуком ФНЧ на сигнал у вигляді δ -функції, зображену на рис. 2.12. Крім того, ФНЧ є інтегруючим елементом, тобто підсумовує добуток відліків на відповідні їм відгуки (імпульсні характеристики). Такі ФНЧ називають також згладжуючими або оцінюючими. Відновлену за дискретними вибірками функцію (сигнал) безперервного часу $S'(t)$ називатимемо оцінкою початкової функції (сигналу) $S(t)$.

Очевидно, що оцінка $S'(t)$ повинна бути якомога ближчою до початкової функції $S(t)$. У зв'язку з цим можна зробити висновок про те, що процес дискретизації сигналу $S(t)$ не можна розглядати окремо від процесу оцінювання. Необхідно прагнути задовольнити компромісні вимоги до точності оцінювання, простоти дискретизації і оцінювання сигналів, а також до простоти практичної реалізації дискретизованих сигналів, заданих найменшим можливим числом N відлікових точок. На рис. 2.13 зображена структурна схема відновлення безперервного сигналу з дискретизованого.

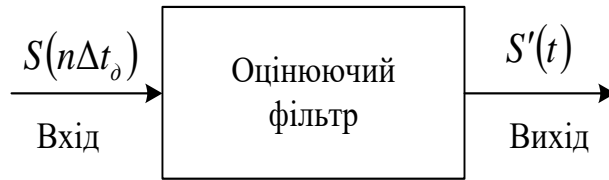


Рис. 2.13. Структурна схема відновлення безперервного сигналу з дискретизованого

Імпульсні характеристики (ІХ) простого RC фільтра нижніх частот зображено на рис. 2.14.

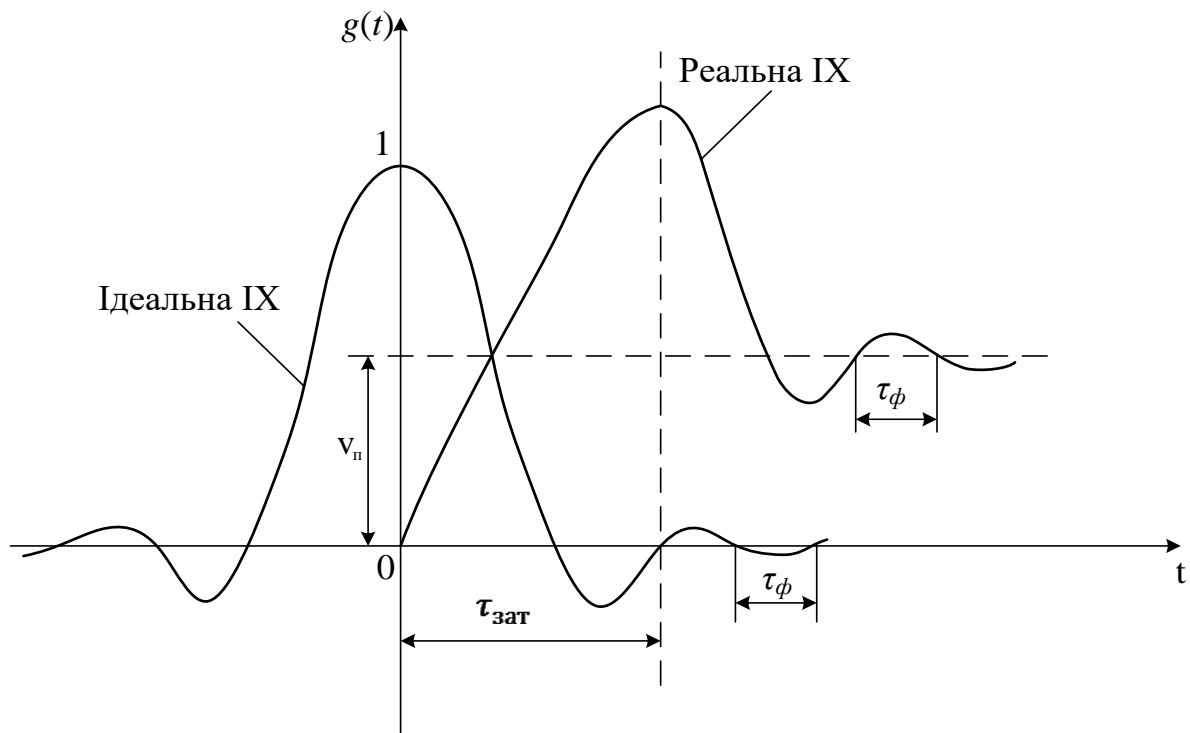


Рис. 2.14. Імпульсні характеристики ФНЧ

Ідеальна імпульсна характеристика ФНЧ знімається при подачі на його вхід сигналу у вигляді δ -функцій (періодичної послідовності), тому виконується умова нормування $\int_{-\infty}^{\infty} \delta(t) dt = 1$, а фільтр вважається ідеальним, тобто його відгук на вхідну дію з'являється на виході миттєво, без затримки. У реального ФНЧ відгук з'являється на виході із затримкою унаслідок інерційності

ланки (ланок) ФНЧ, що складається з реактивних елементів (L, C).

Крім того, реальні відліки, що поступають на ФНЧ, відрізняються від δ -функцій і умови нормування для них не виконуються. Тому реальна імпульсна характеристика ФНЧ буде піднята над віссю часу на величину площі реального відліку. Ця величина називається напругою п'єдесталу U_n .

За імпульсною характеристикою ФНЧ можна визначити його постійну часу $\tau_\phi = RC$ відповідно до рис. 2.14.

Для точного відновлення безперервного сигналу з дискретизованого необхідно виконати умову $\tau_\phi = \Delta t_\delta = 1/2f_m$. Тоді процес відновлення сигналу для ідеального ФНЧ можна пояснити за допомогою рис. 2.15.

Імпульсна характеристика ФНЧ пов'язана з його передавальною комплексною функцією зворотним перетворенням Фур'є

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K(j\omega) e^{j\omega t} d\omega, \quad (2.73)$$

а передавальна функція пов'язана з імпульсною характеристикою прямим перетворенням Фур'є

$$K(j\omega) = \int_{-\infty}^{\infty} g(t) e^{-j\omega t} dt. \quad (2.74)$$

На рис. 2.16 зображена ідеальна і реальна нормовані передавальні характеристики ФНЧ.

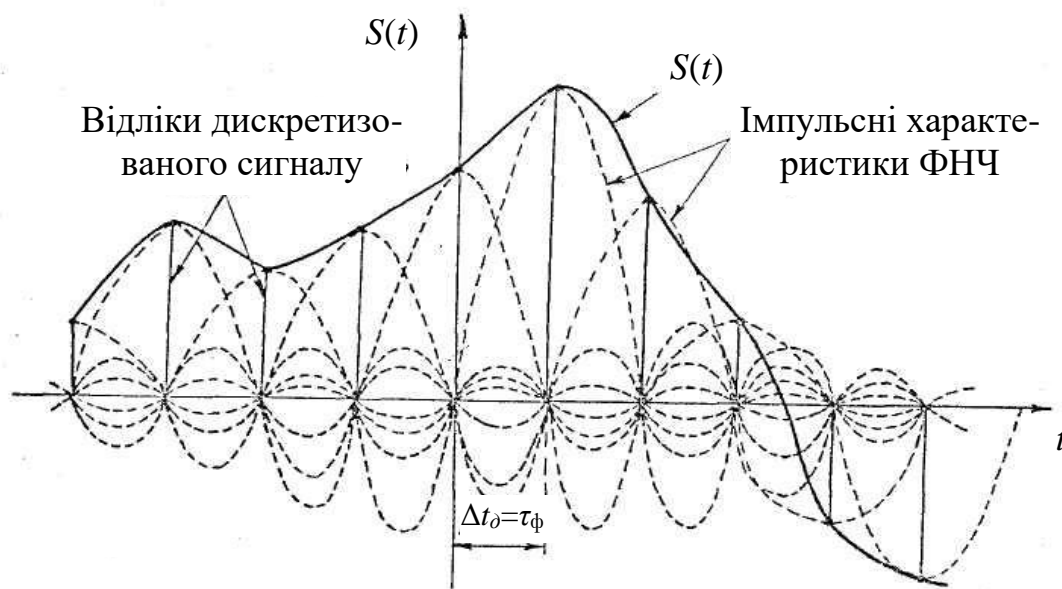


Рис. 2.15. Процес відновлення безперервного сигналу з дискретизованого

За передавальною характеристикою ФНЧ можна визначити його граничну частоту, звану частотою зрізу фільтра $f_{зр}$. У свою чергу ця частота пов'язана з постійною часу ФНЧ виразом $2f_{зр} = 1/\tau_{\phi}$. Таким чином, умовою неспотвореного відновлення безперервного сигналу є

$$\tau_{\phi} = \Delta t_0 = 1/2f_{зр} = 1/2f_m. \quad (2.75)$$

З рис. 2.16 можна зрозуміти, що чим більше відрізнятиметься реальна передавальна характеристика ФНЧ (комплексна залежність коефіцієнта передачі фільтра від частоти) від ідеальної, тим більше в смугу прозорості ФНЧ потраплятиме сусідніх складових спектра дискретизованого сигналу. Таким чином, однією з причин, що призводять до спотворення сигналу при його відновленні, є неідеальність характеристики ФНЧ. Іншою причиною спотворення сигналу є неправильний вибір інтервалу дискретизації. Третьою причиною є відмінність реальних імпульсів, формованих генератором тактових імпульсів

від δ -функцій. Четвертою причиною є використання реальних сигналів $S(t)$, які обмежені в часі, отже, мають нескінченний спектр, а відповідно до теореми Котельникова сигнал, що підлягає дискретизації, повинен мати верхню частоту, що не перевищує f_m . Таким чином, частина спектра початкового сигналу при дискретизації не використовується.

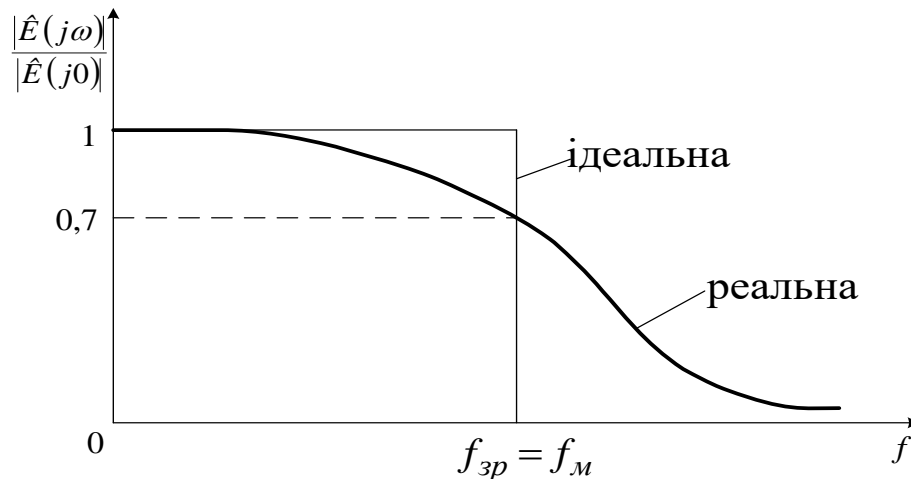


Рис. 2.16. Графіки ідеальної і реальної нормованих передаточних характеристик ФНЧ

Оцінимо помилку при апроксимації довільного сигналу рядом Котельникова. Припустимо, що заданий довільний сигнал $S(t)$ задовольняє умови Діріхле. Тоді його можна подати як

$$S(t) = S_{\text{dc}}(t, \omega_m) + S_{\text{ном}}(t), \quad (2.76)$$

де $S_{\text{dc}}(t, \omega_m)$ - сигнал із спектром, обмеженим значенням ω_m ;

$S_{\text{ном}}(t)$ - сигнал помилки апроксимації сигналу $S(t)$ із спектром, що займає нескінченну смугу частот при $\omega > \omega_m$.

Оскільки спектри вказаних сигналів не перекриваються, то сигнали S_{dc} і $S_{\text{ном}}$ є ортогональними, а їх енергії, тобто квадрати норм сигналів, складаються

$$E_c = \|S\|^2 = \|S_{\partial c}\|^2 + \|S_{ном}\|^2. \quad (2.77)$$

За абсолютну міру помилки при апроксимації сигналу приймають геометричну відстань, що дорівнює нормі сигналу помилки. Якщо відомий енергетичний спектр детермінованого або випадкового сигналу $S(t)$, який дорівнює

$$W_s(\omega) = S^2(\omega) = S(j\omega) \cdot S^*(j\omega), \quad (2.78)$$

то абсолютну помилку апроксимації сигналу рядом Котельникова можна визначити за теоремою Релея за допомогою виразу

$$|S_{ном}|_{абс} = \sqrt{\frac{1}{\pi} \int_{\omega_m}^{\infty} W_s(\omega) d\omega}. \quad (2.79)$$

Відносна помилка апроксимації знаходиться з виразу

$$|S_{ном}|_{від} = \frac{|S_{ном}|_{абс}}{|S_{ном}| + |S_{ном}|_{абс}} \cdot 100\%. \quad (2.80)$$

Таким чином, при дискретизації безперервних сигналів і при їх відновленні важливо правильно вибрати інтервал дискретизації, тривалість дискретизованих імпульсів і частоту зрізу фільтра нижніх частот.

2.5. Випадкові процеси та їх загальні характеристики

Випадковим процесом (ВП) називається така функція часу (зміна напруги, струму в електричному ланцюзі або іншої фізичної величини), значення якої в результаті досліду може бути визначене у будь-який момент часу з імовірністю менше одиниці. Зазначимо, що для детермінованого (регулярного) процесу

(сигналу) всі його значення (параметри) визначаються у будь-який момент часу з імовірністю, що дорівнює одиниці.

Випадковий процес $X(t)$ може приймати різні конкретні форми в результаті досліду, які називаються його реалізаціями $x_i(t)$. Приклади реалізації випадкового процесу наведені на рис. 2.17. Сукупність реалізацій випадкового процесу називається його ансамблем. У деякий фіксований момент часу t_1 різні реалізації процесу $X(t)$ приймають конкретні різні значення $x_1(t_1)$, $x_2(t_2)$, $x_N(t_1)$. При цьому всі реалізації $x_i(t)$ є детермінованими. Сукупність значень $X(t_1)$ є випадковою величиною і називається перетином випадкового процесу.

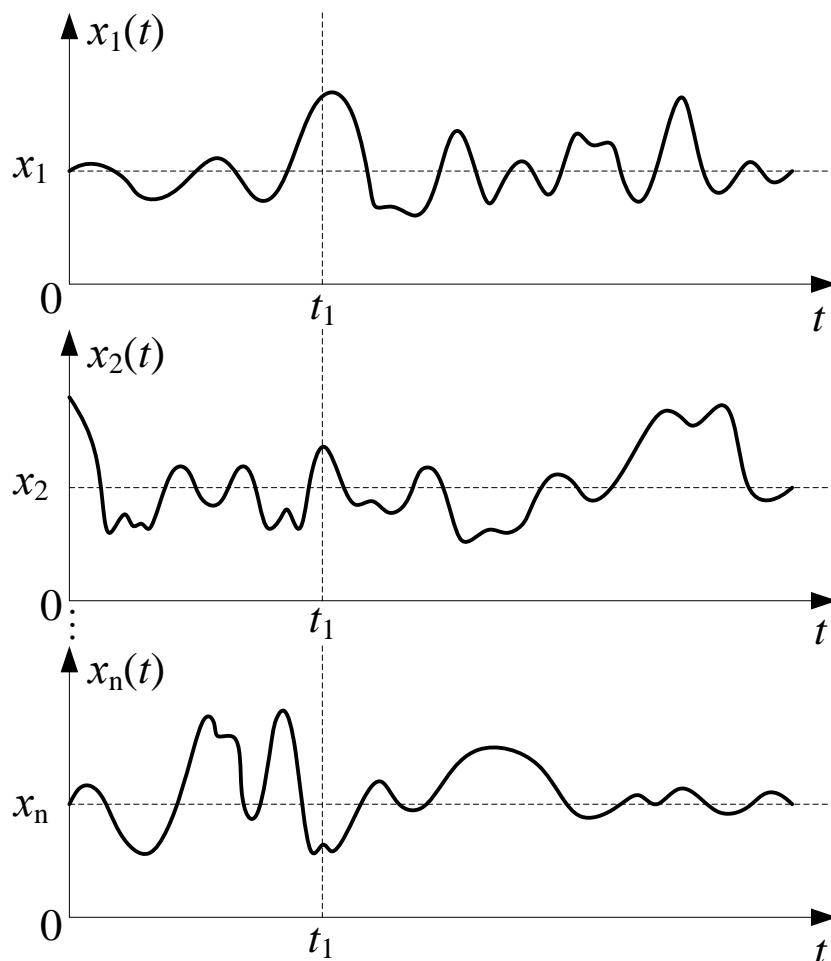


Рис. 2.17. Приклади реалізацій випадкового процесу $X(t)$

2.5.1. Функції розподілу випадкових процесів

Оскільки випадковий процес (сигнал) описується випадковою функцією часу, значення якої при будь-якому значенні аргументу (часу) є випадковими величинами, то його можна задати (описати) тільки в імовірнісному сенсі. З цією метою використовується математичний апарат теорії імовірності, математичної статистики і теорії випадкових процесів.

Якщо визначені будь-які реалізації для n -перетинів випадкового процесу, то його можна описати n -мірною інтегральною функцією розподілу (ІФР) імовірності (інтегральним законом розподілу імовірності), яка записується як

$$F_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P[X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n].$$

Тут вираз $P[X(t_1) \leq x_1]$ означає імовірність того, що випадкова величина $X(t_1)$ для першого перетину випадкового процесу $X(t)$ не перевищує заданого фіксованого значення (рівня) x_1 і т. д.

Якщо існують всі частинні похідні від інтегральної функції розподілу імовірності F_n по кожному з аргументів x_1, x_2, \dots, x_n , то можна визначити n -мірну диференціальну функцію розподілу (ДФР) імовірності випадкового процесу (диференціальний закон розподілу імовірності), звану також функцією щільності розподілу імовірності, яка в кожному з n -перетинів у моменти часу t_1, t_2, \dots, t_n характеризує імовірність попадання значень випадкового процесу в нескінченно малий інтервал $\Delta x = x_k - x_{k-1}$.

Тоді цю функцію можна записати як

$$w_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n F_n}{\partial x^n} = \frac{\partial^n F_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)}{\partial x_1 \partial x_2 \dots \partial x_n}.$$

Функції розподілу імовірності F_n і w_n повністю описують конкретний випадковий процес, проте на практиці достатньо знати двомірні або одномірні функції розподілу імовірності, хоча вони характеризують випадковий процес тільки в двох або в одному його перетинах.

Так, двомірна інтегральна функція $F_2(x_1, x_2; t_1, t_2)$ визначається як імовірність того, що випадкова величина $X(t_1)$ не перевищить значення x_1 , а випадкова величина $X(t_2)$ не перевищить значення x_2 , і записується як

$$F_2(x_1, x_2; t_1, t_2) = P[X(t_1) \leq x_1, X(t_2) \leq x_2].$$

Двовимірна щільність розподілу імовірності визначається як частинна похідна (якщо вона існує) другого порядку від двомірної інтегральної функції розподілу відповідно до виразу

$$\begin{aligned} w_2(x_1, t_1; x_2, t_2) &= \frac{\partial^2 F_2(x_1, t_1; x_2, t_2)}{\partial x_1 \partial x_2} = \\ &= \frac{P[x_1 \leq x(t_1) \leq x_1 + \Delta x; x_2 \leq x(t_2) \leq x_2 + \Delta x]}{\Delta x_1 \Delta x_2}, \end{aligned}$$

де $\Delta x_1 \rightarrow 0, \Delta x_2 \rightarrow 0$.

Двомірні функції розподілу імовірності дозволяють оцінювати розвиток випадкового процесу в часі (між двома перетинами).

Одномірна інтегральна функція розподілу визначається як імовірність того, що у момент часу t_1 випадкова величина $X(t_1)$ не перевищить значення x_1 і записується як

$$F_1(x_1, t_1) = P[X(t_1) \leq x_1].$$

Одномірна щільність розподілу імовірності ВП визначається як частинна похідна від інтегральної функції розподілу з виразу

$$w_1(x_1, t_1) = \frac{\partial F_1(x_1, t_1)}{\partial x_1} = \frac{P[x_1 \leq x(t_1) \leq x_1 + \Delta x]}{\Delta x_1},$$

де $\Delta x_1 \rightarrow 0$.

Вона характеризує імовірність знаходження (попадання) випадкової величини $X(t_1)$ в інтервалі значень випадкового процесу $x_1 + \Delta x$ в перетині $X(t_1)$.

Таким чином, одномірні функції розподілу не дозволяють оцінити розвиток випадкового процесу в часі.

На підставі вищевикладеного можна вказати на такі властивості функцій розподілу випадкових процесів.

Інтегральна функція розподілу імовірності є безрозмірною, невід'ємною і змінюється від 0 до 1.

Диференціальна функція (щільність) розподілу імовірності також є невід'ємною, має розмірність, зворотну величині Δx , і змінюється від 0 до ∞ .

Площа під кривою щільності розподілу імовірності дорівнює

$$\int_{-\infty}^{\infty} w_1(x) dx = 1. \quad (2.81)$$

Вираз (2.81) визначає властивість нормування щільності розподілу імовірності.

2.5.2. Моментні (числові) характеристики випадкових процесів

Хоча випадкові процеси є непередбачуваними в деталях, деякі їх властивості (характеристики) в середньому проявляють достатню регулярність, отже, ними можна оцінювати випадкові процеси. На практиці найбільше значення мають такі моментні (числові) характеристики: математичне очікування (середнє значення постійної складової), середній квадрат (середнє значення змінної складової), середньоквадратичне відхилення, дисперсія, а також функції кореляції випадкового процесу.

Спочатку розглянемо моментні характеристики для нестационарних випадкових процесів. Нестационарними називаються такі процеси, у яких моментні характеристики залежать від вибору моменту часу, тобто від перетинів процесу.

Математичне очікування (перший початковий момент) є середнє значенням перетину $X(t_1)$ у фіксований момент часу t_1

$$\overline{X(t_1)} = M[X(t_1)] = \int_{-\infty}^{\infty} x w_1(x, t_1) dx. \quad (2.82)$$

Тут пряма риса над $X(t_1)$ означає усереднювання за ансамблем реалізацій випадкового процесу у момент часу t_1 .

Середній квадрат (другий початковий момент) визначається з виразу

$$\overline{X^2(t_1)} = M[X^2(t_1)] = \int_{-\infty}^{\infty} x^2 w_1(x, t_1) dx. \quad (2.83)$$

Він характеризує повну потужність (суму постійної і змінної складових) випадкового сигналу. Якщо випадковий процес (сигнал) не містить постійної складової, він називається центрованим і записується як

$$\overset{\circ}{X}(t) = x(t) - \overline{X(t)}.$$

Дисперсія (другий центральний момент) є математичне очікуванням квадрата центрованого процесу. Вона записується виразом

$$D[X(t_1)] = \left[\overset{\circ}{X}(t_1) \right]^2 = \int_{-\infty}^{\infty} [x - \overline{x(t_1)}]^2 w_1(x, t_1) dx. \quad (2.84)$$

Дисперсія характеризує ступінь розсіювання (відхилення) квадратів миттєвих значень, що приймаються окремими реалізаціями у фіксований момент часу t_1 від середнього значення.

Середньоквадратичне значення (відхилення) процесу визначається з виразу $y_x(t_1) = \sqrt{D[x(t_1)]}$ і характеризує середнє значення змінної складової процесу в перетині $X(t_1)$.

Приклади реалізації нестационарних випадкових процесів наведені відповідно на рис. 2.18 (дисперсія постійна, математичне очікування залежить від часу), на рис. 2.19 (математичне очікування постійне, дисперсія залежить від часу),

на рис. 2.20 (математичне очікування і дисперсія залежать від часу).

Стационарним (у вузькому розумінні) називається випадковий процес, у якого всі функції розподілу імовірності (будь-якого порядку) і відповідно всі інші моментні характеристики не залежать від початку відліку часу.

Стационарним (у широкому розумінні) називається випадковий процес, у якого математичне очікування і дисперсія не залежать від часу, а функція кореляції залежить від інтервалу між двома перетинами $\tau = t_2 - t_1$, але не від самих значень t_1 і t_2 . Із стаціонарності у вузькому розумінні випливає стаціонарність у широкому розумінні, але не навпаки.

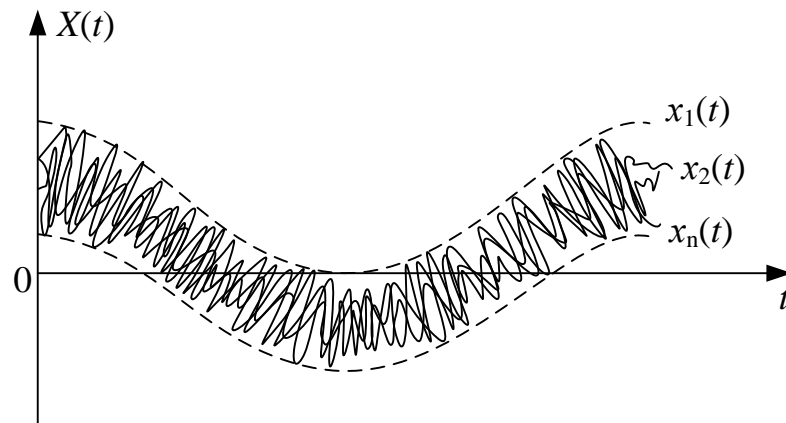


Рис. 2.18. Реалізації нестационарного випадкового процесу
($D_x = \text{const}$, $m_x = \text{var}$)

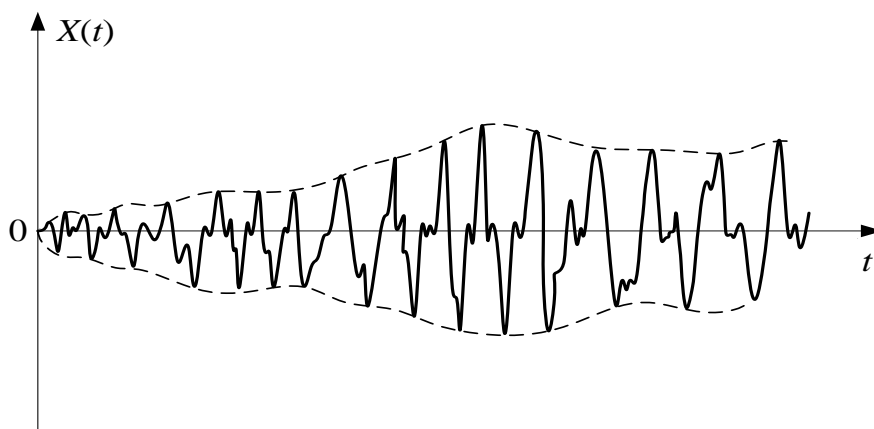


Рис. 2.19. Реалізації нестационарного випадкового процесу
($m_x = \text{const}$, $D_x = \text{var}$)

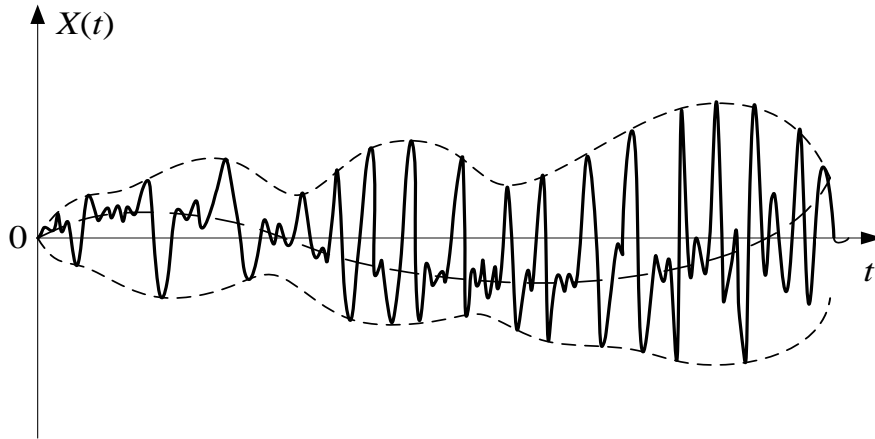


Рис. 2.20. Реалізації нестационарного випадкового процесу
 $(m_x = \text{var}, D_x = \text{var})$

Отже, для випадкових процесів, стаціонарних у широкому розумінні, наведені вище вирази для моментних характеристик з урахуванням їх незалежності від часу можна записати без фіксованих моментів часу в такому вигляді:

$$M[X] = m_x = \bar{X} = \int_{-\infty}^{\infty} xw(x)dx - \text{математичне очікування};$$

$$M[X^2] = \overline{X^2} = \int_{-\infty}^{\infty} x^2w(x)dx - \text{середній квадрат};$$

$$D[X] = M[x - \bar{X}]^2 = D\left[\overset{\circ}{X}\right] = \int_{-\infty}^{\infty} [x - m_x]^2 w(x)dx - \text{дисперсія};$$

$$y_x = \sqrt{D(x)} - \text{середньоквадратичне відхилення}.$$

Приклад реалізації стаціонарного випадкового процесу для випадку, коли $m_x = 0$; $3y_x = 1B$, наведений на рис. 2.21.

На практиці часто використовується не весь ансамбль, а окремі реалізації випадкового процесу, тривалість яких може бути великою, але лише умовно дорівнювати нескінченності. У цьому випадку вводиться властивість ергодичності випадкового процесу. Ергодична властивість стаціонарних процесів полягає в

тому, що будь-яка моментна характеристика процесу, отримана в результаті усереднювання за ансамблем реалізацій, повинна дорівнювати (з імовірністю скільки завгодно близькою до одиниці) аналогічній характеристиці отриманій за однією нескінченно протяжною в часі реалізацією процесу усереднюванням за часом.

Для того щоб випадковий процес був ергодичним, перш за все він повинен бути стаціонарним у широкому розумінні.

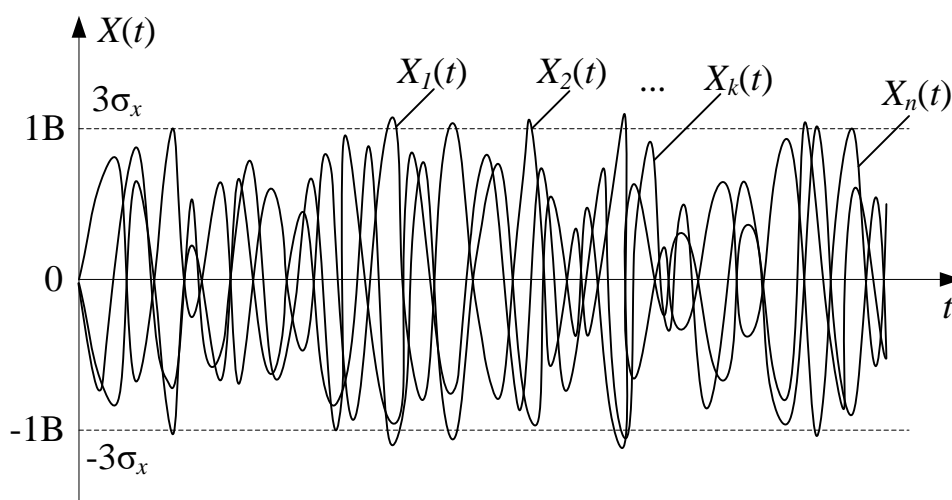


Рис. 2.21. Приклад реалізацій стаціонарного випадкового процесу згідно з правилом 3σ ($m_x = 0$; $3\sigma_x = 1V$)

Для стаціонарних ергодичних процесів (випадкових сигналів) моментні характеристики визначаються такими виразами:

- математичне очікування (постійна напруга, струм)

$$m_x = \tilde{X}(t) = M[X(t)] = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt;$$

- середній квадрат (повна потужність сигналу)

$$\tilde{X}^2(t) = M[X^2(t)] = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} x^2(t) dt;$$

- дисперсія (потужність змінної складової сигналу)

$$D[X(t)] = M[X - m_x]^2 = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} [x(t) - m_x]^2 dt;$$

- середньоквадратичне відхилення (середнє значення змінної складової напруги або струму)

$$y_x = \sqrt{D[x(t)]}.$$

Хвиляста лінія в наведених вище виразах означає усереднювання в часі, а T – інтервал спостереження реалізації випадкового процесу.

З умови ергодичності ВП виходить його стаціонарність, зворотнє ж утвердження не завжди виконується.

2.5.3. Приклади деяких випадкових процесів

Розглянемо випадкові процеси (сигнали), що найчастіше зустрічаються на практиці.

2.5.3.1. Сукупність гармонійних коливань з випадковою амплітудою

Цей випадковий нестационарний процес описується виразом

$$u(t) = U(t) \cos(\omega_0 t + \varphi_0) = U(t) \cos \Phi(t),$$

у якого кутова частота ω_0 і початкова фаза φ_0 є постійними величинами, а амплітуда U_0 приймає випадкові рівноімовірні значення в інтервалі від 0 до U_m . Приклад реалізації такого процесу наведено на рис. 2.22.

Визначимо одномірну щільність імовірності $w(x, t_1)$ для фіксованого моменту часу t_1 , враховуючи властивість нормування функції щільності імовірності (вираз (2.82)). Тоді маємо $w(x, t_1) = 1/U_m \cos \Phi(t_1)$.

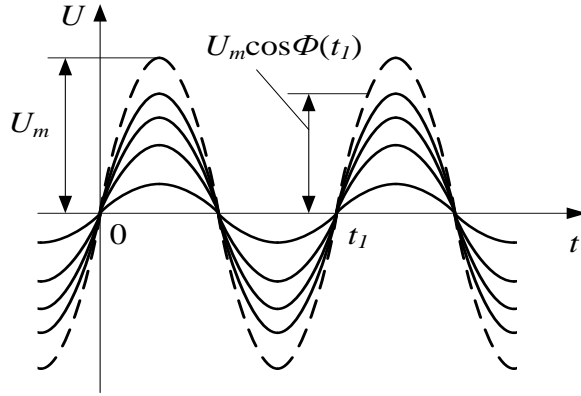


Рис. 2.22. Реалізація гармонійних коливань з випадковою амплітудою

Графіки диференціальної та інтегральної функцій розподілу імовірності гармонійних коливань з випадковою амплітудою наведені на рис. 2.23.

Математичне очікування для нестационарного ВП відповідно до виразу (2.82) записується як

$$\begin{aligned} \overline{U(t_1)} = m_u = M[u(t_1)] &= \int_0^{U_M \cos \Phi(t_1)} u \cdot 1/U_M \cos \Phi(t_1) du = \\ &= \frac{1}{U_M \cos \Phi(t_1)} \int_0^{U_M \cos \Phi(t_1)} u du = \frac{1}{2} U_M \cos \Phi(t_1). \end{aligned}$$

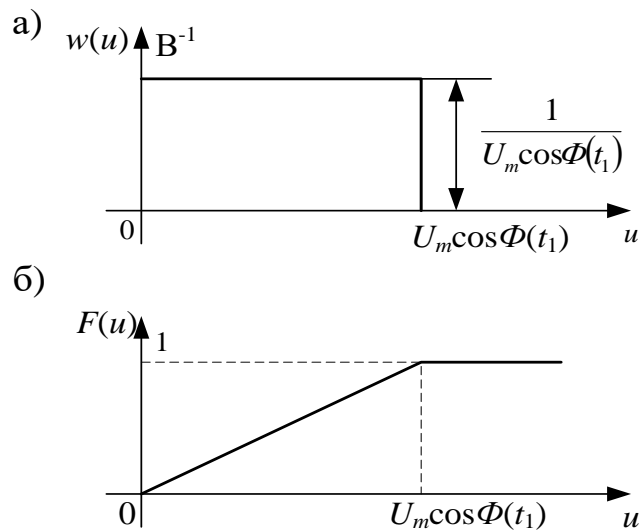


Рис. 2.23. Графіки: а) ДФР; б) ІФР

Оскільки математичне очікування не дорівнює нулю і залежить від часу, то даний випадковий процес є нецентрованим і нестационарним.

Середній квадрат випадкового сигналу визначається з виразу (2.83)

$$\begin{aligned} \overline{U^2(t_1)} &= M[u^2(t)] = \int_{-\infty}^{\infty} u^2 \frac{1}{U_M} \cos \Phi(t_1) du = \\ &= \frac{1}{U_M \cos \Phi(t_1)} \int_0^{U_M \cos \Phi(t_1)} u^2 du = \frac{1}{3} U_M^2 \cos \Phi(t_1) \end{aligned}$$

і також залежить від часу.

Дисперсія випадкового сигналу знаходиться з виразу (2.84)

$$\begin{aligned} D[U(t_1)] &= M[u(t_1) - m_u(t_1)]^2 = \frac{1}{3} U_M^2 \cos^2 \Phi(t_1) - \\ &- \frac{1}{4} U_M^2 \cos^2 \Phi(t_1) = \frac{1}{12} U_M^2 \cos^2 \Phi(t_1) \end{aligned}$$

і є функцією часу.

2.5.3.2. Сукупність гармонійних коливань з випадковими фазами

Такий процес, реалізації якого наведені на рис. 2.24, можна записати як $u(t) = U_0 \cos[\omega_0 t + \varphi(t)]$. Його амплітуда U_0 і кутова частота ω_0 залишаються постійними, а початкова фаза $\varphi(t)$ приймає випадкові рівноімовірні значення в інтервалі $-\pi \leq \varphi(t) \leq \pi$. Тоді на підставі виразу (2.81) вираз для щільності імовірності початкової фази прийме вигляд $w(\varphi) = 1/2\pi$, графік рівномірного розподілу якого наведений на рис. 2.24.

Для k -тої реалізації випадкового сигналу вираз можна записати як $u_k(t) = U_0 \cos(\omega_0 t + \varphi_k)$. Оскільки повна фаза коливань, що дорівнює $\Phi(t) = \omega_0 t + \varphi$, є випадковою величиною і

рівномірною в інтервалі $\omega_0 t - \pi \leq \varphi(t) \leq \omega_0 t + \pi$, то щільність імовірності її визначиться з виразу $w(\varphi) = 1/2\pi$.

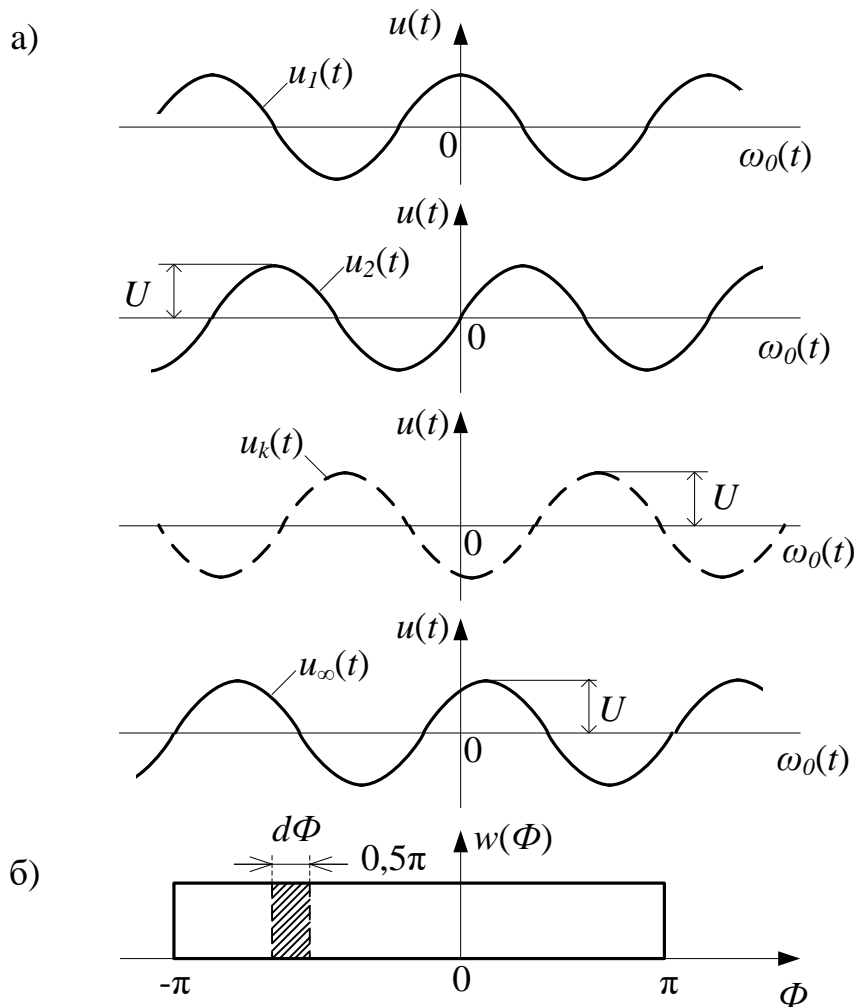


Рис. 2.24. Реалізації гармонійних коливань:
а) з випадковими фазами; б) ДФР ВП

Для знаходження одномірної щільності розподілу імовірності миттєвих значень напруги $w(u)$ випадкового сигналу $u(t)$ скористаємося рис. 2.25, а.

Виділимо інтервал $U + \Delta u$ і визначимо імовірність того, що миттєве значення сигналу опиниться в інтервалі $U + \Delta u$, відповідному проміжку часу Δt . Тоді можна записати $P[U \leq U(t) \leq U + \Delta u] = w(u)du$. З рис. 2.25, а очевидно, що ця імовірність співпадає з імовірністю попадання випадкової фази Φ

коливання в один з двох заштрихованих фазових інтервалів і дорівнює

$$P[\Phi \leq \Phi(t) \leq \Phi + \Delta\Phi] = 2w(\Phi)d\Phi.$$

Отже, для періоду T коливання можна записати рівність

$$w(u)du = 2w(\Phi)d\Phi - (1/2\pi)d\Phi.$$

Звідси отримаємо вираз для функції $w(u)$ з урахуванням того, що щільність імовірності величина ненегативна

$$w(u) = \frac{1}{\pi \left| \frac{du}{d\Phi} \right|}.$$

Оскільки

$$\left| \frac{du}{d\Phi} \right| = \left| \frac{d}{d\Phi} U_0 \cos\Phi \right| = |U_0 \sin\Phi| = \sqrt{U_0^2 (1 - \cos^2\Phi)} = \sqrt{U_0^2 - u^2},$$

то остаточно отримаємо $w(u) = \frac{1}{\pi \sqrt{U_0^2 - u^2}}$ для $-U_0 \leq u \leq U_0$.

Графіки функції розподілу імовірності $w(u)$ і $F(u)$ для даного випадкового процесу наведені на рис. 2.25, б.

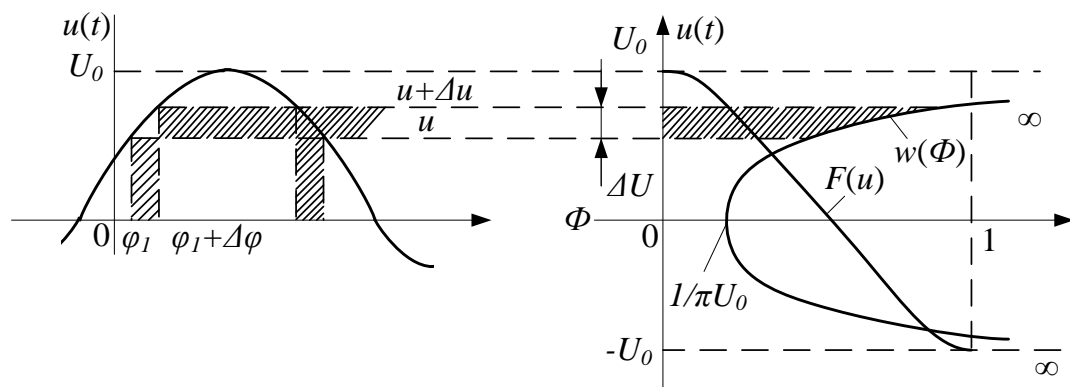


Рис. 2.25. До знаходження: а) одномірної ДФР; б) ДФР та ІФР випадкового процесу

З цього виразу виходить, що функція $w(u)$ змінюється згідно із законом $\arcsin(u)$. Інтегральна функція розподілу імовірності гармонійного сигналу (напруги) з випадковою початковою фазою визначається за формулою

$$F(u) = \int_{-\infty}^u w(u) du = P[U(t) < u] = \begin{cases} 0, & u \leq -U_0; \\ \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{u}{U_0}, & |u| \leq U_0; \\ 1, & u > U_0. \end{cases}$$

Таким чином, функція щільності імовірності $w(u)$ для гармонійного сигналу з випадковою фазою від значення $1/\pi U_0$ при початкових відхиленнях амплітуди прагне до нескінченності при прагненні u до U_0 . Це пов'язано з тим, що якщо виконати велику серію дослідів, кожного разу випадковим чином вибираючи значення u з вказаної області, то величина $U_0 \cos[\omega_0 t + \Phi(t)]$ частіше прийматиме значення, близькі до $\pm U_0$, ніж близькі до 0. Слід зазначити, що даний випадковий процес є центрованим і стаціонарним.

2.5.3.3. Гаусівський (нормальний) випадковий процес

У теорії і техніці зв'язку дуже широко користуються поняттям гаусівського (нормального) закону розподілу. Цим законом описується значне число флуктуаційних завад, а також сума деяких сигналів і флуктуаційних завад (шумів), що діють у каналах зв'язку.

Нормальний випадковий стаціонарний процес описується одномірною функцією щільності розподілу імовірності такого вигляду [46]:

$$w(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{(x - m_x)^2}{2\sigma_x^2}\right], \quad (2.85)$$

де m_x - постійна складова процесу, $m_x = \int_{-\infty}^{\infty} xw(x) dx$;

σ_x^2 - середнє значення квадрата змінної складової процесу (дисперсія), $\sigma_x^2 = \int_{-\infty}^{\infty} (x - m_x)^2 w(x) dx$.

Інтегральна функція розподілу імовірності (ІФР) нормального стаціонарного випадкового процесу не може бути виражена через елементарні функції, як це мало місце з гармонійним сигналом. Враховуючи зв'язок між інтегральною і диференціальною функціями розподілу імовірності, ІФР можна записати як

$$P(a < x < b) = F(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_a^b \exp\left[-\frac{(x - m_x)^2}{2\sigma_x^2}\right] dx. \quad (2.86)$$

Оскільки для флукуаційної завади (шуму) позитивні і негативні миттєві значення напруги (струму) з'являються з однаковою імовірністю, то $m_x = 0$, дисперсія σ_x^2 дорівнює середній потужності завади (шуму) P_n , а ефективне значення завади дорівнює

$$U_{ne} = \sqrt{P_n} = \sigma_x.$$

Тоді вираз (2.86) можна записати як

$$\begin{aligned} P(a < x < b) &= F(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_a^b \exp\left[-\frac{x^2}{2\sigma_x^2}\right] dx = \\ &= \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_0^b \exp\left[-\frac{x^2}{2\sigma_x^2}\right] dx - \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_0^a \exp\left[-\frac{x^2}{2\sigma_x^2}\right] dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{b\sigma_x} \exp\left[-\frac{y^2}{2}\right] dy - \frac{1}{\sqrt{2\pi}} \int_0^{a\sigma_x} \exp\left[-\frac{y^2}{2}\right] dy = \\ &= F\left(\frac{b}{\sigma_x}\right) - F\left(\frac{a}{\sigma_x}\right) = F(u_1) - F(u_2), \end{aligned}$$

де $F(u)$ - функція, звана інтегралом імовірності (інтеграл Лапласа), $F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u \exp\left[-\frac{y^2}{2}\right] dy$;

u - нова змінна інтеграції, $u = x\sigma_i$.

Значення інтеграла імовірності знаходяться табуляцією функції $F(u)$ і приводяться в математичних довідниках (дод. 3).

При знаходженні значення ІФР можна також скористатися функцією Крампа (інтеграл помилок) [24]

$$\Phi(u) = \frac{2}{\sqrt{2\pi}} \int_0^u \exp\left[-\frac{y^2}{2}\right] dy. \quad (2.87)$$

Тоді отримаємо

$$P(a < x < b) = \frac{1}{2} \left[\Phi\left(\frac{b}{\sigma_x}\right) - \Phi\left(\frac{a}{\sigma_x}\right) \right]. \quad (2.88)$$

Функція Крампа також табулюється, і її значення приводяться в довідниках (дод. 1). Зазначимо, що існує зв'язок між функціями $F(u)$ і $\Phi(u)$ в такому вигляді:

$$\Phi(u) = 2F(u) - 1.$$

Функція $\Phi(u)$ є непарною, тобто $\Phi(-u) = -\Phi(u)$, крім того, вона задовольняє рівності $\Phi(\infty) = 1$ і $\Phi(0) = 0$.

Якщо потрібно знайти імовірність того, що перешкода перевищить деякий пороговий рівень γ , то на підставі виразу (2.88), врахувавши в ньому $b/\sigma_x = \infty$; $a = \gamma$, отримаємо

$$P(x > \gamma) = \frac{1}{2} \left[\Phi(\infty) - \Phi\left(\frac{\gamma}{\sigma_x}\right) \right] = \frac{1}{2} \left[1 - \Phi\left(\frac{\gamma}{\sigma_x}\right) \right].$$

На рис. 2.26 наведені графіки реалізації випадкових процесів, що підкоряються нормальному закону розподілу імовірності, а також відповідні їм ДФР і ІФР з різними моментними характеристиками.

Слід зазначити, що наведені вище функції розподілу імовірності не характеризують поведінку функцій $X(t)$ в часі. Так, з рис. 2.26 випливає, що реалізація (в) за своєю щільністю імовірності $w(x)$ і, отже, за значеннями m_x і σ_x нічим не відрізняється від реалізації (а), показаної на тому самому рисунку.

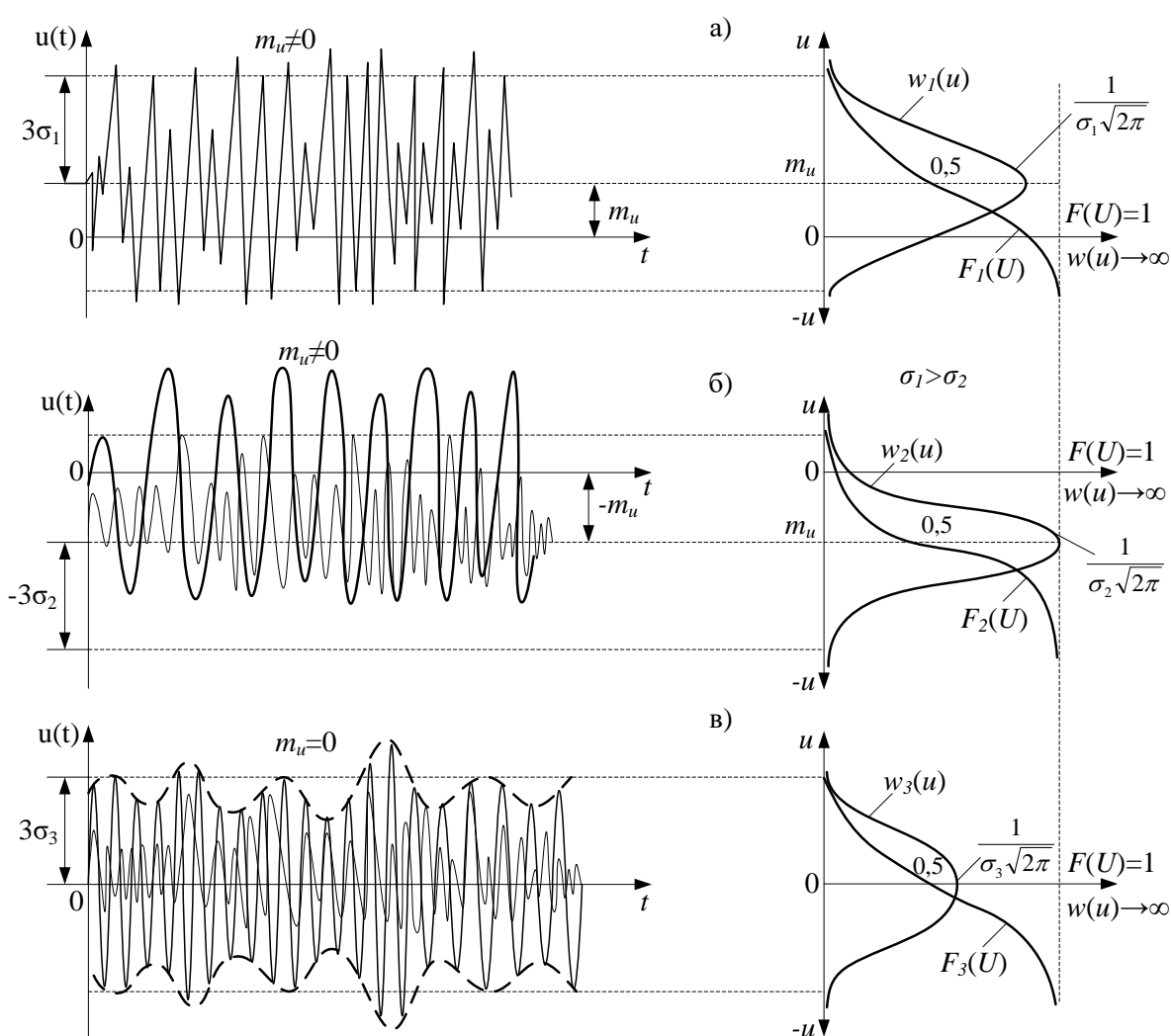


Рис. 2.26. Графіки реалізації випадкових гаусівських процесів

Якщо потрібно знайти імовірність того, що випадкова величина x гаусівського процесу відрізняється від свого середнього значення μ за абсолютною величиною не більш ніж на ε , то скориставшись формулою (2.88), можна записати

$$P(x - \mu < \varepsilon) = 2F\left(\frac{\varepsilon}{\sigma_x}\right) - 1.$$

При $\varepsilon = \sigma_x$ отримуємо $P=0,68268$, при $\varepsilon = 2\sigma_x$ маємо $P=0,95450$, а якщо $\varepsilon = 3\sigma_x$, то $P=0,99730 \approx 1$. Таким чином, випадкова величина x з нормальним розподілом практично не приймає значень, які б відрізнялися від середнього значення за абсолютною величиною більш ніж на $3\sigma_x$. Це твердження називається правилом «трьох сигм» і пояснюється за допомогою рис. 2.27.

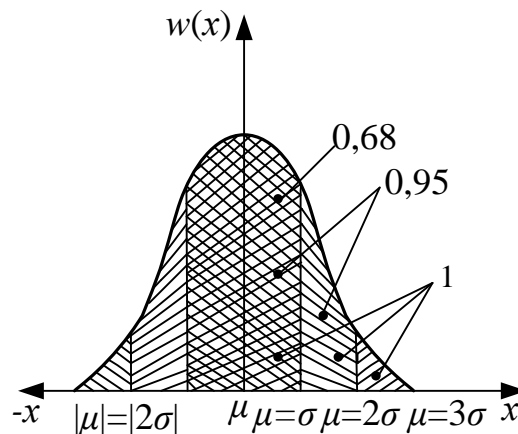


Рис. 2.27. До пояснення правила «трьох сигм»

2.5.3.4. Сума гармонійних реалізацій з випадковими фазами нормального гаусівського шуму

При визначенні виразу для ДФР сумарного процесу припустимо, що нормальний шум описується щільністю імовірності (2.85) при $m_x=0$, а гармонійна напруга описується щільністю розподілу імовірності

$$w(u) = \frac{1}{\pi\sqrt{A^2 - x^2}}$$

і що ці два процеси незалежні (некорельовані). Тоді приріст сумісної імовірності попадання гармонійного коливання в інтервал $(x, x + \Delta x)$ і шуму в інтервал $(y, y + \Delta y)$ можна записати в такому вигляді:

$$w(x, y)dxdy = \frac{dxdy}{\pi\sigma_x\sqrt{2\pi(A^2 - x^2)}} \exp\left(-\frac{y^2}{2\sigma_x^2}\right). \quad (2.89)$$

Після заміни змінних $z = y + x$ у виразі (2.89) та інтегрування його отримаємо

$$\int_{-A}^A w(x, y)dxdy = w(z)dz = \frac{dz}{\pi\sigma_x\sqrt{2\pi}} \int_{-A}^A \frac{\exp\left[-\frac{(z-x)^2}{2\sigma_x^2}\right]}{\sqrt{A^2 - x^2}} dx.$$

Звідси отримаємо вираз для $w(z)$

$$w(z) = \frac{1}{\pi\sigma_x\sqrt{2\pi}} \int_{-A}^A \frac{\exp\left[-\frac{(z-x)^2}{2\sigma_x^2}\right]}{\sqrt{A^2 - x^2}} dx.$$

Даний вираз є інтегралом згортки ДФР з двох складових випадкових процесів, графічне зображення якого наведено на рис. 2.28.

Для обчислення інтеграла згортки при деякому значенні $z = z_1$ будується графік функції $w(x)$, а також функції $w(z_1 - x)$, що є графіком функції $w(y)$, дзеркально переверненим відносно початку координат і зміщеним на величину z_1 . Потім будується графік добутку ординат $w(x)w(z_1 - x)$ і чисельним методом

визначається площа під цією функцією. Отримана площа є щільністю імовірності $w(z_1)$.

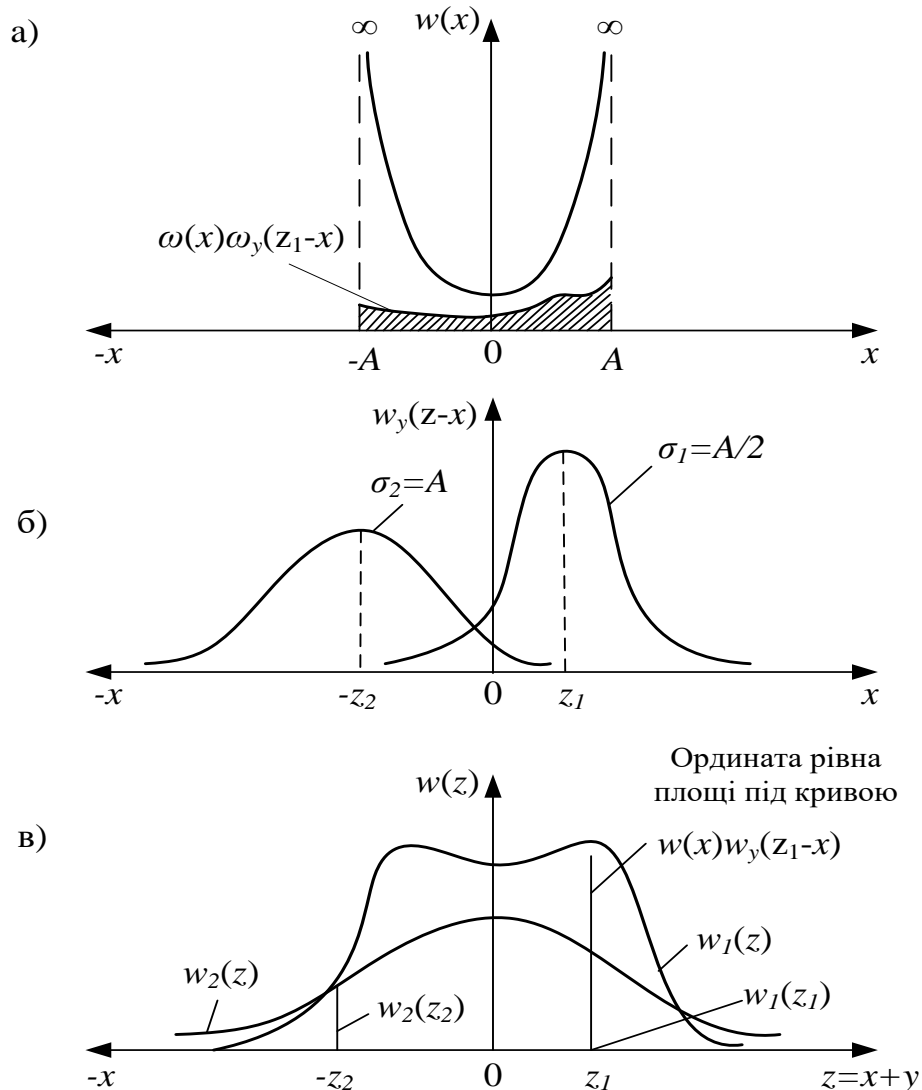


Рис. 2.28. До пояснення інтеграла згортки ДФР двох випадкових процесів

Виконавши аналогічні обчислення для інших значень z_i і з'єднавши знайдені точки плавною кривою, отримаємо результуючу функцію сумісної щільності імовірності двох випадкових процесів $w(z)$. Вираз (2.89) остаточно приводиться до такого вигляду:

$$w(z) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left[-\frac{2Z^2 + A^2}{4\sigma_x^2}\right] \left[I_0\left(\frac{A^2}{4\sigma_x^2}\right) I_0\left(\frac{zA}{\sigma_x^2}\right) + 2 \sum_{n=1}^{\infty} (-1)^n I_n\left(\frac{A^2}{4\sigma_x^2}\right) I_{2n}\left(\frac{zA}{\sigma_x^2}\right) \right],$$

де $I_0(k), I_n(k)$ - модифіковані функції Бесселя 1-го роду нульового і n -го порядків. На практиці при розрахунку обмежуються значеннями функції Бесселя 1-го порядку ($n=1$).

На рис. 2.29 наведені графіки ДФР або щільності імовірності розподілу сумарного процесу (гармонійних коливань з випадковою фазою і гаусівським шумом) для змінної $k = z/\sigma_x$ при різних співвідношеннях $A^2/(2\sigma_x)^2$. Щільність імовірності розподілу ненормованої змінної $z = k\sigma_x$ розраховується за формулою

$$w(z) = \frac{1}{\sigma_x} w(z/\sigma_x).$$

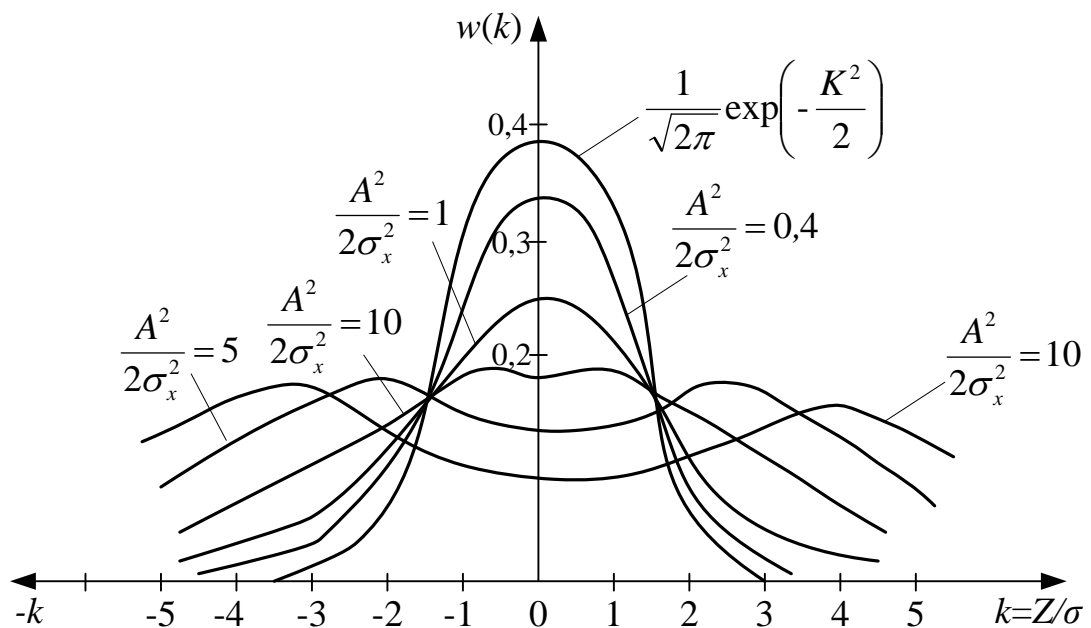


Рис. 2.29. Графіки ДФР суми гармонійних коливань з випадковими фазами і гаусівським шумом

Аналіз графіків, наведених на рис. 2.29 показує, що при значеннях відношення $A^2/(2\sigma_x)^2 \leq 1$ щільність імовірності розподілу амплітуд сумарного процесу підкоряється нормальному закону, тобто функція розподілу $w(k)$ має один максимум при $k = 0$ і переважаючий вплив тут має шумовий процес. При $1 < A^2/(2\sigma_x)^2 \leq 10$ відбувається деформація кривої, тобто перерозподіл щільності імовірності із збільшенням останньої при відхиленнях змінної $k = z/\sigma_x$ в обидва боки від її нульового значення. Таким чином, має місце переважаючий вплив на щільність розподілу імовірності гармонійного процесу з випадковою фазою. Функції розподілу безперервних випадкових величин наведено у дод. 4.

2.5.3.5. Розподіл Пуассона

Часто на практиці доводиться аналізувати процес, який являє собою послідовність моментів (точок), розташованих випадковим чином, наприклад, на осі часу. Такі точки можуть відповідати різним подіям, наприклад моментам часу надходження заявок на обслуговування або моментам часу появи відмов у якій-небудь системі та ін. Точковий випадковий потік використовується в задачі розподілу викликів на телефонній станції протягом доби (рис. 2.30).

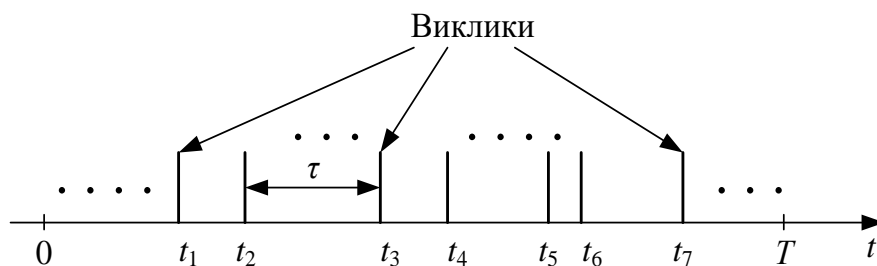


Рис. 2.30. Точковий випадковий потік викликів в інтервалі T

Загальне число викликів протягом доби є випадковою величиною. Для кожного часового інтервалу T шляхом спостережень можна встановити середнє число викликів $\lambda \tau = \mu$

(математичне очікування), де λ – інтенсивність (частота) викликів в одиницю часу. Тоді імовірність появи k -викликів $P_k(T)$ на інтервалі $(0, T)$ часто визначається формулою Пуассона [18]

$$P_k(T) = \frac{(\lambda \tau)^k}{k!} e^{-\lambda \tau} = \frac{(\mu)^k}{k!} e^{-\mu}. \quad (2.90)$$

З цього виразу випливає, що імовірність відсутності на цьому інтервалі викликів дорівнюватиме

$$P_0(T) = e^{-\lambda \tau} = e^{-\mu}.$$

Вирази для функцій розподілу імовірності випадкових дискретних величин отримаємо таким чином. Хай деяка випадкова величина X може приймати одне з m можливих значень $\{x_0, x_1, \dots, x_i, \dots, x_{m-1}\}$, де i - порядковий номер відповідного дискретного значення. Позначимо через $P(X = x_i) = P(x_i)$ імовірність того, що випадкова величина X прийме конкретне значення x_i . Приклад характеру розподілу цієї випадкової величини (ВВ) показаний на рис. 2.31. При цьому сума імовірності всіх можливих значень ВВ завжди дорівнює одиниці, оскільки сукупність подій $\{x_i\}$ утворює повну групу.

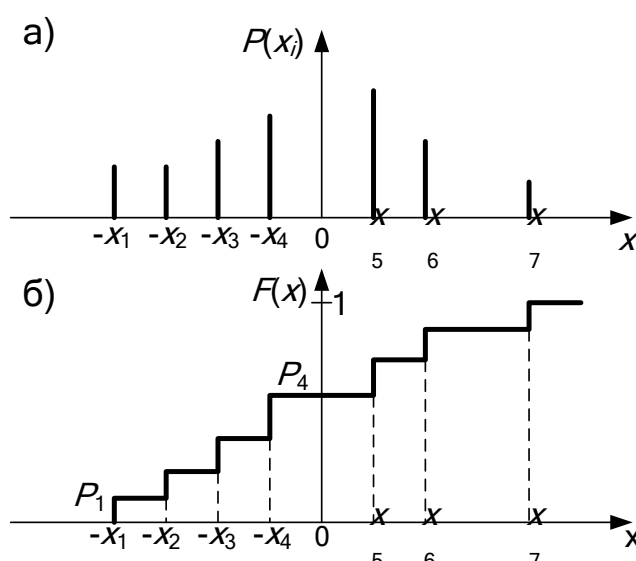


Рис. 2.31. Розподіл імовірності дискретної випадкової величини (а) та її інтегральна функція розподілу (б)

Виберемо на осі абсцис рис. 2.31, а деяку довільну точку x , яку розглядатимемо як незалежну змінну. Тоді одномірна ІФР випадкової величини X може бути записана як

$$F(x) = P(X < x) = \sum_i P(x_i) 1(x - x_i), \quad (2.91)$$

де $1(x - x_i)$ - одинична функція, що починається в точці x_i .

Вид $F(x)$ зображений на рис. 2.31, б. Оскільки підсумовування у виразі (2.91) проводиться по всіх точках, то ІФР дискретної випадкової величини має ступінчастий характер зі стрибками в точках можливих значень випадкової величини.

Якщо використовувати δ -функцію, то можна визначити щільність розподілу імовірності для дискретного розподілу в такому вигляді:

$$w(x) = \frac{dF(x)}{dx} = \sum_i P(x_i) \delta(x - x_i).$$

Ця функція також має дискретний характер. Функції розподілу деяких дискретних випадкових величин наведено у дод. 5.

2.5.3.6. Експоненціальний розподіл

Для випадкових точкових процесів з розподілом Пуассона вираз (2.90) величина τ , що визначає інтервал між викликами (рис. 2.32), описується експоненціальним розподілом з одномірною щільністю розподілу імовірності (ЩРІ):

$$w_1(\tau) = \lambda e^{-\lambda\tau}, \tau \geq 0. \quad (2.92)$$

Графік цієї функції наведено на рис. 2.32, а.

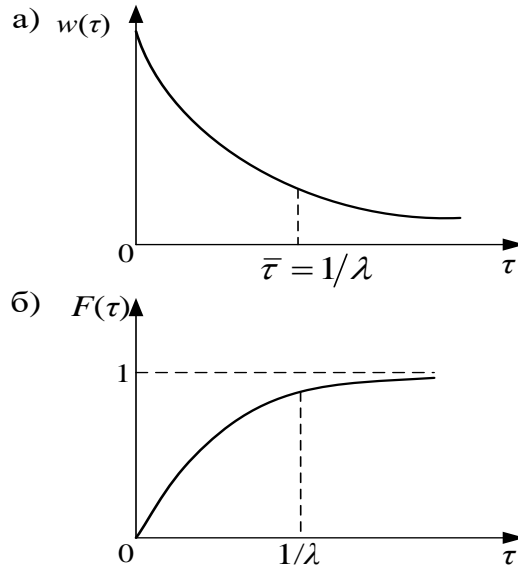


Рис. 2.32. Графіки функцій ЩРІ (а) і ІФР (б) для експоненціального розподілу дискретної величини

Тоді вираз для одновірної ІФР (рис. 2.32, б) запишеться як

$$F_1(\tau) = \int_0^{\tau} \lambda e^{-\lambda x} dx = 1 - e^{-\lambda \tau}. \quad (2.93)$$

Таким чином, за допомогою виразів (2.92) і (2.93) можна знайти, наприклад, імовірність того, що інтервал між сусідніми викликами виявиться рівним або менше деякого значення τ_0 . Ці вирази також використовуються при визначенні ЩРІ інтервалу між двома відмовами (несправностями) в роботі пристрою.

2.5.4. Кореляційні функції детермінованих і випадкових процесів

Для дослідження характеристик сигналів при їх перетворенні в різних функціональних елементах і пристроях засобів зв'язку, окрім спектрального аналізу, широко застосовується кореляційний аналіз. Він дозволяє без аналізу спектра сигналу визначати його основні характеристики і параметри, а також швидкість (частоту) зміни (убування або зростання) якого-небудь параметра сигналу.

Слово «кореляція» означає зв'язок між сигналом $S(t)$ і його копією $S(t-\tau)$, зрушеною щодо початкового сигналу на інтервалі τ , або між якими-небудь перетинами випадкового процесу.

Оскільки детерміновані сигнали є реалізаціями випадкових процесів, то спочатку розглянемо кореляційні функції детермінованих сигналів.

2.5.4.1. Кореляційні функції детермінованих сигналів

Кореляційною функцією детермінованого сигналу називається функція, що визначає зв'язок між сигналом і його копією, зсунутою по осі часових зсувів на інтервал τ .

Для неперіодичних детермінованих сигналів кореляційна функція, що називається також автокореляційною (АКФ), визначається з виразу

$$B_H(\tau) = \int_{-\infty}^{\infty} S(t)S(t-\tau)dt = S(t)S(t+\tau)dt. \quad (2.94)$$

Вираз (2.94) указує на те, що АКФ сигналу при зсуві копії $S(t+\tau)$ вправо на τ або вліво на $-\tau$ відносно початку відліку часового зсуву $\tau = 0$ є парною і симетричною. Приклади побудови АКФ для одиночного прямокутного відеоімпульсу і одиночного прямокутного радіоімпульсу показані на рис. 2.33, а, б.

Математично АКФ описується виразами

$$B_H(\tau) = U_0^2 \tau_i \left(1 - \frac{|\tau|}{\tau_i} \right) \quad \text{для відеоімпульсу,}$$

$$B_H(\tau) = \frac{U_0^2}{2} \tau_i \left(1 - \frac{|\tau|}{\tau_i} \right) \cos \omega_0 \tau \quad \text{для радіоімпульсу.}$$

Таким чином, при $\tau=0$ значення АКФ дорівнює енергії сигналу на опорі 1 Ом. На рис. 2.33, в показані також АКФ для пачки з трьох відеоімпульсів прямокутної форми.

Для періодичних детермінованих сигналів кореляційна функція визначається з виразу

$$B_n(\tau) = \lim_{T_1 \rightarrow \infty} \frac{1}{T_1} \int_{-T_1/2}^{T_1/2} S(t)S(t-\tau)dt = \lim_{T_1 \rightarrow \infty} \frac{1}{T_1} \int_{-T_1/2}^{T_1/2} S(t)S(t+\tau)dt,$$

де параметр T_1 означає інтервал спостереження сигналу, який для періодичного сигналу може дорівнювати його періоду T . Тоді вираз (2.94) можна записати як

$$B_n(\tau) = \frac{1}{T_1} \int_{-T_1/2}^{T_1/2} S(t)S(t-\tau)dt = \frac{1}{T_1} \int_{-T_1/2}^{T_1/2} S(t)S(t+\tau)dt. \quad (2.95)$$

З виразу (2.95) виходить, що АКФ періодичного сигналу є парною і симетричною, а при $\tau = 0$ її значення дорівнює потужності сигналу на опорі 1 Ом.

Приклади побудови АКФ для деяких видів періодичних детермінованих сигналів показані на рис. 2.34. З даного рисунка виходить, що АКФ періодичного сигналу є також періодичною функцією з періодом T .

Математично АКФ на періоді періодичного сигналу описується виразами

а) для послідовності прямокутних відеоімпульсів:

$$B_n(\tau) = U_0^2 \left(1 - \frac{|\tau|}{\tau_i} \right);$$

б) для послідовності прямокутних радіоімпульсів:

$$B_n(\tau) = \frac{U_0^2}{2} \left(1 - \frac{|\tau|}{\tau_i} \right) \cos \omega_0 \tau;$$

в) для гармонійних сигналів вигляду $U(t) = U_0 \cos(\omega_0 t + \varphi_0)$ або $U(t) = U_0 \sin(\omega_0 t + \varphi_0)$:

$$B_n(\tau) = \frac{U_0^2}{2} \cos \omega_0 \tau.$$

Функція кореляції визначає зв'язок між шириною спектра сигналу і часовим інтервалом τ_k , так званим інтервалом кореляції. Під інтервалом кореляції розуміємо максимальний зсув τ_k , у межах якого АКФ зменшується до нуля при зсуві копії неперіодичного сигналу на всю тривалість сигналу вправо або вліво відносно $\tau = 0$.

Зв'язок між кореляційною функцією $B(\tau)$ і енергетичним спектром сигналу $S^2(\omega)$ описується виразами, аналогічними до перетворень Фур'є:

$$B(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S^2(\omega) e^{j\omega\tau} d\omega, \quad (2.96)$$

$$S^2(\omega) = \int_{-\infty}^{\infty} B(\tau) e^{-j\omega\tau} d\tau. \quad (2.97)$$

Тоді справедливе твердження, згідно з яким чим ширше енергетичний спектр $S^2(\omega)$ сигналу $S(t)$, тим менше інтервал кореляції τ_k , тобто тим менше зв'язок сигналу $S(t)$ із зсунутою відносно нього копією $S(t \pm \tau_k)$.

Відповідно, чим більше інтервал кореляції τ_k , тим менше ширина енергетичного спектра $S^2(\omega)$ сигналу, тобто тим сильніше зв'язок (кореляція) між сигналом $S(t)$ і його копією $S(t \pm \tau_k)$. З виразів (2.96) і (2.97) виходить, що АКФ не залежить від ФЧХ сигналу $S(t)$, отже, різним по формі сигналам, що мають однакові амплітудні спектри, відповідають однакові кореляційні функції.

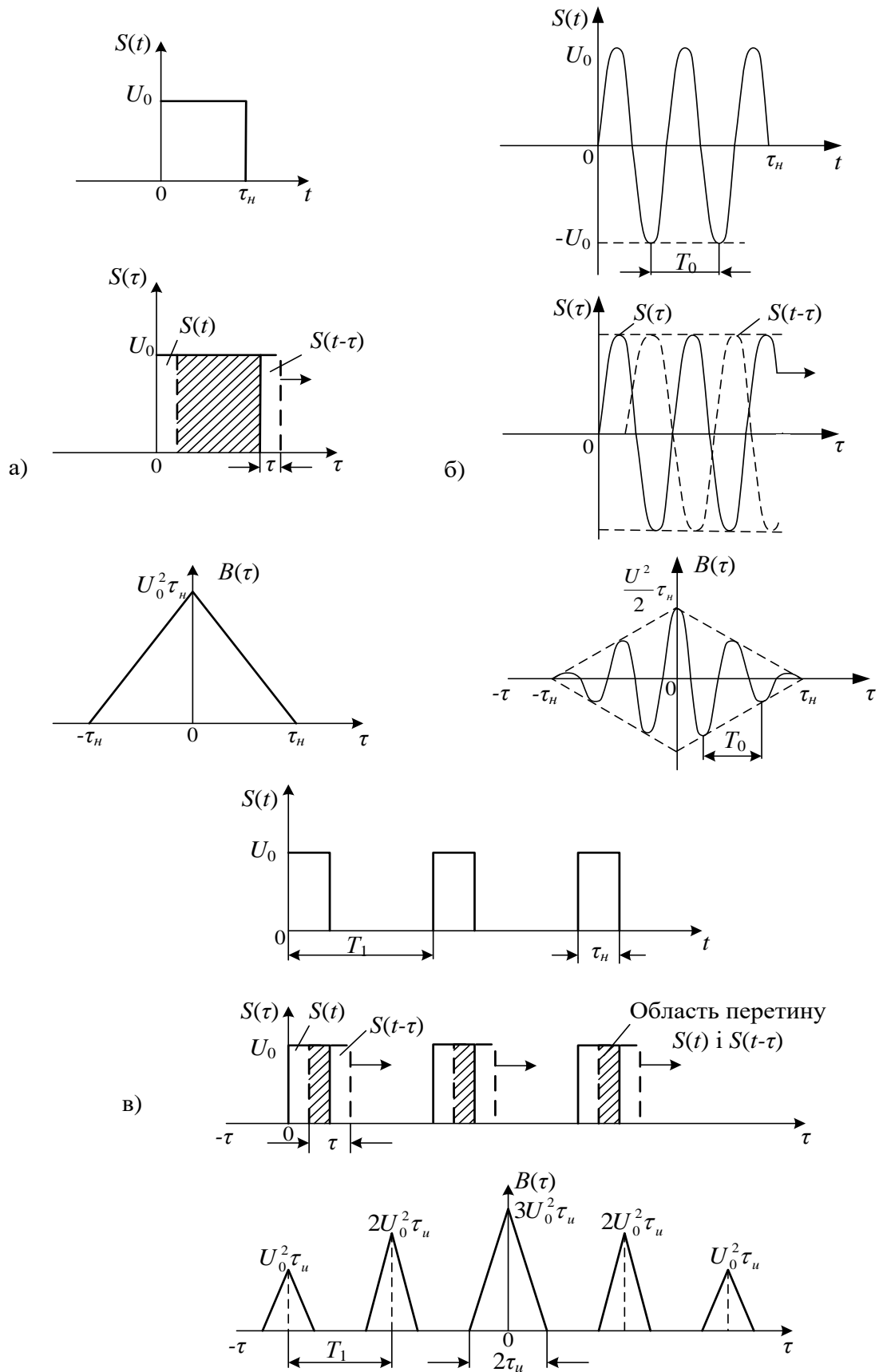


Рис. 2.33. Приклади побудови АКФ неперіодичних сигналів

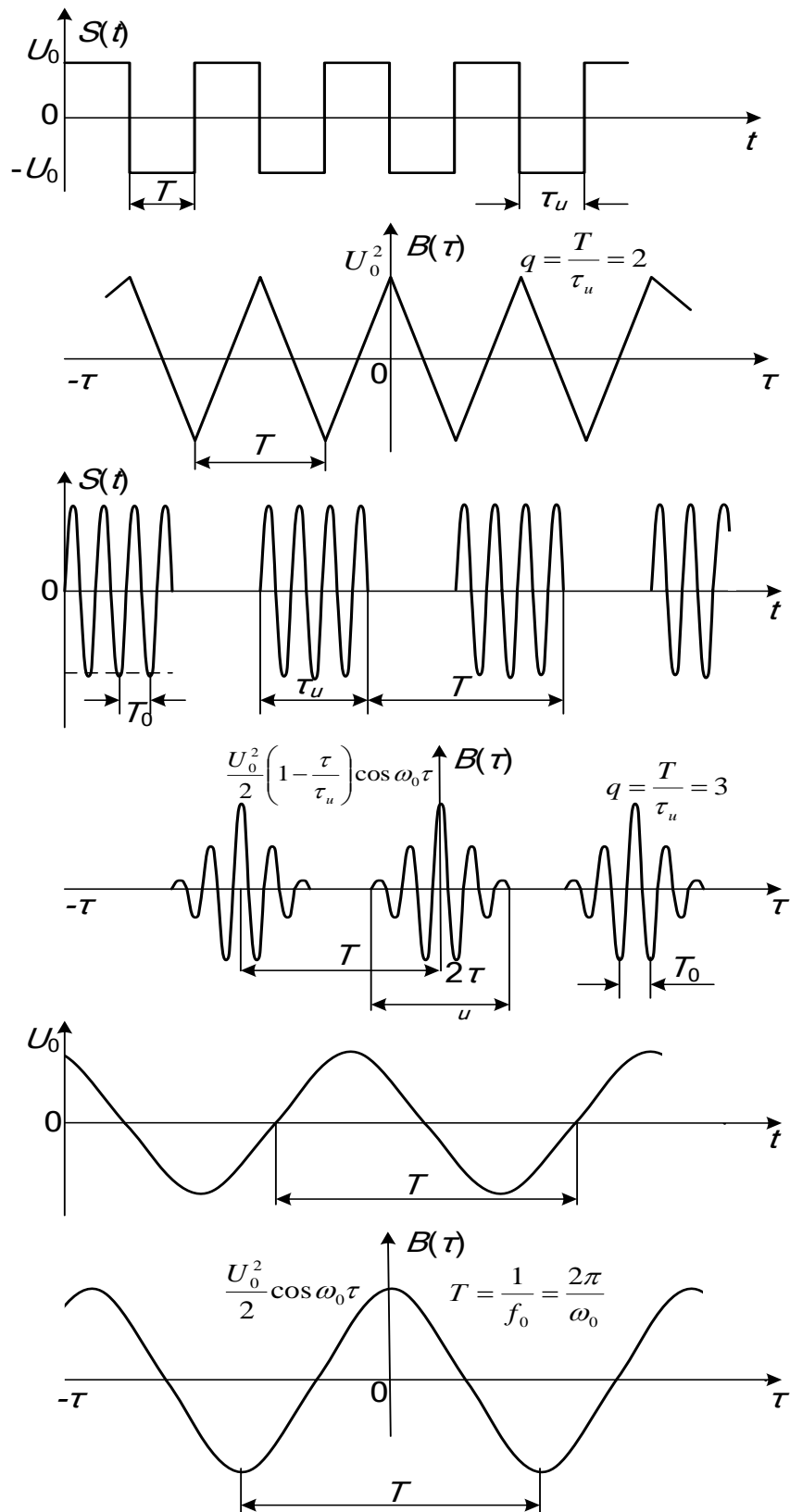


Рис. 2.34. Приклади побудови АКФ для деяких видів періодичних сигналів

2.5.4.2. Кореляційні функції випадкових процесів (сигналів)

Кореляційна (автокореляційна) функція випадкового процесу $X(t)$ встановлює ступінь статистичного зв'язку миттєвих значень процесу, узятих у різні моменти часу. Для центрованого нестационарного випадкового процесу вона визначається як математичне очікування добутку двох значень процесу у двох різних його перетинах при t_1 і t_2 :

$$\begin{aligned} R_X(t_1, t_2) &= M \{ [X(t_1) - \overline{X(t_1)}][X(t_2) - \overline{X(t_2)}] \} = \overline{X(t_1)X(t_2)} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1 - \overline{X(t_1)}][x_2 - \overline{X(t_2)}] \cdot w(x_1, x_2; t_1, t_2) dx_1 dx_2, \end{aligned} \quad (2.98)$$

де $w(x_1, x_2; t_1, t_2)$ - двовимірна щільність імовірності розподілу миттєвих значень випадкового процесу $X(t)$. Автокореляційна функція для нецентрованого нестационарного процесу $X(t)$, що називається коваріаційною, визначається з виразу

$$\begin{aligned} K_X(t_1, t_2) &= M[X(t_1)X(t_2)] = \overline{X(t_1)X(t_2)} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 w(x_1, x_2; t_1, t_2) dx_1 dx_2. \end{aligned} \quad (2.99)$$

З виразів (2.98) і (2.99) випливає, що кореляційна і коваріаційна функції випадкового нестационарного процесу залежать як від вибору його перетину (моментів часу t_1, t_2), так і від різниці моментів часу $\tau = t_2 - t_1$.

Між кореляційною і коваріаційною функціями існує такий зв'язок:

$$R_X(t_1, t_2) = K_X(t_1, t_2) = \overline{X(t_1)X(t_2)}. \quad (2.100)$$

Звідси випливає, що $R_X(t_1, t_2) \leq K_X(t_1, t_2)$.

При $t_2 - t_1 = 0$ вираз (2.100) визначає дисперсію випадкового процесу у момент часу $t = t_1$:

$$R_X(t_1) = K_X(t_1) - \overline{[X(t_1)]^2} = D_x(t_1).$$

Для стаціонарного випадкового процесу $X(t)$ в широкому розумінні автокореляційна функція визначається з виразу

$$\begin{aligned} R_X(\tau) &= M\{[X(t_1) - \bar{X}][X(t_2) - \bar{X}]\} = [X(t_1) - \bar{X}] \times \\ &\times [X(t_2) - \bar{X}] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [X(t_1) - \bar{X}][X(t_2) - \bar{X}] w_2(x_1, x_2) dx_1 dx_2, \end{aligned} \quad (2.101)$$

а коваріаційна функція знаходиться з виразу

$$\begin{aligned} K_X(\tau) &= M[X(t_1)X(t_2)] = \overline{X(t_1)X(t_2)} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 w_2(x_1, x_2) dx_1 dx_2. \end{aligned} \quad (2.102)$$

З виразів (2.101) і (2.102) випливає, що кореляційна і коваріаційна функції стаціонарного випадкового процесу не залежать від вибраних перетинів (моментів часу t_1 і t_2), а залежать від різниці $\tau = t_2 - t_1$ моментів цих перетинів. При $t_2 - t_1 = \tau = 0$ маємо

$$R_X(0) = \int_{-\infty}^{\infty} [x_1 - \bar{X}]^2 w_1(x_1) dx_1 = D_X = \sigma_X^2,$$

тобто кореляційна функція даного процесу, визначена для одного будь-якого перетину, дорівнює дисперсії процесу D_X або квадрату середньоквадратичного відхилення значень випадкового процесу від середнього значення.

Дисперсію можна також знайти з виразу

$$D_X = K_X(0) - (\bar{X})^2.$$

Для ергодичного стаціонарного випадкового процесу вираз АКФ записується як

$$R_X(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} [x(t) - m_X][x(t \pm \tau) - m_X] dt,$$

де m_X – математичне очікування випадкового процесу.

Вираз для коваріаційної функції процесу має вигляд:

$$K_X(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t \pm \tau) dt = R_X(\tau) + m_X^2.$$

Нормована АКФ стаціонарного ергодичного випадкового процесу визначається з виразу:

$$r_X(\tau) = R_X(\tau)/R_X(0),$$

а нормована коваріаційна функція знаходиться з виразу:

$$k_X(\tau) = K_X(\tau)/K_X(0).$$

Інтервал кореляції τ_k для випадкового процесу визначається найбільшим часовим інтервалом між двома перетинами процесу $\tau_k = \tau_{\max} = t_2 - t_1$, у межах якого значення функції кореляції зменшується до певного значення, наприклад до $0,1R_X(0)$.

2.5.4.3. Взаємні кореляційні функції різних випадкових процесів (сигналів)

Для оцінки статистичного зв'язку двох випадкових процесів $X(t)$ і $Y(t)$ вводиться поняття взаємної кореляційної функції (ВКФ).

Для нестационарних центрованих випадкових процесів вираз ВКФ має вигляд

$$R_{XY}(t_1, t_2) = \overline{X(t_1)Y(t_2)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1 - \overline{X(t_1)}][y_2 - \overline{Y(t_2)}] \times \\ \times w_2(x_1, y_2; t_1, t_2) dx_1 dy_2. \quad (2.103)$$

Вираз для взаємної коваріаційної функції записується як

$$K_{XY}(t_1, t_2) = \overline{X(t_1)Y(t_2)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 y_2 w_2(x_1, y_2; t_1, t_2) dx_1 dy_2 = \\ = R_{XY}(t_1, t_2) + \overline{X(t_1)Y(t_2)}, \quad (2.104)$$

де $w_2(x_1, y_2; t_1, t_2)$ - двовимірна щільність розподілу імовірності миттєвих значень випадкових процесів $X(t)$ і $Y(t)$.

Нормовані ВКФ визначаються з виразів

$$r_{XY}(t_1, t_2) = R_{XY}(t_1, t_2) / \sigma_X(t_1) \sigma_Y(t_2), \quad (2.105)$$

$$k_{XY}(t_1, t_2) = K_{XY}(t_1, t_2) / \sigma_X(t_1) \sigma_Y(t_2), \quad (2.106)$$

де $\sigma_X(t_1)$ і $\sigma_Y(t_2)$ - відповідно середньоквадратичне значення випадкових процесів $X(t)$ і $Y(t)$ в моменти часу t_1 і t_2 .

Для стаціонарних у широкому розумінні випадкових процесів вирази (2.103), (2.104), (2.105) і (2.106) запишуться як

$$R_{XY}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1 - \overline{X}][y_2 - \overline{Y}] w_2(x_1, y_2) dx_1 dy_2, \quad (2.107)$$

$$K_{XY}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 y_2 w_2(x_1, y_2) dx_1 dy_2 = R_{XY}(\tau) + \overline{X_1 Y_2}, \quad (2.108)$$

$$r_{XY}(\tau) = R_{XY}(\tau) / \sigma_X \sigma_Y, \quad (2.109)$$

$$k_{XY}(\tau) = K_{XY}(\tau) / \sigma_X \sigma_Y. \quad (2.110)$$

При $t_2 - t_1 = \tau = 0$ маємо $R_{XY}(0) = K_{XY}(0) = \sigma_X \sigma_Y$, $r_{XY}(0) = 1$, $k_{XY}(0) = 1 + K_{XY}(\tau) / \sigma_X \sigma_Y$.

Враховуючи, що $R_{XY}(\tau) = R_{XY}(-\tau)$, отримаємо $-1 \leq r_{XY}(\tau) \leq 1$.

Для ергодичних стаціонарних незалежних випадкових процесів вирази (2.107) і (2.108) запишуться як

$$R_{XY}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} [x(t) - m_x][y(t - \tau) - m_y] dt, \quad (2.111)$$

$$K_{XY}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)y(t - \tau) dt = R_{XY}(\tau) + m_X m_Y, \quad (2.112)$$

$$K_{YX}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} y(t)x(t - \tau) dt = R_{YX}(\tau) + m_X m_Y, \quad (2.113)$$

де $m_X m_Y$ – відповідно математичне очікування випадкових процесів $X(t)$ і $Y(t)$.

Для ергодичного стаціонарного випадкового процесу виконуються умови

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T R_X(\tau) d\tau = 0; \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T R_{XY}(\tau) d\tau = 0;$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T K_X(\tau) d\tau = m_X^2; \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T R_{XY}(\tau) d\tau = m_X m_Y.$$

Інтервал кореляції для взаємної функції кореляції визначається максимальним значенням часового зсуву $\tau_k = \tau_{max} = t_2 - t_1$ випадкового процесу $X(t)$ відносно процесу $Y(t)$ (або навпаки), при якому значення ВКФ досягає, наприклад, 0,1 від свого максимуму.

2.5.4.4. Зв'язок між кореляційною функцією стаціонарного випадкового процесу і його спектральною щільністю потужності

Для ергодичного стаціонарного випадкового процесу існує зв'язок між кореляційною функцією і його спектральною щільністю потужності (СЩП), визначуваний перетвореннями Фур'є за теоремою Вінера-Хінчина [19]:

$$G_X(\omega) = \int_{-\infty}^{\infty} R_X(\tau) e^{-j\omega\tau} d\tau, \quad (2.114)$$

$$R_X(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_X(\omega) e^{j\omega\tau} d\omega, \quad (2.115)$$

де $G_X(\omega)$ - спектральна щільність середньої потужності випадкового процесу, усереднена за всіма реалізаціями на інтервалі T (енергетичний спектр), $G_X(\omega) = \lim_{T \rightarrow \infty} \frac{|X_T(j\omega)|^2}{T}$.

Спектральна щільність середньої потужності процесу пов'язана з його потужністю (дисперсією) виразом

$$P_{X \sim} = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_X(\omega) d\omega = D_X = \sigma_X^2. \quad (2.116)$$

Кореляційна функція для випадкового процесу $X(t)$ у вигляді нормального (гаусівського) «білого» шуму з нульовим математичним очікуванням, часто використовуваного при розрахунках характеристик і параметрів систем передачі інформації, знаходиться з виразу

$$R_X(\tau) = \frac{N_0}{4\pi} \int_{-\infty}^{\infty} e^{j\omega\tau} d\omega = N_0 \delta(\tau), \quad (2.117)$$

де $\frac{N_0}{2}$ - двостороння спектральна щільність середньої потужності процесу, $\frac{N_0}{2} = G_X(\omega)$. При цьому враховано, що $\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega\tau} d\omega = \delta(\tau)$, а спектральна щільність нормального «білого» шуму рівномірна на всіх частотах ($-\infty < \omega < \infty$) і дорівнює $G_X(\omega) = N_0/2$.

Аналіз виразу (2.117) показує, що для «білого» шуму з нескінченним за частотою і рівномірним спектром середньої потужності кореляційна функція $R_X(\tau)$ дорівнює нулю при всіх значеннях часового зсуву τ , окрім $\tau = 0$, при якому $R_X(0) \rightarrow 0$, оскільки $\delta(0) \rightarrow \infty$. Отже, випадковий процес $X(t)$ з нескінченним спектром має інтервал кореляції, що прагне до нуля.

Подібний шум, що має голчасту структуру з нескінченно вузькими випадковими викидами в часі, іноді називають дельта-корельованим процесом. Дисперсія білого шуму, визначена з виразу

$$D_X = R_X(0) = \frac{N_0}{2} \delta(0),$$

є нескінченною (рис. 2.35).

Кореляційна функція для стаціонарного гаусівського процесу з нульовим середнім значенням і рівномірним одностороннім розподілом енергетичного спектра N_0 в межах до Ω_H визначається виразом

$$R_1(\tau) = \frac{N_0 \Omega_H}{2\pi} \text{sinc}(\Omega_H \tau) = N_0 F_B \text{sinc}(\Omega_H \tau), \quad (2.118)$$

де Ω_H - гранична частота фільтра нижніх частот, $\Omega_H = 2\pi F_H$.

Випадковий процес, у якого $0 < F_H < 60$ кГц, називається низькочастотним.

Для високочастотного стаціонарного гаусівського процесу з нульовим середнім значенням і рівномірним двостороннім розподілом енергетичного спектра $N_0/2$ в межах від 0 до ω_B визначається виразом

$$R_2(\tau) = \frac{N_0 \omega_B}{2\pi} \frac{\text{sinc}(\omega_B \tau)}{\omega_B \tau} = N_0 f_B \frac{\text{sinc}(\omega_B \tau)}{\omega_B \tau}, \quad (2.119)$$

де $\omega_B = 2\pi f_B$, $\omega_B \gg \Omega_H$.

У високочастотного випадкового процесу $f_B > 60$ кГц.

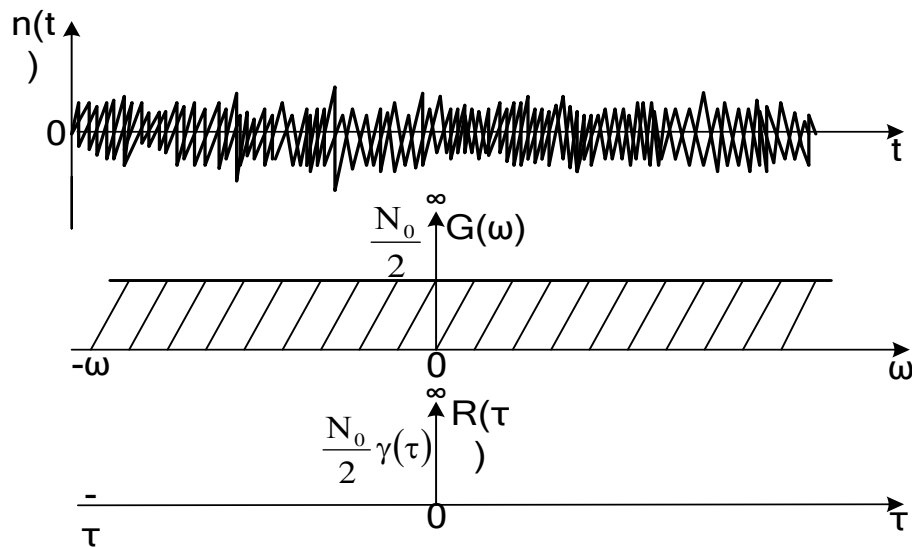


Рис. 2.35. Дисперсія і кореляційна функція «білого» шуму

Кореляційна функція вузькосмугового гаусівського процесу з нульовим середнім значенням і рівномірним одностороннім розподілом енергетичного спектра N_0 в межах $\omega_0 - \Omega_1 \leq \omega \leq \omega_0 + \Omega_2$ знаходиться з виразу

$$R_3(\tau) = \frac{N_0 \Delta\omega}{\pi} \text{sinc}\left(\frac{\Delta\omega\tau}{2}\right) \cos \omega_0 \tau, \quad (2.120)$$

де $\Delta\omega$ - смуга пропускання вузькосмугового фільтра, $\Delta\omega = \omega_2 - \omega_1$;

ω_0 - резонансна частота вузькосмугового фільтра, на виході якого визначається кореляційна функція.

Вузькосмуговим називається фільтр, у якого резонансна частота f_0 багато більше за смугу пропускання, тобто $f_0 \gg \Delta f$,

$\omega_0 = \frac{\omega_2 + \omega_1}{2}$, $\Delta\omega = \omega_2 - \omega_1$. Графіки нормованих кореляційних функцій з урахуванням виразів (2.118)-(2.120) наведені на рис. 2.36.

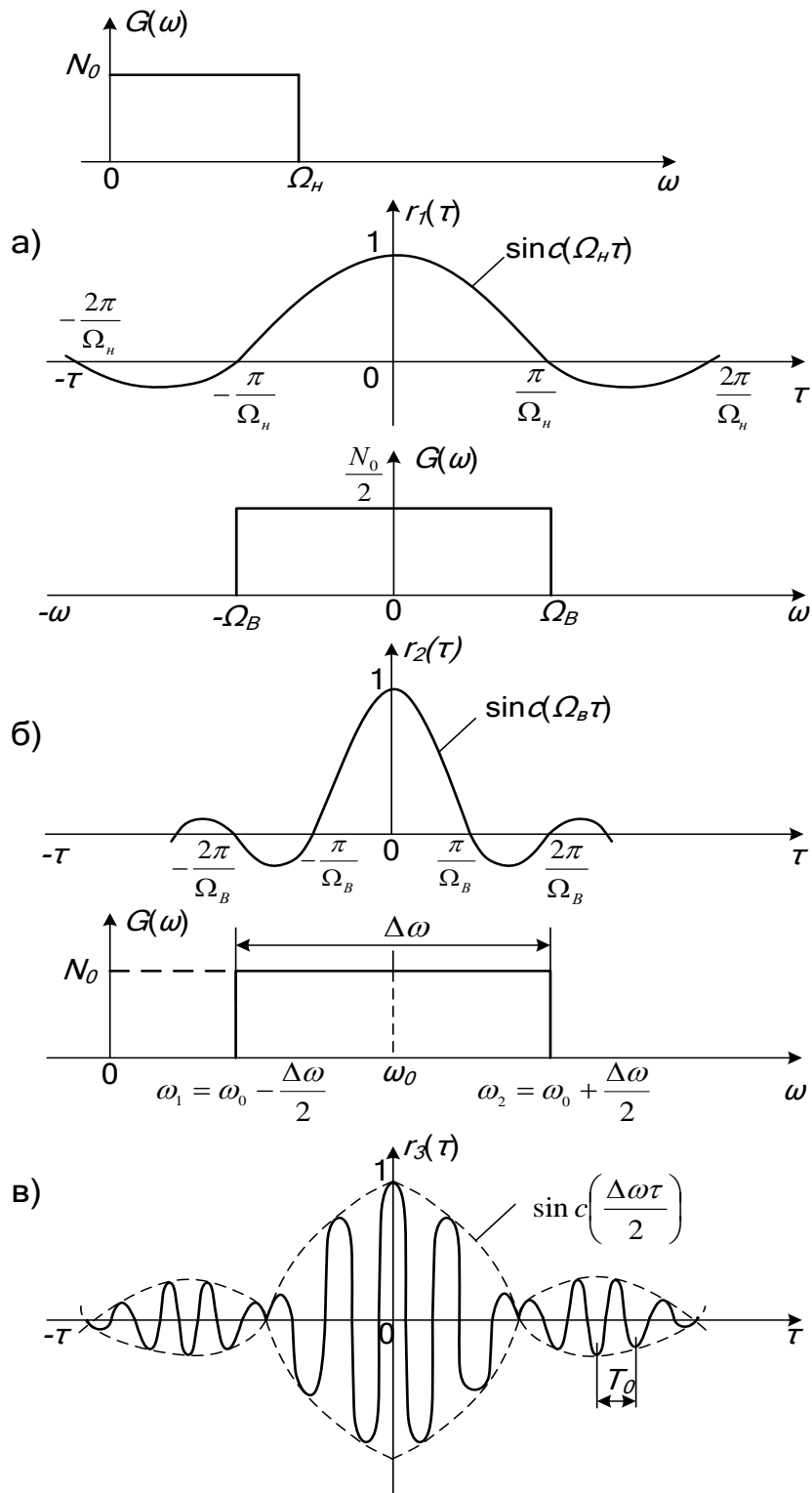


Рис. 2.36. Графіки нормованих кореляційних функцій деяких випадкових процесів

За ними можна визначити інтервал кореляції τ_k , тобто порівняти випадкові процеси за швидкістю зміни в часі їх миттєвих значень (напруги або струму).

На рис. 2.37 показана результуюча взаємкореляційна функція суміші детермінованого гармонійного сигналу $U_c(t) = U_0 \cos \omega_0 t$ з адитивним стаціонарним гаусівським шумом $n(t)$.

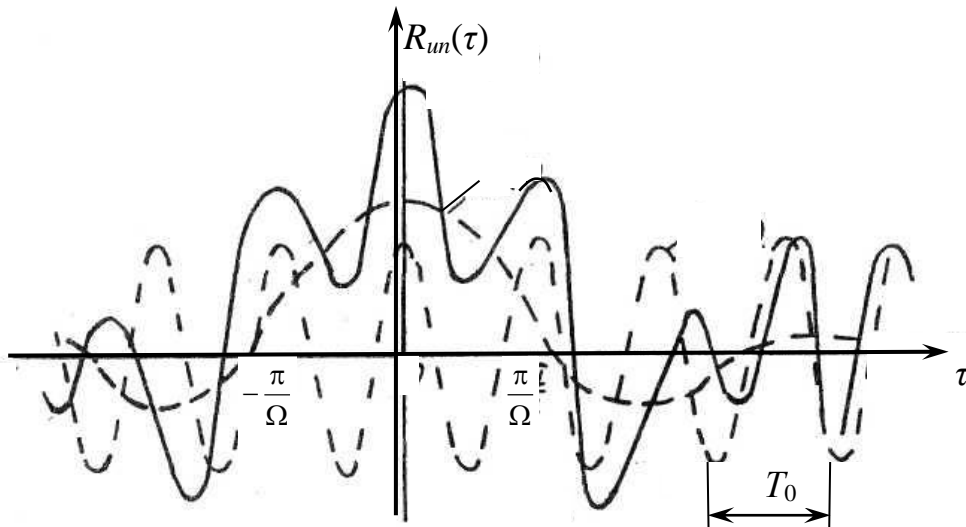


Рис. 2.37. Результуюча ВКФ гармонійного сигналу з адитивним стаціонарним гаусівським шумом

З рис. 2.37 випливає, що після закінчення інтервалу кореляції τ_k кореляційна функція суміші цих сигналів визначається в основному автокореляційною функцією гармонійного сигналу $R_c(\tau)$, що указує на існуюче ослаблення за межами цього інтервалу зв'язку між двома процесами унаслідок загасання кореляційної функції шуму $R_{ш}(\tau)$.

2.6. Аналітичний вузькосмуговий сигнал

2.6.1. Математичні моделі і характеристики аналітичного сигналу

Для багатьох практичних додатків буває корисним подати сигнал $x(t)$ у вигляді процесу зі змінюваною амплітудою обвідної $A(t)$ і повною фазою $\psi(t)$

$$x(t) = A(t) \cos \psi(t).$$

Так, при амплітудній модуляції інформація закладається в $A(t)$, при кутовій в $-\psi(t)$. Подання в загальному випадку є неоднозначним, тобто один і той самий сигнал $x(t)$ може бути поданий нескінченною множиною пар $\{A(t), \psi(t)\}$. З кожної такої пари може бути утворений новий сигнал $x'(t) = A(t) \sin \psi(t)$, у деякому розумінні сполучений із сигналом $x(t)$. Відомо, що $A(t) = \sqrt{x^2(t) + [x'(t)]^2}$, $\psi(t) = \arctg \frac{x'(t)}{x(t)}$. Комплексний сигнал

$\dot{x}(t) = x(t) + jx'(t) = A(t)e^{j\psi(t)}$ є зручним відображенням дійсного сигналу $x(t) = \text{Re}(\dot{x}(t))$.

Із співвідношення $x'(t) = A(t) \sin \psi(t) = A(t) \cos[\psi(t) - \pi/2]$ не впливає (при довільних функціях), що амплітудні спектри $\dot{x}(t)$ та $x(t)$ однакові, а фазові спектри розрізняються зсувом на $\pi/2$.

Сигнал $x(t)$ без постійної складової у вигляді нескінченної суми гармонійних складових для $\omega_k > 0$

$$x(t) = \sum_{(k)}^{\infty} a_k \cos(\omega_k t + \varphi_k) = \sum_{(k)} \frac{a_k}{2} \left[e^{j(\omega_k t + \varphi_k)} - e^{-j(\omega_k t + \varphi_k)} \right] \quad (2.121)$$

і забажаємо, щоб у сполученому з ним сигналі $\hat{x}(t)$ всі частотні складові мали ті ж амплітуди a_k , і при цьому фази одержали зсув, що дорівнюватиме величині фази $-\pi/2$:

$$\begin{aligned} x'(t) &= \sum_{(k)}^{\infty} a_k \cos\left(\omega_k t + \varphi_k - \frac{\pi}{2}\right) = \sum_{(k)} a_k \sin(\omega_k t + \varphi_k) = \\ &= \sum_{(k)} \frac{a_k}{2j} \left[e^{j(\omega_k t + \varphi_k)} - e^{-j(\omega_k t + \varphi_k)} \right] = \hat{x}(t). \end{aligned} \quad (2.122)$$

Зіставивши (2.121) і (2.122), бачимо, що в області додатних частот спектри сигналів $x(t)$ і $\hat{x}(t)$ розрізняються множителем

$e^{j(\omega_k t + \varphi_k)}$, а в області від'ємних частот – множителем $e^{-j(\omega_k t + \varphi_k)}$. Процес $\hat{x}(t)$ називають сигналом, сполученим за Гілбертом з $x(t)$. Передатна функція кола, яке перетворює $x(t)$ на $\hat{x}(t)$ (тобто перетворювача Гілберта (ПГ) (див. рис. 2.38), має вигляд

$$\dot{K}(f) = -j \cdot \text{sign}(f),$$

де $\text{sign}(f)$ – знакова функція, $\text{sign}(f) = \begin{cases} 1, & f > 0, \\ -1, & f < 0. \end{cases}$

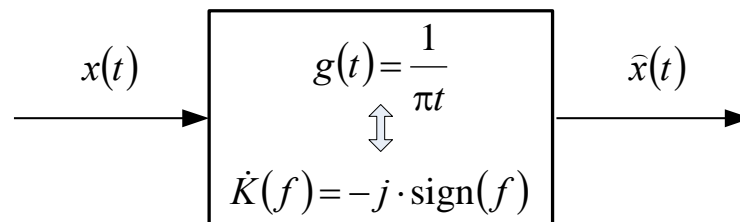


Рис. 2.38. Умовне позначення кола перетворювача Гілберта

Імпульсна характеристика кола визначається зворотним перетворенням Фур'є від $\dot{K}(f)$

$$\begin{aligned} g(t) &= \int_{-\infty}^{\infty} \dot{K}(f) e^{j\omega t} df = 2 \int_0^{\infty} \sin(\omega t) df = 2 \lim_{a \rightarrow 0} \int_0^{\infty} e^{-af} \sin(\omega t) df = \\ &= 2 \lim_{a \rightarrow 0} \frac{2\pi t}{a^2 + (2\pi t)^2} = \frac{1}{\pi t}. \end{aligned}$$

Спектральна щільність $\dot{S}_{\hat{x}}(f)$ для сполученого сигналу $\hat{x}(t)$ зв'язана зі спектром Фур'є $\dot{S}_x(f)$ сигналу $x(t)$ співвідношенням

$$\dot{S}_{\hat{x}}(f) = \dot{S}_x(f) \cdot \dot{K}(f) = -j \cdot \text{sign}(f) \cdot \dot{S}_x(f). \quad (2.123)$$

Вираз можна записати і так: $\dot{S}_x(f) = j \cdot \text{sign}(f) \cdot \dot{S}_{\hat{x}}(f)$. Звідси випливає, що коло, яке здійснює зворотне перетворення Гілберта, передатна функція або імпульсна характеристика кола подачі, яка видає на своєму виході сигнал $x(t)$, при подачі до входу $\hat{x}(t)$ обернені за знаком характеристикам, вказаним на рис. 2.38. Відгук лінійного стаціонарного ПГ $\hat{x}(t)$ можна визначити як згортку функцій $x(t)$ і $g(t)$. Він має вигляд

$$\hat{x}(t) = x(t) \otimes g(t) = \int_{-\infty}^{\infty} x(\tau) g(t - \tau) d\tau = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau. \quad (2.124)$$

Аналогічно можна записати

$$x(t) = \hat{x}(t) \otimes \left(-\frac{1}{\pi t} \right) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\hat{x}(\tau)}{t - \tau} d\tau. \quad (2.125)$$

Вираз (2.124) називають прямим, а (2.125) – оберненим перетворенням Гілберта. Разом їх називають парою перетворень Гілберта.

Зазначимо, що для довільних сигналів $x(t)$ і $\hat{x}(t)$ перетворення Гілберта не реалізуються, тому що вони вимагають імпульсної характеристики кола, визначеної не тільки при $t > 0$, але і при $t < 0$, а для реалізованого кола $g(t) = 0$ і при $t < 0$. Якщо все ж таки ПГ таких сигналів необхідне, його реалізують приблизно, з деякою затримкою t_0 , свідомо відкидаючи гілки $g(t)$, які розташовуються лівіше точки $t = -t_0$ і правіше точки $t = t_0$.

Затримка сигналу повинна бути врахована в інших пристроях, які працюють синхронно з ПГ. Слід зазначити, що похибка перетворення, пов'язана з урізанням імпульсної характеристики ПГ, може виявитися неприпустимо великою. Однак перетворення Гілберта просто реалізуються, якщо сигнал $x(t)$ можна подати через вузькосмугові квадратурні компоненти.

Комплексний сигнал $\dot{x}(t) = x(t) + j \cdot \widehat{x}(t)$, отриманий на основі ПГ, називається *аналітичним*. Розглянемо його основні властивості.

1. Аналітичний сигнал є природним узагальненням комплексного подання гармонійного сигналу на випадок сигналу загального вигляду. Справді, якщо $x(t) = A \cos(\omega t + \varphi)$, то $\dot{x}(t) = A \cos(\omega t + \varphi) + j A \sin(\omega t + \varphi) = A(t) e^{j(\omega t + \varphi)}$. При цьому повна фаза дорівнює $\psi(t) = \omega t + \varphi$, а миттєва частота відповідно $\omega(t) = \frac{d\psi(t)}{dt} = \omega$.

2. У спектрі аналітичного сигналу $\dot{x}(t) = x(t) + j \cdot \widehat{x}(t)$ містяться тільки додатні частоти. Дійсно, з виразів (2.124) і (2.125) видно, що

$$\dot{x}(t) = x(t) + j \cdot \widehat{x}(t) = \sum_{(k)} a_k e^{j(\omega_k t + \varphi_k)} \quad \text{при } \omega_k > 0. \quad (2.126)$$

У загальному випадку, коли спектр $\dot{S}_x(f)$ сигналу $x(t)$ є суцільним, то, врахувавши вираз (2.123), маємо

$$\dot{S}_{\widehat{x}}(f) = \dot{S}_x(f) + j \cdot \dot{S}_{\widehat{x}}(f) = \begin{cases} 0, & f < 0, \\ 2\dot{S}_x(f), & f > 0. \end{cases} \quad (2.127)$$

Аналогічно в спектрі комплексно-сполученого аналітичного сигналу $x^*(t) = x(t) + j \cdot \widehat{x}(t)$ містяться тільки від'ємні частоти

$$\dot{S}_{x^*}(f) = \begin{cases} 0, & f > 0, \\ 2\dot{S}_x(f), & f < 0. \end{cases}$$

3. Скалярний добуток сигналів $x(t)$ і $\widehat{x}(t)$ в просторі Гілберта дорівнює нулю (сигнали ортогональні)

$$\begin{aligned}
(\mathbf{x}, \widehat{\mathbf{x}}) &= \int_{-\infty}^{\infty} x(t)\widehat{x}(t)dt = \int_{-\infty}^{\infty} \dot{S}_x(f)S_{\widehat{x}}^*(f)df = \\
&= \int_{-\infty}^{\infty} S_x^2(f)\text{sign}(f)df = 0.
\end{aligned}
\tag{2.128}$$

4. При загальному фазовому зсуві всіх частотних компонент сигналу $x(t)$ на кут θ аналітичний сигнал збільшується на $e^{j\theta}$. У цьому легко переконатися, додавши до степені e виразу (2.126) до кожного доданку загальний фазовий зсув θ :

$$\dot{x}_\theta(t) = \sum_{(k)} a_k e^{j(\omega_k t + \varphi_k + \theta)} = e^{j\theta} \dot{x}(t).
\tag{2.129}$$

З виразу (2.129) природним чином впливає алгоритм загального фазового зсуву всіх частотних компонент дійсного сигналу $x(t)$:

$$\begin{aligned}
x_\theta(t) &= \text{Re}[\dot{x}_\theta(t)] = \text{Re}[e^{j\theta} \dot{x}(t)] = \text{Re}[(\cos\theta + j \cdot \sin\theta) \times \\
&\times (x(t) + j\widehat{x}(t))] = x(t) \cdot \cos\theta - \widehat{x}(t) \cdot \sin\theta.
\end{aligned}
\tag{2.130}$$

5. Перетворення частоти (транспонування спектра) сигналу $x(t)$ зводиться до множення аналітичного сигналу $\dot{x}(t)$ на $e^{j\Delta\omega t}$.

З цього впливає вираз

$$\dot{x}_{\Delta\omega}(t) = \sum_{(k)} a_k e^{j[(\omega_k + \Delta\omega)t + \varphi_k]} = e^{j\Delta\omega t} \dot{x}(t).
\tag{2.131}$$

З виразу (2.131) впливає формула транспонування спектра дійсного сигналу $x(t)$ такого вигляду:

$$x_{\Delta\omega}(t) = \text{Re}[\dot{x}_{\Delta\omega}(t)] = x(t) \cdot \cos\Delta\omega t - \widehat{x}(t) \cdot \sin\Delta\omega t.
\tag{2.132}$$

Формула (2.132) використовується при фазокомпенсаційному методі односмугової модуляції (коли $\Delta\omega = \omega_0$ – частота несучої), при корекції частотних зсувів, обумовлених неточним сполученням синтезаторів частот на передачі і на прийомі або доплерівськими зсувами частоти, і при демодуляції сигналу.

Операції, пов'язані з ПГ, стають реалізованими й істотно спрощуються, якщо сигнал $x(t)$ можна представити через квадратурні компоненти $A_c(t)$, $A_s(t)$ при заданій частоті (як правило, в спектрі сигналу) ω_0 :

$$x(t) = A_c(t) \cdot \cos \omega_0 t - A_s(t) \cdot \sin \omega_0 t, \quad (2.133)$$

причому верхня частота в спектрі сигналів $A_c(t)$, $A_s(t)$

$$F_B < f_0. \quad (2.134)$$

Умову (2.134) будемо називати умовою “вузкосмуговості в розширеному розумінні”. Найчастіше каналні сигнали в системах зв'язку мають властивість вузкосмуговості у вузькому розумінні

$$F_B \ll f_0. \quad (2.135)$$

Це означає, що квадратурні компоненти $A_c(t)$, $A_s(t)$ змінюються повільно (порівняно з $\cos \omega_0 t$ або $\sin \omega_0 t$) і сигнал (вираз (2.133)) має вигляд квазігармонічного сигналу (рис. 2.39). Саме в цьому випадку обробка сигналу (через обробку низькочастотних квадратурних компонент) технічно простіша і реалізується точніше.

Очевидно, що вираз (2.123) можна записати як

$$x(t) = A(t) \cdot \cos[\omega_0 t + \varphi(t)] = \operatorname{Re} \left[\dot{A}(t) \cdot e^{j\omega_0 t} \right], \quad (2.136)$$

де $\dot{A}(t)$ – комплексна амплітуда, $\dot{A}(t) = A(t) e^{j\varphi(t)} = A_c(t) + jA_s(t)$.

Оскільки множення $\dot{A}(t)$ на $e^{j\omega_0 t}$ означає перенесення спектра сигналу $\dot{A}(t)$ нагору на величину f_0 , то очевидно, що при виконанні умови (2.135) спектр сигналу $\dot{x}(t) = \dot{A}(t)e^{j\omega_0 t}$ містить тільки додатні частоти $f_{min} > 0$, $f_{max} < f + F_B$ і може розглядатися як аналітичний. Тоді сполучений, за Гілбертом, сигнал набуде вигляду

$$\hat{x}(t) = \text{Im}[\dot{A}(t)e^{j\omega_0 t}] = A_c(t)\sin\omega_0 t + A_s(t)\cos\omega_0 t. \quad (2.137)$$

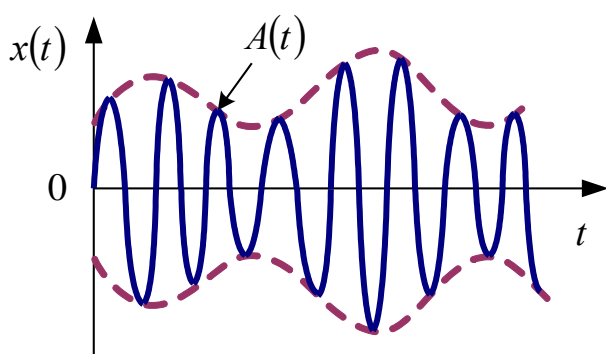


Рис. 2.39. Характерний вигляд вузькосмугового процесу (сигналу)

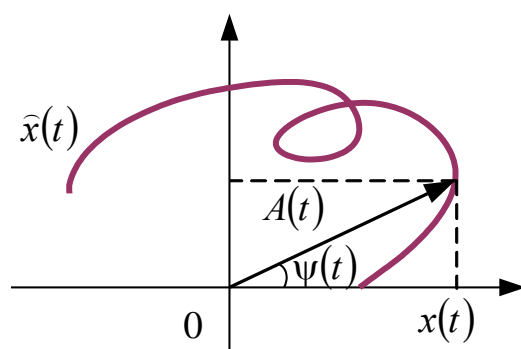


Рис. 2.40. Геометричне зображення аналітичного сигналу зі змінною обвідного $A(t)$

і змінною миттєвою частотою

$$\omega(t) = \frac{d\psi(t)}{dt}$$

З порівняння виразів (2.136) і (2.137) видно, що перетворення Гілберта сигналу $A_c(t)\cos\omega_0 t$ можна записати як

$$H[A_c(t)\cos\omega_0 t] = A_c(t)\sin\omega_0 t,$$

а перетворення Гілберта сигналу $A_s(t)\sin\omega_0 t$ – в такому вигляді:

$$H[A_s(t)\sin\omega_0 t] = -A_s(t)\cos\omega_0 t. \quad (2.138)$$

Формально вирази (2.137) і (2.138) можна інтерпретувати так, що перетворення Гілберта здійснюється над функціями $\cos \omega_0 t$ і $\sin \omega_0 t$

$$H[\cos \omega_0 t] = \cos\left(\omega_0 t - \frac{\pi}{2}\right) = \sin \omega_0 t,$$

$$H[\sin \omega_0 t] = \sin\left(\omega_0 t - \frac{\pi}{2}\right) = -\cos \omega_0 t$$

при незмінності $A_c(t)$ і $A_s(t)$.

З виразів (2.136) і (2.137) видно, що обвідну сигналу має вигляд

$$A(t) = |\dot{x}(t)| = \sqrt{x^2(t) + \hat{x}^2(t)} = \sqrt{A_c^2(t) + A_s^2(t)}, \quad (2.139)$$

а його повна фаза дорівнює

$$\Psi(t) = \varphi_0(t) + \arg \dot{A}(t) = \varphi_0(t) + \arg \operatorname{tg} \frac{A_s(t)}{A_c(t)}. \quad (2.140)$$

Миттєва частота сигналу (швидкість зміни фази) записується як

$$\begin{aligned} \omega(t) &= \frac{d\Psi(t)}{dt} = \omega_0 + \frac{d}{dt} \left[\operatorname{arctg} \frac{A_s(t)}{A_c(t)} \right] = \\ &= \omega_0 + \frac{A'_s(t)A_c(t) - A'_c(t)A_s(t)}{A_c^2(t)A_s^2(t)}. \end{aligned} \quad (2.141)$$

На комплексній площині (рис. 2.40) аналітичний сигнал $\dot{x}(t) = \dot{A}(t) \cdot e^{j[\omega_0 t + \varphi(t)]}$ у загальному випадку відображається вектором обертання, що має змінну довжину $A(t)$. Його кутова швидкість обертання $\omega(t)$ змінюється в часі.

При поданні вузькосмугового в розширеному розумінні сигналу через квадратурні компоненти алгоритм (2.136) реалізується досить просто:

$$x_{\theta}(t) = A_{c,\theta}(t)\cos\omega_0 t - A_{s,\theta}(t)\sin\omega_0 t,$$

де $A_{c,\theta}(t) = A_c(t)\cos\theta - A_s(t)\sin\theta$; $A_{s,\theta}(t) = A_s(t)\cos\theta - A_c(t)\sin\theta$.

Алгоритм перетворення частоти (2.131) можна подати як

$$x_{\Delta\omega}(t) = A_{c,\Delta\omega}(t)\cos\omega_0 t - A_{s,\Delta\omega}(t)\sin\omega_0 t,$$

де $A_{c,\Delta\omega}(t) = A_c(t)\cos\Delta\omega t - A_s(t)\sin\Delta\omega t$;

$$A_{s,\Delta\omega}(t) = A_s(t)\cos\Delta\omega t - A_c(t)\sin\Delta\omega t.$$

Використовуючи поняття сполученого, за Гілбертом, сигналу, Л.М. Фінк ввів дуже важливе для теорії і практики електрозв'язку поняття ортогональності розуміння пари сигналів $x(t)$ і $y(t)$ [48].

Сигнали $x(t)$ та $y(t)$ ортогональні в посиленому розумінні в просторі L_2 , якщо одночасно подані набуває вигляду

$$(x, y) = \int_{-\infty}^{\infty} x(t)y(t)dt = 0, \quad (x, \hat{y}) = \int_{-\infty}^{\infty} x(t)\hat{y}(t)dt = 0. \quad (2.142)$$

Корисність поняття ортогональності в посиленому розумінні пояснюється тим, що сигнали $x(t)$ і $y(t)$, пройшовши через канал зв'язку, часто одержують випадкові фазові зсуви θ_x і θ_y . При виконанні умов (2.142) ортогональність сигналів зберігається при довільних фазових зсувах θ_x і θ_y .

Дійсно, подамо сигнали $x(t)$ і $y(t)$ через обвідну $A(t)$ і повну фазу як

$$x(t) = A_x(t)\cos\psi_x(t), \quad y(t) = A_y(t)\cos\psi_y(t).$$

Сигнали, що одержали фазові зсуви θ_x і θ_y , запишемо як, враховуючи вираз (2.142),

$$x_{\theta_x}(t) = \cos \theta_x x(t) - \sin \theta_x \hat{x}(t), \quad y_{\theta_y}(t) = \cos \theta_y y(t) - \sin \theta_y \hat{y}(t).$$

Ортогональність сигналів $x_{\theta_x}(t)$ і $y_{\theta_y}(t)$ виражається як

$$(x_{\theta}, y_{\theta}) = \int_{-\infty}^{\infty} x_{\theta_x}(t) y_{\theta_y}(t) dt = 0 \quad (2.143)$$

і буде забезпечена тільки в тому випадку, якщо сигнали $x(t)$ і $y(t)$ ортогональні в посиленому розумінні (умова (2.142)).

Часто використовувана в техніці зв'язку система ортогональних у посиленому розумінні сигналів, визначених на інтервалі $-T/2; T/2$, будується на основі тригонометричних функцій кратних частот (ансамблю сигналів)

$$\{\cos \omega_1 t, \cos 2\omega_1 t, \dots, \cos k\omega_1 t, \dots\}. \quad (2.144)$$

Звичайно в ансамблі (2.144) міститься кінцева кількість компонент N . Якщо ансамбль (2.144) доповнити елементами $\sin k\omega_1 t$, то отримаємо ансамбль ортогональних сигналів розмірності $2N$. Якщо кожен елемент останнього ансамблю доповнити протилежним елементом $(-\cos k\omega_1 t, -\sin k\omega_1 t)$, то отримаємо часто використовуваний при передачі дискретних повідомлень ансамбль сигналів, який називають біортогональним. Кількість елементів цього ансамблю $4N$.

2.6.2. Імовірнісні характеристики обвідної і фази вузькосмугового випадкового гаусівського процесу

Для визначення ЩІ обвідної A і фази φ вузькосмугового випадкового процесу будемо виходити із спільної щільності

імовірності $\omega(A_c, A_s)$ квадратурних компонент A_c і A_s у фіксованих перетинах ВП. Вважаючи процес гаусівським, ЩВ набуває наступного вигляду:

$$\omega(A_c, A_s) = \omega(A_c)\omega(A_s) = \tag{2.145}$$

$$= \frac{1}{\sqrt{2\pi\sigma_c^2}} e^{-\frac{(A_c - m_c)^2}{2\sigma_c^2}} \cdot \frac{1}{\sqrt{2\pi\sigma_s^2}} e^{-\frac{(A_s - m_s)^2}{2\sigma_s^2}},$$

де $m_c, m_s, \sigma_c^2, \sigma_s^2$ – МО і дисперсії квадратурних компонент.

Подання виразу (2.145) припускає, що квадратурні компоненти A_c, A_s – незалежні випадкові величини, які завжди можна забезпечити. Дійсно, відомо, що гаусівський вектор комплексної обвідної $\dot{A}(t)$ можна так розташувати в декартовій системі координат, що його проєкції (A_c, A_s) виявляться некорельованими (отже, і статично незалежними при гаусівському розподілу).

Для спрощення аналізу обмежимося випадком, коли дисперсії квадратурних компонентів дорівнюють $\sigma_c^2 = \sigma_s^2 = \sigma^2$. Тоді з виразу (2.145) випливає

$$\omega(A_c, A_s) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ - \left(\frac{A_c^2 + A_s^2}{2\sigma^2} - \frac{m_c^2 + m_s^2}{2\sigma^2} + \frac{A_c m_c + A_s m_s}{\sigma^2} \right) \right\}. \tag{2.146}$$

Перейдемо від декартових координат A_c, A_s до полярних координат $\varphi = \arctg\left(\frac{A_s}{A_c}\right)$. Тоді $A_c = A \cos \varphi, A_s = A \sin \varphi$.

Приріст площі в декартових координатах дорівнює $dA_x dA_c$ (див. рис. 2.41, а), а приріст площі тієї ж величини в полярних координатах запишеться формулою $d\varphi dA$ (див. рис. 2.41, б).

Порівнюючи імовірності елементарних подій у двох системах координат, одержуємо

$$\omega(A_c, A_s) dA_c dA_s = \omega(A_c, A_s) d\varphi dA = \omega(A, \varphi) dA d\varphi,$$

звідки випливає, що $\omega(A, \varphi) = w(A \cos \varphi, A \sin \varphi) A$.

Враховуючи вираз (2.145), одержимо вираз для $\omega(A, \varphi)$ у такому вигляді:

$$\omega(A, \varphi) = \frac{A}{2\pi\sigma^2} e^{\left(-\frac{A^2}{2\sigma^2} - \frac{m_c^2 + m_s^2}{2\sigma^2}\right)} \cdot e^{\frac{A}{2\sigma^2} (m_c \cos \varphi + m_s \sin \varphi)}. \quad (2.147)$$

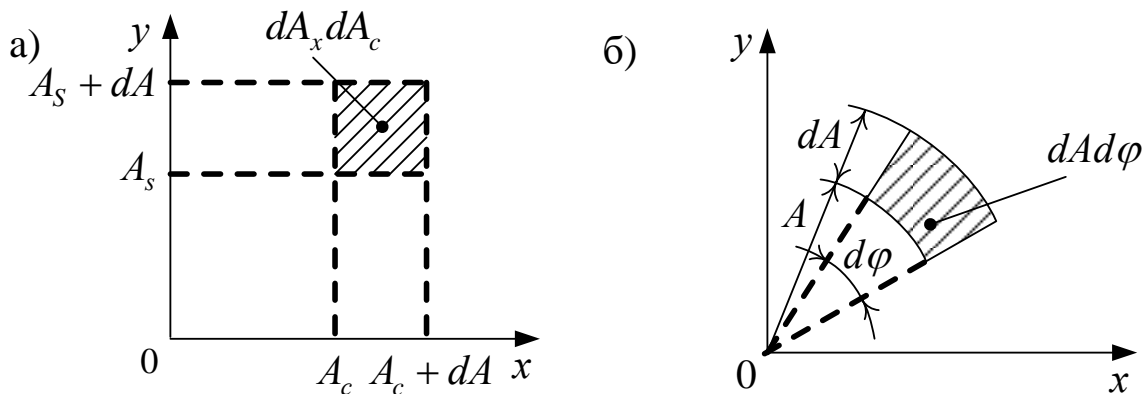


Рис. 2.41. Визначення приросту площі у декартових (а) і полярних (б) координатах

Одномірну щільність імовірності обвідної отримаємо, інтегруючи вираз (2.147) за всіма можливими значеннями φ на інтервалі періодичності гармонійної функції $(-\pi; +\pi)$:

$$\begin{aligned}\omega(A) &= \int_{-\pi}^{\pi} \omega(A, \varphi) d\varphi = \\ &= \frac{A^2}{2\sigma^2} e^{-\frac{A^2}{2\sigma^2} - \frac{m_c^2 + m_s^2}{2\sigma^2}} \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\frac{A^2}{\sigma^2} \sqrt{m_c^2 + m_s^2} \cos(\varphi - \theta)} d\varphi,\end{aligned}\quad (2.148)$$

де $\theta = \operatorname{arctg} \frac{m_s}{m_c}$.

Інтегрування у виразі (2.147) можна виконувати за допомогою табличного інтеграла

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{C \cos(\varphi - \theta)} d\varphi = I_0(C),$$

де $I_0(C)$ – модифікована функція Бесселя нульового порядку (при $C = 0$ маємо $I_0(C) = 1$ і при зростанні $|C|$ функція безупинно зростає).

Після інтегрування одержимо результат

$$\omega(A) = \frac{A}{\sigma^2} e^{-\frac{A^2 + m_c^2 + m_s^2}{2\sigma^2}} I_0\left(\frac{A}{\sigma^2} \sqrt{m_c^2 + m_s^2}\right).\quad (2.149)$$

Формула (2.149) виражає узагальнений закон Релея або закон Райса [19].

Вузькосмуговий випадковий процес $X(t)$ з регулярною частиною $X_p(t)$ з відмінним від нуля МО можна подати, як

$$X(t) = (\bar{A}_c + m_c) \cos \omega_0 t - (\bar{A}_s + m_s) \sin \omega_0 t = \overline{X(t)} + X_p(t).\quad (2.150)$$

Звідси видно, що амплітуда регулярної частини сигналу буде дорівнювати

$$A_p = \sqrt{m_c^2 + m_s^2}. \quad (2.151)$$

Враховавши вираз (2.149) запишемо розподіл Райса в такому вигляді:

$$\omega(A) = \frac{A}{\sigma^2} e^{-\frac{A^2 + A_p^2}{2\sigma^2}} \cdot I_0\left(\frac{AA_p}{\sigma^2}\right), \quad A \geq 0. \quad (2.152)$$

Ця залежність при різних параметрах $\frac{A_p}{\sigma}$ зображена на рис. 2.42, а.

Одноименну щільність імовірності фази одержимо, інтегруючи вираз (2.148) за всіма можливими значеннями A (від 0 до ∞).

$$\omega(\varphi) = \frac{1}{2\pi} e^{-\frac{A_p^2}{2\sigma^2}} + \frac{A_p \cos(\varphi - \theta)}{\sqrt{2\pi\sigma^2}} \left(1 - Q\left(\frac{A_p \cos(\varphi - \theta)}{2\sigma^2}\right)\right) \cdot e^{-\frac{A_p^2 \sin^2(\varphi - \theta)}{2\sigma^2}},$$

де $-\pi \leq \varphi - \theta \leq +\pi$.

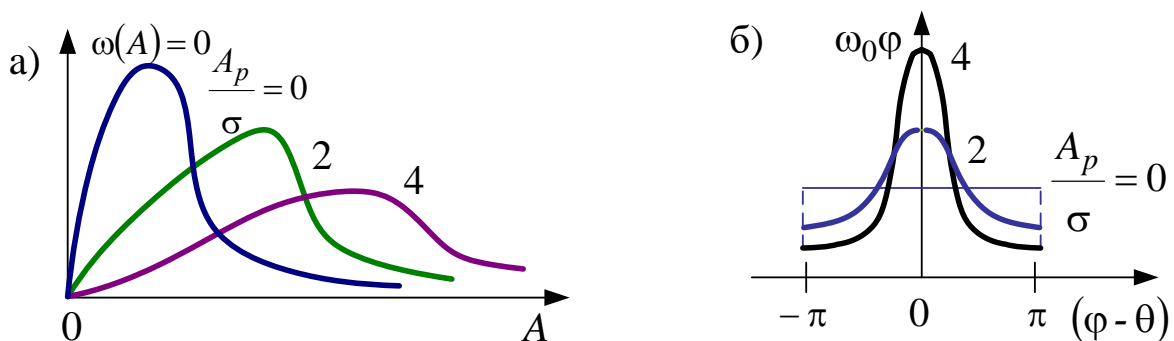


Рис. 2.42. Розподіл амплітуд (а) і фаз (б) райсового вектора

Графік цієї функції зображений на рис. 2.42, б при різних параметрах $\frac{A_p}{\sigma}$. За відсутності регулярної частини сигналу ($m_c = m_s = A_p = 0$) з огляду на те, що $I_0(0) = 1$, отримаємо з виразу (2.149) розподіл амплітуди, за Релеєм,

$$\omega(A) = \frac{A}{\sigma^2} e^{-\frac{A^2}{2\sigma^2}}, \quad A \geq 0. \quad (2.153)$$

Фаза сигналу в цьому випадку має рівномірний розподіл

$$\omega(\varphi) = \begin{cases} \frac{1}{2\pi}, & -\pi \leq \varphi \leq \pi, \\ 0, & |\varphi| > \pi. \end{cases} \quad (2.154)$$

У розрахунках часто переходять до розподілу безрозмірного значення $y = \frac{A}{\sigma}$. Тоді з умови $\omega(y)dy = \omega(A)dA$ знаходимо розподіл за Релеєм для нормованої амплітуди у такому вигляді:

$$\omega(y) = ye^{-\frac{y^2}{2}}, \quad y \geq 0. \quad (2.155)$$

Розподіли (2.153) і (2.152) наведені відповідно на рис. 2.42, а і 2.42, б.

Розподіл нормативної амплітуди відповідно до виразу (2.155) можна подати за допомогою рис. 2.43.

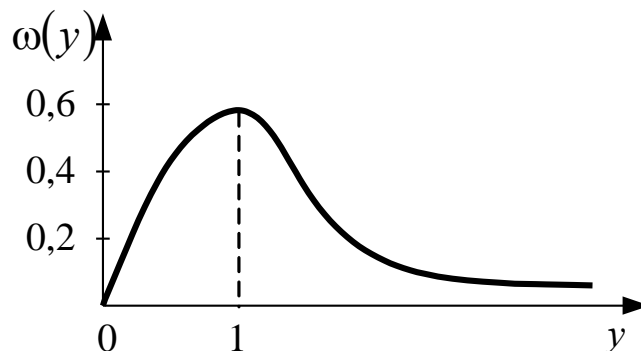


Рис. 2.43. Розподіл амплітуди безрозмірного вектора Релея

Графіки спектральної щільності потужності $G(\omega)$ і функції кореляції (нормованої) $r(\tau)$ вузькосмугового "квазібілого" шуму зображені на рис. 2.36, в.

2.7. Марковські процеси

Способи подання як випадкових, так і детермінованих процесів можна розподілити на прямі і непрямі. До прямих способів описання належить аналітичне подання реалізацій процесу часовими функціями на основі функціонального зв'язку значень сигналу в даний момент із його значеннями в минулому. Те саме стосується і різних функціональних перетворень, відображуваних алгебраїчними, різницевиими, інтегродиференційними рівняннями, в яких функціями часу є детерміновані або випадкові процеси.

При описанні випадкових процесів у попередніх пунктах ми користувалися непрямими характеристиками випадкового процесу: інтегральною функцією розподілу, щільністю імовірностей, моментами першого і другого порядку, дисперсією, функцією кореляції, інтервалом кореляції й іншими усередненими статистичними характеристиками. При непрямих способах описання імовірнісні закономірності відображають стійкість середніх результатів. Непряме описання ВП знаходить основне застосування в задачах аналізу.

У задачах синтезу ВП прямий спосіб зводиться до визначення алгоритму відтворення його реалізацій за заданими імовірнісними характеристиками (непрямого описання). При цьому ВП можна подати на основі опису поведінки динамічної породжуювальної системи (ПС), яка зображена на рис. 2.44, а, на вхід якої впливає "білий" шум $N(t)$, а на виході формується процес $X(t)$ з необхідними характеристиками. Інакше кажучи, прямий спосіб задає алгоритм генерування процесу, а непрямий – імовірнісні характеристики процесу. Як прямі, так і непрямі способи опису випадкових процесів широко застосовуються в задачах теорії зв'язку.

Нехай виходом джерела є неперервна функція $x(t)$, що є реалізацією деякого випадкового процесу. Для опису $X(t)$

скористаємося динамічною породжуювальною системою, (рис. 2.44, а), на вхід якої подається стаціонарний "білий" гаусівський шум $N(t)$ зі спектральною щільністю потужності N_0 . Випадковий процес $X(t)$ на виході найпростішої динамічної системи (інтегруючого RC-ланцюжка) першого порядку (рис. 2.44, б) визначається стохастичним диференціальним рівнянням стану.

$$\frac{dx(t)}{dt} = -\alpha x(t) + \alpha N(t) \quad \text{або} \quad \frac{dx(t)}{dt} + \alpha x(t) = \alpha N(t), \quad (2.156)$$

де α – відома випадкова величина, $\alpha = 1/RC$;

$N(t)$ – стаціонарний "білий" гаусівський шум з відомими статистичними характеристиками:

$$M[N(t)] = 0; \quad M[N(t)N(t-\tau)] = \frac{N_0}{2} \delta(t). \quad (2.157)$$

Загальний розв'язок лінійного рівняння, що складається з розв'язку однорідного рівняння (яке визначає вільні коливання) і окремого розв'язку (при заданій реалізації $n(t)$), має вигляд

$$x(t) = x(t_0) \cdot e^{-\frac{t}{RC}} + \frac{1}{RC} \int_0^t e^{-\frac{t-z}{RC}} n(z) dz,$$

де $x(t_0)$ – початкова умова.

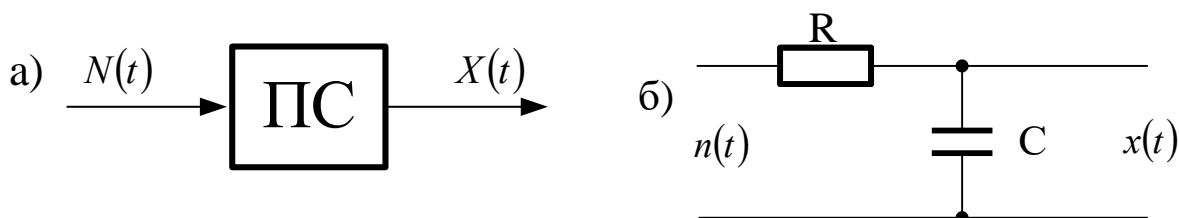


Рис. 2.44. Породжуювальна система (фільтр) (а) і динамічна система 1-го порядку під впливом реалізації $n(t)$ БГШ (б)

Можна показати, що розглянутий ВП $X(t)$ є гаусівським і марковським, а його СЩП записується як

$$G_x(f) = \frac{N_0}{2} |K(j\omega)|^2 = \frac{\alpha^2 N_0}{2(\alpha^2 + \omega^2)}, \quad (2.158)$$

де $K(j\omega)$ – передатна функція інтегруючого RC ланцюжка,

$$K(j\omega) = \frac{1}{1 + j\omega RC}.$$

На рис. 2.45, а зображений графік СЩП, визначений згідно з виразом (2.158) на додатних частотах. Спектр (2.158) відповідає КФ експоненціально-згасаючого типу

$$R_x(\tau) = \int_{-\infty}^{\infty} G_x(f) \cdot e^{j\omega\tau} df = \sigma^2 e^{-\alpha|\tau|}, \quad (2.159)$$

де σ^2 - дисперсія, $\sigma^2 = R_x(0) = \alpha N_0 / 4$.

Графік залежності (2.159) при $\tau \geq 0$ зображений на рис. 2.45, б. На основі рівняння стану (2.156) можна скласти схему пристрою (обчислювача) для формування випадкового процесу з заданими КФ і СЩП (рис. 2.46).

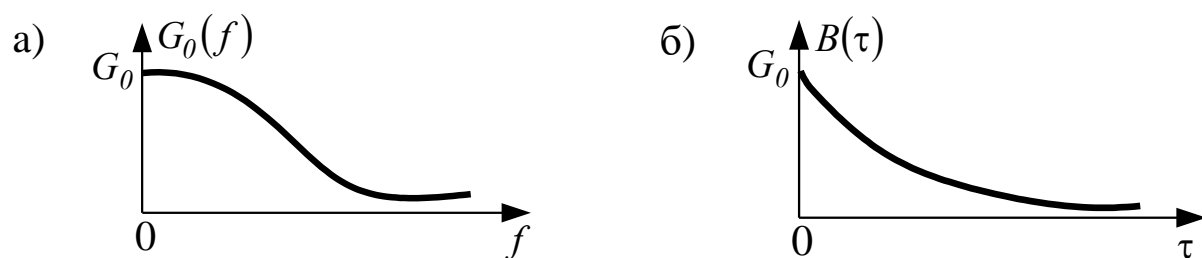


Рис. 2.45. Залежність СЩП на виході динамічної системи 1-го порядку (а) і ФК (б) випадкового процесу $X(t)$

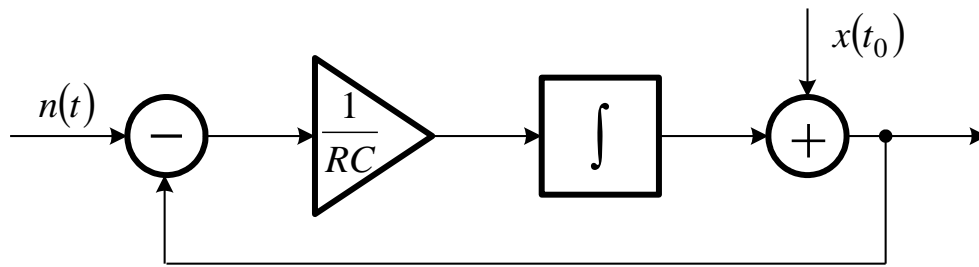


Рис. 2.46. Схема обчислювача для формування ВП на виході динамічної системи 1-го порядку при подачі на вхід гаусівського "білого" шуму

Як модель джерела, що породжує вузькосмуговий випадковий сигнал, можна використовувати випадковий процес $X(t)$ на виході коливального контура (рис. 2.47) при дії на його вході "білого" шуму з тими ж характеристиками, що й у попередньому прикладі.

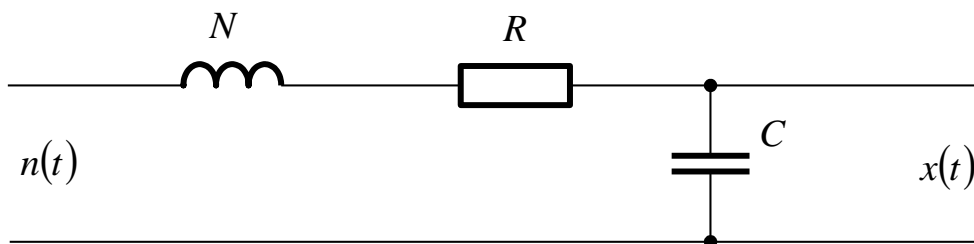


Рис. 2.47. Послідовний коливальний контур під впливом реалізації БГШ $n(t)$

Функції (2.158) і (2.159) випадкового процесу $X(t)$ (який є гаусівським і марковським) визначаються тепер на основі системи стохастичних рівнянь стану

$$\left. \begin{aligned} \frac{dx(t)}{dt} &= x_1(t) \\ \frac{dx_1(t)}{dt} &= 2\alpha x_1(t) - \omega_0^2 x(t) + \omega_0^2 N(t) \end{aligned} \right\}. \quad (2.160)$$

Це відповідає диференціальному рівнянню 2-го порядку (контура), де $2\alpha = R/L$, $\omega_0^2 = 1/LC$. Для контура з малим загасанням вважаємо, що $\omega_0^2 \gg \alpha^2$; ФК процесу $X(t)$ має експоненціально-косинусну форму

$$R_x(\tau) = \sigma^2 e^{-\alpha|\tau|} \left(\cos \omega_0 \tau + \frac{\alpha}{\omega_0} \sin \omega_0 |\tau| \right), \quad (2.161)$$

де σ^2 - дисперсія процесу, $\sigma^2 = \frac{N_0 \omega_0^2}{8\alpha}$.

Спектральна щільність потужності визначається виразом

$$G_x(f) = \int_{-\infty}^{\infty} R_x(\tau) \cdot e^{-j\omega\tau} d\tau = \frac{N_0 \omega_0^4}{2((\omega - \omega_0)^2 + \alpha^2)((\omega + \omega_0)^2 + \alpha^2)}.$$

Гаусівські і марковські ВП, сформовані на основі стохастичних диференціальних рівнянь (2.160), часто використовуються як типові моделі телеметричних повідомлень, телевізійних і мовних повідомлень.

Однією з корисних для практики досліджень моделей випадкових процесів є процес Маркова або випадковий процес без наслідку. Марковським процесом називають такий ВП, для якого умовну щільність розподілу імовірностей можна записати як

$$w_n \left(\frac{x_n}{x_1, x_2, \dots, x_{n-1}} \right) = w_2 \left(\frac{x_n}{x_{n-1}} \right), \quad (2.162)$$

де x_k - значення випадкового процесу в момент t_k .

Рівність (2.162) означає, що значення ВП в обраний момент часу t_n залежить тільки від того, яке значення ВП прийняв у попередній момент часу t_{n-1} , і не залежить від того, які значення він приймав у попередні моменти часу.

Таким чином, вся інформація про марковський процес міститься у двомірній ЩРІ. Врахувавши вираз (2.162) і виходячи з визначення умовної ЩРІ для марковського процесу, n -мірну ЩРВ можемо визначити як

$$w_n(x_1, x_2, \dots, x_n) = w_1(x) \prod_{k=2}^n w_2\left(\frac{x_k}{x_{k-1}}\right). \quad (2.163)$$

Якщо ЩРІ залежить тільки від інтервалу часу між сусідніми моментами $t_k - t_{k-1}$, то процес Маркова є стаціонарним. При $x_1 = x_2 = \dots = x_n = x$ для стаціонарного марковського процесу маємо

$$w_1(x_1, x_2, \dots, x) = w_1(x) \left[\frac{w_2(x, x)}{w_1(x)} \right]^{n-1}. \quad (2.164)$$

Замінімо вираз ЩРІ марковського процесу при $n = 3$ на

$$\begin{aligned} w_3(x_1, x_2, x_3, t_1, t_2, t_3) &= \\ &= w_1(x_1, t_1) w_2\left(\frac{x_2 t_2}{x_1, t_1}\right) w_2\left(\frac{x_3, t_3}{x_2, t_2}\right). \end{aligned} \quad (2.165)$$

Інтегруючи за x_2 ліву і праву частини рівності (2.165), одержимо

$$\begin{aligned} w_2(x_1, x_3, t_1, t_3) &= \\ &= w_1(x_1, t_1) \int_{-\infty}^{\infty} w_2\left(\frac{x_2 t_2}{x_1, t_1}\right) w_2\left(\frac{x_3, t_3}{x_2, t_2}\right) dx_2. \end{aligned} \quad (2.166)$$

З огляду на визначення ЩРІ

$$w_2(x_1, x_3, t_1, t_3) / w_1(x_1, t_1) = w_2\left(\frac{x_3, t_3}{x_1, t_1}\right). \quad (2.167)$$

Вираз $w_2(x_1, x_3, t_1, t_3) / w_1(x_1, t_1) = w_1(x_3, t_3 / x_1, t_1)$ запишемо як

$$w_2\left(\frac{x_3, t_3}{x_1, t_1}\right) = \int_{-\infty}^{\infty} w_2\left(\frac{x_2 t_2}{x_1, t_1}\right) \cdot w_2\left(\frac{x_3, t_3}{x_2, t_2}\right) dx_2. \quad (2.168)$$

Рівняння називається рівнянням Смолуховського, або узагальненим рівнянням Маркова [47]. Для стаціонарного марковського процесу вираз (2.168) записується як

$$w_2\left(\frac{x_3, t_3 - t_1}{x_1, t_1}\right) = \int_{-\infty}^{\infty} w_2\left(\frac{x_2 t_2 - t_1}{x_1, t_1}\right) \cdot w_2\left(\frac{x_3, t_3 - t_2}{x_2, t_2}\right) dx_2.$$

З рівняння Смолуховського може бути отримане диференціальне рівняння переходу неперервного марковського процесу. Врахувавши введені позначення $t_2 = t$, $t_1 = t - \Delta t$, $t_3 = T$, $x_1 = x$, $x_3 = y$, $x_2 = z$, диференціальне рівняння набуде вигляду

$$\begin{aligned} & \frac{\partial w_2\left(\frac{y, T}{x, t}\right)}{\partial t} + A_1(y, T) \frac{\partial w_2\left(\frac{y, T}{x, t}\right)}{\partial x} + \\ & + \frac{1}{2} A_2(y, T) \frac{\partial^2 w_2\left(\frac{y, T}{x, t}\right)}{dx^2} = 0, \end{aligned} \quad (2.169)$$

де

$$A_1(y, T) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{-\infty}^{\infty} (z - y) w_2\left(\frac{z, T}{y, T - \Delta t}\right) dz; \quad (2.170)$$

$$A_2(y, T) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{-\infty}^{\infty} (z - y)^2 w_2\left(\frac{z, T}{y, T - \Delta t}\right) dz. \quad (2.171)$$

Рівняння (2.169) називається першим рівнянням Колмогорова [21].

Щільність розподілу імовірностей марковського процесу задовольняє і друге рівняння Колмогорова

$$\frac{\partial w_2\left(\frac{y, T}{x, t}\right)}{\partial T} = -\frac{\partial}{\partial y} \left[A_1(y, T) \cdot w_2\left(\frac{y, T}{x, t}\right) \right] + \quad (2.172)$$

$$+ \frac{1}{2} \frac{\partial^2}{\partial y^2} \left[A_2(y, T) \cdot w_2\left(\frac{y, T}{x, t}\right) \right].$$

Друге рівняння Колмогорова відоме також як рівняння Фоккера – Планка.

При розгляді неперервних випадкових процесів використання марковської моделі пов'язане з визначенням кроку по осі часу: $\Delta t = t_n - t_{n-1}$. При $\Delta t \rightarrow 0$ марковський процес переходить у пуассонівський [18].

Висновки до глави 2

1. Повідомлення, сигнали і завади в системах електричного зв'язку перш за все визначаються деякими функціями часу (процесами), які є детермінованими або випадковими.

2. Повідомлення, сигнали і завади як вектори (точки) в лінійному просторі можна описати через набір координат у заданому базисі (сукупності лінійно незалежних одиничних векторів). Найчастіше базис є ортонормованим.

3. Для ТЕЗ найбільший інтерес при відображенні сигналів являє n -мірний простір Евкліда R_n , безконечномірний простір Гілберта $L_2(T)$ і дискретний простір Хеммінга 2_n . У цих просторах вводиться поняття скалярного добутку двох векторів (x, y) .

4. Будь-яку безперервну функцію часу як елемент L_2 можна подати узагальненим рядом Фур'є за заданим ортонормованим базисом. Усічений узагальнений ряд Фур'є дає опис безперервної функції часу з певною похибкою.

5. Окремим випадком узагальненого ряду Фур'є є звичайний ряд Фур'є для періодичних функцій, коли як базисні використовуються гармонійні функції кратних частот. Ряд Фур'є можна подати в дійсній і комплексній формі, вводячи в розгляд негативні частоти. Періодична функція характеризується лінійчатим спектром.

6. Спектр, за Фур'є, неперіодичної функції часу $x(t)$ є суцільним. Він визначається спектральною щільністю (СЩ) $\dot{S}_x(f)$. Функція $x(t)$ і СЩ $\dot{S}_x(f)$ зв'язані лінійно парою перетворень Фур'є. Модуль $|\dot{S}_x(f)| = S_x(f)$ визначає амплітудний спектр сигналу $x(t)$, а $\arg \dot{S}_x(f)$ визначає фазовий спектр.

7. Часова функція кореляції (ФК) детермінованого сигналу з розмірністю енергії $B_E(\tau)$ пов'язана зі спектральною щільністю енергії (СЩЕ) $W_x(f)$ парою перетворень Фур'є. Аналогічно зв'язані часова ФК сигналу $B(\tau)$ з розмірністю потужності і спектральна щільність середньої потужності (СЩП) $G_x(f)$. $B_E(0)$ визначає енергію сигналу, а $B(0)$ – його середню потужність.

8. Добуток ефективної тривалості простого сигналу τ_e і ефективної смуги його частот F_e – величина порядку одиниці.

9. Сигнал $x(t)$ з фінітним спектром F_B подається в узагальнений ряд Фур'є за базисною системою функцій $\left\{ \frac{\sin \omega_B(t - k\Delta)}{\omega_B(t - k\Delta)} \right\}$. Координатами розкладання є часові відліки $x(k\Delta)$ в перетинах, кратних інтервалу дискретизації $\Delta \leq 1/2F_B$. Наближене подання такого сигналу на інтервалі T визначається усіченим рядом з числом координат $x(k\Delta)$, що дорівнює базі сигналу $B=2F_B T$.

10. Випадкові процеси (ВП) найчастіше описуються непрямим шляхом через щільність імовірності (ЩІ) і інтегральну функцію розподілу (ІФР). Грубішими характеристиками ВП є його математичне очікування $m_1(t) = \overline{X(t)}$, дисперсія $\sigma^2(t) = \overline{X^2}$ і ФК $B(\tau) = \overline{\dot{X}(t)\dot{X}(t-\tau)}$.

11. У теорії і практиці електрозв'язку велику роль відіграють стаціонарні в широкому розумінні ВП, визначувані незалежною від часу одновимірною ЩІ і двовимірною ЩІ, залежною від різниці вибраних перетинів τ . Такі процеси часто мають властивості ергодичності: характеристики, отримані усереднюванням у часі однієї реалізації, приблизно дорівнюють характеристикам, отриманим усереднюванням за реалізаціями ВП.

12. Функція кореляції і спектральна щільність середньої потужності випадкового процесу зв'язані парою перетворень Вінера-Хінчина. Ефективна ширина спектра ВП F_e обернено пропорційна інтервалу кореляції ВП $\tau_{кор}$.

13. Подання центрованого ВП в узагальнений ряд Фур'є з некорельованими координатами (розкладання Карунена-Лоева) називається канонічним розкладанням. Для ВП з фінітним спектром канонічне розкладання визначається рядом Котельникова.

14. Двовимірною щільністю імовірності повністю визначаються гаусівський і простий марковський випадкові процеси. Некорельованість перетинів гаусівського ВП означає і їх незалежність.

15. Вузкополосні дійсні процеси $x(t)$ (як детерміновані, так і випадкові) зручно виражати через квадратурні складові, обвідну і фазу, через комплексну обвідну. Уявна $\hat{x}(t)$ і дійсна $x(t)$ частини комплексного (аналітичного) вузькосмугового сигналу $\dot{x}(t)$ зв'язані парою перетворень Гілберта.

16. Спектр аналітичного (комплексного) сигналу $\dot{x}(t) = x(t) + j\hat{x}(t)$ лежить повністю в області позитивних частот, а сигналу $\dot{x}^*(t) = x(t) - j\hat{x}(t)$ – в області від'ємних частот.

17. Обвідна вузькосмугового гаусівського ВП при рівній дисперсії квадратурних компонент розподілена за узагальненим законом Релея (або законом Райса). У стаціонарного (центрованого) гаусівського ВП обвідна розподілена за законом Релея, а фаза – рівномірно на відрізку $[-\pi/2; \pi/2]$.

18. Безперервні і дискретні джерела повідомлень, сигналів і завад (процесів, що розвиваються в часі) можна описати або прямим способом (через часові функції їх реалізації) або непрямим способом (через різні спектральні характеристики, а для ВП – через функції розподілу).

19. Зручною моделлю для дослідження дискретних джерел є модель випадкового стаціонарного синхронного двійкового сигналу з нульовим МО і трикутною ФК.

Контрольні питання та задачі

2.1. Як визначається похибка подання сигналу $x(t)$ як елементу простора $L_2(T)$ усіченим узагальненим рядом Фур'є?

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

2.3. Задані на інтервалі $(-T/2; T/2)$ сигнали $x(t) = \sqrt{2}\cos(2\pi t/T)$ і $y(t) = \sqrt{2}\cos(4\pi t/T)$ трактуються як елементи простору Гілберта $L_2(T)$. Доведіть, що: а) норми цих сигналів дорівнюють 1; б) їх скалярний добуток дорівнює 0; в) відстань між ними дорівнює $\sqrt{2}$.

2.4. Доведіть, що в n -мірному просторі Хеммінга 2_n існує лише один ортонормований базис, утворений n векторами: $(1, 0, 0, \dots, 0)$, $(0, 1, 0, \dots, 0)$, \dots , $(0, 0, 0, \dots, 1)$.

2.5. Знайдіть спектральні компоненти періодичної послідовності прямокутних відеоімпульсів тривалістю τ з періодом проходження $T=2\tau$ за спектральною щільністю Фур'є одиночного імпульсу.

2.6. Покажіть, що якщо спектральна щільність, за Фур'є, $\dot{S}_y(f) = \dot{S}_x(f)\dot{S}_g(f)$ сигналу $y(t)$ дорівнює добутку спектральних щільностей сигналів $x(t)$ і $g(t)$, то сам сигнал $y(t)$ визначається часовою згорткою сигналів $x(t)$ і $g(t)$: $y(t) = x(t) \otimes g(t)$. Аналогічно якщо $y(t) = x(t)g(t)$, то $\dot{S}_y(f) = \dot{S}_x(f) \otimes \dot{S}_g(f)$.

2.7. Користуючись визначенням δ -функції, знайдіть спектральну щільність, за Фур'є, постійної складової $x(t) = a$.

2.8. Дано прямокутний імпульс $x(t) = A$, $t \in [-\tau_u/2; \tau_u/2]$. Знайдіть його ФК $B_E(\tau)$ (з розмірністю енергії) і $B(\tau)$ (з розмірністю потужності), а також СЩЕ $W_x(f)$ і СЩП $G_x(f)$.

2.9. Чому дорівнює максимальний інтервал дискретизації мовного і телевізійного сигналу, якщо верхню частоту в спектрі сигналу прийняти рівною відповідно 4 кГц і 6,5 МГц?

2.10. З яких міркувань інтервал дискретизації Δt_δ безперервної функції часу з фінітним спектром на відрізку $[-F_B; +F_B]$ часто вибирають менше, ніж $1/2F_B$?

2.11. Дискретне стаціонарне двійкове джерело описується простим ланцюгом Маркова з матрицею перехідної імовірності

$\begin{vmatrix} P(1|1') & P(1|0') \\ P(0|1') & P(0|0') \end{vmatrix}$, де $P(a_i|a'_j)$ – імовірність появи символу a_i за

умови, що йому передуює символ a_j . Нехай $P(1|1')=0,9$; $P(1|0')=0,7$.
Визначити: імовірність $P(0|1')$, $P(0|0')$; безумовну імовірність передачі символів $P(0)$ і $P(1)$; імовірності передачі ланцюжків з трьох символів $(1, 1, 0)$ і $(0, 1, 0)$.

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

2.13. Покажіть, що випадковий вузькосмуговий процес $X(t) = A_c(t)\cos\omega_0t - A_s(t)\sin\omega_0t$ з незалежними квадратурними компонентами $A_c(t)$ і $A_s(t)$ стаціонарний лише тоді, коли $\overline{A_c(t)} = \overline{A_s(t)} = 0$ і $\overline{A_c^2} = \overline{A_s^2} = \sigma^2 = \text{const}$.

ГЛАВА 3

МОДУЛЬОВАНІ СИГНАЛИ В СИСТЕМАХ ПЕРЕДАЧІ ІНФОРМАЦІЇ

3.1. Загальні положення

Модуляція - це процес змінення одного чи кількох параметрів переносника відповідно до змінення сигналу, який діє на нього. Як переносник широко використовують гармонічні коливання, періодичну послідовність імпульсів і вузькосмуговий випадковий процес. Сигнал, який діє на керований параметр переносника, називається *модулюючим*.

Параметри переносника, які змінюються в часі під впливом модулюючого сигналу, називаються *інформаційними*, оскільки в їх змінних міститься інформація, що передається. Фізичний процес змінення параметрів переносника і є модуляцією. Таким чином, будь-який модулятор (рис. 3.1) має два входи: один - для переносника, інший - для модулюючого сигналу.

Параметри, які залишаються незмінними, є постійними ознаками сигналу. Вони можуть бути використані на прийомі для розрізнення сигналу від завад. У багатьох випадках модулювання сигнал можна подати як добуток двох функцій:

$$s(t) = f(t)M[u_m(t)], \quad (3.1)$$

де $f(t)$ - функція, яка зображує несучу коливання (переносника);

$M(t)$ - модуляційна функція, яка виражає дію сигналу, що передається, тобто $u_m(t)$, на несучу $f(t)$.

Якщо для зображення несучої вибирається *аналітичний* сигнал, то для кожної модуляційної функції $M(t)$ існує комплексний модульований сигнал $s(t)$; його дійсна і уявна частини відповідають реально існуючому модульованому сигналу, а модуль визначає обвідну. У випадку, коли несучою є *гармонічне* коливання $A_0 e^{j\omega_0 t}$, модуляційна функція виражає діяння відеосигналу $u_m(t)$ на амплітуду (частоту або фазу) несучої.

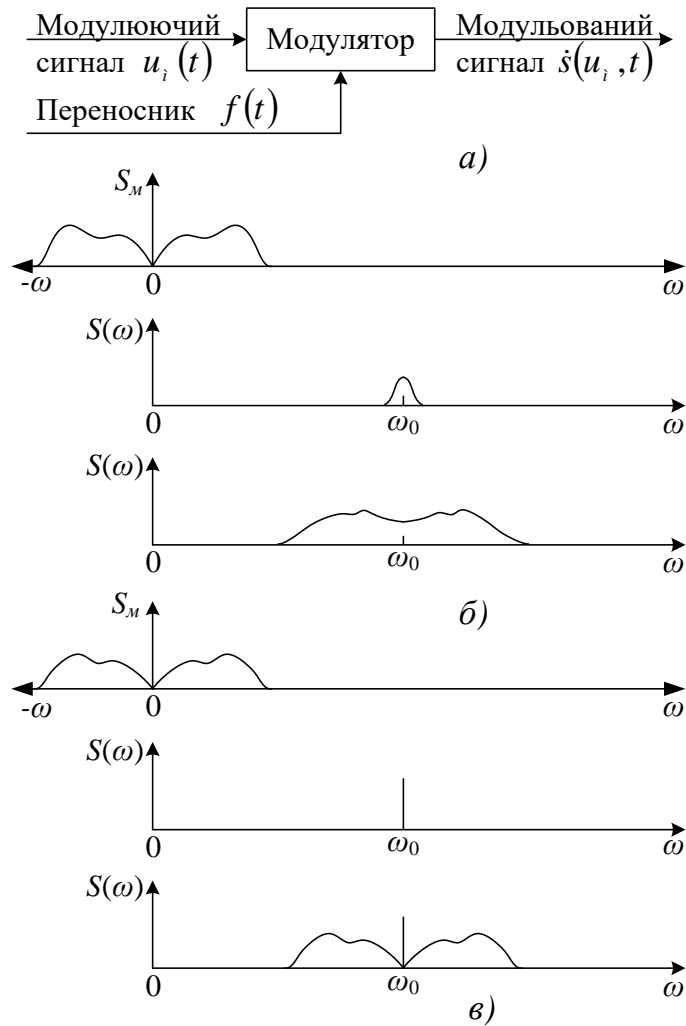


Рис. 3.1. Модулятор (а) і зсув спектра при модуляції: загальний випадок аналітичної несучої (б) та гармонічної несучої (в)

Спектр модульованого коливання (вираз (3.1)) згідно з теоремою про спектр добутку визначається згорткою:

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_f(\nu) S_m(\omega - \nu) d\nu, \quad (3.2)$$

де $S_f(\nu)$ - спектральна щільність сигналу $s(t)$;

ν - частота модулюючого сигналу $M[u_m(t)]$;

S_m - спектральна щільність модульованого сигналу.

Звідси випливає, що процес модуляції призводить до складного перетворення спектра сигналу. Якщо несучою є вузькосмугове коливання, то модуляція спричиняє розширення

спектра і перенесення його в область навколо несучої частоти (рис. 3.1, б). Якщо несуча - чиста синусоїда, то має місце просто зсув спектра (рис. 3.1, в).

3.2. Основні види аналогової амплітудної модуляції

Одним із видів аналогової модуляції (тобто первинний сигнал і сигнал-переносник неперервні) є безпосереднє передавання та амплітудна модуляція (АМ), різновиди якої - балансна (БМ) і односмугова (ОМ) модуляції.

Безпосереднє передавання. Найпростішим сигналом для передавання неперервного повідомлення $s(t)$ є сигнал, пропорційний $u_m(t)$:

$$s(t) = Au_m(t), \quad (3.3)$$

де A - якась стала.

Такий сигнал відповідає формулі (3.1), якщо взяти $f(t) = A$. Прикладом безпосереднього передавання повідомлень є звичайний телефонний зв'язок по проводах.

Амплітудна модуляція (АМ). Для цього виду модуляції

$$s(t) = A_0 e^{j\omega_0 t}, \quad M[u_m(t)] = 1 + mu_m(t), \quad (3.4)$$

де m - коефіцієнт модуляції.

Тоді модульований сигнал запишеться так:

$$\dot{s}(t) = A_0 [1 + mu_m(t)] e^{j\omega_0 t} = A_0 [1 + mu_m(t)] (\cos \omega_0 t + i \sin \omega_0 t). \quad (3.5)$$

Цей вираз дає зображення реального амплітудно-модульованого сигналу (АМ-сигналу):

$$s(t) = \text{Re } \dot{s}(t) = A_0 [1 + mu_m(t)] \cos \omega_0 t. \quad (3.6)$$

Спектр сигналу в загальному випадку визначається як перетворення Фур'є від $s(t)$:

$$S(j\omega) = \int_{-\infty}^{\infty} s(t)e^{-j\omega t} dt = A_0 \int_{-\infty}^{\infty} \cos \omega_0 t e^{-j\omega t} dt + mA_0 \int_{-\infty}^{\infty} u(t) \cos \omega_0 t e^{-j\omega t} dt.$$

Врахувавши, що $\cos \omega t = \frac{1}{2}(e^{j\omega t} + e^{-j\omega t})$, а $\int_{-\infty}^{\infty} e^{j\omega t} dt = 2\pi\delta(\omega)$,

отримаємо

$$S(\omega) = \pi A_0 \delta(\omega - \omega_0) + \frac{mA_0}{2} S_i(\omega_0 + \omega) + \frac{mA_0}{2} S_i(\omega_0 - \omega), \quad (3.7)$$

де S_i - спектр повідомлення, яке передається.

Звідси бачимо, що при АМ відбувається перенесення спектра повідомлення на частоту ω_0 (рис. 3.1, в). Ширина спектра сигналу F при АМ удвічі більша за ширину спектра повідомлень F_m , тобто $F = 2F_m$.

При модуляції одним тоном, тобто синусоїдальним сигналом, коли $u(t) = \cos \Omega_m t$,

$$\begin{aligned} s(t) &= A_0(1 + m \cos \Omega_m t) \cos \omega_0 t = \\ &= A_0 \cos \omega_0 t + \frac{mA_0}{2} \cos(\omega_0 + \Omega_m) + \frac{mA_0}{2} S_i(\omega_0 - \Omega_m). \end{aligned} \quad (3.8)$$

При амплітудній модуляції гармонічного переносника приріст амплітуди переносника пропорційний миттєвим значенням модулюючого сигналу $u_m(t)$, тобто

$$A(t) = A_0 + \Delta A(t) = A_0 + au_m(t), \quad (3.9)$$

де a - коефіцієнт пропорційності; частота і фаза гармонічного переносника залишаються сталими.

Часову діаграму АМ-сигналу наведено на рис. 3.2, з якого видно, що з миттєвим значенням $u_m(t)$ амплітуда $A(t)$ то підвищується до A_{max} , одержуючи при цьому приріст $\Delta A_+ = A_{max} - A_0 = au_{mmax}$, то зменшується до A_{min} , одержуючи

приріст $\Delta A_- = A_0 - au_{m\min}$. Зверніть увагу, що амплітуда повторює форму модулюючого сигналу $u_m(t)$. В АМ-сигналі амплітуда $A(t)$ є *обвідною* високочастотного заповнення $\cos(\omega_0 + \psi_0)$.

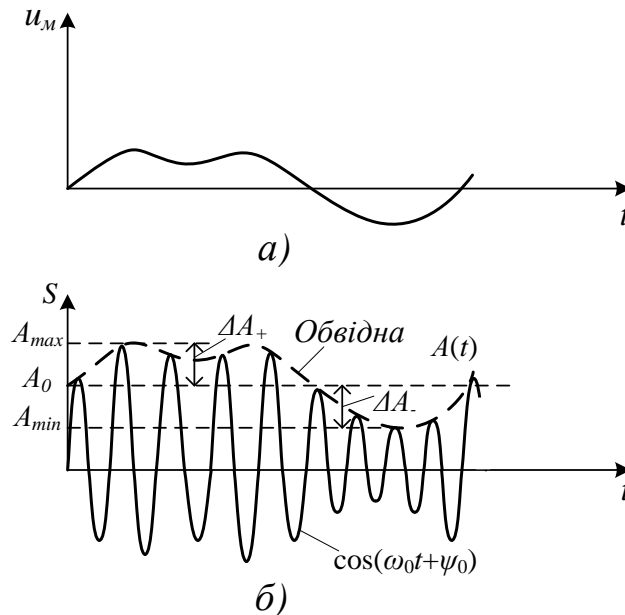


Рис. 3.2. Амплітудна модуляція сигналу:
а) модулюючий сигнал $u_m(t)$; б) АМ-сигнал

Для математичного опису АМ-сигналу у формулу (3.9) замість коефіцієнта a , що залежить від конкретної схеми модулятора, вводять *коефіцієнт модуляції* $m_{AM} = \Delta A_{\text{сер}} / A_0$, який дорівнює *відносному значенню приросту амплітуди*. Тут $\Delta A_{\text{сер}} = (A_+ + A_-) / 2$ - середньоарифметичне значення приросту амплітуди АМ-сигналу. Оскільки середнє значення амплітуди АМ-сигналу за час модуляції $A_0 = (A_{\text{max}} + A_{\text{min}}) / 2$, то коефіцієнт модуляції

$$m_{AM} = \Delta A_{\text{сер}} / A_0 = (A_{\text{max}} - A_{\text{min}}) / (A_{\text{max}} + A_{\text{min}}). \quad (3.10)$$

Отже, *коефіцієнт модуляції* - це *відношення різниці між максимальним і мінімальним значеннями амплітуд АМ-сигналу до суми цих значень*. Досить часто коефіцієнт модуляції

визначається у відсотках, але при всіх розрахунках АМ-сигналів користуються коефіцієнтом модуляції m_{AM} у відносних одиницях.

Для симетричного модулюючого сигналу $u_m(t)$ (тобто середнє значення амплітуди якого дорівнює нулю) АМ-сигнал також симетричний: $\Delta A_+ = \Delta A_- = \Delta A$ і

$$m_{AM} = \Delta A / A_0 ; \quad (3.11)$$

отже, коефіцієнт модуляції дорівнює відношенню максимального приросту амплітуди до амплітуди переносника. Він характеризує глибину амплітудної модуляції і може змінюватись у межах $0 \leq m_{AM} \leq 1$.

Приклад 3.1. Визначити коефіцієнт модуляції АМ-сигналу, числову діаграму якого в лінійному масштабі зображено на рис. 3.2, б.

Для визначення коефіцієнта модуляції скористаємося формулою (3.10). При розрахунках знати абсолютне значення амплітуд не обов'язково. Обчислення можна зробити в умовних одиницях. За 1 умов. од. візьмемо 1 мм на вертикальній осі амплітуд. Тоді $A_{min} = 4$ умов. од. $A_{max} = 12$ умов. од. і

$$m_{AM} = (A_{max} - A_{min}) / (A_{max} + A_{min}) = (12 - 4) / (12 + 4) = 0,5 \cdot 100 = 50\%.$$

Амплітудна модуляція гармонічним коливанням. У найпростішому випадку модулюючий сигнал $u_i(t)$ є гармонічним коливанням із частотою $\Omega_m \ll \omega_0$ і початковою фазою ψ . При цьому

$$s_{AM}(u_m, t) = A_0 [1 + m_{AM} \cos(\Omega_m t + \psi) \cos(\omega_0 t + \psi_0)] \quad (3.12)$$

є аналітичним виразом (математичною моделлю) однотонального АМ-сигналу, тобто модульованого одним гармонічним коливанням тональної частоти. На рис. 3.3, а, б зображено часові діаграми однотонального АМ-сигналу при різних значеннях m_{AM} . На рисунку дуже добре видно симетричність модуляції та

характерні спотворення при *перемодуляції*, коли форма обвідної вже не повторює форму модулюючого гармонічного коливання (рис. 3.3, в).

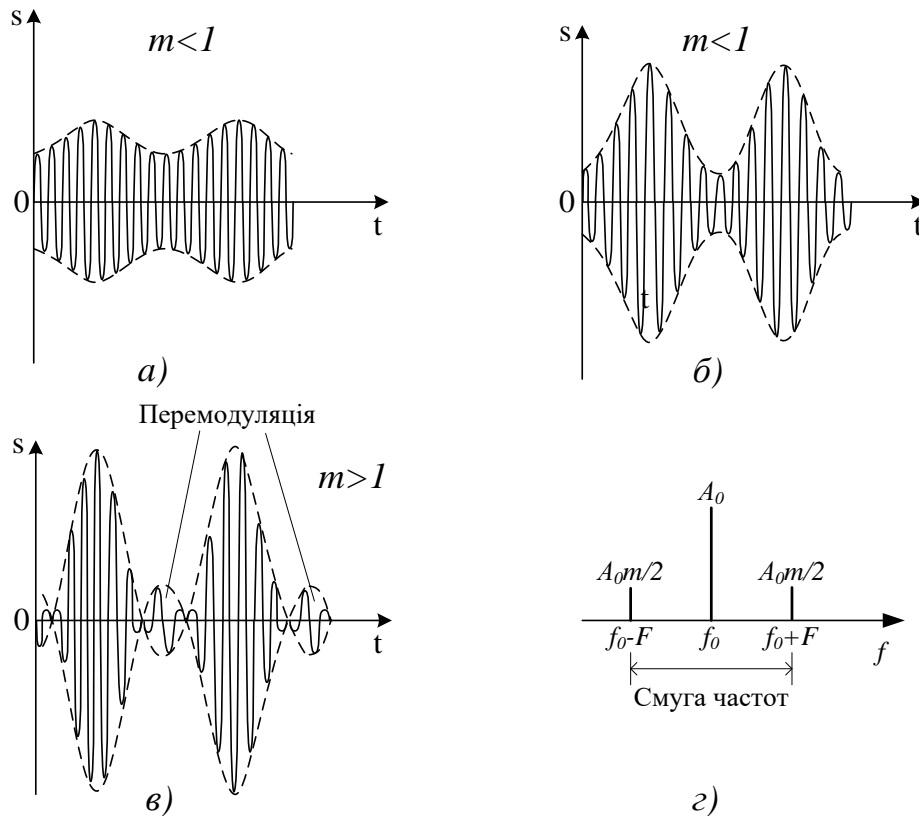


Рис. 3.3. Однотональна амплітудна модуляція: а) з малим коефіцієнтом модуляції; б) глибока модуляція; в) перемодуляція; г) спектральна діаграма АМ-сигналу для однотональної модуляції

Однотональний АМ-сигнал можна подати також у вигляді суми гармонічних складових. Якщо використати тригонометричну формулу добутку косинусів $\cos \alpha \cos \beta = 0,5[\cos(\alpha + \beta) + \cos(\alpha - \beta)]$, із формули (3.12) отримаємо

$$\begin{aligned}
 s_{AM}(u_m, t) = & A_0 \cos(\omega_0 t + \psi_0) + \\
 & + 0,5 A_0 \cos m_{AM} \cos[(\omega_0 + \Omega_m)t + \psi_0 + \psi] + \\
 & + 0,5 A_0 \cos m_{AM} \cos[(\omega_0 - \Omega_m)t + \psi_0 - \psi].
 \end{aligned} \quad (3.13)$$

Із формули (3.13) випливає, що однотональний АМ-сигнал має три гармонічні спектральні складові з частотами f_0 (переносника); $f_0 + F$ (верхньою бічною); $f_0 - F$ (нижньою бічною).

Спектральна діаграма однотонального АМ-сигналу, побудована за формулою (3.13), є симетричною відносно частоти переносника f_0 (рис. 3.3, з). Амплітуди бічних коливань однакові, і навіть для $m_{AM} = 1$ не перевищують половини амплітуди переносника A_0 .

Амплітудна модуляція при складному модулюючому сигналі. Гармонічні модулюючі сигнали і відповідно однотональний АМ-сигнал практично зустрічаються рідко. У більшості випадків модулюючі первинні сигнали є складними функціями часу. Аналітичний вираз АМ-сигналу можна подати у вигляді формули

$$s_{AM}(u_m, t) = A_0 [1 + m_{AM} u_m(t)] \cos(\omega_0 + \psi_0). \quad (3.14)$$

Спектр АМ-сигналу при складному модулюючому сигналі якісно визначається з таких міркувань. Будь-який складний сигнал $u_m(t)$ можна подати у вигляді скінченної (чи нескінченної) суми гармонічних складових, якщо скористатись для цього рядом чи інтегральним перетворенням Фур'є. Кожна гармонічна складова сигналу $u_m(t)$ з частотою Ω_i сприяє появі в спектрі АМ-сигналу двох бічних складових з частотами $f_0 \pm F_i$, а множина гармонічних складових модулюючого сигналу $\sum F_i$ - множині бічних складових з частотами $\sum (f_0 \pm F_i)$. Для наочності такі перетворення спектра для АМ наведено на рис. 3.4.

З рис. 3.4 видно, що в спектрі складного АМ-сигналу, крім складової з частиною переносника f_0 , містяться групи *верхніх* ($f_0 + \Delta F$) та *нижніх* ($f_0 - \Delta F$) бічних коливань, що утворюють відповідно верхню та нижню бічні смуги частот АМ-сигналу. При цьому верхня бічна смуга частот є масштабною копією як дискретного, так і неперервного спектра модулюючого сигналу, зсунутого за частотою на величину f_0 . Нижня бічна смуга частот

також повторює спектральну діаграму (спектральну щільність) сигналу $u_M(t)$, але з дзеркальним розташуванням частот (у зворотному порядку) відносно частоти переносника f_0 .

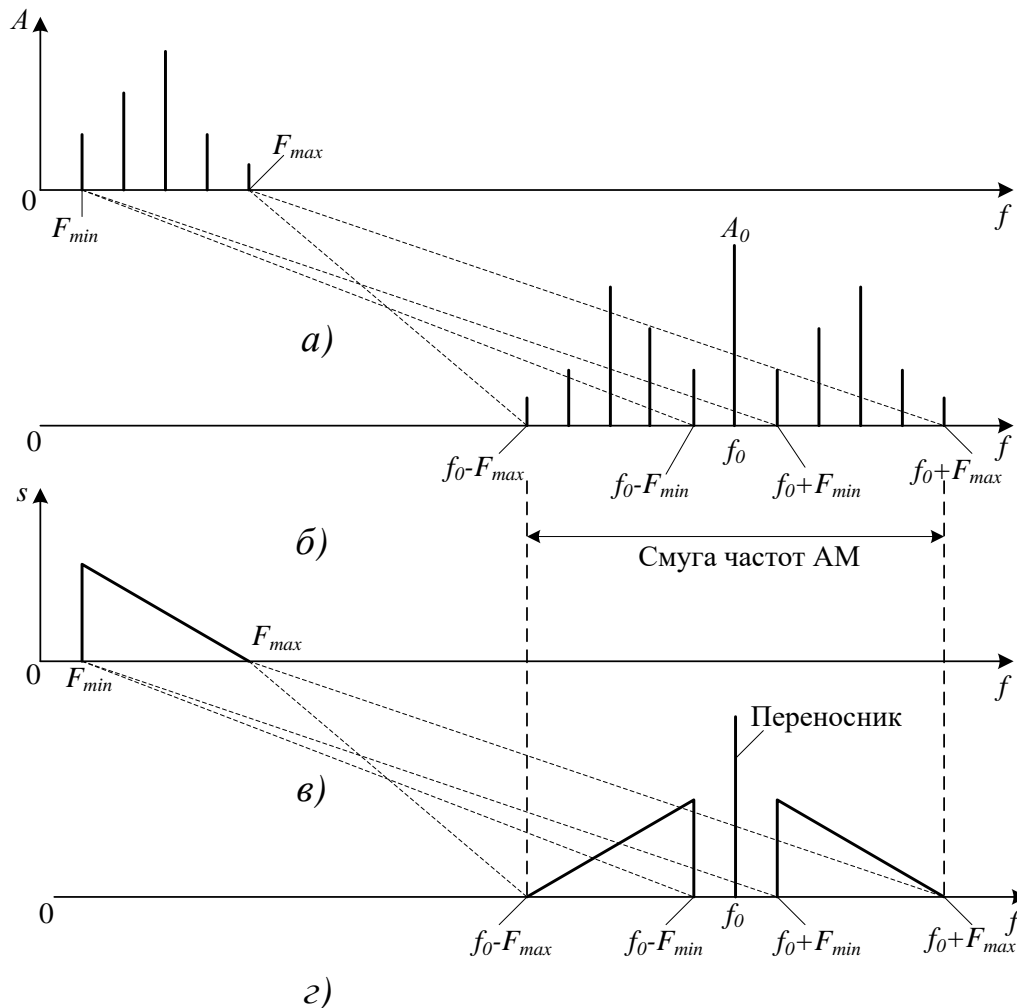


Рис. 3.4. Спектральні діаграми: а) модулюючого сигналу; б) АМ-сигналу для багатоканальної модуляції; в) модулюючого сигналу з неперервним спектром; г) АМ-сигналу для модуляції сигналом із неперервним спектром

Із зазначеного можна дійти важливого висновку: *ширина спектра АМ-сигнула F_{AM} дорівнює подвоєному значенню максимальної частоти F_{max} спектра модулюючого сигналу, тобто*

$$F_{AM} = 2F_{max}. \quad (3.15)$$

З цього виразу випливає, що амплітуда модульованого сигналу змінюється від $A_{min} = A_0(1 - m_{AM})$ до $A_{max} = A_0(1 + m_{AM})$, а потужність сигналу від $P_{min} = P_n(1 - m_{AM})^2$ до $P_{max} = P_n(1 + m_{AM})^2$, де $P_n = \frac{A_0^2}{2}$ - потужність несучого коливання. Середня потужність АМ-сигналу дорівнює

$$P_{сер} = \frac{A_0^2 T}{2T} \int_0^T (1 + m_{AM} \cos \Omega_m t) dt = P_n \left(1 + \frac{m_{AM}^2}{2} \right). \quad (3.16)$$

При $m_{AM} = 1$ $P_{max} = 4P_n$, а $P_{сер} = 1,5P_n$; відношення середньої потужності до максимальної дорівнює 0,375, що свідчить про суттєвий недолік амплітудної модуляції, тобто неповне використання потужності передавача.

Балансна модуляція (БМ). Крім звичайної АМ, застосовується передавання АМ-сигналу без несучої - балансна модуляція. Для цього виду модуляції

$$f(t) = A_0 e^{j\omega_0 t}, M[u(t)] = u(t), \quad (3.17)$$

тоді

$$\begin{aligned} \dot{s}(t) &= A_0 e^{j\omega_0 t} u(t) = A_0 u(t) \cos \omega_0 t + j A_0 u(t) \sin \omega_0 t; \\ s(t) &= \text{Re } \dot{s}(t) = A_0 u(t) \cos \omega_0 t. \end{aligned} \quad (3.18)$$

Спектр сигналу при БМ

$$S(\omega) = \frac{1}{2} A_0 [S_i(\omega_0 + \omega) + S_i(\omega_0 - \omega)]. \quad (3.19)$$

Тут знаходяться тільки дві бічні смуги, оскільки несуча відсутня.

При **односмуговій модуляції (ОМ)** передається тільки одна бічна смуга. При передаванні *верхньої бічної смуги*

$$f(t) = A_0 e^{j\omega_0 t}, M[u_M(t)] = u_M(t) + ju_M^*(t); \quad (3.20)$$

$$\begin{aligned} \dot{s}(t) &= A_0 [u_M(t) \cos \omega_0 t - u_M^*(t) \sin \omega_0 t] + \\ &+ jA_0 [u_M(t) \sin \omega_0 t + u_M^*(t) \cos \omega_0 t]; \quad (3.21) \\ s(t) &= \operatorname{Re} \dot{s}(t) = A_0 [u_M(t) \cos \omega_0 t - u_M^*(t) \sin \omega_0 t]. \end{aligned}$$

Спектр сигналу ОМ буде дорівнювати

$$S(\omega) = \frac{A_0 m}{2} S_i(\omega_0 + \omega). \quad (3.22)$$

Дійсно, якщо розкласти функції $u_M(t)$ і $u_M^*(t)$ у ряд Фур'є

$$\begin{aligned} u_M(t) &= \sum_{k=0}^{\infty} C_k \cos(\Omega_k t + \varphi_k); \\ u_M^*(t) &= \sum_{k=0}^{\infty} C_k \sin(\Omega_k t + \varphi_k) \end{aligned}$$

і врахувати, що $\cos x$ і $\sin x$ є парою перетворення Гілберта, то отримаємо

$$s(t) = \sum_{k=0}^{\infty} C_k \cos[(\omega_0 + \Omega_k)t + \varphi_k] + j \sum_{k=0}^{\infty} C_k \sin[(\omega_0 + \Omega_k)t + \varphi_k].$$

Таке зображення є аналітичним для усіх $\omega_0 > 0$. Із заміною модуляційної функції $M[u_M(t)]$ на спряжену $M^*[u_M(t)] = u_M(t) - ju_M^*(t)$ форма сигналу $s(t)$ відповідає *нижній бічній смузі*.

Спектри БМ- і ОМ-сигналів можна отримати зі спектра АМ-сигналу, якщо з нього вилучити або складову на частоті переносника для БМ-сигналу, або складову на частоті переносника та одну з бічних смуг (верхню чи нижню) для ОМ. Такі перетворення спектра АМ-сигналу показані на рис. 3.5.

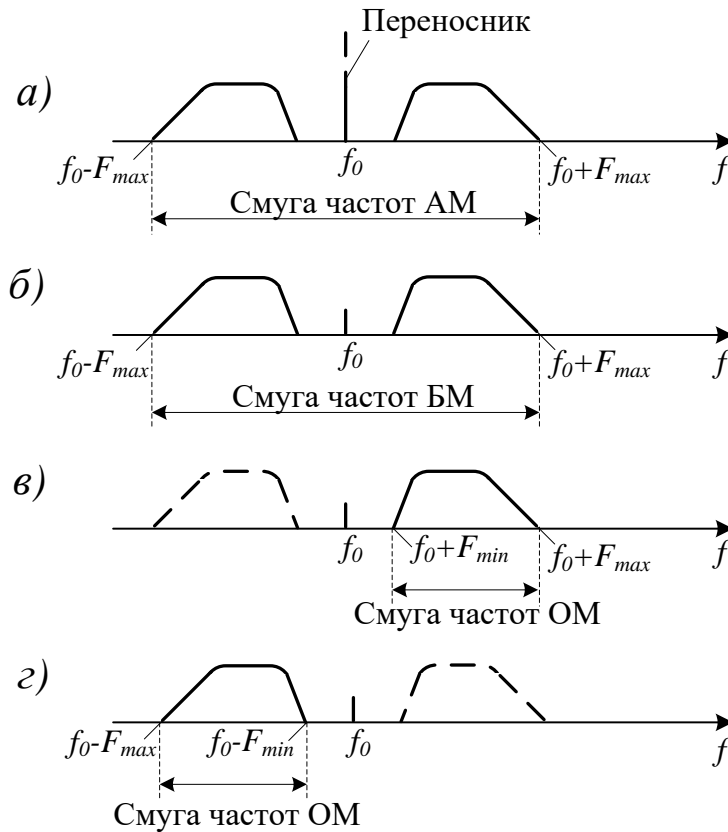


Рис. 3.5. Спектральні діаграми: а) АМ-сигналу; б) БМ-сигналу; в) ОМ-сигналу (верхня бічна смуга); г) ОМ-сигналу (нижня бічна смуга)

Важливою перевагою БМ- і ОМ-сигналів є підвищення швидкості ефективності використання потужності передавача, що покращує відповідно якість приймання таких сигналів. Крім того, при ОМ у два рази зменшується ширина спектра модульованого сигналу, що дозволяє вдвічі збільшити кількість сигналів у заданій смузі частот. Тому ОМ широко застосовується в багатоканальному зв'язку з частотним розділенням.

3.3. Кутова модуляція

До кутових видів модуляції належать частотна (ЧМ) та фазова (ФМ) модуляції. Для частотної модуляції відхилення частоти модульованого сигналу від частоти переносника ω_0 змінюється пропорційно миттєвим значенням модулюючого сигналу:

$$\omega(t) = \omega_0 + \Delta\omega_\delta u_M(t), \quad (3.23)$$

де $\Delta\omega_\delta$ - коефіцієнт пропорційності, який називають *девіацією частоти* (від лат. *deviato* - відхилення); він дорівнює максимальному відхиленню частоти модульованого сигналу від частоти переносника ω_0 .

Зміну частоти ЧМ-сигналу графічно подано на рис. 3.6, де зазначено девіацію частоти $\Delta\omega_\delta$, яка відповідає найбільшому відхиленню частоти вниз: $\Delta\omega_\delta = \Delta\omega_-$, оскільки $\Delta\omega_+ < \Delta\omega_-$. У виразі (3.23) величина $u_M(t)$ нормована, тобто $|u_M(t)| \leq 1$.

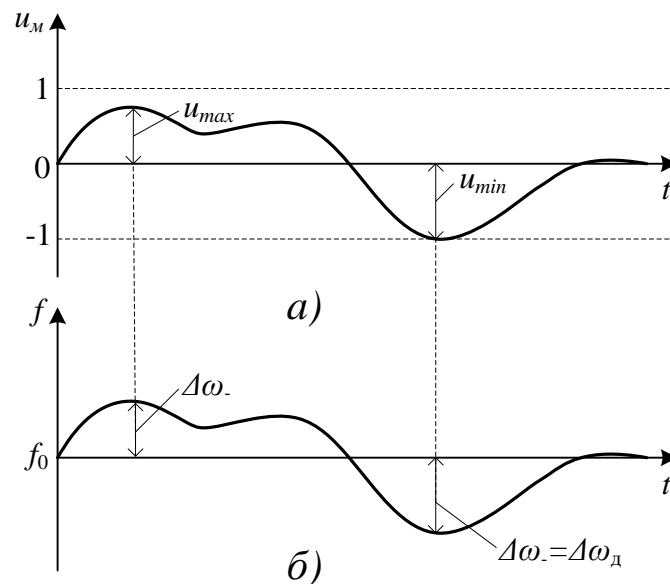


Рис. 3.6. Миттєва частота ЧМ-сигналу: а) модулюючий сигнал; б) зміна миттєвої частоти

Девіація частоти є одним із головних параметрів частоти модуляторів і може набувати значень від одиниць герц до сотень мегагерців у модуляторах різного призначення при обов'язковому виконанні умови $\Delta\omega_\delta \ll \Delta\omega_0$.

Для фазової модуляції відхилення (зсув) фази від лінійної $\cos(\omega_0 t + \psi_0)$ змінюється пропорційно миттєвим значенням модулюючого сигналу:

$$\psi(t) = \omega_0 t + \psi_0 + \Delta\omega_\delta u_m(t). \quad (3.24)$$

Фізичний зміст цього коефіцієнта пояснює рис. 3.7, де зображені модулюючий сигнал і повна фаза ФМ-сигналу. Якщо $u_m(t)$ зростає, то і повна фаза зростає, причому швидше, ніж за лінійним законом. Якщо ж сигнал $u_m(t)$ спадає, маємо зменшення швидкості зростання фази $\psi(t)$. Абсолютна величина відхилення (зсуву) фази від лінійного закону найбільша, коли $u_m(t)$ досягає екстремальних значень.

На рис. 3.7, б зазначені максимальні відхилення фази вгору $\Delta\varphi_+$ та вниз $\Delta\varphi_-$. Найбільше відхилення фази від лінійної і є девіацією фази $\Delta\varphi_\delta$ для ФМ. У прикладі, що наведений на рис. 3.7, $\Delta\varphi_\delta = \Delta\varphi_+$. Вимірюється $\Delta\varphi_\delta$ в радіанах і може набирати значень від одиниць до десятків тисяч радіан.

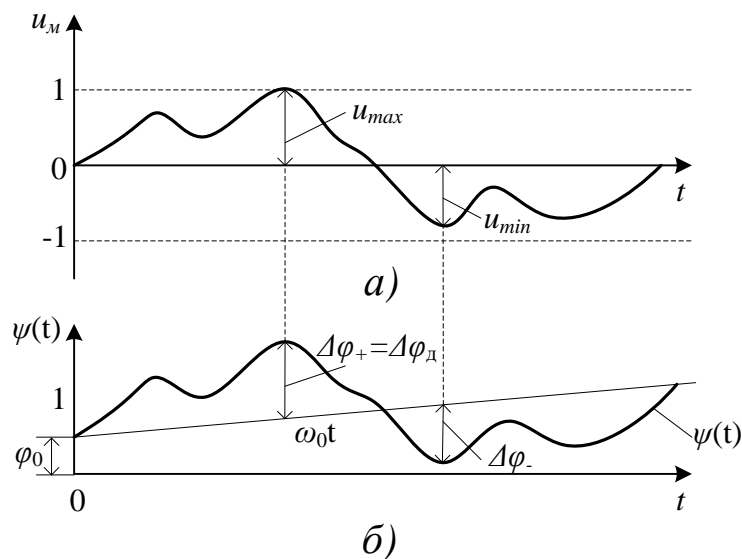


Рис. 3.7. Повна фаза ФМ-сигналу: а) модулюючий сигнал; б) зміна повної фази

Аналітичний вираз (математичну модель) ФМ-сигналу можна записати так:

$$s_{\text{ФМ}}(u_m, t) = A_0 \cos[\omega_0 t + \psi_0 + \Delta\omega_\delta u_m(t)]. \quad (3.25)$$

У випадку кутової модуляції (ЧМ та ФМ) модульована функція має вигляд

$$M[u_m(t)] = e^{j\varphi(t)}. \quad (3.26)$$

При синусоїдальній несучій $f(t) = A_0 e^{j\omega_0 t}$ модульований сигнал матиме такий вираз:

$$\dot{s}(t) = A_0 e^{j[\omega_0 t + \varphi(t)]} = A_0 \cos[\omega_0 t + \varphi(t)] + jA_0 \sin[\omega_0 t + \varphi(t)]. \quad (3.27)$$

Реальний сигнал

$$s(t) = \text{Re } \dot{s}(t) = A_0 \cos[\omega_0 t + \varphi(t)] = A_0 \cos \psi(t). \quad (3.28)$$

Це звичайне зображення сигналу з кутовою модуляцією. Згідно з виразом (3.28) повна фаза високочастотного коливання дорівнює

$$\psi(t) = \omega_0 t + \varphi(t), \quad (3.29)$$

а миттєва частота коливання змінюється за законом похідної від $\psi(t)$, тобто

$$\omega(t) = \frac{d\psi(t)}{dt} = \omega_0 + \frac{d\varphi(t)}{dt}. \quad (3.30)$$

Навпаки, у разі змінення частоти за законом $\omega(t)$ (вираз (3.30)) фаза коливання

$$\psi(t) = \int \omega(t) dt. \quad (3.31)$$

У випадку фазової модуляції $\varphi(t) = \varphi_0 + \Delta\varphi_m u(t)$, де $\Delta\varphi_m$ - девіація фази. Тоді на підставі виразів (3.28) і (3.29)

$$\varphi(t) = \varphi_0(t) + \Delta\varphi_m u(t) + \varphi_0; \quad (3.32)$$

$$s(t) = A_0 \cos[\omega_0 t + \Delta\varphi_m u(t) + \varphi_0]. \quad (3.33)$$

При частотній модуляції за законом первинного сигналу, який характеризує повідомлення, змінюється частота несучого коливання:

$$\omega = \omega_0 + \Delta\omega_\partial u(t), \quad (3.34)$$

де $\Delta\omega_\partial$ - амплітуда частотного відхилення (девіація частоти).

Повна фаза коливання при цьому дорівнює

$$\psi(t) = \int_0^t \omega(t) dt + \varphi_0 = \omega_0 t + \Delta\omega_\partial \int_0^t u(t) dt + \varphi_0. \quad (3.35)$$

Тоді ЧМ-сигнал має вигляд

$$s(t) = A_0 \cos\left[\omega_0 t + \Delta\omega_\partial \int_0^t u(t) dt + \varphi_0\right]. \quad (3.36)$$

При модуляції одним тоном, коли $u(t) = \cos \Omega_m t$, сигнали при ФМ і ЧМ за формою мають однаковий вигляд:

$$s(t) = A_0 \cos(\omega_0 t + m \cos \Omega_m t + \varphi_0), \quad (3.37)$$

де m - індекс модуляції: при фазовій модуляції $m = \Delta\varphi_m$, а при частотній модуляції $m = \frac{\Delta\omega_\partial}{\Omega_m}$.

Для визначення спектра сигналу замінимо у формулі (3.37) косинус суми двох кутів за відомими тригонометричними формулами, взявши для спрощення $\varphi_0 = 0$:

$$s(t) = A_0 \cos(m \sin \Omega_m t) \cos \omega_0 t - A_0 \sin(m \sin \Omega_m t) \sin \omega_0 t. \quad (3.38)$$

З теорії бesselевих функцій відомі такі співвідношення [4]:

$$\sin(m \sin \Omega_M t) = 2 \sum_{k=1}^{\infty} I_{2k-1}(m) \sin(2k-1)\Omega_M t, \quad (3.39)$$

$$\cos(m \sin \Omega_M t) = I_0(m) + 2 \sum_{k=1}^{\infty} I_{2k}(m) \cos 2k\Omega_M t, \quad (3.40)$$

де $I_k(m)$ - бесселева функція першого роду k -го порядку від аргументу m (рис. 3.8).

Після підстановки виразів (3.39) і (3.40) у вираз (3.38) отримаємо

$$\begin{aligned} s(t) = & A_0 I_0(m) \cos \omega_0 t + \sum_{k=1}^{\infty} A_0 I_k(m) \cos(\omega_0 + k\Omega_M)t + \\ & + \sum_{k=1}^{\infty} (-1)^k A_0 I_k(m) \cos(\omega_0 - k\Omega_M)t. \end{aligned} \quad (3.41)$$

Таким чином, навіть при синусоїдальних ЧМ і ФМ отримаємо теоретично нескінченний спектр. Він складається з несучої ω_0 і двох бічних смуг $\omega_0 \pm k\Omega$. Амплітуда несучої $A_0 I_0(m)$ при ЧМ і ФМ, на відміну від АМ, залежить від модулюючого коливання. При деяких значеннях m ($m=2, 3, 4, 5$) вона може дорівнювати нулю. Амплітуда бічних частот $A_k = A_0 I_k(m)$. Однак практично ширина спектра ЧМ- і ФМ-сигналів обмежена. На рис. 3.9 наведено спектр сигналу з кутовою модуляцією одним тоном при $m=5$. Як бачимо, амплітуди бічних частот швидко зменшуються зі збільшенням номера гармоніки k . При $k > m$ складові спектра малі, тому ними можна знехтувати. Практично ширина спектра сигналу при кутовій модуляції дорівнює $F = 2(m+1)F_m$, де $F_m = \frac{\Omega}{2\pi}$ - частота модулюючого коливання.

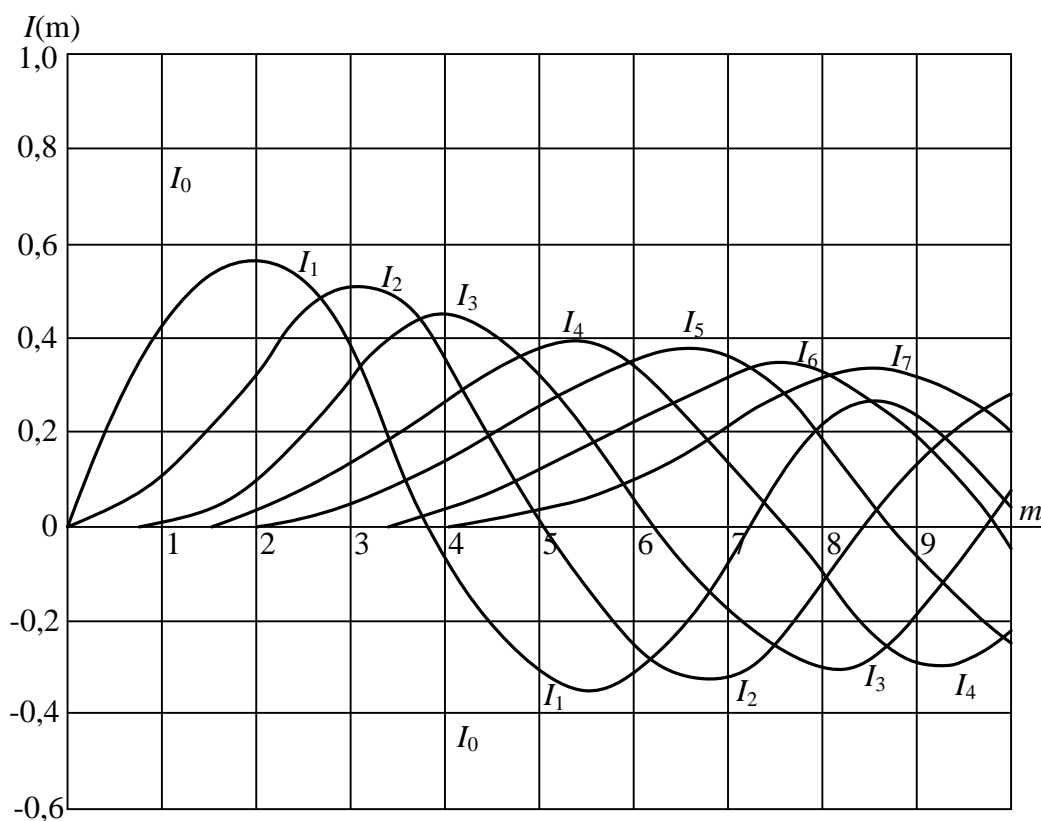


Рис. 3.8. Графіки функції Бесселя 1 роду n-го порядку

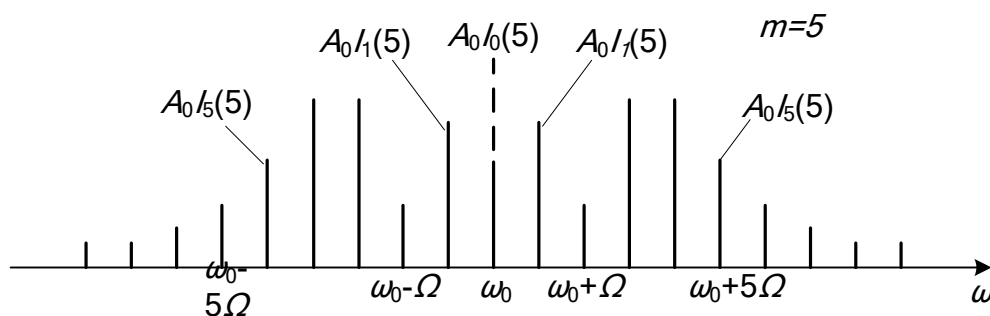


Рис. 3.9. Спектр сигналу з кутовою модуляцією при $m = 5$

Різницю між ФМ і ЧМ можна краще усвідомити, розглядаючи ФМ- і ЧМ-сигнали при модуляції повідомлення $u(t) = \cos \Omega_m t$. У цьому випадку

$$s_{\Phi M}(t) = A_0 \cos(\omega_0 t + \Delta\varphi_m \cos \Omega_m t + \varphi_0); \quad (3.42)$$

$$\begin{aligned}
s_{\text{ЧМ}}(t) &= A_0 \cos(\omega_0 t + \Delta\varphi_m \int \cos \Omega_m t dt + \varphi_0) = \\
&= A_0 \cos\left(\omega_0 t + \frac{\Delta\omega_m}{\Omega} \sin \Omega_m t + \varphi_0\right).
\end{aligned}
\tag{3.43}$$

Порівнюючи вирази (3.42) і (3.43), визначимо еквівалентне значення зміни фази при ЧМ:

$$\Delta\varphi_{\text{еквЧМ}} = \frac{\Delta\omega_m}{\Omega_m}.
\tag{3.44}$$

З отриманого виразу випливає, що при ЧМ зміна фази сигналу обернено пропорційна частоті модулюючого повідомлення, тоді як при ФМ $\Delta\varphi$ є величиною, яка залежить не від частоти модулюючого повідомлення, а від амплітуди:

$$\Delta\omega_{\text{еквФМ}} = \Delta\varphi_m \Omega_m.
\tag{3.45}$$

Отже, при ЧМ девіація частоти не залежить від модулюючого повідомлення, а при ФМ девіація частоти прямо пропорційна частоті модулюючого сигналу. Залежності (3.44) і (3.45) проілюстровані на рис. 3.10.

Різниця між ЧМ і ФМ полягає і в тому, що ширина спектра ФМ-коливання залежить від ширини спектра модулюючого повідомлення, тоді як ширина спектра при ЧМ практично не залежить від ширини спектра модулюючого повідомлення, якщо $\Delta\omega_m \gg F_\epsilon$, і дорівнює

$$\Delta f_{\text{ЧМ}} \approx \frac{2\Delta\omega_m}{2\pi}.$$

Функції Бесселя мають цікаву закономірність: чим вищий порядок k функції Бесселя, тим при більших значеннях аргументу t спостерігається її максимум, але якщо $k > t$, значення функції Бесселя є малою величиною. З цього випливає, що малими будуть і складові спектра, а тому ними можна нехтувати.

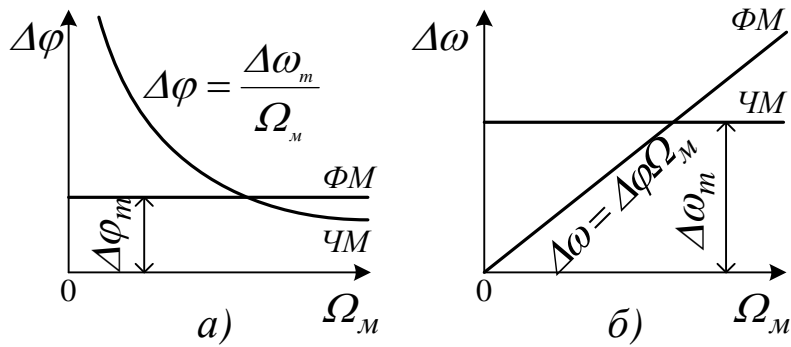


Рис. 3.10. Залежність фази (а) і частоти (б) модульованих коливань при ФМ і ЧМ

Ширина спектра кутових модуляцій залежить від того, з якою амплітудою ми відкидаємо складові спектра. Практично вважають, що можна нехтувати тими спектральними складовими, номери яких $k > m + 1$ (значення амплітуди їх менше за 5% амплітуди переносника). З цього випливає, що ширина спектра сигналів з кутовими модуляціями дорівнює

$$F_{ЧМ} = F_{ФМ} \approx 2(m + 1)F_m, \quad (3.46)$$

де F_m - частота гармонічного модулюючого сигналу.

Інколи вважають, що для забезпечення передавання модулюючого сигналу з більш високою точністю потрібно передавати спектральні складові із значеннями амплітуди до 1% амплітуди переносника. Тоді ширина кутових модуляцій

$$F_{ЧМ} = F_{ФМ} \approx 2(m + \sqrt{m} + 1)F_m.$$

Якщо $m < 0,6$, то кутова модуляція вважається *вузькосмуговою*, її ширина спектра порівнянна з шириною спектра амплітудної модуляції.

Якщо $m \gg 1$, то кутова модуляція є *широкосмуговою*, з формули (3.46) випливає, що ширина смуги частот приблизно дорівнює подвоєній девіації частоти. Кутові модуляції, особливо широкосмугові, мають більшу завадостійкість, ніж амплітудна модуляція, тому вони знаходять широке застосування у системах зв'язку для якісного передавання повідомлень. Однак слід

пам'ятати, що при цьому значно розширюється смуга частот модульованого сигналу.

Якщо модулюючий сигнал є складним, тобто містить у собі ряд гармонічних складових, то спектр ЧМ і ФМ можна знайти описаним раніше способом, використавши вираз для косинуса кількох аргументів. Спектр модульованого сигналу при цьому буде дуже складним і міститиме різні комбінаційні частоти. Але і в цьому разі загальна смуга частот модульованого сигналу наближено обчислюється за формулою (3.46), куди замість F_m , підставляється максимальна частота спектра модулюючого сигналу; індекс модуляції обчислюється на цій максимальній частоті.

У випадку малого індексу модуляції спектр ЧМ- і ФМ-сигналів, як і для АМ, має тільки три складові

$$s(t) = A_0 \cos \omega_0 t + \frac{mA_0}{2} \cos(\omega_0 + \Omega_m)t - \frac{mA_0}{2} \cos(\omega_0 - \Omega_m)t. \quad (3.47)$$

Це безпосередньо випливає з виразу (3.41), якщо врахувати, що при $m \ll 1$ $\sin(m \sin \Omega_m t) \approx m \sin \Omega_m t$, а $\cos(m \sin \Omega_m t) \approx 1$.

Порівняння виразів (3.8) і (3.47) показує, що розбіжність спектрів сигналів при амплітудній і кутовій модуляції полягає тільки в зсуві фаз коливання нижньої бічної частоти на 180° відносно його положення при АМ. Ця розбіжність є суттєвою та ілюструється векторними діаграмами, зображеними на рис. 3.11.

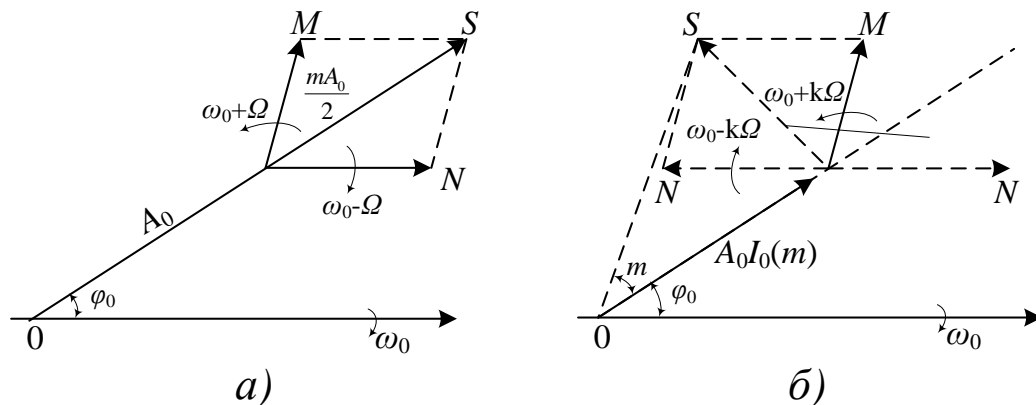


Рис. 3.11. Векторні діаграми АМ-сигналу (а) та ЧМ-сигналу ($m < 1$) (б)

Досить часто використовується *односмугова* кутова модуляція. Якщо функція $\dot{\varphi}(t)$ аналітична, тобто

$$\dot{\varphi}(t) = \varphi(t) + i\varphi^*(t),$$

то сигнал

$$\begin{aligned} \dot{s}(t) &= f(t)M[u_m(t)] = A_0 e^{j[\varphi(t) + j\varphi^*(t)]} = \\ &= A_0 e^{-\varphi^*(t)} \sin[\omega_0 t + \varphi(t)] \end{aligned} \quad (3.48)$$

також є аналітичною функцією при $\omega_0 > 0$. Він не містить від'ємних частот, хоча і має нескінченний спектр в області додатних частот:

$$s(t) = \text{Re } \dot{s}(t) = A_0 e^{-\varphi^*(t)} \cos[\omega_0 t + \varphi(t)]. \quad (3.49)$$

Вираз (3.47) визначає новий модулюючий сигнал, який є варіантом сигналу односмугової кутової модуляції. Щоб довести це, розглянемо частотну модуляцію одним тоном $u(t) = \sin \Omega_m t$. Для цього випадку функція $\varphi(t)$ і її перетворення Гілберта набувають вигляду

$$\varphi(t) = \Delta\omega_\delta \int u_m(t) dt = m \sin \Omega_m t, \quad \varphi^*(t) = -m \cos \Omega_m t,$$

де m - індекс модуляції, $m = \frac{\Delta\omega_\delta}{\Omega_m}$.

Модулююча функція в такому разі

$$M[u_m(t)] = e^{j[\varphi(t) + j\varphi^*(t)]} = \sum_{k=0}^{\infty} \frac{m^k}{k!} e^{ik\Omega_m t},$$

а модулюючий сигнал

$$\dot{s}(t) = M[u_m(t)] = e^{j\omega_0 t} = \sum_{k=0}^{\infty} \frac{m^k}{k!} e^{j(\omega_0 + k\Omega_m)t}.$$

Звідси бачимо, що спектр модульованого сигналу складається з одної бічної смуги частот. Односмуговий ЧМ-сигнал можна отримати зі звичайного ФМ-сигналу шляхом перетворення Гілберта (наприклад, фазовим зсувом на $\pi/2$) і модуляції амплітуди за експоненціальним законом. Тоді обмеження такого сигналу в приймачі відновить нижню бічну смугу частот і дозволить застосувати для детектування звичайний дискримінатор.

3.4. Дискретна модуляція гармонічного переносника

Дискретна модуляція (маніпуляція) є окремим випадком модуляції гармонічного переносника, коли модулюючий сигнал $u_m(t)$ - дискретний. При дискретній модуляції закодоване повідомлення $u(t)$ - послідовність кодових символів $\{a_i\}$ - перетворюється на послідовність елементів сигналу $\{s_i\}$. Останні відрізняються від кодових символів тільки електричним представленням. Дискретна модуляція може також полягати у взаємодії кодових символів $\{a_i\}$ з переносником $f(t)$. Така дискретна модуляція аналогічна неперервній.

У процесі модуляції один з параметрів переносника змінюється за законом, який визначається кодом. У разі безпосереднього передавання переносником може бути постійний струм, змінними параметрами якого є величина і напрямок. Звичайно за переносник, як і при неперервній модуляції, використовується змінний струм (гармонічне коливання). У цьому випадку можна отримати амплітудну (АМн), частотну (ЧМн) і фазову (ФМн) дискретні модуляції.

Дискретну модуляцію часто називають *маніпуляцією*, а пристрій, що здійснює дискретну модуляцію (дискретний модулятор), називають *маніпулятором* або *генератором сигналів*. На рис. 3.12 наведені графіки сигналів при різних видах маніпуляції.

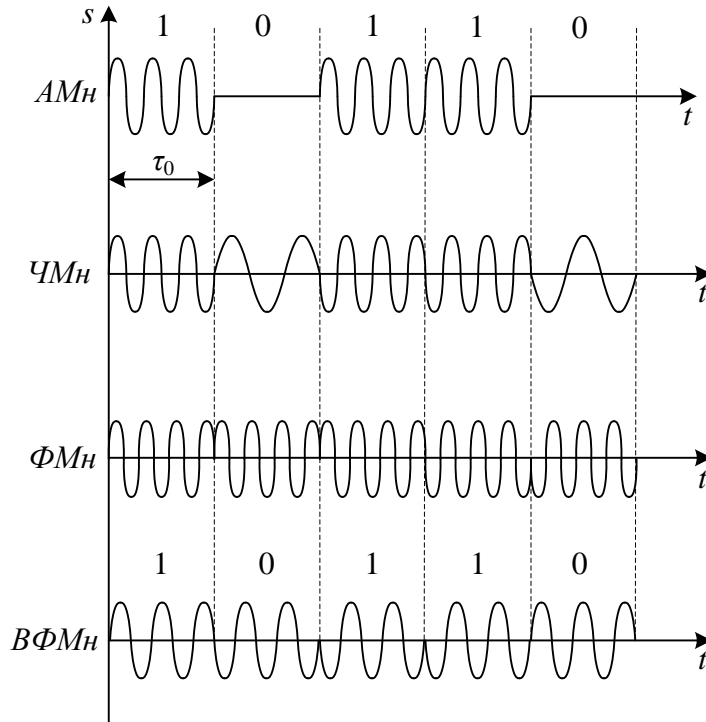


Рис. 3.12. Сигнали при різних видах дискретної модуляції

При АМн символу 1 відповідає передавання несучого коливання впродовж часу τ_0 (посилка), а символу 0 - відсутність коливання (пауза). При ЧМн передавання несучого коливання з частотою f_1 відповідає символу 1, а передавання коливання з частотою f_0 відповідає символу 0. При ФМн змінюється фаза несучої на 180° на кожному переході від 1 до 0 і від 0 до 1.

Крім цих видів маніпуляції, застосовується *відносна фазова модуляція* (ВФМн). На відміну від ФМн, у системі ВФМн фаза несучого коливання змінюється на 180° при передаванні символів 1 і залишається незмінною при передаванні символів 0.

Маніпуляція кожної даної послідовності здійснюється відносно попередньої. Очевидно, таким чином можна змінювати будь-який параметр несучого коливання: змінюючи частоту, отримаємо *відносну частотну маніпуляцію* (ВЧМн), а змінюючи амплітуду, - *відносну амплітудну маніпуляцію* (ВАМн).

Розглянемо спектри сигналів при деяких видах дискретної модуляції. Вважатимемо, що модуляція виконується двійковим повідомленням $u(t)$, що є періодичною послідовністю імпульсів з періодом $T = 2\tau_0$.

Амплітудна маніпуляція. АМн-сигнал можна записати у вигляді

$$s(t) = A_0 u(t) \sin(\omega_0 t + \varphi_0), \quad (3.50)$$

де періодична функція $u(t)$ на інтервалі $T = 2\tau_0$ дорівнює

$$u(t) = \begin{cases} 1 & \text{при } 0 < t < \tau_0, \\ 0 & \text{при } -\tau_0 < t < 0. \end{cases} \quad (3.51)$$

Подано $u(t)$ через ряд Фур'є:

$$u(t) = \frac{1}{2} + \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin k\Omega_M t. \quad (3.52)$$

Тоді АМн-сигнал запишеться у вигляді

$$\begin{aligned} s(t) &= A_0 \left(\frac{1}{2} + \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin k\Omega_M t \right) \sin(\omega_0 t + \varphi_0) = \\ &= \frac{1}{2} A_0 \sin(\omega_0 t + \varphi_0) + \frac{A_0}{2\pi} \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \cos(\omega_0 t - k\Omega_M t + \varphi_0) - \\ &- \frac{A_0}{2\pi} \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \cos(\omega_0 t + k\Omega_M t + \varphi_0). \end{aligned} \quad (3.53)$$

Спектр АМн-сигналу, побудований за формулою (3.53), показано на рис. 3.13, а. Він складається з несучого колювання з амплітудою $\frac{1}{2} A_0$ двох бічних смуг, спектральні складові яких мають амплітуди

$$A_k = \frac{A_0}{2\pi} \frac{1 - \cos k\pi}{k} = \frac{A_0}{k\pi} \sin \frac{k\pi}{2}. \quad (3.54)$$

Обвідна спектра дискретного АМн-сигналу виражається

формулою

$$A(\omega) = \frac{A_0 \tau_0}{2} \frac{\sin \frac{1}{2}(\omega - \omega_0)\tau_0}{\frac{1}{2}(\omega - \omega_0)\tau_0}, \quad (3.55)$$

тобто спектр є зсунутим на частоту ω_0 спектром одиночного імпульсного сигналу $u(t)$.

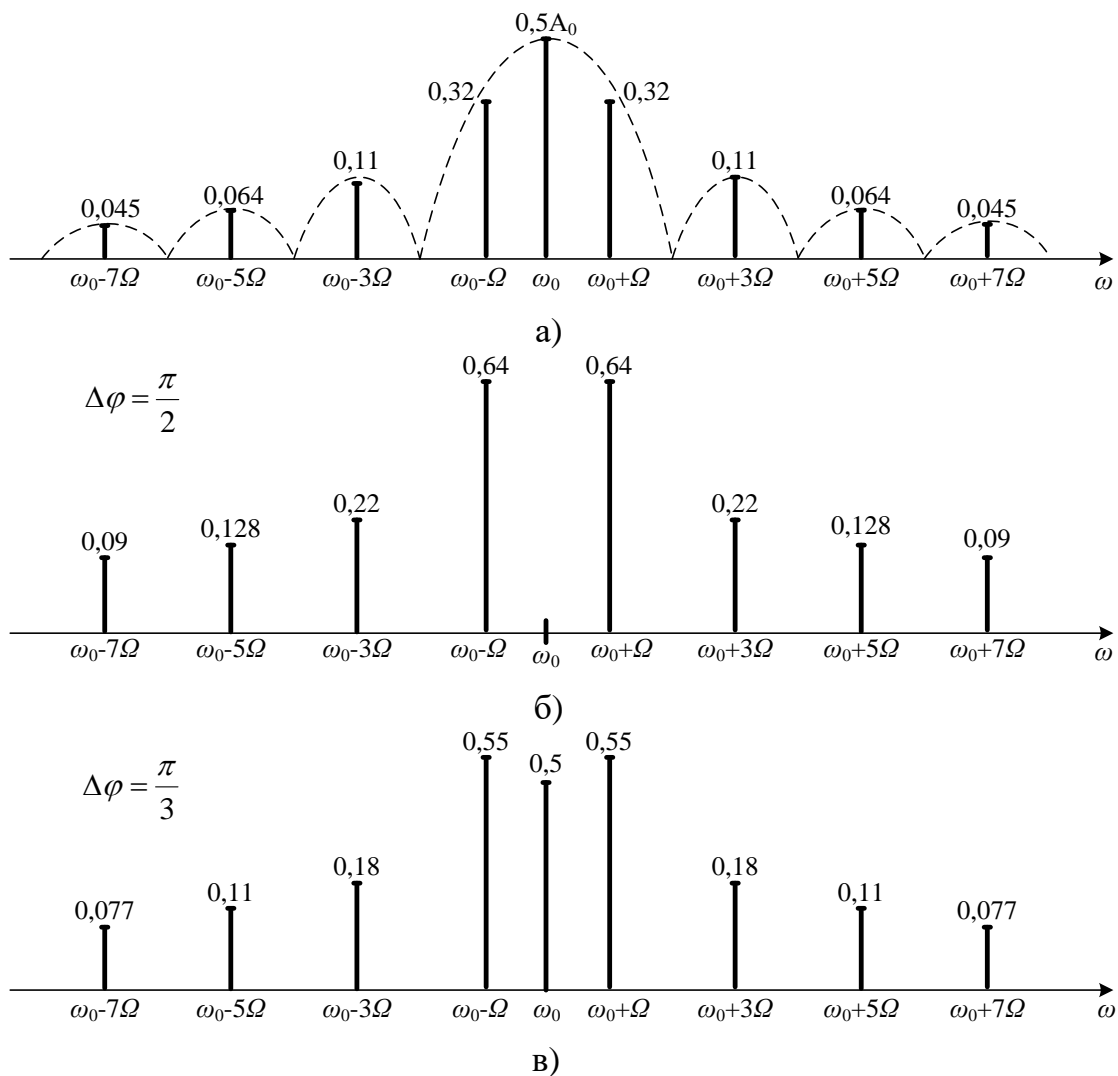


Рис. 3.13. Спектр сигналу при амплітудній (а) і фазовій (б, в) маніпуляціях

Фазова маніпуляція. ФМН-сигнал можна записати у вигляді

$$\begin{aligned} s(t) &= A_0 \sin(\omega_0 t + \Delta\varphi u(t) + \varphi_0) = \\ &= A_0 \cos[\Delta\varphi u(t)] \sin(\omega_0 t + \varphi_0) + A_0 \sin[\Delta\varphi u(t)] \cos(\omega_0 t + \varphi_0). \end{aligned} \quad (3.56)$$

Періодична функція, яка визначає закон змінення фази на інтервалі $T = 2\tau_0$, виражається формулою

$$u(t) = \begin{cases} 1 \text{ при } 0 < t < \tau_0, \\ -1 \text{ при } -\tau_0 < t < 0. \end{cases} \quad (3.57)$$

Підстановка виразу (3.57) у вираз (3.56) дає

$$s(t) = A_0 \cos \Delta\varphi \sin(\omega_0 t + \varphi_0) + u(t) \sin \Delta\varphi \cos(\omega_0 t + \varphi_0). \quad (3.58)$$

Подамо $u(t)$ через ряд Фур'є:

$$u(t) = \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin k\Omega_M t.$$

Тоді ФМН-сигнал запишеться у вигляді

$$\begin{aligned} s(t) &= A_0 \cos \Delta\varphi \sin(\omega_0 t + \varphi_0) + \frac{2A_0}{\pi} \sin \Delta\varphi \cos(\omega_0 t + \varphi_0) \times \\ &\times \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin k\Omega_M t = A_0 \cos \Delta\varphi \sin(\omega_0 t + \varphi_0) - \\ &- \frac{A_0}{\pi} \sin \Delta\varphi \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin(\omega_0 t - k\Omega_M t + \varphi_0) + \\ &+ \frac{A_0}{2} \sin \Delta\varphi \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin(\omega_0 t - k\Omega_M t + \varphi_0). \end{aligned} \quad (3.58)$$

Спектр ФМн-сигналу для різноманітних значень девіації фази $\Delta\varphi$, побудований за формулою (3.58), наведений на рис. 3.13, б, в. Він складається з несучого колювання і двох бічних смуг. Амплітуда несучого колювання залежить від $\Delta\varphi$:

$A(\omega_0) = A_0 \cos \Delta\varphi$ і при $\Delta\varphi = \pm \frac{\pi}{2}$ дорівнює нулю. Амплітуди спектральних складових у бічних смугах також залежать від $\Delta\varphi$.

Із збільшенням $\Delta\varphi$ від 0 до $\frac{\pi}{2}$ (див. рис. 3.13, б, в) амплітуда несучого колювання зменшується до нуля, а амплітуди бічних частот збільшуються.

Коли $\Delta\varphi = \frac{\pi}{2}$, вся енергія ФМн-сигналу зосереджується тільки в бічних смугах. Як і при АМ, обвідна дискретного спектра бічних частот є зсунутою на частоту ω_0 спектром одиночного імпульсного сигналу $u(t)$, помноженим на $\sin \Delta\varphi$:

$$A(\omega) = A_0 \tau_0 \sin \Delta\varphi \frac{\sin \frac{1}{2}(\omega - \omega_0)\tau_0}{\frac{1}{2}(\omega - \omega_0)\tau_0}. \quad (3.59)$$

Аналогічно визначається спектр сигналу при частотній маніпуляції.

3.5. Імпульсна модуляція

У системах зв'язку з імпульсною модуляцією переносником інформації служить періодична послідовність імпульсів однакової форми (рис. 3.14, а)

$$f(t) = \sum_{k=-\infty}^{\infty} A_0 u(t - t_k), \quad (3.60)$$

де $u(t)$ - нормована функція, яка характеризує форму імпульсу;

A_0 - амплітуда імпульсу;
 t_k - початок переднього фронту k -го імпульсу, $t_k = kT_i + t_0$;
 T_i - період прямування імпульсів;
 t_0 - початок відліків послідовності.

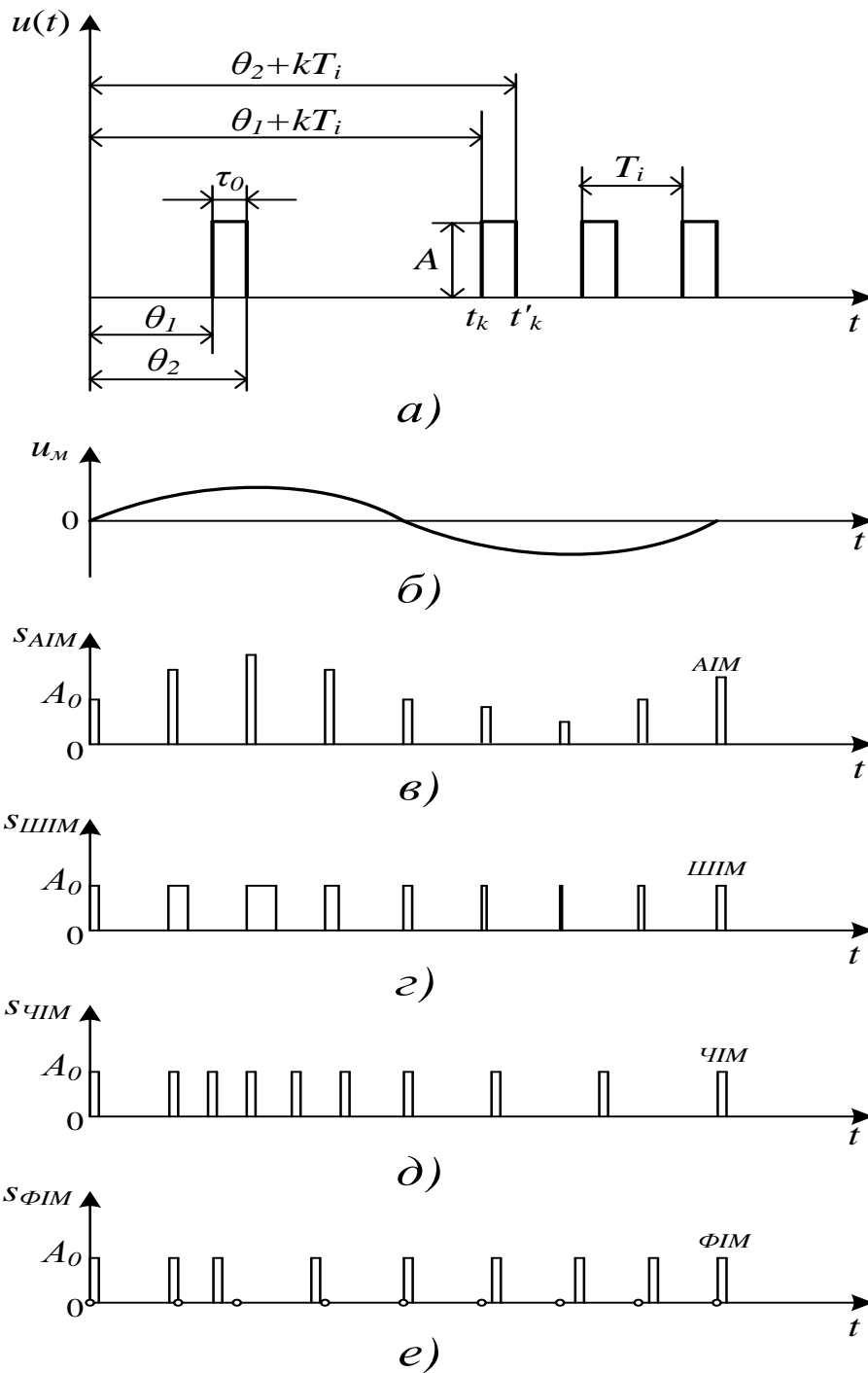


Рис. 3.14. Параметри періодичної послідовності прямокутних імпульсів (а), модулюючий сигнал (б) та модульовані сигнали при різних видах імпульсної модуляції (в-е)

При модуляції один з параметрів послідовності змінюється відповідно до модульованого сигналу. Так, при *амплітудно-імпульсній модуляції* (АІМ) змінюється амплітуда імпульсу (рис. 3.14, в):

$$A(t) = A_0 + \Delta A u(t). \quad (3.61)$$

При *широтно-імпульсній модуляції* (ШІМ) змінюється тривалість імпульсу (рис. 3.14, г):

$$\tau(t) = \tau_0 + 2\Delta\tau_m u(t), \quad (3.62)$$

де $\Delta\tau_m$ - максимальне відхилення фронту імпульсів в один бік.

При *фазово-імпульсній модуляції* (ФІМ) змінюється зсув імпульсів відносно тактових точок kT_i , (рис. 3.14, е):

$$t_k = \theta(t) = kT_i + \Delta\tau_m u(t). \quad (3.63)$$

При *частотно-імпульсній модуляції* (ЧІМ) відповідно до повідомлення, що передається, змінюється частота прямування імпульсів (рис. 3.14, д).

Як при ФІМ, імпульси зміщуються відносно тактових точок, але за іншою закономірністю. Різниця ФІМ і ЧІМ аналогічна різниці між ФМ і ЧМ синусоїдального переносника.

Періодичну послідовність прямокутних імпульсів можна записати у вигляді

$$f(t) = \sum_{-\infty}^{\infty} A_0 u(t - t_k) = \begin{cases} A_0 & \text{при } \theta_1 + kT_i < t < \theta_2 + kT_i, \\ 0 & \text{при } \theta_2 + kT_i < t < \theta_1 + (k+1)T_i. \end{cases} \quad (3.64)$$

Її можна подати рядом Фур'є:

$$f(t) = \frac{1}{2} \sum_{k=-\infty}^{\infty} A_k e^{jk\omega_1 t},$$

$$\text{де } A_k = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \exp(-j\omega_1 t) dt, \quad \omega_1 = \frac{2\pi}{T}.$$

У нашому випадку

$$A_k = \frac{2A_0}{T_i} \int_{-\tau_0/2}^{\tau_0/2} e^{-jk\omega_i t} dt = \frac{2A_0\tau_0 \left\{ \exp\left[-j\frac{k\omega_i\tau_0}{2}\right] - \exp\left[j\frac{k\omega_i\tau_0}{2}\right] \right\}}{ik\omega_i T_i} =$$

$$= \frac{2A_0\tau_0}{T_i} \frac{\sin \frac{k\omega_i\tau_0}{2}}{\frac{k\omega_i\tau_0}{2}},$$
(3.65)

тоді

$$f(t) = \frac{A_0\tau_0}{T_i} \sum_{k=-\infty}^{\infty} \frac{\sin \frac{k\omega_i\tau_0}{2}}{\frac{k\omega_i\tau_0}{2}} e^{jk\omega_i(t-t_0)} =$$

$$= \frac{A_0\tau_0}{T_i} \left[1 + 2 \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_i\tau_0}{2}}{\frac{k\omega_i\tau_0}{2}} \cos k\omega_i(t-t_0) \right],$$
(3.66)

де $\tau_0 = \theta_2 - \theta_1$; $t_0 = \frac{\theta_1 + \theta_2}{2}$; $\omega_i = \frac{2\pi}{T_i}$.

Спектр амплітуд періодичної послідовності прямокутних імпульсів наведений на рис. 3.15. Амплітуди спектральних компонентів A_k визначаються модулями спектральної щільності

$|A_k|$ (вираз (3.65)) на гармоніках частоти повторення $\omega_i = \frac{2\pi}{T_i}$.

Форма обвідної частотного спектра періодичної послідовності залежить від форми окремого імпульсу. Зі збільшенням періоду повторення інтервал частот між сусідніми спектральними компонентами скорочується, їх число збільшується, а амплітуда кожної компоненти зменшується із збереженням сталого співвідношення між ними. У разі необмеженого

збільшення T_i періодична послідовність вироджується в одиночний імпульс, а лінійний спектр стає суцільним.

Спектр періодичної послідовності радіоімпульсів отримується зі спектра послідовності відеоімпульсів перенесенням шкали частот на несучу частоту ω_0 і доповненням отриманого спектра його дзеркальним відображенням.

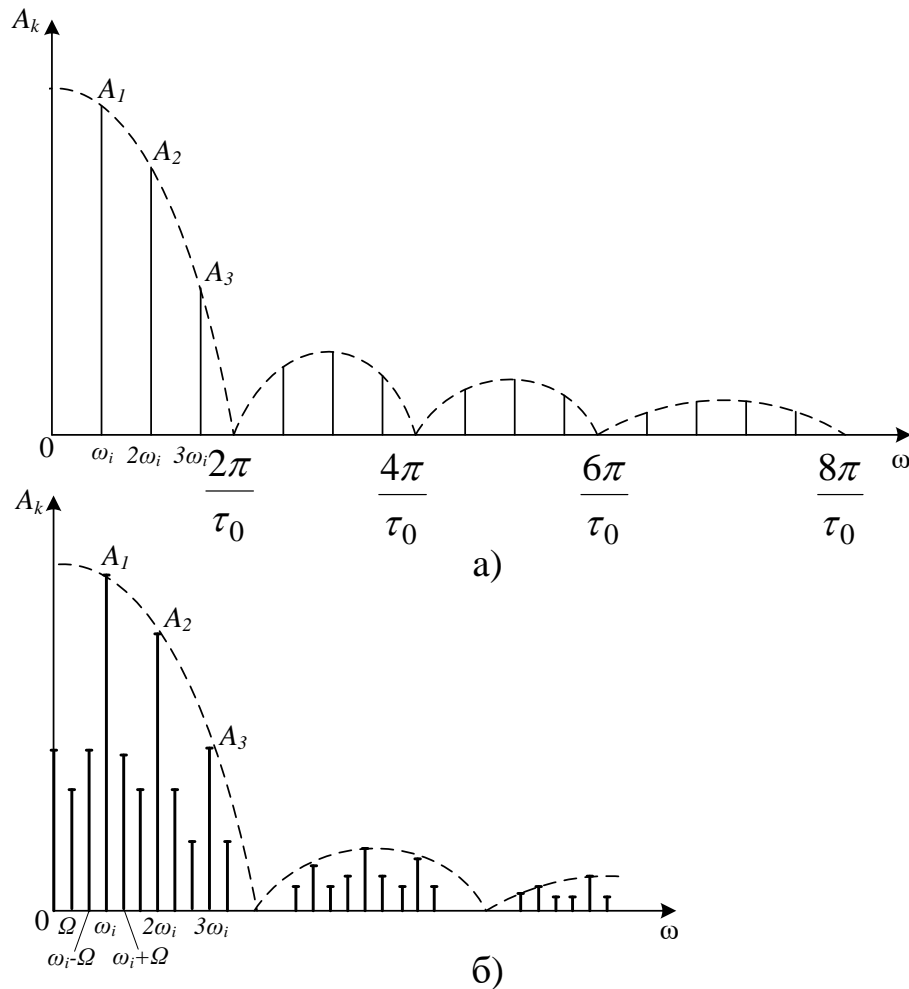


Рис. 3.15. Спектр періодичної послідовності прямокутних імпульсів (а) та спектр АІМ-сигналу (б)

При модуляції параметри, які входять у вирази (3.64) і (3.66), є функціями часу: $A_0 = A(t)$; $\theta_1 = \theta_1(t)$; $\theta_2 = \theta_2(t)$. Модульована послідовність буде тепер уже неперіодичною функцією, яка деформована відносно початкової:

$$s(t) = \sum_{k=-\infty}^{\infty} A(t)u(t-t_k) = \begin{cases} u(t) \text{ нпу } \theta_1(t) + kT_i < t < \theta_2(t) + kT_i, \\ 0 & \text{нпу } \theta_2(t) + kT_i < t < \theta_1(t) + (k+1)T_i, \end{cases} \quad (3.67)$$

або

$$\begin{aligned} s(t) &= u(t) \frac{\theta_2(t) - \theta_1(t)}{T_i} = \\ &= 1 + 4 \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_1}{2} [\theta_2(t) - \theta_1(t)]}{k\omega_i [\theta_2(t) - \theta_1(t)]} \cos k\omega_i \left[t - \frac{\theta_1(t) + \theta_2(t)}{2} \right]. \end{aligned} \quad (3.68)$$

Остання формула визначає частотний спектр деформованої послідовності імпульсів. Для отримання спектрів каналів при різних видах модуляції у формулу (3.68) необхідно підставити відповідний вираз для модульованого сигналу.

Для прикладу знайдемо спектр при АІМ. У разі модуляції одним тоном $u(t) = \sin \Omega(t)$ і $A = A_0(1 + m \sin \Omega t)$, інші параметри послідовності незмінні: $\theta_2 - \theta_1 = \tau_0$, $\theta_1 = 0$ і $\theta_2 = \tau_0$. Після підстановки цих значень для частотного спектра АІМ-сигналу маємо

$$\begin{aligned} s(t) &= \frac{A_0\tau_0}{T_i} \left[1 + 2 \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_i\tau_0}{2}}{\frac{k\omega_i\tau_0}{2}} \right] \cos k\omega_i \left(t - \frac{t_0}{2} \right) + \frac{mA_0\tau_0}{T_i} \sin \Omega_m t + \\ &+ \frac{mA_0\tau_0}{T_i} \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_i\tau_0}{2}}{\frac{k\omega_i\tau_0}{2}} \sin \left[(k\omega_i + \Omega_m)t - \frac{k\omega_i\tau_0}{2} \right] + \\ &+ \frac{mA_0\tau_0}{T_i} \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_i\tau_0}{2}}{\frac{k\omega_i\tau_0}{2}} \sin \left[(k\omega_i - \Omega_m)t - \frac{k\omega_i\tau_0}{2} \right]. \end{aligned} \quad (3.69)$$

На рис. 3.15 наведено графік спектра АІМ-сигналу. Він показує, що при АІМ модулюється за амплітудою кожна складова спектра немодульованої послідовності імпульсів як ізольована несуча. У спектрі є низькочастотне модулююче повідомлення $u(t)$ з частотою Ω_m ; отже, демодуляція при АІМ може бути здійснена за допомогою фільтра нижніх частот, який пропускає низькочастотні коливання $u(t)$.

Аналогічно визначається спектр й інших видів імпульсної модуляції. Для обчислення спектра при ФІМ у формулу (3.68) слід підставити вираз (3.63), яким визначається зміна положення імпульсу відповідно до повідомлення, що передається, а при ШІМ - вираз (3.62), за яким визначається зміна тривалості імпульсу.

При *імпульсно-кодовій модуляції* (ІКМ) передавання окремих значень сигналу зводиться до передавання окремих груп імпульсів. Ці групи передаються одна за одною через відносно великі проміжки часу порівняно з тривалістю окремих імпульсів. Кожна кодова група імпульсів є регулярним неперіодичним сигналом, спектр якого може бути обчислений на підставі перетворень Фур'є (див. гл.8).

Ширина спектра послідовності імпульсів практично не залежить від частоти повторення ω_r і визначається, головним чином, шириною спектра одного імпульсу. У разі наявності модуляції будь-якого виду спектр розширюється незначно внаслідок бічних частот крайніх складових спектра немодульованих імпульсів. Тому робоча смуга частот, яка зайнята імпульсними сигналами, практично залежить не від виду модуляції, а від тривалості та форми імпульсу.

3.6. Модуляція шумової несучої

За переносники можна використовувати не тільки періодичні коливання, але і вузькосмуговий випадковий процес. Такі переносники знаходять практичне застосування, наприклад, в оптичних системах зв'язку з некогерентним випромінюванням, де сигналом є вузькосмуговий гаусівський шум.

Вузькосмуговий випадковий процес можна зобразити як квазігармонічне коливання

$$f(t) = u(t) \cos[\omega_0 t + \varphi(t)] = u(t) \cos \psi(t), \quad (3.70)$$

обвідна $u(t)$ якого повільно змінюється з фазою $\psi(t)$. При амплітудній модуляції відповідно до повідомлення, що передається, змінюється обвідна $u(t)$, при фазовій модуляції - фаза $\varphi(t)$, при частотній - миттєва частота $\omega(t) = \frac{d\psi}{dt}$.

Розглянемо амплітудну модуляцію шумової несучої. Вираз для модульованої несучої можна записати у вигляді

$$y(t) = [1 + mu(t)]f(t), \quad (3.71)$$

де $f(t)$ - переносник;

$u(t)$ - модулююча функція (відеосигнал);

m - коефіцієнт модуляції.

Припускають, що модулюючий процес $u(t)$ є стаціонарним нормальним процесом з середнім значенням, яке дорівнює нулю: $u(t) = 0$. Процеси $f(t)$ і $u(t)$ незалежні. За цих умов функція кореляції $B_y(\tau)$ модульованої за амплітудою шумової несучої буде

$$\begin{aligned} B_y(\tau) &= \overline{y(t)y(t+\tau)} = \frac{\overline{[1 + mu(t)]f(t)[1 + mu(t+\tau)]f(t+\tau)}}{\overline{[1 + mu(t) + mu(t+\tau) + m^2u(t)u(t+\tau)]f(t+\tau)}} = \\ &= [1 + m^2 B_u(\tau)] B_f(\tau), \end{aligned} \quad (3.72)$$

де $B_i(\tau)$ - функція кореляції, що характеризує модулюючу функцію;

B_f - функція кореляції, що характеризує сигнал-переносник.

Тепер знаходимо енергетичний спектр

$$G_y(\omega) = 2 \int_0^{\infty} B_y(\tau) \cos \omega \tau d\tau = 2 \int_0^{\infty} B_f(\tau) \cos \omega \tau d\tau + \quad (3.73)$$

$$+ 2m^2 \int_0^{\infty} B_f(\tau) \cos \omega \tau d\tau.$$

Перший інтеграл дає енергетичний спектр шумової несучої $G_j(\omega)$. Для другого інтеграла на основі теореми про спектр добутків маємо

$$\int_0^{\infty} B_i(\tau) B_f(\tau) \cos \omega \tau d\tau = \frac{1}{2} \int_{-\infty}^{\infty} G_f(\nu) G_i(\omega - \nu) d\nu.$$

Остаточно спектр модульованої несучої дорівнює

$$G_y(\omega) = G_j(\omega) + m^2 \int_{-\infty}^{\infty} G_f(\tau - \omega) G_i(\omega - \nu) d\nu.$$

Таким чином, спектр, модульований за амплітудою шумової несучої, отримують суперпозицією спектра несучої і згорткою цього спектра зі спектром сигналу повідомлення, яке передається, зсунутого в область високих частот на величину ω_0 . Аналогічно визначаються функція кореляції та енергетичний спектр при ФМ і ЧМ.

Застосування "шумових" сигналів дозволяє послабити вплив завмирання в каналах з багатопроменевим розповсюдженням радіохвиль. Пояснимо це на простому прикладі. Нехай на вхід приймача надходять сигнали двох променів $\xi(t)$ і $\xi(t - \tau)$ зі зсувом на час τ . Потужність результуючого сигналу, яка визначається за достатньо великий час T ,

$$P = \overline{[\xi(t) + \xi(t - \tau)]^2} = 2[P_0 + B_\xi(\tau)],$$

де $B_\xi(\tau)$ - функція кореляції сигналу;

P_0 - його середня потужність.

Функція кореляції шуму зменшується зі збільшенням τ і тим швидше, чим ширший його спектр. Відповідно при достатньо великій ширині спектра можна вважати, що $B_{\xi}(\tau) \approx 0$ і $P \approx 2P_0$, тобто середня потужність прийнятого сигналу, незважаючи на завмирання, залишається приблизно постійною.

Застосування в якості переносника реалізацій *реального шуму* пов'язано з труднощами, що виникають при формуванні і прийманні таких сигналів. Тому на практиці знайшли застосування шумоподібні сигнали, які не є випадковими; вони формуються за певним алгоритмом. Однак їх статистичні властивості близькі до властивостей шуму: енергетичний спектр майже рівномірний, а функція кореляції має вузький основний пік і невеликі бічні викиди.

У теперішній час відомі методи формування шумоподібних сигналів, які у разі великої бази ($2TF$) дозволяють незалежно відтворювати їх на приймальному і передавальному кінцях і відповідають вимогам синхронізації цих сигналів.

Широко застосовуються *дискретні сигнали*, що будуються на розбитті інформаційної послідовності тривалістю T на N бінарних елементів тривалістю $\tau = \frac{T}{N}$. Таке розбиття дозволяє отримати

сигнал тривалістю T зі смугою $F \approx \frac{1}{\tau}$ і значенням бази $2TF \gg 1$.

Послідовності бінарних елементів утворюють коди, які вибирають так, щоб забезпечити задані властивості сигналу. За допомогою модуляції або гетеродинування формується високочастотний сигнал, який передається по каналу. Часто при цьому використовується модуляція фази на два положення: 0 і π .

Функція кореляції дискретних сигналів при достатньо великому числі елементів N має головний максимум, зосереджений в області $-\tau \leq \Delta\tau \leq \tau$, і бічні пелюстки, які мають порівняно малий рівень (рис. 3.15). Ця функція дуже нагадує функцію автокореляції відрізка шуму зі смугою F .

У системах зв'язку, де використовуються шумоподібні (складові) сигнали, кожний елемент повідомлення передається не одним, а кількома елементами сигналу, що несуть (повторюють) одну й ту саму інформацію. Число N може сягати кількох сотень і

навіть тисяч. Як буде показано в подальшому, це дає змогу реалізувати накопичення сигналу і забезпечити високу завадостійкість навіть тоді, коли рівень сигналу нижчий за рівень завад.

Деякі дискретні сигнали будуються на основі *лінійних рекурентних послідовностей*. Ці сигнали мають добрі кореляційні властивості та порівняно нескладну практичну реалізацію. Структура сигналів носить випадковий характер, хоча спосіб їх формування цілком регулярний. Неперервні ФМ-сигнали, побудовані на рекурентних послідовностях, можуть мати ідеальну автокореляційну функцію.

Серед лінійних рекурентних послідовностей особливе місце займають *псевдовипадкові M-послідовності Хаффмена*. Вони являють собою сукупність N періодично повторюваних символів d_i , кожний з яких може набувати одного з двох значень: +1 або -1, що визначається взятим з протилежним знаком добутком двох або більшого числа (але завжди парного) попередніх сигналів [39]:

$$d_i = -d_{i-n}d_{i-m}\dots d_{i-e}d_{i-k}, \text{ причому}$$

$$n > m > \dots > e > k \geq 1; i = (n+1) \div N. \quad (3.74)$$

Якщо вибрати початкову послідовність d_1, d_2, \dots, d_n , то на підставі виразу (3.74) можна утворити неповторну елементарну послідовність $\{d_i\}$ із N символів, де

$$N = 2^n - 1. \quad (3.75)$$

Вона вміщуватиме всі комбінації n символів з двох елементів: +1 і -1, окрім комбінацій, які складаються з одних від'ємних одиниць. Внаслідок цього кожна послідовність $\{d_i\}$ вміщує 2^{n-1} додатних одиниць і $2^{n-1} - 1$ від'ємних одиниць. Тому

$$\sum_{i=1}^N d_i = 1. \quad (3.76)$$

При $i > N$ символи повторюються в тому самому порядку, тобто при будь-якому цілому

$$d_{i+pN} = d_i. \quad (3.77)$$

Розглянемо найпростіший приклад: $n = 2$. У цьому випадку $d_i = d_{i-2}d_{i-1}$. Виберемо початкову послідовність $d_1 = -1, d_2 = +1$. Тоді послідовність буде такою: $\dots -1, +1, +1, -1, +1, +1, -1, +1; \dots$ Вона складає всі можливі комбінації з двох символів: $-1, +1, +1, +1, +1, -1$, окрім забороненої комбінації $-1, -1$. Отримана послідовність $-1, +1, +1$ повторюється через $N = 2^2 - 1 = 3$ символи. При $n = 3$ можна утворити дві послідовності з $N = 2^3 - 1 = 7$ символів за правилами $d_i = -d_{i-3}d_{i-2}$ і $d_i = -d_{i-3}d_{i-1}$. Майже кожному цілому числу n відповідає кілька чисел k , при яких за правилом (3.74) утворюється послідовність.

З виразу (3.77) випливає, що число N є максимальним періодом нескінченної послідовності Хаффмена. Можуть утворитися також послідовності меншого періоду. Максимальне число різних послідовностей максимального періоду для будь-якого n дорівнює

$$M = \frac{1}{2} \varphi(2^n - 1), \quad (3.78)$$

де $\varphi(x)$ - функція Ейлера (функція Ейлера $\varphi(x)$ дорівнює кількості цілих чисел, включаючи одиницю, менших за число x і взаємно простих з x . Наприклад, якщо x дорівнює простому числу, то $\varphi(q) = q - 1$).

Бінарні псевдовипадкові послідовності Хаффмена мають ряд чудових властивостей. Нормована функція автокореляції в неперервному режимі роботи має головний максимум, який дорівнює одиниці, і однакові за величиною бічні пелюстки, які

дорівнюють $-1/N$. Функція взаємної кореляції для різних послідовностей дорівнює $-1/M$. В імпульсному режимі роботи рівень бічних пелюсток не перевищує $1/\sqrt{N}$. Будь-які послідовності при заданому n відрізняються як порядком чергування символів $+1$ і -1 , так і максимальним значенням бічних пелюсток. При цьому можна вказати послідовність, максимальний рівень бічних пелюсток якої буде найменшим серед можливих послідовностей для заданого n . Генерування псевдовипадкових послідовностей Хаффмена порівняно просто здійснюється за допомогою регістрів зсуву.

Крім сигналів Хаффмена, практичне застосування знаходять й інші види дискретних сигналів: сигнали Пелі-Плоткіна, послідовність символів Лежандра, коди Баркера, багатофазні коди Френка [36]. Можливі, нарешті, різні варіанти складних сигналів.

У радіолокації широко застосовуються сигнали з *лінійними змінами частоти усередині імпульсу* (ЛЧМ). Пояснюється це тим, що ЛЧМ-сигнали мають добрі кореляційні властивості і прийняти їх можна за допомогою погоджених фільтрів.

Шумоподібний сигнал може піддаватися модуляції всіма відомими способами. При амплітудній модуляції змінюються амплітуди всіх елементів. У разі частотної модуляції варіанти сигналу відрізняються середньою частотою, а фазової - різницею фаз між елементами двох посилок.

Специфічним видом модуляції, яка властива тільки широкосмуговим системам зв'язку, є *структурна модуляція* або *модуляція за формою* сигналу (див. гл. 9). У цьому випадку як варіанти сигналу використовуються коливання, побудовані на однакових елементах, але по-різному розташовані. Наприклад, двійкове передавання можна здійснити за допомогою сигналів виду

$$s_1(t) = f_1(t) \cos \omega_0 t, s_2(t) = f_2(t) \cos \omega_0 t, \quad (3.79)$$

де $f_1(t)$ і $f_2(t)$ - функції які набувають значень ± 1 відповідно до заданої кодової послідовності і задовольняють умову

$$\int_0^T f_1(t) f_2(t) dt \approx 0. \quad (3.80)$$

У разі виконання останньої умови маємо систему майже ортогональних сигналів. Як функції $f_1(t)$ і $f_2(t)$, що задовольняють умову (3.80), можна використовувати розглянуту вище двійкову псевдовипадкову послідовність Хаффмена $f_1(t) = f(t)$ і цю ж послідовність зі зсувом у часі: $f_2(t) = f(t + \Delta t)$.

Аналогічно будуються багатопозиційні широкосмугові системи зі структурною модуляцією. У цьому випадку застосовується ансамбль шумоподібних сигналів $f_i(t) (i = 1, 2, \dots, m)$. При цьому, обов'язково, різниця між сигналами має бути достатньою для розділення на прийомі. З цієї точки зору великий інтерес являють протилежні й ортогональні сигнали.

3.7. Енергетичний спектр модульованих сигналів

До цих пір ми розглядали модуляцію несучого коливання детермінованим процесом $u(t)$, який відображає певне повідомлення або окрему його реалізацію. Сукупність можливих повідомлень є деяким випадковим процесом. Так, при передаванні мови або музики статистичні властивості повідомлень, що передаються, дуже близькі до властивостей нормального випадкового процесу. Важливими характеристиками коливання, модульованого випадковим процесом, є функція кореляції та енергетичний спектр.

Слід підкреслити, що модульований сигнал є нестационарним випадковим процесом навіть тоді, коли модулюючі процеси (повідомлення) стаціонарні. Енергетичний спектр нестационарного випадкового процесу визначається за допомогою двократного усереднення - за множиною і за часом. Спочатку визначається усереднена за часом кореляційна функція, а потім оберненим перетворенням Фур'є - шуканий енергетичний спектр.

Розглянемо випадок, коли передаване повідомлення $u(t)$ являє собою стаціонарний процес з середнім значенням $\overline{u(t)}$, а переносник - гармонічні коливання $A_0 \cos \omega_0 t$.

При амплітудній модуляції

$$s(t) = A_0 [1 + m u(t)] \cos \omega_0 t,$$

де m - середньоквадратичне значення коефіцієнта модуляції.

Функція кореляції модульованого сигналу

$$\begin{aligned} B(t, \tau) &= \overline{s(t)s(t+\tau)} = A_0^2 \overline{[1 + m u(t)][1 + m u(t+\tau)]} \times \\ &\times \overline{\cos \omega_0 t \cos(\omega_0 t + \tau)} = 0,5 A_0^2 [1 + m^2 B_i(\tau)] [\cos(2\omega_0 t + \tau) + \cos \omega_0 \tau], \end{aligned} \quad (3.81)$$

де $B_i(t)$ - функція кореляції повідомлення $u(t)$, що передається.

Як бачимо, функція $B(t, \tau)$ залежить від часу, що вказує на нестационарність модульованого сигналу. Після усереднення за часом отримаємо

$$B(\tau) = \overline{B(t, \tau)} = \frac{1}{2} A_0^2 [1 + m^2 B_i(\tau)] \cos \omega_0 \tau.$$

Застосовуючи до $B(\tau)$ перетворення Фур'є, знаходимо енергетичний спектр сигналу при АМ:

$$\begin{aligned} G(\omega) &= 2 \int_0^{\infty} B(\tau) \cos \omega \tau d\tau = A_0^2 \int_0^{\infty} [1 + m^2 B_i(\tau)] \cos \omega_0 \tau \cos \omega \tau d\tau = \\ &= \frac{\pi A_0^2}{2} \delta(\omega - \omega_0) + \frac{m^2 A_0^2}{4} G_i(\omega - \omega_0). \end{aligned}$$

Таким чином, спектр гармонічного коливання, модульованого за амплітудою випадковим процесом, складається з несучого коливання з частотою ω_0 і зсунутого на ω_0 спектра повідомлення $u(t)$, яке передається.

Сигнал при кутовій модуляції (ЧМ і ФМ) можна записати в загальному вигляді:

$$s(t) = A_0 \cos[\omega_0 t + \varphi(t)].$$

При ФМ $\varphi(t) = \Delta\varphi u(t)$, а при ЧМ $\varphi(t) = \Delta\omega \int u(t) dt$. Тут $\Delta\varphi$ і $\Delta\omega$ - середньоквадратичні значення девіації відповідно фази і частоти.

Функція кореляції модульованого сигналу

$$\begin{aligned} B(t, \tau) &= \overline{s(t)s(t+\tau)} = A_0^2 \overline{\cos[\omega_0 t + \varphi(t)] \cos[\omega_0(t+\tau) + \varphi(t+\tau)]} = \\ &= \frac{A_0^2}{2} \{ \overline{\cos[2\omega_0 t + \omega_0 \tau + \varphi(t+\tau)]} + \overline{\cos[\omega_0 \tau + \varphi(t+\tau) - \varphi(t)]} \}. \end{aligned}$$

При усередненні за часом перший доданок обертається в нуль. Другий доданок не залежить від часу t , тому

$$B(\tau) = \overline{B(t, \tau)} = \frac{A_0^2}{2} \overline{\cos[\omega_0 \tau + \varphi(t+\tau) - \varphi(t)]}.$$

Позначимо різницю $\varphi(t+\tau) - \varphi(t) = \theta[\tau]$ і за відомою формулою подамо косинус суми двох кутів у вигляді

$$\overline{\cos(\omega_0 \tau + \theta)} = \overline{\cos \omega_0 \tau} \overline{\cos \theta} - \overline{\sin \omega_0 \tau} \overline{\sin \theta}.$$

Середні за множиною значення косинуса і синуса θ можна знайти, якщо відомий закон розподілу ймовірностей повідомлень $u(t)$. Якщо $u(t)$ підпорядковується нормальному закону, то $\theta(\tau)$, яке є лінійним перетворенням від $u(t)$, також матиме нормальний розподіл з нульовим середнім значенням і дисперсією σ_θ^2 . Легко переконатися, що в цьому випадку

$$\overline{\cos \theta} = e^{-\frac{\sigma_\theta^2}{2}}; \quad \overline{\sin \theta} = 0.$$

Отже, усереднена за часом функція кореляції сигналу при кутовій модуляції

$$B(\tau) = \frac{A_0^2}{2} \cos \omega_0 \tau e^{-\frac{\sigma_\theta^2}{2} \tau^2}. \quad (3.82)$$

Дисперсію процесу $\theta(\tau)$ можна виразити через функцію кореляції або енергетичний спектр повідомлення $u(t)$. Дійсно,

$$\sigma_\theta^2 = \overline{[\varphi(t+\tau) - \varphi(t)]^2} = 2[B_\varphi(0) - B_\varphi(\tau)],$$

де $B_\varphi(t)$ - функція кореляції процесу $\varphi(t)$.

При ФМ $\varphi(t) = \Delta\varphi u(t)$, тому $B_\varphi(\tau) = \Delta\varphi^2 B_i(\tau)$; при ЧМ $\varphi(t) = \Delta\omega\psi$, де $\psi = \int u(t)dt$, тому $B_\varphi(\tau) = \Delta\omega^2 B_\psi(\tau)$. Потім можна визначити енергетичний спектр модульованого сигналу шляхом перетворення Фур'є від функції $B(\tau)$ (вираз (3.82)).

3.8. Цифрові методи модуляції

Для передавання цифрового первинного сигналу по каналу використовують різні несучі $f(t)$. Нижче ми розглянемо тільки гармонічну несучу.

Цифрова амплітудна модуляція (ЦАМ). Канальний сигнал при ЦАМ (лінійній модуляції) можна записати як

$$u_{AM}(t) = \left(U_0 + K_{AM} \sum_{n=0}^{\infty} b_n^{(k)} \nu(t - nT) \right) \cos(\omega_0 t + \varphi_0), \quad (3.83)$$

де K_{AM} - девіація амплітуди;

U_0 - стала складова.

Спектр цього сигналу містить несучу і дві бічні смуги, кожна з яких повторює спектр первинного сигналу $\hat{b}_o(t)$. Формування сигналу та його детектування можна здійснити

параметричними і нелінійними схемами. Проте з отриманої оцінки первинного сигналу $\hat{b}_o(t)$ потрібно вилучити на окремих тактових інтервалах T оцінки кодових символів $\hat{b}_n(t)$. Для цього треба мати вирішуючий пристрій, який знаходить ці оцінки.

На рис. 3.16 для окремих блоків введені позначення: АД - амплітудний детектор (когерентний або некогерентний), ВП - вирішуючий пристрій, який з тактовим інтервалом T приймає рішення про кодові символи $b_n^{(k)}$. Ці рішення фіксуються в запам'ятовуючому пристрої ЗП.

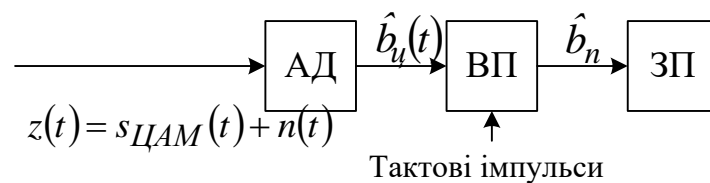


Рис. 3.16. Схема детектування сигналу ЦАМ і прийняття рішення про кодовий символ

Зазначимо, що для прийому дискретних повідомлень слід забезпечити не тільки синхронізацію за частотою несучої ω_0 і за фазою сигналу (при когерентному детектуванні), але і тактову, а також циклічну (з урахуванням методів передавання дискретних повідомлень по реальних каналах) синхронізацію. Через наявність адитивного шуму в каналі прийому повідомлень за схемою рис. 3.16 не є найкращим.

Якщо у вираз (3.83) $U_0 = 0$, то маємо сигнал ЦАМ без несучої (ЦБАМ). У цьому випадку можливе або синхронне (когерентне) детектування, або некогерентне детектування з відновленням несучої в місці прийому. Методи ЦАМ дозволяють за квадратурною схемою (рис. 3.17) організувати передавання і приймання двох незалежних дискретних повідомлень. Тут K – кодер вхідних сигналів; Γ_Ω – генератор гармоніки; Γ_{ω_0} – генератор несучого коливання; φ – фазообертач; $+$ – перемножувач; БВМ – блок визначення модуля; \int – інтегратор; ВП – вирішуючий пристрій; ДК – декодер вихідних сигналів.

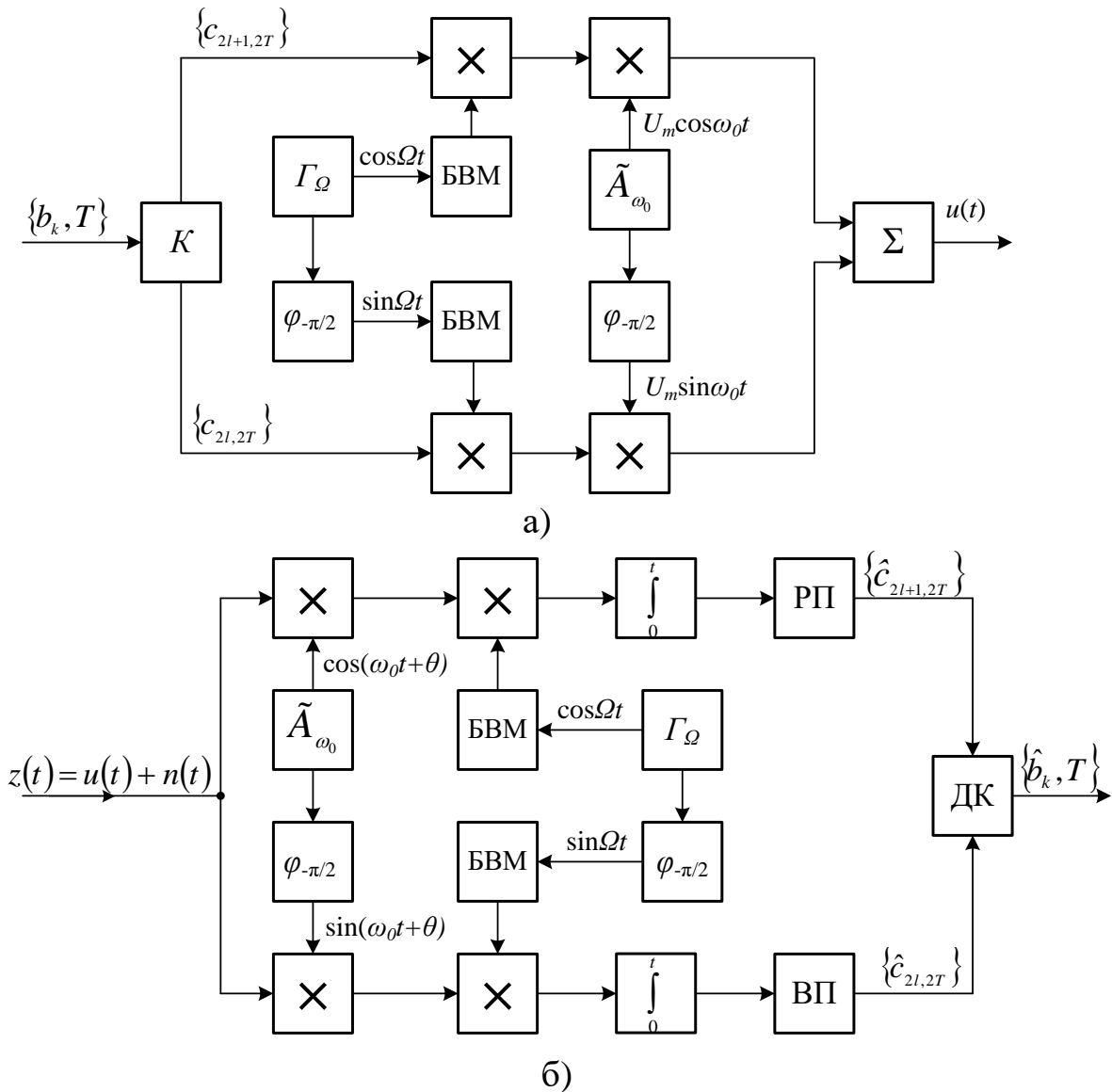


Рис. 3.17. Квадратурні схеми: а) формування двійкового сигналу; б) детектування двійкового сигналу

Сигнал, який передається в цифровій квадратурній амплітудній модуляції (ЦКАМ), тобто коли вектори сигналів розташовані в квадратній решітці,

$$u_{\text{ЦКАМ}}(t) = b_{y1}(t) \cos(\omega_0 t + \varphi_0) + b_{y2}(t) \sin(\omega_0 t + \varphi_0). \quad (3.84)$$

Приєм в цьому випадку може бути тільки когерентним. Якщо підставити у вираз (3.84) $b_{y2}(t) = \hat{b}_{y1}(t)$, то можна за квадратурною схемою організувати й односмугову ЦАМ.

Цифрова фазова модуляція (ЦФМ). Канальний сигнал при ЦФМ можна записати у вигляді

$$\begin{aligned}
 u_{\text{ЦФМ}}(t) &= U_m \cos\left(\omega_0 t + K_{\text{ФМ}} \sum_{n=0}^{\infty} b_n^{(k)} \nu(t - nT)\right) = \\
 &= U_m \cos\left(K_{\text{ФМ}} \sum_{n=0}^{\infty} b_n^{(k)} \nu(t - nT)\right) \cos\omega_0 t - \\
 &\quad - U_m \sin\left(K_{\text{ФМ}} \sum_{n=0}^{\infty} b_n^{(k)} \nu(t - nT)\right) \sin\omega_0 t,
 \end{aligned} \tag{3.85}$$

де U_m - амплітуда каналного сигналу;

$K_{\text{ФМ}}$ - девіація фази.

Вид цього сигналу, реалізація модулятора і детектора суттєво спрощуються при використанні прямокутних імпульсів $\nu(t)$ з одиничною амплітудою і тривалістю T . У цьому випадку відсутня міжсимвольна інтерференція (МСІ), і замість виразу (3.85) при $b_n = \pm 1$ маємо

$$u_{\text{ЦФМ}}(t) = U_m \cos\left(\frac{\Delta\theta}{2}\right) \cos\omega_0 t - U_m \sum_{n=0}^{\infty} b_n^{(k)} \nu(t - nT) \sin\left(\frac{\Delta\theta}{2}\right) \sin\omega_0 t, \tag{3.86}$$

де $\Delta\theta$ - різниця фаз для двох позицій коду, $\Delta\theta = 2K_{\text{ФМ}}$.

Спектральний склад сигналу (вираз (3.86)) за змістом не відрізняється від сигналу (виразу (3.83)). Відповідно ЦФМ і когерентне детектування в розглянутому випадку реалізуються так само, як для ЦАМ, зокрема за квадратурною схемою (ЦКАМ). Зауважимо, що коли різниця фаз при двопозиційному коді $\Delta\theta = \pi$ (використовуються протилежні сигнали), то несуча в спектрі сигналу ЦФМ зникає.

У роботі [41] розглядаються також методи відносної (різничевої) ЦФМ. У цьому випадку можна, крім когерентного, використовувати і некогерентний демодулятор. На практиці широко розповсюджені багатопозиційні системи ЦФМ, коли початкова фаза несучої набуває не двох, а m значень.

У техніці передавання даних широко застосовуються як багатопозиційні ($m > 2$) системи ЦАМ (багаторівневі системи з лінійною модуляцією) та ЦФМ (з нелінійною модуляцією), так і їх змішані варіанти (ЦАФМ). Амплітудно-фазові діаграми деяких систем сигналів наведені на рис. 3.18 (знаком «+» позначається початок координат). Для підвищення якості передавання (мінімізації середньої ймовірності помилкового приймання в каналі з шумом) намагаються підібрати таку сигнально-кодову конструкцію, щоб сигнальні точки дозволених кодових комбінацій знаходились одна від одної на максимально можливій відстані. Часто конструкція рис. 3.18, б (чотирьохпозиційна ЦФМ або система ФМ-4), яка має широке застосування в техніці зв'язку, реалізується двійковою ЦКАМ (з фізичної точки зору ЦКАМ є фазовою модуляцією).

Це можна зробити таким перекодуванням: непарні тактові послілки тривалості T двійкової вхідної послідовності подають на один вхід квадратурного модулятора; на інший вхід квадратурного модулятора подають парні послілки вхідної двійкової послідовності; послілки передають по каналу тривалістю інтервалу $T_k = 2T$. Зберігши швидкість передавання інформації $R_i = \frac{\log_2 4}{T_k} = \frac{2}{2T} = \frac{1}{T}$, можна підвищити завадостійкість системи, використовуючи деякі методи кодування.

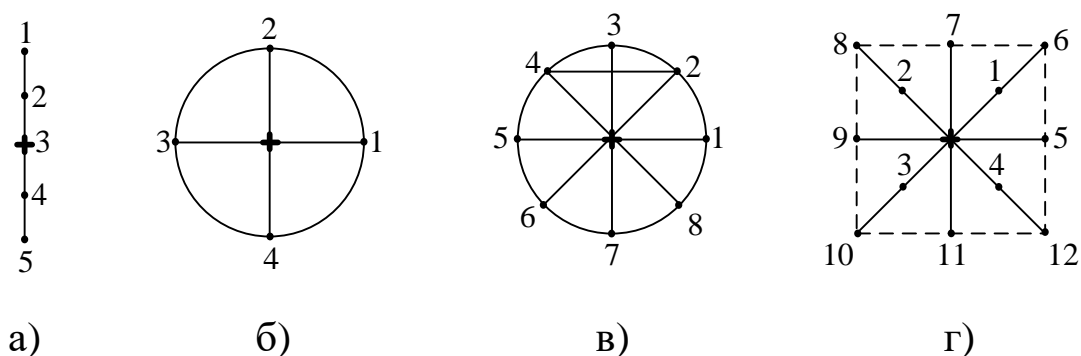


Рис. 3.18. Розташування сигнальних точок у 5-позиційній ЦАМ (а), у 4-позиційній ЦФМ (ФМ-4) (б), у 8-позиційній ЦФМ (в), у 12-позиційній ЦАФМ (г)

Цифрова частотна модуляція (ЦЧМ). Якщо частотна модуляція реалізується вибором одного з m незалежних

гармонічних сигналів, то в загальному випадку при кожному перемиканні (з i -тої на j -ту позицію; $i, j \in \overline{1, m-1}$) виникає розрив фази каналного сигналу. Дійсно, якщо i -тий гармонічний сигнал

$$u_i(t) = U_m \cos(\omega_i t + \varphi_i), \quad (3.87)$$

то в момент комутації $t_k = kT$ маємо

$$u_i(t_k) = U_m \cos(\omega_i kT + \varphi_i), \psi_i(t_k) = \omega_i kT + \varphi_i.$$

При комутації j -го генератора

$$u_j(t_k) = U_m \cos(\omega_j kT + \varphi_j), \psi_j(t_k) = \omega_j kT + \varphi_j.$$

На практиці смуговий фільтр, який обмежує ширину спектра сигналу, як правило, ввімкнений, і стрибки фази призводять до перехідного процесу у фільтрі. Внаслідок цього виникає паразитна амплітудна модуляція сигналу, і пік-фактор сигналу (відношення його пікової і середньої потужностей) збільшується. Крім того, при використанні m незалежних генераторів для забезпечення ортогональності системи сигналів необхідне рознесення частот $\Delta f = i/T$ ($i = 1, 2, \dots$), тобто мінімальне рознесення $\Delta f_{min} = 1/T$.

З метою звуження спектра і збереження мінімального пік-фактора сигналу потрібно забезпечити неперервність зміни миттєвої фази сигналу при менших значеннях Δf_{min} .

Частотну модуляцію з неперервною фазою скорочено позначають ЧМНФ. У системах ЧМНФ миттєва частота сигналу змінюється за законом

$$\omega(t) = \omega_0 + K_{ЧМ} \sum_{n=0}^{\infty} b_n^{(k)} v_u(t - nT), \quad (3.88)$$

де $v_u(t)$ - «частотний» імпульс,
а каналний сигнал –

$$u_{\text{ЧМНФ}}(t) = U_m \left(\cos \omega_0 t + K_{\text{ЧМ}} \sum_{n=0}^{\infty} b_n^{(k)} v(t - nT + \varphi_0) \right), \quad (3.89)$$

де $v(t)$ - «фазовий» імпульс, $v(t) = \int v_x(t_1) dt_1$;

φ_0 - початкова фаза;

$K_{\text{ЧМ}}$ - девіація частоти.

Якщо $v_u(t)$ - прямокутний імпульс одиничної висоти, тоді фазовий імпульс

$$v(t) = \int_0^1 dt_1 = t, \quad t \in [0, T].$$

Для забезпечення більш «гладкого» змінення фази і частоти (відповідно звуження спектра сигналу) на практиці (наприклад у цифрових системах мобільного стільникового зв'язку за загальноєвропейським стандартом GSM) використовують гаусівську форму «частотного» імпульсу та інтегральну форму «фазового»:

$$v_u(t) = \frac{1}{\sqrt{2\pi\tau_i^2}} \exp \left[-\frac{\left(t - \frac{T}{2}\right)^2}{2\tau_i^2} \right], \quad v(t) = F \left(\frac{t - \frac{T}{2}}{\tau_i} \right),$$

де τ_i - величина, пропорційна ефективній тривалості частотного імпульсу.

У подальшому для спрощення аналізу вважатимемо, що «частотний» імпульс $v_u(t)$ є прямокутним з одиничною амплітудою і тривалістю T .

Запишемо сигнал з виразу (3.88) на відрізку $[0, T]$ при передаванні i -тої позиції символу:

$$u_i(t) = U_0 \cos(\omega_i t + \varphi_{0,k}), \quad k = 1, 2, 3, \dots, \quad (3.90)$$

де $\omega_i = \omega_0 + \frac{2\pi}{T}i$;

$\varphi_{0,k}$ - початкова фаза до даного (k -го) тактового інтервалу.

Здійснюючи ЧМНФ, можна забезпечити (при коефіцієнті модуляції $m = 0,5$) ортогональність сигналів (вираз (3.90)) при частотному зсуві:

$$\frac{\Delta\omega_{min}}{2\pi} = \Delta f_{min} = \frac{1}{2T}. \quad (3.91)$$

Дійсно, частотний зсув

$$(u_i, u_j) = U_0^2 \int_0^T \cos(\omega_i t + \varphi_{0,k}) \cos(\omega_j t + \varphi_{0,k}) dt,$$

або

$$(u_i, u_j) = \frac{U_0^2}{2} \cdot \frac{\sin(\omega_i - \omega_j)T}{(\omega_i - \omega_j)T}. \quad (3.92)$$

При отриманні цього результату враховано, що доданок, який відповідає сумарній частоті $\omega_i + \omega_j$, малий. Мінімальне рознесення частот, при якому права частина виразу (3.92) обертається в нуль, визначається зі співвідношення $2\Delta f\pi T = \pi$, звідки випливає вираз (3.91).

Цифрову ЧМ з неперервною фазою і параметром вираз (3.91), тобто $\Delta f_{min} = \frac{1}{2T}$, називають *модуляцією з мінімальним (частотним) зсувом* – ММЗ (*minimum shift keying MSK*).

Покажемо, що коефіцієнт модуляції в цій системі $m = 0,5$. Визначимо коефіцієнт частотної модуляції m як відношення девіації частоти $K_{ЧМ} = \Delta\omega/2$ (максимальне відхилення від середньої частоти) до частоти модулюючого сигналу типу «точка» $\Omega_m = \pi/T [F = 1/(2T)]$:

$$M = K_{\text{ЧМ}} / \Omega_m = \frac{\Delta\omega T}{2\pi} = \Delta f T. \quad (3.93)$$

З урахуванням виразів (3.91) і (3.93) отримаємо, що для ММЗ коефіцієнт модуляції дійсно дорівнює 0,5. При використанні для ЧМ сигналів виразу (3.87), ортогональних у підсилювальному значенні, мінімальний індекс $m = 1$. Відсутність стрибків фази в системах ЧМНФ сприятливо позначається на формі амплітудного спектра сигналу. При $m = 0,5$ амплітудний спектр сигналу ЧМНФ дуже вузький і зосереджений поблизу частоти несучої. При значеннях $M > 1$ амплітудний спектр сигналів ЧМНФ стає широким.

Цифрову частотну модуляцію можна реалізувати різними способами, наприклад управлінням частоти генератора гармонічних сигналів за законом (3.88). При цьому початкова фаза опорного сигналу на n -му тактовому інтервалі

$$\varphi_{0n} = \pi M \sum_{l=0}^{n-1} b_l + \varphi_0. \quad (3.94)$$

Для систем ММЗ з коефіцієнтом $m = 0,5$ широко застосовується квадратурний метод модуляції з зсувом модулюючих функцій. Для обґрунтування такої модуляції зобразимо сигнал виразу (3.89) при ММЗ $[K_{\text{ЧМ}} = \pi/(2T)]$ на k -му і $(k+1)$ -му тактових інтервалах як

$$\begin{aligned} u(t)_{\text{ММЗ}_{k,k+1}} &= U_m \cos \left[\omega_0 t + \frac{\pi}{2T} b_k v(t - kT) + \frac{\pi}{2T} b_{k+1} v(t - (k+1)T) + \varphi_{0k} \right] = \\ &= U_c(t) \cos \omega_0 t - U_s(t) \sin \omega_0 t, \quad t \in [kT, (k+2)T], \end{aligned} \quad (3.95)$$

де U_c, U_s - косинусна і синусна складові сигналу;

$$\varphi_{0k} = \frac{\pi}{2} \sum_{i=0}^{k-1} b_i, \quad (3.96)$$

а квадратурні компоненти

$$\begin{aligned}
 U_c(t) &= U_m \cos \left\{ \omega_{0k} + \frac{\pi}{2T} b_k \nu \left[(t - kT) + \frac{\pi}{2T} b_{k+1} \nu [t - (k+1)T] \right] \right\}, \\
 U_s(t) &= U_m \sin \left\{ \omega_{0k} + \frac{\pi}{2T} b_k \nu \left[(t - kT) + \frac{\pi}{2T} b_{k+1} \nu [t - (k+1)T] \right] \right\}.
 \end{aligned} \tag{3.97}$$

Розіб'ємо інформаційний потік $\{b_k\}$ на два потоки, які відповідають парним і непарним індексам k : $\{b_{2l}\}$ і $\{b_{2l+1}\}$. З урахуванням виразу (3.96) при $k = 2l$ маємо $\varphi_{0,2l} = \pi \frac{1}{2} 2m$ (m – ціле число). Тоді, користуючись формулами косинуса і синуса суми кутів, можна записати

$$\begin{aligned}
 U_{s,l}(t) &= U_m \cos \pi m \times \\
 &\times \left\{ \frac{\pi}{2T} b_{2l} \nu (t - 2lT) + \frac{\pi}{2T} b_{2l+1} \nu [t - (2l+1)T] \right\}.
 \end{aligned} \tag{3.98}$$

На відрізку $[2lt; (2l+1)T]$ формула (3.98) набуває вигляду

$$U_{s,l}(t) = U_m \cos \pi m b_{2l} \sin \left[\frac{\pi}{2T} (t - 2lT) \right].$$

На відрізку $[(2l+2)T]$, маючи на увазі, що перший доданок під знаком синуса дорівнює $\frac{\pi}{2} b_{2l}$, і використовуючи формулу $\sin(\pm \pi/2 + \psi) = \pm \cos \psi$, отримаємо

$$U_{s,l}(t) = U_m \cos \pi m b_{2l} \cos \left\{ \frac{\pi}{2T} [t - (2l+1)T] \right\}.$$

Позначимо $c_{2l} = \cos \pi m b_{2l}$, що можна інтерпретувати як деяке перекодування. Таким чином, на відрізку $[2lt; (2l+1)T]$ для $U_{s,l}(t)$ маємо єдиний вираз

$$U_{s,l}(t) = U_m c_{2l} \left\{ \sin \left[\frac{\pi}{2T} (t - 2lT) \right] \lambda(t - 2lT) + \right. \\ \left. + \cos \left[\frac{\pi}{2T} (t - (2l + 1)T) \right] \lambda[t - (2l + 1)T] \right\}, \quad (3.99)$$

де λt - дискретна функція, $\lambda t = \begin{cases} 1, t \in [0, T] \\ 0, t \notin [0, T] \end{cases}$.

Формулу (3.99) можна подати більш компактно:

$$U_{s,l}(t) = U_m c_{2l} \left| \sin \frac{\pi}{2T} (t - 2lT) \right| \lambda \left(\frac{t - 2lT}{2} \right). \quad (3.100)$$

Розглянемо випадок $k = 2l + 1$. Тоді з урахуванням виразу (3.96)

$$\varphi_{0,2l+1} = \pi m + \frac{\pi}{2} b_{2l}.$$

Підставивши в $U_{c,l}(t)$ значення $k = 2l + 1$, отримаємо

$$U_{c,l}(t) = U_m \sin \left(\pi m + \frac{\pi}{2} b_{2l} \right) \sin \left\{ \frac{\pi}{2T} b_{2l+1} \nu [t - (2l + 1)T] + \right. \\ \left. + \frac{\pi}{2T} b_{2l+2} \nu [t - (2l + 2)T] \right\}.$$

На відрізку $[(2l + 1)T; (2l + 2)T]$ останній вираз набирає вигляду

$$U_{c,l}(t) = -U_m \cos \pi m \cdot b_{2l} b_{2l+1} \sin \left\{ \frac{\pi}{2T} [t - (2l + 1)T] \right\}.$$

На відрізку $[(2l + 2)T; (2l + 3)T]$ отримаємо

$$U_{c,l}(t) = -U_m \cos \pi m \cdot b_{2l} b_{2l+1} \cos \left\{ \frac{\pi}{2T} [t - (2l + 2)T] \right\}.$$

Позначимо

$$-\cos \pi m \cdot b_{2l} b_{2l+1} = -c_{2l} b_{2l+1} = c_{2l+1},$$

що можна інтерпретувати як деяке перекодування. Тоді на відрізку $[(2l+2)T; (2l+3)T]$ для $U_{c,l}(t)$ отримаємо єдиний вираз

$$U_{c,l}(t) = -U_m c_{2l+1} \left| \sin \frac{\pi}{2T} [t - (2l+1)T] \right| \lambda \left| \frac{t - (2l+1)T}{2} \right|. \quad (3.101)$$

З урахуванням формул (3.100) і (3.101) весь сигнал ММЗ можна записати так:

$$\begin{aligned} u_{\text{ММЗ}}(t) = & -U_m \sum_{l=0}^{\infty} c_{2l} \left| \sin \frac{\pi}{2T} (t - 2lT) \right| \lambda \left| \frac{t - 2lT}{2} \right| \sin \omega_0 t + \\ & + U_m \sum_{l=0}^{\infty} c_{2l+1} \left| \sin \frac{\pi}{2T} [t - (2l+1)T] \right| \lambda \left| \frac{t - (2l+1)T}{2} \right| \cos \omega_0 t. \end{aligned} \quad (3.102)$$

На рис. 3.17, а дана квадратурна схема, яка формує сигнал ММЗ згідно з виразом (3.102). Позначення для окремих блоків такі: K - кодер, який перетворює інформаційний потік $\{b_{k,T}\}$ двійкових символів, що прямують з тактовим інтервалом T ($b_k \in [-1;1]$), на два інформаційних потоки двійкових символів $\{c_{2l,2T}\}$ і $\{c_{2l+1,2T}\}$, що прямують з тактовим інтервалом $2T$ ($c_l \in [-1;1]$); Γ_{Ω} - генератор гармонічного сигналу з частотою $F = 1/(4T)$; $\varphi_{-\pi/2}$ - фазообертач на $-\pi/2$; $БВМ$ - блок визначення модуля; Γ_{ω_0} - генератор гармонічного сигналу (несучої) з частотою ω_0 ; $ДК$ - декодер, який перетворює два інформаційних потоки на один. При реалізації схеми враховано, що

$$\sin \left\{ \frac{\pi}{2T} [t - (2l+1)T] \right\} = \cos \left\{ \frac{\pi}{2T} (t - 2lT) \right\}.$$

Існують різноманітні способи демодуляції сигналу ММЗ.

1. Частотне детектування без урахування неперервності фази. При такому способі відбувається оцінювання частоти сигналу протягом одного тактового інтервалу T . Оскільки початкова фаза $\varphi_{0,n}$, при цьому не визначається, чи можливе тільки некогерентний прийом.

2. Когерентне детектування з відслідкуванням фази сигналу. Якщо попередні символи демодульовані без помилок, можливе передбачення початкової фази у відповідності з виразом (3.94). При цьому можливе когерентний прийом, і ймовірність похибки відповідає прийому двійкових ортогональних сигналів з поправкою на точність оцінки початкової фази.

3. Квадратурна когерентна демодуляція протягом двох тактових інтервалів $2T$ (рис. 3.17, б). Демодуляція елемента сигналу з номером $n = 2l$ провадиться на інтервалі $[2lT; 2(l+1)T]$ по синусній гілці, а демодуляція елемента сигналу з номером $n = 2l + 1$ - на інтервалі $[(2l+1)T; (2l+3)T]$ по косинусній гілці.

На рис. 3.17, б введені такі позначення: *ВП* - вирішуючий пристрій; *ДК* - декодер, який здійснює об'єднання символів $\{\hat{c}_{2l,2T}\}$ і $\{\hat{c}_{2l+1,2T}\}$ в єдиний потік, їх декодування і видачу інформаційного потоку $\{\hat{b}_{k,T}\}$; \int - блок інтегрування. Сусідні елементи інформаційної послідовності виділяються в двох гілках обробки, при цьому інтегрування в обох гілках здійснюється з зсувом на T . Інтегратор верхньої гілки здійснює інтегрування в межах $[kT; (k+2)T]$, інтегратор нижньої гілки - у межах $[(k+1)T; (k+3)T]$. Після виявлення знаку напруги в блоках *ВП* інтегратори зводяться до нульових початкових умов. Вхідні корисні сигнали в кожній з гілок обробки

$$s_{ЦЧМ}(t)_1 = \pm \gamma Um |\sin \Omega_m t| \cos(\omega_0 t + \theta_{кан}),$$

$$s_{ЦЧМ}(t)_2 = \pm \gamma Um |\cos \Omega_m t| \sin(\omega_0 t + \theta_{кан}),$$

де γ і $\theta_{кан}$ - коефіцієнт передачі і фазовий зсув у каналі відповідно.

Отже, в кожній з гілок на інтервалі $2T$ аналізується (з метою прийняття рішення про інформаційний символ c_{2l} або c_{2l+1}) система протилежних сигналів. Енергія сигналу по кожній гілці обробки

$$E = \gamma^2 U_m^2 \int_0^{2T} (\sin \Omega_m t \cos \omega_0 t)^2 dt = \gamma^2 U_m^2 \int_0^{2T} (\cos \Omega_m t \sin \omega_0 t)^2 dt = \frac{\gamma^2 U_m^2}{2} T,$$

тобто збігається з енергією гармонічного сигналу амплітуди γU_m на інтервалі $(0; T)$. Тому ймовірність похибки виходить такою самою, як при оптимальному когерентному прийому двійкових протилежних сигналів з тактовим інтервалом T (ФМ2 - двопозиційна фазова модуляція). Але для системи ФМ2 через стрибки фази на початку тактових інтервалів необхідна більша смуга частот, ніж для ММЗ.

3.9. Класифікація видів модуляції

Велику кількість різних видів модуляції, використовуваних у системах телекомунікації, можна класифікувати за певними ознаками. Так, якщо модулюючий сигнал $S_m(t)$ і переносник $u_n(t)$ безперервні, то модуляція також називається безперервною або аналоговою. При цьому переносник повинен бути гармонійним.

Якщо модулюючий сигнал є дискретним або цифровим, а переносник гармонійним, то модуляція називається дискретною або маніпуляцією.

Якщо в якості переносника вибрати періодичну послідовність імпульсів (відео або радіо), а модулюючий (інформаційний) сигнал буде безперервним, такий вид модуляції називається імпульсним.

Класифікація видів модуляції наведена на рис. 3.19.

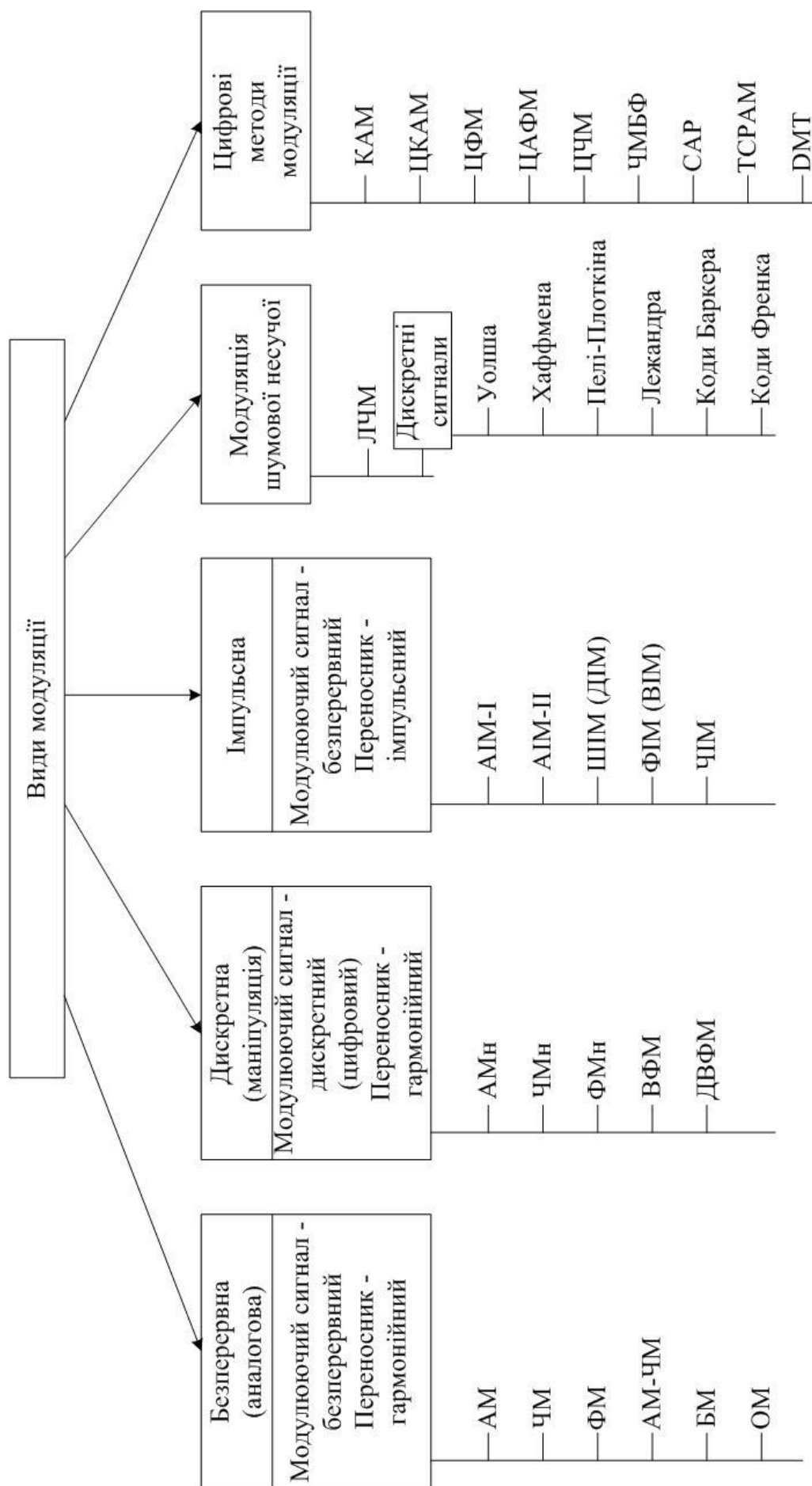


Рис. 3.19. Класифікація видів модуляції

Тут також вказані види модуляції безперервним сигналом шумової несучої, причому в якості несучого сигналу використовуються як безперервний шумоподібний вузькосмуговий процес, так і дискретні випадкові послідовності.

У телекомунікаційних системах, особливо для бездротового доступу, почали широко застосовуватися цифрові методи модуляції, які дозволяють при заданій завадостійкості підвищувати швидкість передачі інформації в дискретному каналі зв'язку [39].

Останнім часом для забезпечення доступу абонентів до телекомунікаційних і інформаційних ресурсів Інтернету застосовуються цифрові абонентські лінії (ЦАЛ), що забезпечують доступ із швидкістю більше 1 Мбіт/с. З цією метою були спеціально розроблені технології передачі xDSL (Digital Subscriber Line), характеристики яких регламентовані Рекомендаціями МСЕ-Т G.991...G.993 [43]. Дані технології передбачають використання таких видів модуляції, як КАМ, САР (Carrierless Amplitude and Phase modulation), яку можна розглядати як різновид КАМ-модуляції, а також ТС РАМ і DMT.

Структурна схема САР-модулятора наведена на рис. 3.20.

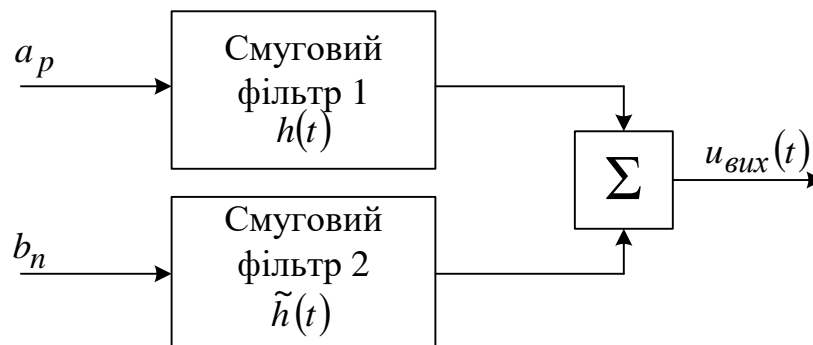


Рис. 3.20. Структурна схема САР-модулятора

Модулятор складається з двох смугових фільтрів, імпульсні реакції яких $h(t)$ і $\tilde{h}(t)$ зв'язані за Гілбертом. Ці фільтри збуджуються дискретними послідовностями інформаційних сигналів, що поступають з тактовою частотою $f_0 = 1/T$:

$$S(t) = \sum_{p=-\infty}^{\infty} [a_p h(t - pT) + b_p \tilde{h}(t - pT)]. \quad (3.103)$$

Якщо виконуються умови $h(t) = p(t) \cos \omega_n t$ та $\tilde{h}(t) = p(t) \sin \omega_n t$, то вираз (3.103) матиме вигляд

$$S(t) = \sum_{p=-\infty}^{\infty} [a_p p(t - pT) \cos \omega_n t + b_p p(t - pT) \sin \omega_n t]. \quad (3.104)$$

Модуляція ТСПАМ (Trellis Coded Pulse Amplitude Modulation) або амплітудно-імпульсна модуляція з решітчастим (трелісним) кодуванням дозволяє одночасно зі збільшенням швидкості передачі виявляти і виправляти помилки за рахунок використання згорткового кодування.

Модуляцію ДМТ (Discrete Multi Tone) називають також дискретною багатотонною передачею. У радіосистемах такий вид модуляції має назву ортогонального мультиплексування з розділенням за частотою (OFDM, Orthogonal Frequency Division Multiplexing). Цифрові системи передачі з ортогональними гармонійними сигналами (ЦСП ОГС) мають ряд специфічних переваг перед іншими типами ЦСП з паралельною передачею сигналів. Число незалежних несучих може досягати декількох тисяч. Такий вид модуляції можна ефективно використовувати в адаптивних системах зв'язку, кодуючи певні підгрупи частот різними способами і аналізуючи стан каналу зв'язку. При дії завад на яку-небудь частину ортогональних сигналів їх виключають на приймальному боці, а передача і прийом здійснюються на інших частотах, де забезпечується необхідне відношення сигнал/шум.

Слід зазначити, що поняття імпульсно-кодова модуляція (ІКМ), диференційно-імпульсно-кодова модуляція (ДІКМ), дельта-модуляція (ДМ), адаптивна диференційно-імпульсно-кодова модуляція (АДІКМ) не зовсім точно відображають поняття модуляції, тому ці види перетворень сигналів на цифрову форму розглянуті в гл. 8.

Висновки до глави 3

1. Унаслідок лінійності амплітудної модуляції спектр АМ сигналу повторює форму спектра повідомлення і займає смугу, що в два рази перевищує максимальну частоту первинного (модуючого) сигналу.

2. Багато перетворень сигналів у системах електричного зв'язку (перетворення і множення частоти, генерація коливань, модуляція і демодуляція, випрямлення та ін.) пов'язано з трансформацією спектра, отже, вони можуть бути виконані або в лінійних системах (ланцюгах) із змінними параметрами (параметричні ланцюги), або в нелінійних ланцюгах.

3. Спектр середньої потужності АМ-сигналу при модуляції випадковим процесом повторює форму спектра середньої потужності первинного сигналу, але не залежить від щільності вірогідності модульованого процесу.

4. З енергетичних міркувань і міркувань економії смуги частот каналу являють інтерес системи з балансною модуляцією (АМ без несучої) або системи з однією бічною смугою (ОБП).

5. Сигнали ОБП можна формувати шляхом модуляції двох гармонійних несучих, зрушених за фазою на $\pi/2$, за допомогою первинного сигналу і його перетворення за Гілбертом.

6. Схема квадратури передачі і прийому сигналів дозволяє одній несучій передавати повідомлення від двох незалежних джерел.

7. Фазова і частотна модуляції є різновидами кутової модуляції (КМ).

8. Спектр сигналів з КМ при великих індексах модуляції ширший і складніший за спектр АМ.

9. Спектр середньої потужності сигналу КМ при модуляції випадковим процесом має складніший характер, ніж при АМ, і залежить від щільності імовірності модульованого процесу.

10. У режимі сильного сигналу завадостійкість КМ при великих індексах модуляції значно більше, ніж при АМ.

11. Часові і спектральні характеристики цифрового первинного сигналу істотно залежать від вибору аналогового сигналу переносника дискретних повідомлень.

12. Система двійкової частотної модуляції з мінімальним зсувом (ММЗ) (з індексом модуляції 0,5) дозволяє досягти завадостійкості системи з протилежними сигналами (ФМ-2), але займає вужчу смугу частот.

13. Модульований сигнал при імпульсному переноснику займає ширшу смугу частот, ніж при гармонійній несучій.

Контрольні питання та задачі

3.1. Диференціальна крутизна параметричного елемента змінюється згідно із законом $S = S_0 + S_1 \cos \omega_y t$, а вхідний сигнал $u_{ex} = U(t) \cos(\omega_i t + \varphi_i)$. Знайти крутизну перетворення частоти як відношення амплітуди струму частоти $\Delta\omega = \omega_y - \omega_i$ до амплітуди вхідного сигналу.

3.2. На нелінійний елемент з вольт-амперною характеристикою (ВАХ) $i = a_0 + a_1(u - E) + a_2(u - E)^2$ діє напруга $u = E + U_1 \cos \omega_1 t + U_2 \cos \omega_2 t$. Знайти амплітуди всіх частотних складових струму.

3.3. На нелінійний елемент з ВАХ апроксимованою ламаною прямою подається (при кутовій відсічці $\pi/2$) гармонійний сигнал з амплітудою 2 В. Крутизна ВАХ $S = 10 \text{ мА/В}$. Знайти амплітуди постійної складової струму I_0 , першої і другої гармонік I_1 і I_2 . Для порівняння знайти величини I_0, I_1, I_2 методом трьох ординат.

3.4. Прямокутний імпульс тривалістю 1 мс модулює за амплітудою гармонійну несучу частоти $f_0 = 1 \text{ МГц}$ і амплітуди 1 В. Знайти спектр первинного і вторинного модульованого сигналу.

3.5. Знайти статистичну модуляційну характеристику АМ зсувом у ланцюзі бази транзистора при апроксимації характеристики $i_k = f(U_{be})$ ламаною прямою.

3.6. На лінійний детектор при куті відсічки $\theta = \pi/2$ подається АМ сигнал $U_{AM} = [1 + 0.5 \cos(5 \times 10^3 t) \times \cos(5 \times 10^6 t)]$. Крутизна характеристики нелінійного елемента $S = 10 \text{ мА/В}$.

Знайти напругу на навантаженні, утворену паралельним з'єднанням опору $R = 10$ кОм і ємності C , що задовольняє необхідні умови роботи детектора.

3.7. У чому різниця між двома різновидами кутової модуляції: фазовою і частотною?

3.8. У чому особливості систем цифрової частотної модуляції з безперервною фазою, зокрема системи з ММЗ?

3.9. В якості імпульсної несучої використовується періодична послідовність прямокутних імпульсів тривалістю $\tau = 10$ мкс та періодом $T = 125$ мкс. Несуча модулюється за амплітудою мовним сигналом з верхньою частотою спектра $F_B = 3400$ Гц. Яку ширину спектра займає сигнал АІМ? Скільки мовних сигналів можна передати на імпульсній несучій шляхом часового розподілу каналів?

3.10. Двійкова частотна модуляція реалізується шляхом почергового вибору одного з двох незалежних гармонійних сигналів з частотами f_0 і f_1 . Модель первинного сигналу – випадкова стаціонарна двійкова синхронна послідовність. Знайти функцію кореляції і спектральну щільність середньої потужності ЧМ-сигналу.

3.11. На вхід лінійного детектора при куті відсічки $\pi/2$ поданий однотональний АМ-сигнал $U_{AM}(t) = U_0(1 + m \cos(\Omega_m t) \cos(\omega_0 t))$ і стаціонарний "квазібілий" (у смузі частот) гаусівський шум з дисперсією σ_u^2 . Знайти математичне очікування (МО) і дисперсію вхідного коливання, а також імовірність того, що вихідний НЧ сигнал (при великому відношенні сигнал – шум) перевищує середньоквадратичне значення шуму σ_u .

ГЛАВА 4

МАТЕМАТИЧНІ МОДЕЛІ КАНАЛІВ ЗВ'ЯЗКУ

4.1. Загальні відомості про канали зв'язку

Як було зазначено вище, під каналом зв'язку розуміється сукупність засобів, призначених для передачі сигналів (повідомлень) між різними точками системи зв'язку. Під «засобом» розуміють і технічні пристрої, і лінію зв'язку – фізичне середовище, у якому поширюється сигнал між пунктами зв'язку. Канал зв'язку можна подати як послідовне з'єднання пристроїв (блоків) і лінії зв'язку, що виконують різні функції в загальній системі зв'язку.

Дискретна система зв'язку зображена на рис. 1.8, і при необхідності аналізу окремих її блоків схему можна деталізувати. Наприклад, можна врахувати ряд пристроїв (реалізованих електричними колами), розміщених у проміжних пунктах лінії зв'язку між передавачем і приймачем (наприклад, підсилювальні пристрої, комутуюча апаратура, розташовані вздовж ліній дальнього провідного зв'язку або ретранслятори радіорелейних ліній зв'язку).

У даній главі не розглядаються питання електричного сполучення окремих блоків, що утворюють канал зв'язку. У першу чергу обґрунтовується вибір функціональних блоків каналу, при якому була б забезпечена найбільша ефективність передачі повідомлень (інформації). У залежності від розв'язуваних завдань під каналом зв'язку також можна розуміти різну сукупність блоків телекомунікаційної системи, що у ході вирішення завдання вважається заданою.

Класифікацію каналів зв'язку зручно розглядати з використанням різних ознак. У залежності від призначення систем зв'язку канали зв'язку поділяють на телеграфні, фототелеграфні, телефонні, звукового віщання, передачі даних, телевізійні, телеметричні, змішані і т. п. У залежності від того, чи поширюються сигнали між пунктами зв'язку у вільному просторі або по направляючим лініям розрізняють канали радіо- (прямої видимості, супутникові і т. п.) і провідного зв'язку (повітряні,

кабельні, волоконно-оптичні лінії зв'язку, хвилевідні СВЧ тракти і т. п.) У залежності від форми сигналів на вході і виході каналу розрізняють канали безперервні, дискретні, цифрові, змішані (наприклад дискретно-безперервні).

Розрізняють канали з зосередженими параметрами, у яких сигнали на вході і виході описуються функціями часу тільки одного скалярного параметра (наприклад часу t), і канали з розподіленими параметрами, у яких сигнали на вході і (або) виході описуються функціями більше одного скалярного параметра (наприклад часу t і просторових координат x , y , z). Такі сигнали називають *електромагнітними полями*.

Істотною є класифікація каналів електричного зв'язку за діапазоном (смугою) частот який вони займають. Так, на сучасних симетричних кабельних лініях зв'язку застосовують сигнали, що займають смуги частот у діапазоні, обмеженому зверху частотою до сотень кілогерців. Додаткові заходи щодо забезпечення симетрії кабельних пар дозволяють збільшити верхню межу використовуваного діапазону частот до одиниць мегагерц. У даний час коаксіальні кабелі пропускають сигнали в діапазоні частот до сотень мегагерц. На повітряних провідних лініях використовують частоти не вище 150 кГц, тому що на вищих частотах на цих лініях дуже відчутна заважаюча дія адитивних завад і різко зростає загасання в лінії.

Радіозв'язок здійснюється за допомогою електромагнітних хвиль, що поширюються в частково обмеженому (наприклад землею й іоносферою) просторі. Зараз у радіозв'язку застосовують частоти приблизно від $3 \cdot 10^3$ до $3 \cdot 10^{12}$ Гц. Цей діапазон поділяють відповідно до десяткової класифікації, наведеної в табл. 4.1 [42].

У цій таблиці в дужках зазначені нестандартні, але використовувані назви діапазонів хвиль. Діапазон дециміліметрових хвиль підходить до діапазону інфрачервоних хвиль. Завдяки створенню і широкому впровадженню квантових генераторів (лазерів) освоєний і діапазон світлових хвиль (оптичний діапазон). Практично в оптоволоконних лініях зв'язку використовуються частоти порядку 10^{14} Гц (довжини хвиль 1,55; 1,35; 0,85 мкм).

Таблиця 4.1

Класифікація діапазонів частот у радіоканалах

Найменування хвиль	Діапазон хвиль	Найменування частот	Діапазон частот
Декакілометрові (наддовгі, СВД)	100...10 км	ДНЧ дуже низькі	3...30 кГц
Кілометрові (довгі, ДВ)	10...1 км	НЧ низькі	30...300 кГц
Гектометрові (середні, СВ)	1000...100 м	СЧ середні	300...3000 кГц
Декаметрові (короткі, КВ)	100...10 м	ВЧ високі	3...30 МГц
Метрові (ультра-короткі, УКВ)	10...1 м	ДВЧ дуже високі	30...300 МГц
Дециметрові	100...10 см	УВЧ ультрависокі	300...3000 МГц
Сантиметрові	10...1 см	НВЧ надвисокі	3...30 ГГц
Міліметрові	10...1 мм	ВВЧ вельмивисокі	30...300 ГГц
Дециміліметрові (субміліметрові)	1...0,1 мм	ГПЧ гіпервисокі	300...3000 ГГц

Для сучасного етапу розвитку техніки зв'язку характерна тенденція до її переходу на усе більш високі частоти. Це викликано рядом причин, зокрема необхідністю підвищувати швидкість передачі повідомлень, можливістю одержати гостроспрямоване випромінювання при невеликих розмірах випромінювачів, меншою інтенсивністю атмосферних і багатьох видів промислових завад у більш високочастотних діапазонах, можливістю застосування завадостійких широсмугових систем модуляції і т. п.

Для теорії електричного зв'язку великий інтерес являє класифікація каналів зв'язку за характером сигналів на вході і виході каналу, відповідно до якої розрізняють:

а) безперервні (за рівнями), на вході і виході яких сигнали безперервні (аналогові). Прикладом може служити канал, заданий між виходом модулятора і входом демодулятора в будь-якій системі зв'язку;

б) дискретні (за рівнями), на вході і виході яких сигнали дискретні. Такі канали задані між точками $a - \hat{a}$ і $b - \hat{b}$ на схемі рис. 1.8;

в) дискретні з боку входу і безперервні з боку виходу або навпаки. Такі канали називаються дискретно-безперервними або напівбезперервними (наприклад канали, задані між точками $a - z$, (див. $b - z$ рис. 1.8).

У будь-якому дискретному або напівбезперервному каналі міститься безперервний канал. Варто пам'ятати, що дискретність і безперервність каналу не пов'язана з характером переданих повідомлень: можна передавати дискретні повідомлення по безперервному каналу і безперервні повідомлення по дискретному. У ТЕЗ аналізуються канали (сигнали) з безперервним і дискретним часом.

4.2. Перетворення випадкових сигналів у лінійних каналах з постійними параметрами

Дослідження перетворень випадкових процесів при їхньому проходженні через динамічні системи (як з регулярними, так і з випадково змінюваними параметрами) пов'язане з розв'язанням задач двох типів: визначення кореляційної функції і спектральної щільності потужності відгуку $Y(t)$ на виході системи, заданої своїми характеристиками, щодо даної кореляційної функції або спектральної щільності потужності вхідного впливу $X(t)$; визначення багатомірного розподілу імовірностей відгуку $Y(t)$ на виході системи за багатомірним розподілом вхідного впливу $X(t)$.

Друга із зазначених задач є загальнішою. З її розв'язку може бути отриманий розв'язок першої задачі. Однак надалі обмежимося розглядом тільки першої задачі і лише вкажемо шляхи розв'язання другої, складнішої задачі. Так, можна стверджувати, що якщо смуга частот F_x , яка зайнята вхідним випадковим процесом $X(t)$, набагато ширша від смуги

пропускання ΔF даної лінійної системи, то розподіл вихідного випадкового процесу $Y(t)$ має тенденцію наближатися до гаусівського.

Дійсно, у стаціонарній детермінованій лінійній системі з фінітною, тобто обмеженою в часі межами $0 \dots \tau_n$, імпульсною характеристикою (ІХ) $g(t)$ відгук на її дію $Y(t)$ можна записати відповідно до інтеграла Дюамеля у такому вигляді:

$$Y(t) = \int_0^{\tau_n} g(t)X(t-\tau) \cdot d\tau = \lim_{\Delta\tau \rightarrow 0} \sum_{k=1}^N g(k\Delta\tau)X(t-k\Delta\tau)\Delta\tau. \quad (4.1)$$

Крок дискретизації $\Delta\tau$ можна вибрати таким, щоб дорівнював інтервалу кореляції вхідного процесу $\tau_{кор} = 1/F_x$. Припустимо, що вхідний процес центрований $\bar{X}(t) = 0$, тоді є центрованим і вихідний процес. Вузька смуга пропускання ΔF означає, що тривалість імпульсної характеристики τ_n велика в порівнянні із $\Delta\tau$.

Перетин вихідного процесу $Y(t)$ в будь-який момент часу t визначається за формулою (4.1) N доданками суми. У цю суму входить безліч некорельованих між собою перетинів процесу $X(t)$. Розподіл імовірностей такої суми відповідно до центральної граничної теореми теорії імовірності (теореми Ляпунова) близький до гаусівського (причому тим ближчий, чим більшою є величина N , обумовлена відношенням $F_x/\Delta F$) [30]. У граничному випадку, якщо на вхід каналу впливає “білий” шум, у якого ширина спектра нескінченна (не співпадаючі в часі відліки не корельовані), а канал має обмежену смугу пропускання, то $N \rightarrow \infty$ і вихідний процес буде строго гаусівським. Зазначена властивість лінійного каналу зберігається і при зміні параметрів каналу.

Використовуючи функції знаходження законів розподілу випадкових величин (випадкових процесів), можна знаходити і розподіл вихідного процесу будь-якого порядку, якщо відомий розподіл вхідного процесу. Однак визначення багатомірних

імовірнісних характеристик відгуку лінійних систем виявляється досить громіздким і складним, незважаючи на те, що для розв'язання цієї задачі розроблений ряд спеціальних прийомів.

Далі визначимо функції кореляції вихідного процесу. Зрозуміло, що для стаціонарних випадкових процесів існує пара перетворень Вінера-Хінчина між функціями кореляцій (ФК) $B_x(\tau)$, $B_y(\tau)$ процесів $X(t)$ і $Y(t)$ і їхніми спектральними щільностями потужності (СЦП): $G_x(f)$, $G_y(f)$. Для стаціонарної лінійної системи і при випадкових стаціонарних впливах ФК вихідного стаціонарного процесу $Y(t)$ визначається з виразу

$$B_y(\tau) = \int_{-\infty}^{\infty} G_x(f) K^2(f) \cdot e^{j\omega\tau} df.$$

Можна також показати, що ФК відгуку детермінованої параметричної системи на стаціонарні входні впливи $X(t)$ визначається формулою

$$B_y(t, \tau) = \int_{-\infty}^{\infty} G_x(f) K^2(j\omega, t) K(-j\omega, t + \tau) e^{j\omega\tau} df, \quad (4.2)$$

тобто у даному випадку вихідний процес - нестаціонарний.

4.3. Перетворення випадкових сигналів у нелінійних каналах з постійними параметрами

Обмеживши розгляд нелінійних перетворень моделлю $y(t) = \varphi[x(t)]$, помітимо, що перетворення $x \rightarrow y$, як правило, однозначне, що не завжди можна сказати про зворотне перетворення $y \rightarrow x$ (наприклад коло із квадратичною характеристикою $y = kx^2$).

При нелінійних перетвореннях виникає трансформація (зміна) спектра входного впливу. Так, якщо на вхід нелінійної системи діє суміш регулярного сигналу й адитивного шуму $X(t) = S(t) + N(t)$ у вузькій смузі частот F_c , що згрупувалася біля

середньої частоти f_0 , то в загальному випадку на виході будуть присутні складові комбінаційних частот трьох видів, що згрупуються біля частот $nf_0 (n = 0, 1, \dots)$: продукти взаємодії складових вхідного сигналу між собою ($c \times c$), продукти взаємодії складових вхідного шуму ($ш \times ш$); продукти взаємодії сигналу і шуму ($c \times ш$). Розділити їх на виході системи, як правило, неможливо.

Якщо відома характеристика $y = \varphi(x)$ нелінійної системи і двовимірний розподіл вхідного впливу $\omega(x_1, x_2; t_1, t_2)$, то основні характеристики вихідного процесу (МО і ФК), у принципі, завжди можна визначити. Так, математичне очікування відгуку записується у вигляді $\bar{Y}(t) = \overline{\varphi(X(t))} = \int_{-\infty}^{\infty} \varphi(x) \omega(x, t) dx$, а його кореляційна функція в такому вигляді [46]:

$$B_y(t, \tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\varphi(x_1) - \bar{Y}(t)] \cdot [\varphi(x_2) - \bar{Y}(t + \tau)] \omega(x_1, x_2; t_1, t_2) dx_1 dx_2.$$

Прямим перетворенням Вінера-Хінчина можна за ФК знайти і спектральну щільність потужності процесу $Y(t)$.

Аналіз проходження випадкових впливів через нелінійні кола сильно спрощується для вузькосмугових впливів, якщо скористатися їхнім квазігармонічним представленням.

4.4. Проходження сигналів через канали зв'язку з випадковими параметрами

Крім розглянутих детермінованих перетворень сигналу в окремих ланках каналу (зокрема у лінії зв'язку або середовищі розповсюдження хвилі) мають місце і випадкові перетворення сигналу. У найпростішому випадку це перетворення зводиться до складання сигналу з незалежним від нього випадковим процесом, який називають адитивною завадою або адитивним шумом. У більш складних каналах до цього додаються випадкові зміни параметрів каналу, у результаті яких навіть під час відсутності

адитивних завад прийнятий сигнал не визначається однозначно переданим.

Розглянемо характерні перетворення сигналу в лінійних каналах (колах) з випадковими параметрами.

У загальному вигляді лінійну систему (або лінійний канал) можна описати випадковою ІХ $G(t, \tau)$, що є випадковою функцією двох аргументів: t (момент спостереження реакції) і τ (час, що пройшов з моменту подачі δ -імпульса на вхід кола).

Лінійний канал, параметри якого підпадають під дію випадкових зовнішніх впливів (температури, тиску, вологості), можна характеризувати також випадковою передатною функцією змінних ω і t

$$K(j\omega, t) = \int_{-\infty}^{\infty} G(t, \tau) e^{-j\omega\tau} d\tau. \quad (4.3)$$

Можна показати [15], що функція кореляції процесу $Y(t)$ на виході каналу з характеристикою (4.3) при подачі на вхід стаціонарного процесу $X(t)$ визначається виразом

$$B_y(t, \tau) = \int_{-\infty}^{\infty} G_x(f) \cdot \Pi_k(f, \tau, t) e^{j\omega t} df, \quad (4.4)$$

де Π_k - системна характеристика каналу,
 $\Pi_k(f, \tau, t) = \overline{K(j\omega, \tau)K(-j\omega, t - \tau)}$.

Для каналу з постійними параметрами маємо $\Pi_k(f, \tau, t) = K(j\omega, \tau)K(-j\omega, t + \tau)$, і з виразу (4.4) випливає вираз (4.2).

Зупинимось докладніше на моделях каналів, з якими найчастіше доводиться зустрічатися на практиці. Узагальнену модель для випадкового вхідного впливу $X(t)$ можна подати як

$$Y(t) = \gamma X(t - \tau), \quad (4.5)$$

де параметри τ (час затримки сигналу) і (або) γ (коефіцієнт загасання сигналу) флюктують у часі.

Як правило, такі флюктуації в провідних лініях зв'язку викликаються змінами зовнішніх умов і відбуваються надзвичайно повільно й у дуже невеликих відносних межах. Це означає, що за час тривалості відлікового інтервалу $\Delta = 0,5F$, де F – ширина спектра сигналу, параметри каналу не встигають помітно змінюватися. У радіоканалах при багатопроменевому поширенні хвиль, у гідроакустичних і в деяких інших каналах флюктуації виражені більш помітно.

Якщо вхідний сигнал вузькосмуговий, його зручно подати в квазігармонічній формі: $X(t) = A(t)\cos[\omega_0 t + \Phi(t)]$, де $A(t)$ і $\Phi(t)$ – функції, що повільно змінюються. Тому при досить малій затримці τ можна в першому наближенні вважати $A(t - \tau) \approx A(t)$ і $\Phi(t - \tau) \approx \Phi(t)$, а вихідний сигнал з виразу (4.5) записати так:

$$Y(t) = \gamma A(t - \tau)\cos[\omega_0(t - \tau) + \Phi(t - \tau)] \approx \gamma A(t) \times \quad (4.6)$$

$$\times \cos[\omega_0(t) + \Phi(t) + \theta] = \gamma \cos \theta X(t) - \gamma \sin \theta \hat{X}(t),$$

де θ – фазовий зсув в каналі, $\theta = \omega_0 \tau$;

$\hat{X}(t)$ – процес, сполучений з $X(t)$, за Гілбертом [44].

Таким чином, при вузькосмуговому сигналі мала затримка зводиться до деякого зсуву фази. Важливо зазначити, що навіть при дуже малих відносних флюктуаціях часу затримки τ фазовий зсув θ (через великі значення ω_0) може змінюватися в дуже великих межах. Для цього досить виконання умови $\Delta\tau \gg \frac{1}{f_0}$, де $\Delta\tau$ – середньоквадратичне відхилення затримки, f_0 – середня частота спектра сигналу. Ця умова в радіоканалах звичайно виконується.

Має місце більш складний випадок, коли сигнал проходить паралельними шляхами від входу каналу до його виходу (рис. 4.1), так що на виході кожного шляху сигнал описується виразом (4.6), але значення γ і τ для різних шляхів різні і до того

ж у невеликих межах флюктують. Такого роду багатопроменеве поширення сигналу характерне для більшості радіо-, гідроакустичних та деяких інших каналів (у тому числі провідних).

Електромагнітні хвилі поширюються, як правило, в неоднорідному середовищі і відбиваються від різних неоднорідностей (відбивачів). Ці неоднорідності можуть бути розподілені усередині відносно невеликого відбиваючого (розсіюючого) об'єму. У цьому випадку різниці ходу (різниці значень τ) для окремих шляхів невеликі. Якщо по такому каналу направити дуже короткий імпульс, то і на його виході імпульс буде також досить коротким. Такий канал прийнято називати *однопроменевим*. Наявність різних шляхів не викликає в цьому випадку істотного розсіювання енергії (розтягнення) сигналу в часі, але призводить до виникнення явища завмирань, яке полягає в більш-менш швидких випадкових змінах передатної функції каналу (мультиплікативна завада).

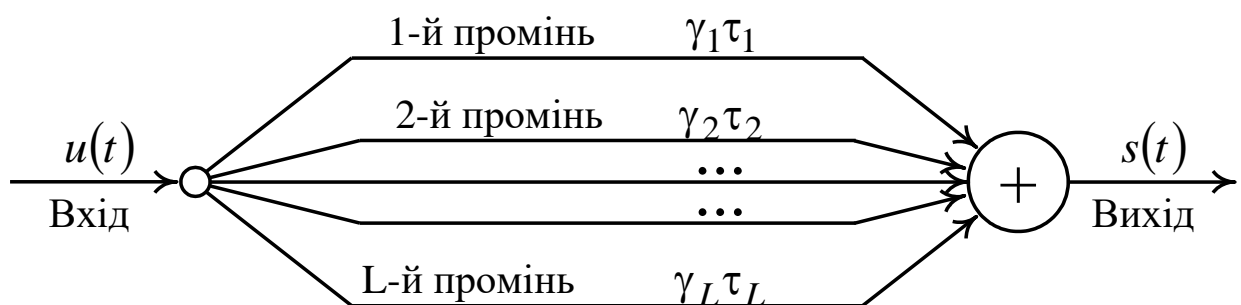


Рис. 4.1. Багатопроменеве поширення сигналу

Для пояснення завмирань розглянемо передачу по каналу (див. рис. 4.1) гармонійного сигналу з одиничною амплітудою $u(t) = \text{Re}(e^{j\omega t})$.

Тоді на виході сигнал можна записати в такому вигляді:

$$s(t) = \text{Re} \left(\sum_{l=1}^L \gamma_l e^{j\omega(t-\tau_l)} \right) = \text{Re} \left(e^{j\omega t} \right),$$

де L – кількість шляхів (променів, що попадають у точку прийому);

γ_l – коефіцієнт передачі по l -му променю;

τ_l – час поширення l -го променя;

$\dot{\gamma}_l$ – комплексний коефіцієнт передачі по l -му променю,

$$\dot{\gamma}_l = \gamma_l e^{-j\omega\tau_l};$$

$\dot{\gamma}$ – комплексна амплітуда вихідного сигналу, що у даному випадку за визначенням дорівнює передатній функції каналу,

$$\dot{\gamma} = \sum_{l=1}^L \gamma_l e^{-j\omega\tau_l}.$$

Передатна функція в загальному випадку залежить від частоти. Якщо врахувати, що внаслідок хаотичних переміщень відбивачів значення γ_l і τ_l флюктують, то $\dot{\gamma}$ залежить також від часу, являючи собою випадкову функцію (мультиплікативну заваду) $\dot{\gamma}(t, \omega)$. У багатьох випадках ця функція флюктує значно швидше, ніж величини γ_l і τ_l .

Важливою характеристикою каналу з завмираннями є розподіл імовірностей комплексної передатної функції $\dot{\gamma}(t, j\omega)$ і у першу чергу її модуля γ . Для визначення цього розподілу подамо $\dot{\gamma}$ в такому вигляді:

$$\dot{\gamma} = \gamma e^{-j\theta} = X(t, \omega) + jY(t, \omega),$$

де γ і θ – відповідно модуль і аргумент передатної функції, які також є випадковими функціями t і ω , $\gamma = |\dot{\gamma}|$;

X і Y – квадратурні складові, $X = \gamma \cos \theta$ і $Y = \gamma \sin \theta$.

З іншого боку, за формулою (4.6)

$$\dot{\gamma} = \sum_{l=1}^L \gamma_l e^{-j\omega\tau_l} = \sum_{l=1}^L \gamma_l e^{j\theta_l} = \sum_{l=1}^L \gamma_l \cos \theta_l + j \sum_{l=1}^L \gamma_l \sin \theta_l.$$

Звідки одержимо $X(t, \omega) = \sum_{l=1}^L \gamma_l \cos \theta_l$; $Y(t, \omega) = -\sum_{l=1}^L \gamma_l \sin \theta_l$.

Оскільки значення $X(t, \omega)$ й $Y(t, \omega)$ утворюються в результаті додавання великої кількості слабо корельованих величин з обмеженими дисперсіями, то до них можна застосувати центральну граничну теорему теорії імовірності Ляпунова і вважати їх нормально розподіленими.

Для випадку, коли усі γ_l одного порядку і фазові зсуви досить великі, легко показати, що X і Y мають однакові дисперсії $\sigma_x^2 = \sigma_y^2 = \sigma^2$, а їхні математичні очікування $m_x = m_y = 0$. Тут одномірний розподіл імовірності γ є

релеєвським, тобто $w(\gamma) = \frac{\gamma}{\sigma^2} \exp\left(-\frac{\gamma^2}{2\sigma^2}\right)$, $\gamma \geq 0$.

Фаза результуючого сигналу θ при цьому розподілена рівномірно на інтервалі $(-\pi, +\pi)$. Дисперсія квадратурних складових σ^2 дорівнює середній потужності сигналу, що надходить. Такі завмирання, як і канали, у яких вони виявляються, називаються *релеєвськими*.

У багатьох каналах завмирання відрізняються від релеєвських. Іноді в одному з підпроменів коефіцієнт передачі γ_l значно більший, ніж в інших, і можна вважати, що, крім дифузно відбитих підпроменів, у місце прийому приходить і регулярний (незавмираючий) промінь. У цьому випадку коефіцієнт передачі каналу $\gamma = \sqrt{X^2 + Y^2}$ підкоряється узагальненому розподілу

Релея $w(\gamma) = \frac{\gamma}{\sigma^2} \exp\left(-\frac{\gamma^2}{2\sigma^2} - q^2\right) I_0\left(\frac{\gamma\sqrt{2q}}{\sigma}\right)$, $\gamma \geq 0$.

Тут $q^2 = \frac{m_x^2 + m_y^2}{2\sigma^2}$ – відношення середніх потужностей регулярної і флукутуючих складових.

У загальному випадку, коли $\sigma_x^2 \neq \sigma_y^2$ і $m_x \neq 0$, $m_y \neq 0$, виходить так званий чотирипараметричний розподіл модуля і фази завмираючого сигналу (загальна гаусівська модель каналу). Відповідні щільності імовірності наведені в роботі [47].

Якщо по однопроменевому каналу з завмираннями передається відносно вузькосмуговий сигнал, а

середньоквадратичне відхилення запізнювання $\Delta\tau$ в окремих підпроменях задовольняє умову

$$\Delta\tau \ll \frac{1}{F_c}, \quad (4.7)$$

де F_c – ширина спектра сигналу,

Зміни початкових фаз на різних частотах ω у спектрі сигналу дорівнюють $\omega\Delta\tau$, майже однакові. При цьому всі складові спектра сигналу завмирають одночасно, тобто їхні амплітуди і фази змінюються однаково. Такі завмирання називаються загальними або гладкими.

Зазначимо, що умова (4.7) може виконатися при $\Delta\tau \gg \frac{1}{f_0}$, тому що в радіоканалах $f_0 \gg F_c$. Якщо ж умова (4.7) не виконується, то в різних областях спектра сигналу процеси завмирань не збігаються (селективні за частотою завмирання). При цьому спостерігаються істотні зміни форми сигналу, що характерно для багатопроменевих каналів радіозв'язку (сигнали, що приходять у точку прийому, утворені відображенням від сильно рознесених у просторі розсіюючих об'ємів).

Швидкість змін у часі комплексного випадкового процесу $\dot{\gamma}(t, j\omega) = X(t, \omega) + jY(t, \omega)$ (при фіксованій частоті), або, як говорять, швидкість завмирань сигналу, характеризується часом кореляції $\tau_{кор}$ квадратурних компонентів $X(t, \omega)$ і $Y(t, \omega)$ або

шириною спектра завмирань $\Delta f_{зав} \approx \frac{1}{\tau_{кор}}$.

4.5. Адитивні завади в каналах зв'язку

У каналах зв'язку адитивні завади викликаються різними причинами і можуть приймати довільні форми, індивідуальні реалізації яких важко врахувати. Саме ці завади частіше викликають незворотні перетворення переданих сигналів. Незважаючи на велику розмаїтість, адитивні завади за їхньою

електричною і статистичною структурою розділяють на три основних класи: флуктуаційні (розподілені за частотою і часом), зосереджені за частотою (квазігармонійні) і зосереджені в часі (імпульсні).

Коли відсутні адитивні завади детерміновані лінійні перетворення сигналу найчастіше оборотні. У присутності навіть дуже слабкої адитивної завади лінійні перетворення виявляються необоротними.

Флуктуаційні завади. З фізичної точки зору адитивні флуктуаційні завади породжуються в системах зв'язку різного роду флуктуаціями, тобто випадковими відхиленнями тих або інших фізичних величин (параметрів) від їхніх середніх значень. Так, джерелом шуму в електричних колах можуть бути флуктуації струму, обумовлені дискретною природою носіїв заряду (електронів, іонів). Дискретна природа електричного струму виявляється в електронних лампах і напівпровідникових приладах у вигляді дробового ефекту (при заданому режимі живлення схем випадково змінюються кількість створюваних носіїв заряду).

Флуктуаційні завади можуть вважатися гаусівськими випадковими процесами, які найчастіше є стаціонарними з нульовими середніми значеннями (принаймні на визначених часових проміжках, котрі називають інтервалами стаціонарності).

Сума великої кількості будь-яких завад від різних джерел внаслідок умов центральної граничної теореми теорії імовірностей Ляпунова також має характер флуктуаційної завади. Багато завад при проходженні через прийомний пристрій часто набувають властивостей нормальної флуктуаційної завади.

Найбільш розповсюдженою причиною шуму в апаратурі зв'язку є флуктуації, обумовлені тепловим рухом носіїв зарядів. Випадковий тепловий рух носіїв заряду в будь-якому провіднику викликає випадкову різницю потенціалів (напруги) на його кінцях. Середнє значення такої напруги дорівнює нулю, а змінна складова проявляється як шум. Тепловий шум на вході приймача являє собою гаусівський випадковий процес з нульовим середнім і спектральною щільністю потужності, що дорівнює [49]

$$G_n(f) = \frac{hf}{2 \left(\exp\left(\frac{hf}{kT}\right) - 1 \right)}, \quad (4.8)$$

де h – постійна Планка, $h = 6,624 \cdot 10^{-34}$ Дж·с;

k – постійна Больцмана, $k = 1,38 \cdot 10^{-23}$ Дж/град;

T – абсолютна температура джерела шуму;

f – поточна частота.

У діапазоні звукових частот і радіочастот виконується нерівність $hf \ll kT$, і тому, розкладаючи у виразі (4.8)

експоненту в ряд, одержуємо $G_n(f) = \frac{kT}{2} = \frac{N_0}{2}$.

Величина N_0 є одnobічною (на позитивних частотах) спектральною щільністю шуму. Специфічним для напівпровідникових приладів є шум, який називають флікер-шумом, він виникає в результаті різного роду поверхневих явищ. Його спектральна щільність у широкому діапазоні частот змінюється за гіперболічним законом (вона пропорційна $1/f$).

Зазвичай на частотах вище 10 кГц флікер-шумами нехтують.

Космічні завади в системах радіозв'язку, викликані радіовипромінюванням Сонця й інших космічних об'єктів, мають характер флуктуаційних шумів.

Зосереджені за спектром завади. До зосереджених за спектром адитивних завад прийнято відносити сигнали сторонніх радіостанцій, навмисні завади, випромінювання генераторів високої частоти різного призначення (промислових, медичних) і т. п. У загальному випадку це модульовані коливання, тобто квазігармонічні коливання зі змінюваними параметрами. В одних випадках ці коливання є неперервними (наприклад, сигнали віщальних і телевізійних радіостанцій), в інших – вони носять імпульсний характер (сигнали радіотелеграфних станцій і систем передачі даних). На відміну від флуктуаційних завад, ширина спектра зосередженої завади в більшості випадків не перевищує смуги пропускання приймача, а в деяких випадках вона набагато вужче цієї смуги. У діапазоні коротких хвиль зосереджені за спектром завади є основними у визначенні якості зв'язку і

вважаються випадковими коливаннями з флуктуаціями фаз і амплітуд (завмираннями), розподіл яких такий же, як у корисних сигналів.

Імпульсні завади. До імпульсних (зосереджених у часі) адитивних завад прийнято відносити завади у вигляді одиночних імпульсів, що йдуть один за одним через такі великі проміжки часу, що перехідні явища в приймачі від одного імпульсу встигають практично закінчитися до моменту приходу наступного імпульсу. До таких завад відносять багато видів атмосферних та індустриальних завад. Зазначимо, що «флуктуаційна завада» і «імпульсна завада» є поняттями відносними.

У залежності від частоти проходження імпульсів одна і та сама завада може впливати як імпульсна на приймач із широкою смугою пропускання і як флуктуаційна на приймач з відносно вузькою смугою пропускання. На практиці імпульсні завади доводиться розглядати як випадковий, відносно широкосмуговий (тим ширший, чим коротші імпульси завади) процес, що складається з окремих рідких, випадково розподілених у часі і за амплітудою імпульсів. Імовірнісні властивості таких завад з достатньою для практичних потреб повнотою описуються розподілом імовірностей амплітуд імпульсів і розподілом часових інтервалів між цими імпульсами. Для останніх часто застосовують модель Пуассона (вираз (2.90)).

Розподіл амплітуди імпульсних завад часто описується логнормальним законом [28]

$$w(A) = \frac{1}{A\sqrt{2\pi\beta^2}} e^{-\frac{(\ln A - \alpha)^2}{2\beta^2}},$$

де α , β^2 – відповідно МО і дисперсія перетвореного ВП;
 $\rho = \ln A$.

Для цього закону характерна наявність істотних значень в області великих амплітуд, що відповідає реальній моделі імпульсної завади.

Зауважимо, що логнормальним законом часто описують повільні (добові, сезонні) флуктуації амплітуд радіосигналів, обумовлені не інтерференційними явищами, а поглинаннями сигналу в середовищі поширення.

4.6. Квантовий шум в оптичному каналі зв'язку

В оптичному діапазоні частот ($\lambda = 30 \dots 0,01$ мкм), де виконується умова $hf \gg kT$, тепловий шум виявляється дуже слабким. Однак у цьому діапазоні при слабких сигналах істотне значення має «квантовий шум», викликаний дискретною природою світлового випромінювання. Відповідно до квантової теорії електромагнітного поля його енергія сигналу випромінюється і поглинається квантами, причому енергія одного такого кванта (фотона) дорівнює hf . В елементарному сигналі тривалості T з високостабільною несучою частотою f (когерентне одномодове випромінювання) і амплітудою U детермінованою може бути тільки середня енергія, яка дорівнює $E_{cp} = mhf$ (m – середня кількість фотонів на інтервалі T). Конкретна ж реалізація елементарного сигналу має енергію $E = nhf$, де n – випадкова кількість реєстрованих фотонів.

У сучасних системах оптичного зв'язку в основному використовується АМ оптичного несучого коливання за амплітудою або інтенсивністю (потужністю).

Ідеальна система оптичного зв'язку при ізохронній передачі двійкових повідомлень ("1" і "0") має такі характеристики [16]:

1. Тактовий інтервал (час передачі біта) постійний і дорівнює T , отже, швидкість передачі інформації $R_i = 1/T$, біт/с.

2. При передачі "1" оптична енергія, випромінювана у вигляді імпульсів за час передачі одного біта, $E_{пер} = nE_{\phi}$, де n – кількість випромінених фотонів, $E_{\phi} = hf$ – енергія одного фотона (кванта), а оптична енергія при передачі 0 дорівнює нулю. Оптична енергія в місці прийому дорівнює на тактовому інтервалі T величині E_{np1} при передачі "1" і нулю при передачі "0" відповідно.

3. Імовірності передачі "1" і "0" $p(1) = p(0) = 0,5$. У цьому випадку усереднену за тривалий час прийняту потужність P_{np} можна виразити через середню потужність, прийняту за час передачі біта при посилю "1". Таким чином, $P_{np} = 0,5$, $P_{np1} = 0,5E_{np1}\left(\frac{1}{T}\right) = 0,5E_{np1}R_u$.

Реальна система оптичного зв'язку відрізняється від ідеальної таким:

1. Час передачі біта інформації не залишається постійним – цей ефект називають *фазовим тремтінням цифрового сигналу*.

2. Випромінювана оптична енергія не залишається строго однією і тією ж. При передачі як кодового символу "1", так і кодового символу "0" має місце шум передавача, що призводить до випадкових змін амплітуди від імпульсу до імпульсу. Крім того, має місце «шум лазера», обумовлений статистичною природою взаємодії між збудженням лазера і створюваним потоком фотонів. Флуктуації прийнятої енергії збільшуються ще більше через зміни загасання в каналі зв'язку. Крім того, з'являються флуктуації енергії на окремих тактових інтервалах у місці прийому, обумовлені статистичною природою взаємодії потоку фотонів (оптичний сигнал) і створеного фотодетектором (звичайно це фотодіод) потоку електронно-діркових пар. Умовно будемо говорити в цьому випадку про шум фотодетектора.

3. Дуже імовірно, що при передачі символу "0" випромінюється малий, але цілком визначений рівень енергії (шум лазера), не враховуючи шуму передавача і каналу. Відношення середньої енергії, прийнятої при передачі символу "0", до середньої енергії при передачі символу "1" характеризується коефіцієнтом $r_3 = \frac{P_{np0}}{P_{np1}}$. При цьому

вважають, що в ідеальній системі $r_3 = 0$, однак, як правило, це не так, особливо якщо лазерне джерело випромінювання зміщене поблизу порога генерації.

4. Кінцева тривалість випромінюваних імпульсів і додаткова часова дисперсія (розсіювання) при їхній передачі по каналу призводять до того, що в практичних системах зв'язку

відбувається накладання сусідніх посилок, тобто виявляється міжсимвольна інтерференція.

Шум лазера, що випромінює потік фотонів, має квантову природу. Імовірність появи точно n фотонів на інтервалі T на передавальній стороні визначається розподілом Пуассона:

$$P_n(T) = \frac{m^n}{n!} e^{-m},$$

де $n = 0, 1, 2, \dots$;

$$m = \nu T ;$$

ν – інтенсивність потоку.

Таким чином, шум лазера – це «квантовий шум», тому що він проявляється у флуктуаціях параметрів сигналу, детермінованого за класичними уявленнями. Цей шум не є адитивним, тому що залежить від самого корисного сигналу. З урахуванням цього в наведеній формулі варто вважати, що при передачі символу "1" маємо $m = m_1$, а при передачі символу "0" одержуємо $m = m_0$. Як зазначалося вище, при передачі символу "0" (відсутність збудження лазера) може спостерігатися визначений, хоча і малий рівень енергії, обумовлений тим, що імовірність не появи фотонів на цьому інтервалі $p(0) = e^{-m_0} \neq 1$, де m_0 – середня кількість шумових фотонів на інтервалі T при відсутності збудження лазера. Із збільшенням середньої потужності випромінюваного сигналу $P_{пер}$ внесок квантового шуму в порівнянні з іншими шумами тракту передачі падає.

Шум фотодетектора має природу, аналогічну до шуму лазера, тому що падаючий на фотодіод стаціонарний світловий потік генерує електронно-діркові пари носіїв заряду як незалежні випадкові події. Якщо за відрізок часу T на фотодіод падає оптична енергія, яка дорівнює в середньому $E_{пр}$, то варто очікувати, що буде створено в середньому N пар носіїв заряду, які визначаються з виразу

$$N = \eta \frac{E_{np}}{E_{\phi}} = \eta E_{np} \frac{\lambda}{hC},$$

де η – квантова ефективність взаємодії, що показує середнє відношення кількості електронно-діркових пар у фотодетекторі до кількості падаючих фотонів ($\eta \leq 1$);

C – швидкість світла;

E_{ϕ} – енергія фотона;

λ – довжина хвилі випромінювання;

h – постійна Планка.

Внаслідок стохастичної природи взаємодії фотонів з фотодетектором дійсна кількість пар носіїв заряду, що генеруються кожним оптичним імпульсом, буде флуктувати навколо середнього значення N . Імовірність того, що кількість створених пар носіїв заряду на інтервалі T дорівнює K , визначається пуасонівським розподілом

$$P_K(T) = \frac{N^K}{K!} e^{-N}, \quad K = 0, 1, 2, \dots$$

Слід зазначити, що в реальних оптичних лініях зв'язку, крім квантового шуму, існують і інші заважаючі чинники (у тому числі адитивні завади), що призводить до необхідності збільшення потужності оптичного сигналу в каналі зв'язку.

4.7. Моделі безперервних каналів зв'язку

Для математичного описання каналу зв'язку необхідно і достатньо вказати множину сигналів, які можуть бути подані на його вхід, і для будь-якого з цієї множини допустимого вхідного сигналу задати випадковий процес (сигнал) на його виході. Задати випадковий процес можна за допомогою його функцій розподілу імовірностей. Так, у неперервному каналі можна задати апріорну щільність розподілу імовірностей (багатомірну) $\omega_n[u]$ вхідного процесу $U(t)$ на інтервалі аналізу T_a і багатомірну

умовну (перехідну) щільність розподілу імовірностей $\omega_n \left[\frac{z}{u} \right]$, тобто щільність реалізації прийнятого випадкового коливання $Z(t) = U(t) + N(t)$ за умови передачі реалізації $U(t)$, де $N(t)$ - адитивна завада (шум).

Точне математичне описання будь-якого реального каналу, зазвичай, досить складне. Замість цього використовують спрощені математичні моделі, які дозволяють виявити всі найважливіші закономірності реального каналу, якщо при побудові моделі враховані найбільш істотні особливості каналу.

Розглянемо найбільш прості і широко використовувані математичні моделі каналів, почавши з неперервних каналів, оскільки вони багато в чому визначають і характер дискретних каналів.

4.7.1. Ідеальний канал без завад

Даний канал відображається лінійним колом з постійною передатною функцією, як правило, зосередженою в обмеженій смузі частот. Припустимі будь-які вхідні сигнали, спектр яких лежить у визначеній смузі частот F_c і які мають обмежену середню потужність P_c (або пікову потужність $P_{\text{пик}}$). Ці обмеження характерні для всіх неперервних каналів, і надалі про них не йтиметься.

В ідеальному каналі вихідний сигнал $S(t)$ при заданому вхідному $u(t)$ є детермінованим і визначається з виразу $S(t) = \gamma u(t - \tau)$, де γ – постійний коефіцієнт передачі каналу, τ – постійна часова затримка. Цю модель іноді використовують для описання кабельних каналів. Однак, строго кажучи, вона не придатна для реальних каналів, у яких неминуче присутні, хоча б і дуже слабкі, адитивні завади.

4.7.2. Канал з адитивним гаусівським шумом

Сигнал на виході такого каналу записується як

$$Z(t) = \gamma u(t - \tau) + N(t) = S(t) + N(t), \quad (4.9)$$

де $N(t)$ – гаусівський адитивний шум з нульовим математичним очікуванням і заданою кореляційною функцією.

Найчастіше розглядається "білий" гаусівський шум (БГШ) або "квазібілий" (з рівномірною спектральною щільністю в смузі спектра сигналу $S(t)$). Часто при аналізі можна τ не враховувати, що відповідає зміні початку відліку часу на виході каналу. Деяке ускладнення моделі (4.9) виходить, якщо коефіцієнти передачі γ і запізнювання τ вважати відомими функціями часу:

$$Z(t) = \gamma(t)u[t - \tau(t)] + N(t).$$

Така модель задовільно описує багато провідних каналів, радіоканали при зв'язку в межах прямої видимості, а також радіоканали з повільними загальними завмираннями, при яких можна надійно передбачати значення γ і τ .

4.7.3. Канал з невизначеною фазою сигналу й адитивним шумом

Ця модель відрізняється від моделі (4.9) тим, що в ній запізнювання є випадковою величиною. Для вузькосмугових сигналів з урахуванням виразу (4.6) вираз (4.9) при постійному γ і випадкових τ можна подати як [44]

$$Z(t) = \gamma[\cos \theta u(t) - \sin \theta \hat{u}(t)] + N(t),$$

де $\hat{u}(t)$ – перетворення Гілберта від $u(t)$;

θ - випадкова фаза, $\theta = \omega_0 \tau$.

Розподіл імовірностей θ вважається заданим, найчастіше рівномірним на інтервалі від 0 до 2π . Ця модель задовільно описує ті самі канали, що і попередня, якщо фаза сигналу в них флукутує. Така флукутація викликається невеликими змінами протяжності каналу, властивостей середовища, у якому проходить сигнал, а також фазовою нестабільністю опорних генераторів.

Однопроменевий гаусівський канал із загальними завмираннями (флуктуаціями амплітуд і фаз сигналу) також описується формулою (4.9), але множник γ , як і фаза θ , вважаються випадковими процесами. Іншими словами, випадковими будуть квадратурні компоненти

$$X = \gamma \cos \theta, \quad Y = \gamma \sin \theta.$$

При змінюванні квадратурних компонентів X , Y у часі прийняте колювання може бути записане як

$$\begin{aligned} Z(t) &= X(t)u(t) - Y(t)\hat{u}(t) + N(t) = \\ &= \gamma(t)[\cos \theta u(t) - \sin \theta \hat{u}(t)] + N(t). \end{aligned} \quad (4.10)$$

Як зазначалося вище, одномірний розподіл коефіцієнта передачі каналу γ може бути релеєвським або узагальненим релеєвським. Такі канали називають відповідно каналами з релеєвськими або узагальненими релеєвськими (або райсівськими) завмираннями. У рамках загальної гаусівської моделі каналу γ має чотиріпараметричний розподіл. Модель однопроменевого каналу з завмираннями досить добре описує багато каналів радіозв'язку в різних діапазонах хвиль, а також деякі інші канали.

Багатопроменевий гаусівський канал із селективними за частотою завмираннями узагальнює модель (4.10):

$$Z(t) = \sum_{n=1}^N \gamma_n(t) [\cos \theta_n u(t - \bar{\tau}_n) - \sin \theta_n \hat{u}(t - \bar{\tau}_n)] + N(t), \quad (4.11)$$

де N – кількість променів у каналі;

$\bar{\tau}_n$ – середній час затримки для n -го променя.

Багатопроменева загальна гаусівська модель добре описує багато каналів радіозв'язку. Для моделі (4.11) умова (4.7) не виконується, якщо під $\Delta\tau$ розуміти інтервал запізнювання між променями.

4.7.4. Канал з міжсимвольною інтерференцією й адитивним шумом

Ця модель каналу зв'язку є окремим випадком умови (4.3), коли $G(t, \tau)$ від t не залежить (або змінюється дуже повільно), тому розсіювання за частотою практично не спостерігається.

Міжсимвольна інтерференція (МСІ) спричиняється розсіюванням сигналу в часі при його проходженні каналом зв'язку. Зокрема на виході багатопроменевого каналу корисний сигнал виявляється деформованим так, що одночасно присутні відгуки каналу на частинах вхідного сигналу, що належать до досить віддалених моментів часу. При передачі дискретних повідомлень це призводить до того, що при прийомі одного символу на вхід прийомного пристрою впливають також відгуки на більш ранні (а іноді і більш пізні) символи, що у цих випадках можуть (при неоптимальних методах прийому) виявляти себе як завади.

Міжсимвольна інтерференція викликається нелінійністю ФЧХ каналу й обмеженістю його смуги пропускання. У радіоканалах причиною МСІ найчастіше є багатопроменеве поширення радіохвиль.

Нехай передавач передає синхронно з тактовим інтервалом T послідовність елементарних сигналів, які відповідають ланцюжку символів $b_{-Q}, b_{-(Q-1)}, \dots, b_{-1}, b_0, b_1, \dots, b_{D-1}, b_D$, причому кожний із символів послідовності обирають з можливого для даного коду набору $0, 1, \dots, m-1$ (m – основа коду).

Позначимо відгук лінійного каналу на елементарний сигнал, що відповідає символу b_r , через $s_r(t)$, $rT \leq t \leq (Q+r+1) \cdot T$, де

$Q = \left\lceil \frac{\Delta\tau}{T} \right\rceil$ – відносна пам'ять каналу, обумовлена цілою частиною

від розподілу часу розсіювання каналу $\Delta\tau$ (тривалості перехідного процесу в каналі) на T . Тоді прийняте колювання $z(t)$ в місці прийому на інтервалі аналізу $T_a = (D+1) \cdot T$ при пошуку рішення щодо символу b_0 можна записати як

$$z(t) = s_0(t) + g_{m.u.}(t) + n(t), \quad (4.12)$$

де $s_0(t)$ – сигнал, обумовлений аналізованим символом b_0 ;

$g_{mi}(t)$ - сигнал міжсимвольної інтерференції, обумовлений символами, переданими до і після аналізованого символу,

$$g_{mi}(t) = g_{зал}(t) + g_{сл}(t) = \sum_{r=-Q, r \neq 0}^D s_r(t - rT);$$

$n(t)$ – адитивний шум у каналі;

$g_{зал}(t)$ – сигнал, який визначає залишковий сигнал МСІ, обумовлений символами, переданими до аналізованого,

$$g_{зал}(t) = \sum_{l=-Q}^{-l} s_r(t - rT);$$

$g_{сл}(t)$ – сигнал, котрий визначає сигнал МСІ, обумовлений символами, переданими після аналізованого, $g_{сл}(t) = \sum_{l=1}^D s_r(t - rT)$.

Чим більша швидкість передачі символів $1/T$ у кожному частотному каналі при заданій його смузі пропускання, тим більша кількість сусідніх з аналізованим символів визначає сигнал $g_{mi}(t)$.

У деяких випадках у моделі (4.12) можна вважати, що елементарні сигнали на прийомі $s_r(t)$ і передачі $u_r(t)$ зв'язані детермінованим (як правило, лінійним) відношенням. Тоді при незначному рівні шумів $n(t)$ у каналі можна здійснити його корекцію, тобто перейти до моделі каналу, що не спотворює. Однак при значних рівнях шумів у каналі з МСІ гранична якість може забезпечити лише оптимальний прийом. При випадкових змінах параметрів каналу функції $s_r(t)$ стають випадковими і модель (4.12) ускладнюється.

4.8. Моделі дискретних каналів зв'язку

Як правило, у системі передачі даних усередині дискретного каналу міститься неперервний канал. Перетворення неперервного сигналу на дискретний і назад здійснює модем. Тому можна

вивести математичну модель дискретного каналу з моделей неперервного каналу при заданому модемі, хоча такий підхід призводить до складних моделей.

Розглянемо прості моделі дискретного каналу, при побудові яких властивості неперервного каналу і модема не враховуються. Необхідно, однак, пам'ятати, що при проектуванні системи зв'язку є можливість варіювати в досить широких межах модель дискретного каналу при заданій моделі неперервного каналу, змінюючи характеристики модема.

Модель дискретного каналу містить безліч заданих можливих сигналів на його вході і розподіл умовних імовірностей вихідного сигналу при заданому вхідному. Тут вхідним і вихідним сигналами є послідовності n кодових символів. Тому для визначення можливих вхідних сигналів досить указати число m різних символів (основа коду), а також тривалість T передачі кожного символу. Будемо вважати значення T однаковим для всіх символів, що виконується в більшості випадків. Тоді для каналу зв'язку «без пам'яті» можна записати [24]

$$P(\hat{b}_j / b_i) = \begin{cases} \frac{P}{m-1}, & i \neq j, \\ 1-p, & i = j, \end{cases} \quad (4.13)$$

де $P(\hat{b}_j / b_i)$ – умовна імовірність того, що при передачі символу b_i був прийнятий символ \hat{b}_j ;

$1-p$ – імовірність первинного прийому символу;

$\frac{P}{m-1}$ – імовірність того, що $m-1$ символів прийняті

помилково.

Термін «без пам'яті» означає, що імовірність помилкового прийому символу не залежить від передісторії, тобто від того, які символи передавалися до нього і як вони були прийняті. Надалі, для скорочення, замість «імовірність помилкового прийому символу» будемо говорити «імовірність помилки».

Із роботи [20] випливає, що імовірність кожного n -мірного вектора помилки в такому каналі може бути визначена з виразу

$$P(E^{[n]}) = \left(\frac{P}{m-1}\right)^l (1-p)^{n-l},$$

де l – кількість ненульових символів у векторі помилки (вага вектора помилки).

У двійковому каналі вага вектора збігається з його нормою.

Імовірність того, що сталося l помилок, розташованих як завгодно впродовж послідовності довжини n , визначається формулою Бернуллі

$$P(l) = C_n^l \left(\frac{P}{m-1}\right)^l (1-p)^{n-l}, \quad (4.14)$$

де C_n^l – біноміальний коефіцієнт, що дорівнює числу різних поєднань l помилок у блоці довжиною n , $C_n^l = \frac{n!}{l!(n-l)!}$.

Цю модель називають також каналом з *біноміальним розподілом*. Вона задовільно описує канал, що виникає при певному виборі модема, якщо в неперервному каналі відсутні завмирання, а адитивний шум у ньому "білий" (або принаймні "квазібілий"). Неважко побачити, що імовірність появи помилок (кратних $l \geq 1$) у двійковій кодовій комбінації довжини n відповідно до моделі (4.14) при $p \ll 1$ буде дорівнювати

$$P_{ош} = \sum_{l=1}^n C_n^l p^l (1-p)^{n-l} = 1 - (1-p)^n \approx np.$$

Імовірності переходів у двійковому симетричному каналі схематично показані у вигляді графа на рис. 4.2.

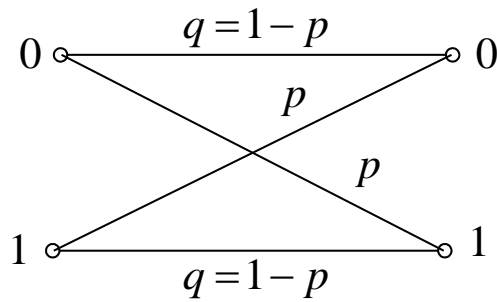


Рис. 4.2. Імовірності переходів у двійковому симетричному каналі

Постійний симетричний канал без пам'яті зі стиранням відрізняється від попередніх тим, що алфавіт на виході каналу містить додатковий $(m+1)$ -й символ, який часто позначають знаком «?». Цей символ з'являється тоді, коли 1-а схема (демодулятор) не може надійно впізнати переданий символ. Імовірність такої відмови від визначення або стирання символу p_c в даній моделі постійна і не залежить від переданого символу. За рахунок введення стирання вдається значно знизити імовірність помилки, іноді її навіть вважають рівною нулю. На рис. 4.3 схематично зображені імовірності переходів у такій моделі.

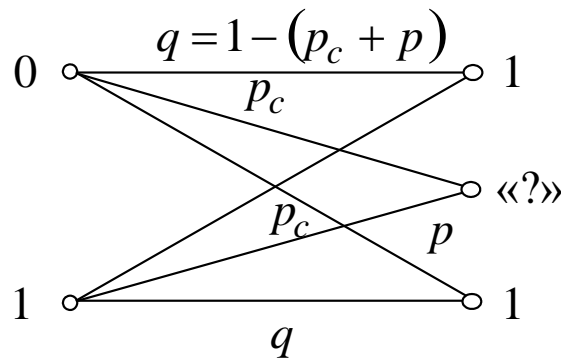


Рис. 4.3. Імовірності переходів у двійковому симетричному каналі зі стиранням

Несиметричний канал без пам'яті характеризується, як і попередні моделі, тим, що помилки виникають у ньому

незалежно, однак імовірності помилок залежать від того, який символ передається.

Так, у двійковому несиметричному каналі імовірність $P(\hat{1}|0)$ прийому символу 1 при передачі символу 0 не дорівнює імовірності $P(0|1)$ прийому 0 при передачі 1 (рис. 4.4).

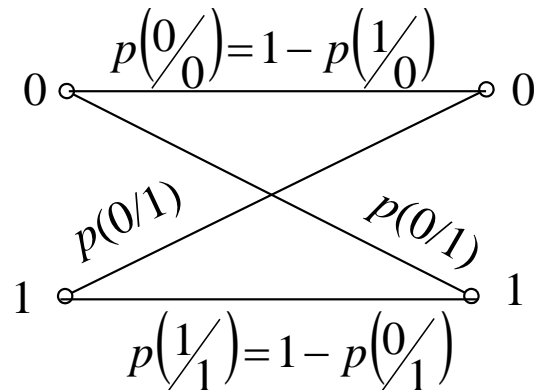


Рис. 4.4. Імовірності переходів у двійковому несиметричному каналі

У цій моделі імовірність вектора помилки залежить від того, яка послідовність символів передається.

4.8.1. Деякі моделі дискретних каналів з пам'яттю

Якщо в постійному симетричному каналі без пам'яті умовна імовірність помилкового прийому $(i + r)$ -го символу за умови, що i -й символ прийнятий помилково, дорівнює безумовній імовірності помилки, то в каналі з пам'яттю вона може бути більшою або меншою за цю величину.

Відхилення розподілу помилок від біноміального (каналу без пам'яті) в реальних каналах викликається різними причинами. Так, дискретним відображенням більшості радіоканалів є канал з пам'яттю внаслідок завмирань.

Іншою причиною можуть бути атмосферні і взаємні завади. Іноді відхилення від біноміального розподілу викликається особливостями методів модуляції і демодуляції. В ущільнених кабельних лініях зв'язку причиною пам'яті вважають комутаційні завади в АТС, що виникають при переключеннях окремих

елементів каналу і по суті вимикають його на короткий час із системи зв'язку.

Найпростішою моделлю двійкового каналу з пам'яттю є марковська, яка визначається матрицею перехідних імовірностей

$$P = \begin{bmatrix} 1 - P_1 & P_1 \\ P_2 & 1 - P_2 \end{bmatrix},$$

де P_1 – умовна імовірність того, що прийнято $(i + 1)$ -й символ помилково, якщо попередній символ прийнято правильно;

$1 - P_1$ – умовна імовірність того, що прийнято $(i + 1)$ -й символ помилково, якщо попередній прийнято помилково;

$1 - P_2$ – умовна імовірність того, що прийнято $(i + 1)$ -й символ правильно, якщо попередній прийнято помилково.

Безумовна (середня) імовірність помилки в розглянутому каналі p повинна задовольняти рівняння

$$p = P_2 P(i_{\text{помил}}) + P_1 P(i_{\text{прав}}) = P_2 p + P_1 (1 - p).$$

Звідки маємо $p = \frac{P_1}{1 + P_1 - P_2}$. Ця модель проста для

використання, однак вона досить неточно відтворює властивості реальних каналів.

Частіше для дискретного каналу з пам'яттю використовується модель Гілберта. Відповідно до цієї моделі канал може знаходитися в двох станах S_1 і S_2 . У стані S_1 помилок не відбувається, а в стані S_2 помилки виникають незалежно з імовірністю p_2 . Переходи з одного стану в інший утворюють просте марковське коло з матрицею переходів

$$S = \begin{bmatrix} 1 - P(S_2 | S_1) & P(S_2 | S_1) \\ P(S_1 | S_2) & 1 - P(S_1 | S_2) \end{bmatrix},$$

де $P(S_2 | S_1)$ – імовірність переходу зі стану S_1 в S_2 ;

$P(S_1 | S_2)$ – імовірність переходу зі стану S_2 в S_1 .

Імовірності перебування каналу в станах S_1 і S_2 відповідно дорівнюють

$$P(S_1) = \frac{P(S_1/S_2)}{P(S_1/S_2) + P(S_2/S_1)}; \quad P(S_2) = \frac{P(S_2/S_1)}{P(S_1/S_2) + P(S_2/S_1)}. \quad (4.15)$$

Безумовна імовірність помилки при цьому запишеться як

$$p = p_2 P(S_2) = p_2 \frac{P(S_2/S_1)}{P(S_1/S_2) + P(S_2/S_1)}.$$

При використанні моделі Гілберта зазвичай припускають, що $p_2 = 0,5$ (тобто цей стан розглядається як обрив зв'язку).

Це добре узгоджується з уявленням про канал, у якому на деяких часових інтервалах через погані умови проходження або дії потужних завад зв'язок порушується, або з уявленням про провідний канал на інтервалі, де діють сильні комутаційні завади або сплески імпульсних завад. Модель Гілберта можна узагальнити, ввівши до розгляду замість двох станів каналів N станів. Але тоді ускладнюється її використання.

Відносно просту модель дискретного каналу з групуванням помилок (з пам'яттю) запропонував Пуртов [19]. У цій моделі використовуються лише два параметри: імовірність помилок p і показник групування α . У моделі Пуртова залежність імовірності $P_{ном}$ появи спотвореної комбінації (з кількістю спотворених елементів ≥ 1) довжини n характеризується як відношення кількості спотворених комбінацій $N_{спотв}(n)$ до загальної кількості переданих комбінацій $N(n)$:

$$P[\geq 1, n] = \lim_{N \rightarrow \infty} \frac{N_{спотв}(n)}{N(n)}.$$

Імовірність $P[\geq 1, n]$ є функцією від n , що не зростає. Відповідно до моделі Пуртова можна записати

$$P[\geq 1, n] \approx n^{1-\alpha} p.$$

Якщо $\alpha = 0$, то $P\{\geq 1, n\} \approx np$, що відповідає біноміальній моделі (дискретному каналу без пам'яті). У цьому випадку немає пакування (групування) помилок.

Найбільше значення α (від 0,5 до 0,7) спостерігається на кабельних лініях зв'язку (короткочасне переривання зв'язку). У радіорелейних лініях (де бувають інтервали з великою інтенсивністю помилок і інтервали з рідкими помилками) $\alpha = 0,3 \dots 0,5$; для деяких ліній короткохвильового радіозв'язку $\alpha = 0,3 \dots 0,4$.

Відповідно до моделі Пуртова-Попова імовірність наявності комбінацій довжиною n з урахуванням t і більше помилок знаходиться з виразу

$$P[\geq 1, n] = \lim_{N \rightarrow \infty} \frac{N_{\text{comb}}[\geq t, n]}{N(n)} \approx \left(\frac{n}{t}\right)^{1-\alpha} p. \quad (4.16)$$

Аналізуючи вираз (4.16), можна зробити висновок, що при заданому n чим більше групування помилок (більше t), тим менша кількість спотворених кодових комбінацій. Це очевидно, бо при одній і тій самій кількості помилок пакування призводить до їхнього зосередження на окремих комбінаціях (кратність t зростає), а кількість спотворених комбінацій зменшується.

Іноді як модель каналу з пам'яттю використовують модель, у якій імовірність вектора помилки $E^{(n)}$ не залежить від переданої послідовності. Імовірність кожного вектора помилки вважається заданою і не визначається його вагою. У багатьох каналах із двох векторів помилки з однаковою вагою більш ймовірним виявляється той, у якому одиниці розташовані близько одна від одної, тобто є тенденція до групування помилок.

Безумовний інтерес являють моделі дискретного каналу, побудовані на основі заданої моделі неперервного каналу, задавання способів модуляції-демодуляції і кодування-

декодування. Однак у загальному вигляді побудувати таку модель важко.

4.8.2. Модель дискретно-безперервного каналу

Дискретно-неперервний (комбінований) канал зв'язку з незалежними символами b_i на вході і неперервним сигналом $z(t)$ на виході описується апіорними імовірностями вхідних символів $P(b_i)$ і перехідними (умовними) щільностями $w_n[z/b_i]$ прийнятої реалізації $z(t)$ (на заданому інтервалі T) за умови передачі символу b_i . Цю щільність називають *функцією правдоподібності*.

Замість функцій правдоподібності дискретно-неперервний канал можна описати апостеріорними імовірностями $P(b_i/z)$ передачі символу b_i при фіксації на прийомі коливання $z(t)$. Відповідно до формули Байєса ці імовірності записуються як

$$P(b_i/z) = \frac{P(b_i)\omega[z/b_i]}{\omega(z)},$$

де $\omega(z)$ - щільність прийнятого коливання,

$$\omega(z) = \sum_{i=0}^{m-1} P(b_i)\omega[z/b_i].$$

Безперервно-дискретний канал описується аналогічно.

Висновки до глави 4

1. Канали зв'язку та електричні ланцюги, що їх реалізують, можна розділити за характером сигналів, що діють на вході і виході, на безперервні, дискретні (цифрові) і дискретно-безперервні (цифро-безперервні) або безперервно-дискретні (безперервно-цифрові).

2. Класифікація каналів (систем, ланцюгів) заснована на властивостях системних операторів, що зв'язують вхід і вихід.

Розрізняють лінійні і нелінійні, стаціонарні і нестаціонарні, зосереджені системи і системи з розподіленими параметрами.

3. В якості інтегральних системних характеристик лінійних ланцюгів найчастіше використовують імпульсну характеристику і передавальну функцію системи, які зв'язані парою перетворень Фур'є.

4. Розрахунок проходження вузькосмугових сигналів через вузькосмугові системи (канали) істотно спрощується, якщо скористатися поняттями комплексних обвідних вхідного сигналу і низькочастотного еквівалента передавальної функції системи.

5. Як при детермінованому, так і при стаціонарному випадковому впливі спектральна щільність середньої потужності на виході лінійної стаціонарної системи дорівнює добутку спектральної щільності середньої потужності вхідного процесу на квадрат модуля передавальної функції системи. Відповідно функція кореляції вихідного процесу визначається згортокою функції кореляції вхідного процесу і тимчасової автокореляційної функції системи.

6. При проходженні випадкових процесів через вузькосмугові лінійні системи вихідний процес має тенденцію до нормалізації незалежно від розподілу вхідного процесу.

7. У теорії та інженерній практиці найчастіше користуються такими моделями лінійного безперервного каналу: неспотворюючий канал з адитивним шумом, канал з невизначеною (випадковою) фазою і адитивним шумом, однопроменевий канал з випадковою фазою і амплітудою (із завмираннями) і адитивним шумом, багатопроменевий канал з адитивним шумом, канал з міжсимвольною інтерференцією (пам'яттю) і адитивним шумом.

8. Для будь-якої моделі дискретного каналу можна визначити поняття випадкового вектора помилки як порозрядну різницю між послідовностями входу і виходу. У двійковому каналі елементи вектора помилки приймають значення 0 і 1. Різні моделі каналів відрізняються розподілом вектора помилки.

9. Найпростіша модель дискретного каналу - симетричний канал без пам'яті (біноміальний канал). Модель дещо ускладнюється введенням символу стирання.

10. Найпростіша модель дискретного каналу з пам'яттю - марковська модель, коли помилки утворюють простий ланцюг Маркова, тобто залежать від того, правильно або помилково прийнятий попередній символ, але не залежать від того, який символ передається.

11. Дискретно-безперервний канал характеризується апріорною імовірністю вхідних символів і функціями правдоподібності (умовною щільністю вихідного сигналу) або апостеріорною імовірністю вхідних символів.

Контрольні питання та задачі

4.1. За якими ознаками можна класифікувати канали зв'язку?

4.2. Який канал називається безперервним, дискретним і дискретно-безперервним?

4.3. Як зв'язані між собою імпульсна і передавальна характеристики лінійної системи?

4.4. Канал характеризується імпульсною характеристикою $g(t, \tau) = g(t)e^{-\alpha\tau}1(t - \tau)$. Знайти відгук каналу $y(t)$, якщо на вхід поданий прямокутний відеоімпульс з амплітудою A і тривалістю T : $x(t) = A[1(t) - 1(t - T)]$.

4.5. Канал має імпульсну характеристику $g(t) = A(1 - e^{-\alpha t})\cos\omega_0 t \cdot 1(t)$. Чи володіє такий канал стійкістю?

4.6. Двопроменевий канал має імпульсну характеристику $g(t) = \gamma_1\delta(t - \tau_1) + \gamma_2\delta(t - \tau_2)$. Знайти модуль його передавальної функції (АЧХ каналу). На яких частотах АЧХ має нульові значення при $\gamma_1 = \gamma_2 = \gamma$?

4.7. Написати інтеграл згортки через квадратурні компоненти комплексних обвідних входу, виходу і каналу. Нарисувати квадратурну електричну схему, що реалізує низькочастотну фільтрацію обвідної вхідного сигналу.

4.8. Які два типи завдань вирішуються при розгляді проходження випадкових впливів через канал зв'язку і його ланки?

4.9. На виході параметричного ланцюга є АМ сигнал $Y(t) = \left[U_0 + k \overset{\circ}{X}(t) \right] \cos(\omega_0 t + \varphi_0)$. Модулюючий процес $\overset{\circ}{X}(t)$

гаусівський, центрований і має функцію кореляції $B(\tau) = e^{-\alpha|\tau|}$.

Знайти одновимірну щільність центрованого процесу $\overset{\circ}{Y}(t)$, функцію кореляції і спектральну щільність потужності процесу $Y(t)$.

4.10. На вхід синхронного детектора з опорною напругою $u_{on}(t) = a \cos(\omega_0 t + \varphi)$ поступає суміш АМ сигналу і стаціонарного гаусівського вузькосмугового флуктуаційного шуму $X(t) = (U_0(1 + mb(t)) + X_c(t)) \cos \omega_0 t + X_s(t) \sin \omega_0 t$ ($b(t)$ – первинний сигнал). Знайти одновимірний розподіл процесу $Y_{н.ч.}(t)$ на виході ФНЧ, його функцію кореляції (якщо спектр вхідного шуму рівномірний у межах смуги частот $f_0 - F_c \leq f \leq f_0 + F_c$), відношення сигнал-шум на виході ФНЧ і на вході детектора.

4.11. На вхід нелінійної схеми з характеристикою $y = x^2$ поступає стаціонарний стандартний гаусівський процес з коефіцієнтом кореляції $R(\tau) = e^{-\alpha|\tau|}$. Знайти одновимірний розподіл вихідного процесу, його математичне очікування, дисперсію, функцію кореляції і спектральну щільність потужності.

4.12. На вхід перемножника з випадковою передавальною функцією $K(t)$ і функцією кореляції $\overline{K(t)K(t+\tau)} = \sigma^2 e^{-\alpha|\tau|}$ поступає стаціонарний "квазібілий" (в межах смуги частот $-F_c, F_c$) низькочастотний шум зі спектральною щільністю N_0 . Знайти функцію кореляції вихідного процесу, його дисперсію і порівняти з дисперсією вхідного шуму.

4.13. Що розуміють під квантовим шумом в оптичних лініях зв'язку?

ГЛАВА 5

ІНФОРМАЦІЙНІ ХАРАКТЕРИСТИКИ ДЖЕРЕЛ ПОВІДОМЛЕНЬ І КАНАЛІВ ЗВ'ЯЗКУ

5.1. Загальні відомості про джерела повідомлень і канали зв'язку

У гл. 1 вже були визначені такі поняття, як інформація, повідомлення, сигнал, кодування і декодування, а в гл. 4 описані деякі моделі дискретних і безперервних каналів зв'язку. Тому в цій главі пригадаємо лише деякі з цих визначень, які безпосередньо знадобляться нам для викладання матеріалу з теорії інформації.

Отже, формально джерело дискретних повідомлень повністю визначається алфавітом $A=(a_0, a_1, \dots, a_{K-1})$ використовуваних символів заданого об'єму K і розподілом імовірностей $P(a)$, заданим на послідовностях $a=(a_1, a_2, \dots, a_n)$, $a_i \in A$ довільної довжини n , складених з символів даного алфавіту A .

Якщо джерело має фіксовану швидкість, на відміну від джерела з регульованою швидкістю, яка визначається кодером, то додатково задається швидкість джерела $v_{дж}$ як число символів, що генеруються таким джерелом за 1 с. Це відповідає окремому випадку, коли тривалість різних символів відрізняється одна від одної.

Кажуть, що джерело *не має пам'яті*, якщо послідовно видавані ним символи взаємно незалежні один від одного. Для повного опису такого джерела достатньо задати лише розподіл імовірності на окремих символах, тобто $P(a_i)$, $a_i \in A$, якщо джерело є *стаціонарним*, тобто розподіли $P(a)$ не залежать від часових зрушень.

Як простий приклад джерела повідомлень можна розглядати друкарський текст якою-небудь мовою. Тоді об'єм алфавіту K співпадатиме з об'ємом алфавіту вибраної мови (26 - для англійської, 32 - для російської), а розподіл імовірності визначатиметься імовірністю різних словосполучень. Очевидно, що це приклад джерела з пам'яттю, оскільки букви в будь-якій

мові не розташовані випадково, а підпорядковані детермінованій або імовірнісній залежності.

Формальний опис дискретного каналу зв'язку зводиться до завдання двох алфавітів на вході і виході каналу: $X=(x_0, x_1, \dots, x_{l-1})$ і $Y=(y_0, y_1, \dots, y_{m-1})$ відповідно різних об'ємів l і m і умовного розподілу імовірності $P(y|x)$, де $x=(x_1, \dots, x_n)$, $y=(y_1, \dots, y_n)$ – послідовності символів з відповідних алфавітів, що мають довільну, але однакову довжину n . Для каналу зв'язку також може додатково задаватися швидкість передачі v_k як число символів, що передаються по каналу зв'язку за 1 с. Канал не має пам'яті, якщо символи приймаються по ньому послідовно незалежно один від одного. У цьому випадку достатньо знати умовний розподіл $P(y_j|x_i)$, $x_i \in X$, $y_j \in Y$ для повного опису такого каналу зв'язку.

Розглянемо деякі важливі окремі випадки дискретного каналу [24].

1. *M-ний симетричний канал (mСК) без пам'яті і без стирання символів (mСК)*. Це канал із співпадаючими об'ємами вхідного і вихідного алфавітів і з умовним розподілом

$$P(y_j|x_i) = \begin{cases} 1-p, & y_j = x_i, \\ p/(m-1), & y_j \neq x_i. \end{cases} \quad (5.1)$$

Рівність (5.1) означає, що всі символи, які передаються по такому каналу зв'язку, приймаються правильно з однаковою імовірністю $1-p$ і кожен символ може перейти в будь-якій інший, відмінний від переданого, також з однією і тією самою імовірністю $p/(m-1)$.

2. *Двійковий симетричний канал без пам'яті і без стирання символів (ДСК або 2СК)*. Це окремий випадок *mСК* при $m=2$.

3. *Двійковий за входом симетричний канал без пам'яті зі стираннями символів*. Це канал без пам'яті з двійковим вхідним алфавітом $X=(x_0, x_1)$ і трійковим вихідним алфавітом $Y=(y_0, y_1, y_2)$ з умовним розподілом

$$P(y_j|x_i) = \begin{cases} 1 - p_c, & i = j, \\ p_c, & j = 2, \\ 0, & i \neq j, \end{cases} \quad (5.2)$$

де p_c – імовірність стирання (або імовірність прийому ненадійного символу).

Цей приклад показує, що не обов'язково об'єми вхідного і вихідного алфавітів повинні співпадати. У даному випадку третій вихідний символ y_2 сформований «штучно» як ненадійно прийнятий символ x_0 або x_1 . (При когерентному оптимальному прийомі сигналів рівної енергії рішення про нього може ухвалюватися, якщо абсолютна величина кореляційного інтеграла у виразі (2.111) менша за деякий заданий поріг).

4. *Двійковий канал з адитивним шумом.* Це канал з двійковими алфавітами на вході і виході, що має в загальному випадку пам'ять. Умовні розподіли в такому каналі зв'язку мають вигляд

$$P(y|x) = P(e = y \oplus x), \quad (5.3)$$

де символ \oplus означає попарне складання по *mod 2* всіх елементів послідовності x і y (нагадаємо, що правила складання по *mod 2* мають вигляд $0 \oplus 0 = 0$, $0 \oplus 1 = 1$, $1 \oplus 0 = 1$, $1 \oplus 1 = 0$);

$P(e)$ – розподіл імовірності, заданий на довільних двійкових послідовностях.

Рівність (5.3) означає, що перехід деякої вхідної послідовності x у вихідну y однозначно визначається так званим *зразком* (або вектором) помилок e , який має одиниці в тих позиціях, де x і y відрізняються, і нулі в решті позицій.

Визначимо формально *кодування* в дискретному каналі як однозначне відображення послідовностей, складених з символів джерела, в послідовності, складені з вхідних символів каналу зв'язку.

Декодуванням називатимемо однозначне відображення послідовностей, складених з вихідних символів каналу, в

послідовності, складені з символів джерела. Тут ми вважаємо, що алфавіти символів джерела й одержувача інформації співпадають.

Довжини послідовностей на вході каналу, відповідні різним послідовностям символів джерела, можуть бути взагалі різними. Тоді кодування називається *нерівномірним*.

Прикладом нерівномірного кодування може служити відомий код Морзе. Дійсно, у цьому коді букві «е» відповідає послідовність з однієї точки «.», а букві «ш» — послідовність з чотирьох тире «----». Зазначемо, що код Морзе – це кодування не з двійковим вхідним алфавітом (крапка і тире), як може показатися на перший погляд, а з трійковим алфавітом: крапка, тире і пропуск.

Якщо довжини всіх послідовностей однакові, то кодування називається *рівномірним*. Легко визначити першу функцію кодування і декодування. Вона полягає в тому, щоб узгоджувати алфавіт джерела повідомлень з алфавітом входу каналу і відповідно алфавіт виходу каналу з алфавітом джерела одержувача повідомлень. Таке узгодження дійсно забезпечує принципову можливість передавати послідовність символів джерела повідомлень до одержувача через дискретний канал зв'язку.

Проте довільно вибране «узгодження» (кодування-декодування) не гарантує ще надійної і швидкої передачі таких повідомлень.

Одна з проблем передавання інформації саме і полягає в тому, щоб знайти потенційні можливості кодування і декодування (за відсутності обмежень на їх складність), коли інформація може бути передана зі скільки завгодно високою правильністю при максимально можливій швидкості її передачі від джерела до одержувача.

У главі наведено регулярні способи кодування-декодування сигналів з урахуванням складності їх реалізації на практиці. Для формулювання і доведення теорем кодування теорії інформації нам потрібно буде викласти деякі нові поняття, які мають і самостійне значення.

5.2. Основні поняття теорії інформації

5.2.1. Часткова кількість інформації

Припустимо що дискретне джерело повідомлень, описане в п. 5.1, видає деяку послідовність символів a . Дамо формальне визначення *часткової кількості інформації* $i(a)$, що міститься в цьому повідомленні, виходячи з таких вимог:

1. Кількість інформації $i(a)$ повинна бути адитивною функцією, тобто для пари взаємно незалежних повідомлень a_1, a_2 вона має дорівнювати сумі кількості інформації в кожному з них, тобто $i(a_1, a_2) = i(a_1) + i(a_2)$.

2. Кількість інформації, що міститься в достовірному повідомленні (що має імовірність $P(a) = 1$), дорівнює нулю.

3. Кількість інформації повинна залежати тільки від імовірності переданого повідомлення, а не від його змісту, тобто $i(a) = f(P(a))$.

4. Кількість інформації повинна бути безперервною функцією від $P(a)$.

5. Кількість інформації є величиною невід'ємною, тобто $0 \leq i(a) < \infty$.

Можна показати, що єдина функція, яка задовольняє всі ці умови, має вигляд [50]

$$i(a) = -\log_c P(a) \geq 0. \quad (5.4)$$

Основа логарифма c у виразі (5.4) може бути вибрана довільною, що впливає лише на одиницю вимірювання кількості інформації.

Якщо $c=2$, то інформація вимірюється у двійкових одиницях або в *бітах*, якщо $c=e$ (як в натуральних логарифмах), то інформація вимірюється в натуральних одиницях або в *натах*, а якщо $c=10$, як в десяткових одиницях, то інформація вимірюється в *дитах*. Із роботи [19] випливає, що вибір основи логарифма не вплине на основні результати теорії інформації, а саме на теорему кодування.

Із співвідношення (5.4) видно, що кількість інформації, яка міститься в повідомленні, тим більше, чим менше імовірність його появи, причому кількість інформації, що міститься в повідомленні про «неможливу» подію, дорівнює нескінченності.

У теорії інформації, що належить до проблем систем зв'язку, не розглядається питання про корисність інформації.

5.2.2. Ентропія джерела повідомлень

Вище визначено часткову кількість інформації, що міститься в деякій послідовності a , виданою джерелом повідомлень. Проте сам факт генерування саме цієї послідовності є випадковою подією, що має імовірність $P(a)$, а отже, випадковою величиною виявляється і кількість інформації $i(a)$. Тому можна поставити питання про середню кількість інформації, що видається деяким джерелом повідомлень, яке можна визначити як математичне очікування випадкової величини $i(a)$.

Якщо поки обмежитися лише послідовностями довжини n , позначивши їх $a^{[n]}$, то відповідно до відомої формули для математичного очікування дискретної випадкової величини отримаємо

$$M \{ i(a^{[n]}) \} = M \{ -\log P(a^{[n]}) \} = -\sum_{i=1}^{K^n} P(a_i^{[n]}) \log P(a_i^{[n]}), \quad (5.5)$$

де підсумовування, як видно, проводиться за всіма можливими послідовностями довжини n з елементами, узятими з алфавіту об'єму K . Для того щоб отримати вичерпну інформаційну характеристику джерела повідомлень, яке взагалі може видавати послідовності необмеженої довжини, потрібно обчислити границю середньої кількості інформації $i(a^{[n]})$, віднесену до одного символу послідовності.

Отримана величина $H(A)$, за Шенноном, називається *ентропією* джерела повідомлень і знаходиться з виразу

$$H(A) = - \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{K^n} P(a_i^{[n]}) \log P(a_i^{[n]}). \quad (5.6)$$

Якщо береться логарифм по основі 2, то $H(A)$ вимірюється в бітах на символ повідомлення. Вираз (5.6), очевидно, матиме сенс лише тоді, коли границя в його правій частині існує. Ця властивість виконується для стаціонарних джерел. Таблиці значень логарифмів приведені у дод. 6.

Зазначимо, що буква A в позначенні ентропії $H(A)$ указує на певне джерело з алфавітом A , причому скорочено опускається вид імовірнісного розподілу $p(a)$.

Якщо джерело повідомлень не має пам'ять, то, використовуючи властивість логарифмічної функції, легко показати, що його ентропія буде дорівнювати

$$H(A) = - \sum_{i=0}^{K-1} P(a_i) \log P(a_i), \quad (5.7)$$

де $P(a_i) i = 0, 1, \dots, K-1$ – імовірності видачі джерелом символів повідомлень $a_i \in A$, причому вони не залежать від номера елемента послідовності, оскільки джерело є стаціонарним.

Визначимо основні властивості ентропії.

1. $H(A) \geq 0$, причому $H(A) = 0$ тоді і тільки тоді, коли одна з послідовностей має одиничну імовірність, а всі інші – нульову.

2. Для будь-якого стаціонарного джерела повідомлень

$$H(A) \leq - \sum_{i=0}^{K-1} P(a_i) \log P(a_i). \quad (5.8)$$

Оскільки вираз (5.8) у правій частині – це ентропія джерела без пам'яті, то дана властивість означає, що пам'ять зменшує ентропію джерела.

3. Для будь-якого стаціонарного джерела повідомлень справедлива нерівність

$$H(A) \leq \log K = H_{\max}(A), \quad (5.9)$$

причому рівність має місце тоді і лише тоді, коли джерело не має пам'яті і всі його символи рівноімовірні.

З властивості 2 витікає, що при доведенні нерівності (5.9) можна обмежитися джерелами без пам'яті. Для доведення властивості 3 розглянемо різницю

$$\begin{aligned} H(A) - \log K &= \sum_{i=0}^{K-1} p(a_i) \left[\log \frac{1}{P(a_i)} - \log K \right] = \\ &= \log e \sum_{i=0}^{K-1} P(a_i) \ln \frac{1}{KP(a_i)}. \end{aligned} \quad (5.10)$$

Далі скористаємося відомою нерівністю $\ln x \leq x - 1$, $x > 0$.

Тоді маємо

$$\begin{aligned} H(A) - \log K &\leq \log e \cdot \sum_{i=0}^{K-1} P(a_i) \left(\frac{1}{KP(a_i)} - 1 \right) = \\ &= \log e \cdot \left(\sum_{i=0}^{K-1} \frac{P(a_i)}{KP(a_i)} - \sum_{i=0}^{K-1} P(a_i) \right) = \log e \cdot (1 - 1) = 0. \end{aligned} \quad (5.11)$$

Рівність у виразі (5.11) матиме місце тільки при $p(a_i) = 1/K$, що і доводить дану властивість.

Скориставшись властивостями 1–3, можна наочно пояснити значення поняття ентропії – це середня інформативність джерела на один символ, що визначає «несподіванку» або «непередбачуваність» видаваних ним повідомлень. Повністю детерміноване джерело, що видає лише одну, наперед відому послідовність, володіє нульовою інформативністю. Навпаки, найбільш «хаотичне» джерело, що видає взаємно незалежні і рівноімовірні символи, володіє максимальною інформативністю.

Ентропія джерела повідомлень тісно пов'язана з поняттям його надмірності, яке формально визначається таким чином:

$$\gamma = \frac{\log K - H(A)}{\log K} \cdot 100\% = \left(1 - \frac{H(A)}{\log K}\right) \cdot 100\% \geq 0. \quad (5.12)$$

Як видно з виразу (5.12), чим більше ентропія джерела, тим менше його надмірність і навпаки.

Відомо, що надмірність природних мов є вельми важливою властивістю, що дозволяє сприймати рукописний або спотворений текст, чути мову у великих акустичних завадах і т. п.

Теорія інформації дозволяє кількісно оцінити ці можливості. Для експериментального обчислення ентропії або надмірності природних мов використовуються статистичні дані про частість, з якою зустрічаються букви тексту і їх поєднання (біграми, тріграми і т. п.).

Так, якщо скористатися так званою статистикою англійської мови 1-го порядку (тобто частістю окремих букв), то ентропія виявляється рівною 4,03 біт/букву, а якщо статистикою порядку, то 3,32 біт/букву. У той же час точнішу оцінку ентропії можна отримати, скориставшись, наприклад, запропонованим ще К. Шенноном «методом випадкового вгадування» [30].

Такий підхід дає верхню оцінку ентропії англійської мови 2 біт/букву і нижню 1 біт/букву. Це дозволяє зробити висновок, що основна надмірність природної мови визначається багатовимірними залежностями між буквами (кореляцією тексту), і для її використання необхідно знати ці залежності.

Аналогічні висновки можна зробити і щодо російської та інших мов, хоча їх ентропії дещо відрізняються одна від одної.

Найбільш простої форми набуває ентропія у разі двійкового джерела повідомлень без пам'яті. Якщо скорочено позначити $P(a_0)=P(0)$ і $P(a_1)=P(1)$, то

$$\begin{aligned} H(A) &= -[P(0)\log P(0) + P(1)\log P(1)] = \\ &= -[P(0)\log P(0) + (1 - P(0))\log(1 - P(0))]. \end{aligned} \quad (5.13)$$

Вид цієї функції показаний на рис. 5.1 для основи логарифма, що дорівнює двом.

Якщо джерело повідомлень має фіксовану швидкість $v_{дж} = \frac{1}{T_{дж}}$, симв/с, то визначимо продуктивність джерела $H'(A)$ як ентропію в одиницю часу, (секунду):

$$H'(A) = v_{дж} H(A). \quad (5.14)$$

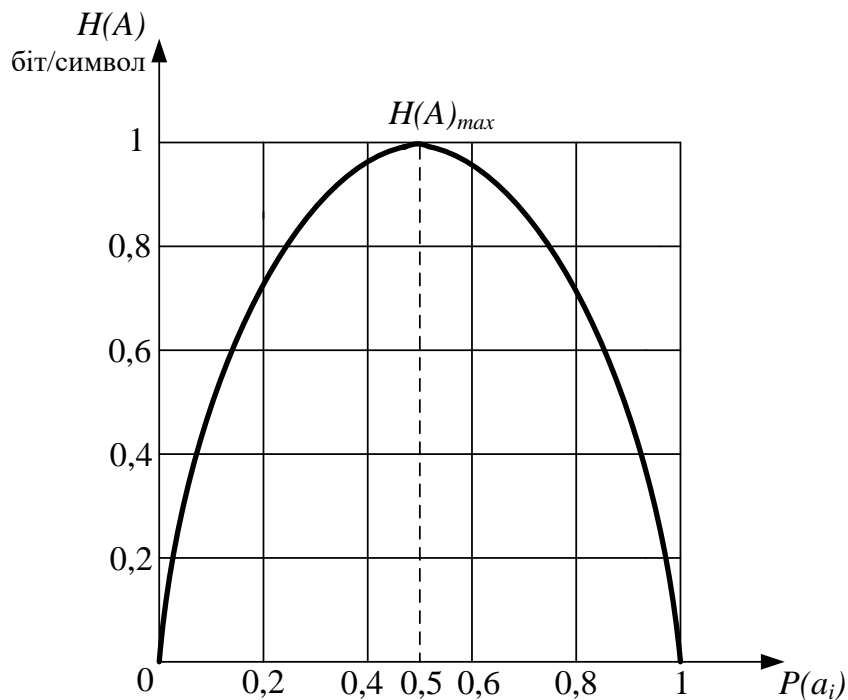


Рис. 5.1. Ентропія двійкового джерела без пам'яті

Максимум виразу (5.14), що дорівнює $R_{дж} = \frac{\log_2 K}{T_{дж}}$, біт/с, зазвичай називають *інформаційною швидкістю* джерела повідомлень.

5.3. Інформаційні характеристики дискретних каналів зв'язку

5.3.1. Кількість інформації, що передається по каналу зв'язку (взаємна інформація)

Розглянемо деякий дискретний канал зв'язку, заданий алфавітом входу і виходу (X, Y) і умовним розподілом $P(y|x)$, $x \in X^n$, $y \in Y^n$, і припустимо, що на вхід каналу

безпосередньо поступають повідомлення від деякого дискретного джерела повідомлень з алфавітом A , співпадаючим з вхідним алфавітом каналу X . Це припущення еквівалентне умові, що сам вхід каналу є дискретним джерелом повідомлень з алфавітом X і розподілом імовірності послідовностей, складених з символів цього алфавіту.

Припустимо, що це джерело має деяку ентропію $H(X)$ і видає послідовність x . Тоді, якщо на виході каналу з'явилася деяка послідовність y , можна поставити таке питання: як розрахувати кількість інформації $i(x|y)$, що міститься в x , за умови, що на виході з'явилася послідовність y ? Відповідь на нього можливо отримати, скориставшись визначенням часткової кількості інформації (вираз (5.4)), але при заміні безумовного розподілу імовірності на умовне, тобто

$$i(x|y) = -\log P(x|y) \geq 0. \quad (5.15)$$

Визначимо умовну ентропію $H(X|Y)$ входу каналу X при відомому виході Y як МО (вираз (5.15)) аналогічно тому, як це було зроблено при знаходженні ентропії джерела з виразу (5.6).

Тоді маємо

$$H(X|Y) = -\lim_n \frac{1}{n} \sum_{i=1}^{l^n} \sum_{j=1}^{m^n} P(x_i^{[n]}, y_j^{[n]}) \log P(x_i^{[n]} | y_j^{[n]}), \quad (5.16)$$

де верхній індекс n у квадратних дужках означає довжину вхідних і вихідних послідовностей.

В окремому випадку каналу без пам'яті легко отримати з виразу (5.16) вираз

$$H(X|Y) = - \sum_{i=0}^{l-1} \sum_{j=0}^{m-1} P(x_i, y_j) \log P(x_i | y_j). \quad (5.17)$$

Умовна ентропія має такі властивості:

$$1) H(X|Y) \geq 0.$$

(Доводиться за визначенням $H(X|Y)$);

2) якщо вхід і вихід каналу зв'язку взаємно однозначні, тобто

$$P(x|y) = \begin{cases} 1, & y = \tilde{y}(x) \in Y^n, \\ 0, & y \neq \tilde{y}(x), \tilde{y}(x_1) \neq \tilde{y}(x_2), \text{ при } x_1 \neq x_2, \end{cases}$$

то $H(X|Y) = 0$;

$$3) H(X|Y) \leq H(X); \quad (5.18)$$

$$4) H(X|Y) = H(X) \quad (5.19)$$

тоді, і лише тоді, коли $P(x|y) = P(x)$ при всіх $x \in X^n$, $y \in Y^n$, тобто коли x і y взаємно незалежні (перевіряється безпосередньо).

Наведені вище властивості дозволяють наочно пояснити значення поняття умовної ентропії $H(X|Y)$. Це середня інформація, що втрачається з кожним символом у каналі зв'язку через завади.

Дійсно, оскільки $H(X|Y) \geq 0$, то інформація через завади завжди втрачається, але ніколи не отримується. Якщо помилки в каналі відсутні, тобто вхідна послідовність переходить в одну і лише в одну вихідну послідовність, то $H(X|Y) = 0$, і втрати інформації відсутні. (Не виключена ситуація, коли вхідна послідовність переходить з деякою імовірністю в різні вихідні, але все-таки за вихідною послідовністю можна однозначно визначити вхідну. Тоді, як і раніше, $H(X|Y) = 0$, і втрати інформації будуть відсутні.)

Нарешті, якщо вихід каналу не залежить від входу, а саме це і означає збіг умовної і безумовної імовірностей, то по каналу зв'язку не передаватиметься ніякої інформації, а оскільки на вхід каналу поступає інформація $H(X)$, то вся вона виявляється втраченою через завади, тобто $H(X|Y) = H(X)$.

Визначимо *кількість інформації*, що передається по каналу зв'язку $I(X, Y)$, або взаємну інформацію між виходом Y і входом X як різницю між кількістю інформації, що поступає на вхід і

дорівнює ентропії входу $H(X)$, і кількістю інформації, втраченою в каналі зв'язку і яка дорівнює умовній ентропії $H(X|Y)$.

Тому можливо записати

$$I(X, Y) = H(X) - H(X|Y). \quad (5.20)$$

Ця величина має такі властивості:

$$1) I(X, Y) = I(Y, X) = H(Y) - H(Y|X),$$

де ентропія виходу $H(Y)$ і умовна ентропія виходу при відомому вході $H(Y|X)$ визначаються абсолютно аналогічно до виразів (5.6) і (5.16) з тією лише різницею, що розподіли, які є там, $P(a_i^{[n]}) = P(x_i^{[n]})$ і $P(x_i^{[n]}|y_j^{[n]})$ повинні бути замінені на $P(y_j^{[n]})$ і $P(y_i^{[n]}|x_j^{[n]})$ відповідно. Ця властивість вимагає доведення [44];

$$2) I(X, Y) \geq 0.$$

(Впливає безпосередньо з властивості 3 умовної ентропії);

3) $I(X, Y) = 0$ тоді і тільки тоді, коли вхід і вихід каналу статистично незалежні, тобто $P(y|x) = p(y)$ при всіх $x \in X^n$, $y \in Y^n$. (Впливає безпосередньо з властивості 4 умовної ентропії);

$$4) I(X, Y) \leq H(X).$$

(Впливає з визначення $I(X, Y)$ і властивості 1 умовної ентропії);

$$5) \text{ якщо в каналі зв'язку відсутні завади, то } I(X, Y) = H(X)$$

(Впливає з властивості 2 умовної ентропії);

$$6) I(X, Y) \leq H(Y).$$

(Впливає з властивості 1 для кількості переданої по каналу інформації і властивості 1 для умовної ентропії);

$$7) I(X, X) = H(X).$$

(Впливає формально безпосередньо з визначення, якщо покласти $x=y$, $X=Y$);

8) будь-які детерміновані або випадкові перетворення вихідних послідовностей каналу, у тому числі й угруповання спостережень, тобто об'єднання декількох послідовностей, що

належать певній підмножині в один символ, не можуть збільшити кількість інформації.

У випадку, якщо ці перетворення взаємно однозначні, то кількість інформації не змінюється. Зазначена властивість впливає з того, що, пройшовши по каналу (перетворювачу), інформація або втрачається (за наявності неоднозначних перетворень), або не змінюється (при однозначних перетвореннях).

Вираз (5.20) для взаємної інформації може бути перетворено до такого вигляду, який іноді зручніший для розрахунків:

$$\begin{aligned}
 I(X, Y) &= \lim_n \frac{1}{n} \left[M \left\{ \log \frac{1}{P(x_i^{[n]})} \right\} - M \left\{ \log \frac{1}{P(x_i^{[n]}/y_j^{[n]})} \right\} \right] = \\
 &= \lim_n \frac{1}{n} \left[M \left\{ \log \frac{P(x_i^{[n]}/y_j^{[n]})}{P(x_i^{[n]})} \right\} \right] = \lim_n \frac{1}{n} \left[M \left\{ \log \frac{P(x_i^{[n]}, y_j^{[n]})}{P(x_i^{[n]})P(y_j^{[n]})} \right\} \right] = \\
 &= \lim_n \frac{1}{n} \sum_{i=1}^{l^n} \sum_{j=1}^{m^n} P(x_i^{[n]}, y_j^{[n]}) \cdot \log \frac{P(x_i^{[n]}, y_j^{[n]})}{P(x_i^{[n]})P(y_j^{[n]})}.
 \end{aligned}$$

В окремому випадку каналу без пам'яті отримуємо такий вираз:

$$I(X, Y) = \sum_{i=0}^{l-1} \sum_{j=0}^{m-1} P(x_i, y_j) \log \frac{P(x_i, y_j)}{P(x_i)P(y_j)}. \quad (5.21)$$

Формула (5.21) підтверджує первинне визначення $I(X, Y)$ як кількості інформації, що передається по каналу зв'язку з завадами. Дійсно, якщо є обрив каналу, відповідний відсутності статистичного зв'язку між входом і виходом, ($P(x_i, y_j) = P(x_i)P(y_j)$), то інформація по каналу не передається (властивість 3).

Якщо в каналі завади відсутні, то можливо записати

$$P(x_i, y_j) = \begin{cases} P(x_i) = P(y_j), & j = i \\ 0 & , \quad j \neq i \end{cases}$$

тобто інформація передається повністю (властивість 5).

Властивість 1 виявляється менш очевидною. Вона показує, що кількість інформації, яка передається з входу на вихід і «ніби» з виходу на вхід, виявляється одним і тим самим. По суті ця властивість і дозволяє визначити величину $I(X,Y)$, що задається співвідношенням (5.20), як взаємну інформацію між X і Y (або, що те саме, між Y і X).

Таким чином, величина $I(X,Y)$ подібно до коефіцієнта кореляції для гаусівських випадкових величин виступає тут як міра *статистичної залежності* між входом X і виходом Y каналу зв'язку. Властивість 1 спрощує розрахунок $I(X,Y)$ в деяких випадках, що визначається такою властивістю умовної ентропії:

9) для двійкового каналу з адитивним шумом (п. 4.7) де $H(E)$ – ентропія завади, що задається співвідношенням (5.6), в якому розподіл $p(a_i^{[n]})$ замінений на розподіл зразків помилок $p(e_i^{[n]})$.

Об'єднавши властивість 9 для умовної ентропії в каналі з адитивним шумом і властивість 1 для кількості передаваної інформації, отримуємо

$$I(X,Y) = H(Y) - H(E). \quad (5.22)$$

Визначення взаємної інформації наочно ілюструється на рис. 5.2.

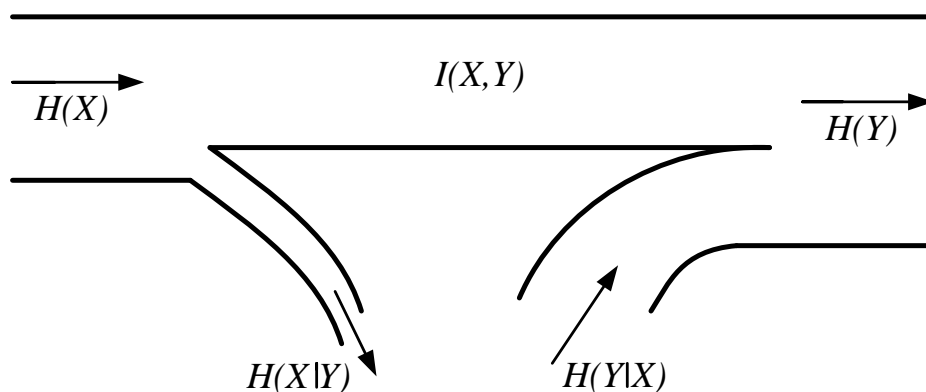


Рис. 5.2. Ілюстрація передачі інформації по каналу з завадами

Якщо для каналу зв'язку задана швидкість передачі v_K , симв/с, то аналогічно до визначення продуктивності джерела можна визначити швидкість передачі інформації по каналу зв'язку $I'(X,Y)$, біт/с,

$$I'(X,Y) = v_K \cdot I(X,Y). \quad (5.23)$$

5.3.2. Пропускна спроможність дискретного каналу зв'язку

Визначимо пропускну спроможність C дискретного каналу зв'язку з адитивними завадами як максимум кількості інформації $I(X,Y)$ за будь-якими розподілами $p(x)$ входу каналу, тобто

$$C = \max_{P(x)} I(X,Y). \quad (5.24)$$

З визначення видно, що пропускна спроможність каналу зв'язку залежить тільки від властивостей самого каналу, тобто вхідного і вихідного алфавітів X, Y і заданого на них умовного розподілу імовірностей $p(y|x)$, $x \in X^n$, $y \in Y^n$, і не залежить від того джерела, яке підключене до входу каналу.

Пропускна спроможність каналу має такі властивості:

$$1) C = \max_{P(x)} \{H(Y) - H(Y|X)\}.$$

(Впливає з властивості 1 кількості інформації);

$$2) C \geq 0.$$

(Впливає з властивості 2 кількості інформації);

$$3) C = 0$$

тоді і тільки тоді, коли вхід і вихід каналу статистично незалежні, тобто має місце «обрив каналу». (Впливає з властивості 3 кількості інформації);

$$4) C \leq \log m.$$

(Впливає з властивості 4 кількості інформації і властивості 3 для ентропії, у даному випадку $K=m$);

5) $C = \log m$ за відсутності завад у каналі зв'язку. (Впливає з властивості 5 кількості інформації і властивості 3 для ентропії);

$$6) C \leq \max_{P(x)} \{H(Y)\} \leq \max_{P(y)} \{H(Y)\}.$$

(Впливає з властивості 6 кількості інформації).

Являє значний інтерес обчислення пропускної спроможності для різних каналів зв'язку з адитивними завадами. У загальному випадку це досить складне, а іноді і просто неможливе для виконання завдання.

Проте для деяких розглянутих раніше моделей каналів це виявляється цілком можливим. Розглянемо деякі із них.

1. m СК без пам'яті.

Скористаємося для розрахунку пропускної спроможності такого каналу формулою, обумовленою властивістю 1,

$$C = \max_{p(x)} \{H(Y) - H(Y|X)\}. \quad (5.25)$$

Покажемо, що умовна ентропія $H(Y|X)$ в m СК не залежить від вхідного розподілу $P(x)$. Дійсно, з її визначення для каналу без пам'яті одержуємо з перетвореного для $H(Y|X)$ з урахуванням виразу (5.17)

$$\begin{aligned} H(Y|X) &= - \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} P(y_j|x_i) P(x_i) \log P(y_j|x_i) = \\ &= - \sum_{i=0}^{m-1} P(x_i) \sum_{j=0}^{m-1} P(y_j|x_i) \log P(y_j|x_i) = \\ &= \sum_{i=0}^{m-1} P(x_i) \left(p \log \frac{p}{m-1} + (1-p) \log(1-p) \right) = \\ &= - \left(p \log \frac{p}{m-1} + (1-p) \log(1-p) \right). \end{aligned} \quad (5.26)$$

Тому вираз (5.25) можна переписати так:

$$C = \max_{p(x)} \{H(Y)\} + p \log \frac{p}{m-1} + (1-p) \log(1-p). \quad (5.27)$$

За властивістю 3 ентропії $H(Y) \leq \log m$, але якщо вибрати рівномірний розподіл на вході, тобто

$$p(x_i) = 1/m, \quad i = 0, 1, \dots, m-1,$$

то одержимо

$$P(y_j) = \sum_{i=0}^{m-1} P(x_i) P(y_j | x_i) = \frac{1}{m} \sum_{i=0}^{m-1} P(y_j | x_i) = \frac{1}{m} \quad (5.28)$$

і, отже,

$$\max_{p(y)} \{H(Y)\} = \log m, \quad (5.29)$$

причому максимум досягається при рівномірному вхідному розподілі імовірностей. Підставляючи вираз (5.29) у вираз (5.27), остаточно одержуємо

$$C = \log m + p \log \frac{p}{m-1} + (1-p) \log(1-p). \quad (5.30)$$

2. Двійковий симетричний канал (2СК) без пам'яті.

Це окремий випадок m -кового СК (m СК) при $m=2$. Підставляючи $m=2$ у вираз (5.30), знаходимо (для основи логарифмів 2), біт/символ,

$$C = 1 + p \log_2 p + (1-p) \log_2(1-p). \quad (5.31)$$

3. Двійковий по входу канал зі стиранням має пропускну спроможність, біт/симв,

$$C = 1 - p_{cm}. \quad (5.32)$$

4. Двійковий канал з пам'яттю й адитивним шумом.

Із властивості 1 для кількості інформації й властивості 5 умовної ентропії одержуємо

$$C = \max_{p(y)} \{H(Y)\} - H(E). \quad (5.33)$$

Аналогічно до того, як це було доведено у виразі (5.28), одержуємо, що при виборі взаємно незалежних і рівноімовірнісних символів x_i символи y_i на виході також виявляються взаємно незалежними й рівноімовірнісними. Тому виконується рівність (5.29), а з урахуванням виразу (5.33) знаходимо, біт/символ,

$$C = 1 - H(E). \quad (5.34)$$

Зазначимо, що ентропія джерела завади E у вигляді послідовності двійкових символів з імовірностями p й $1-p$ задовольняє нерівність, біт/символ,

$$H(E) \leq -p \log_2 p - (1-p) \log_2 (1-p),$$

причому рівність настає лише для джерела завад без пам'яті.

Таким чином, пропускна спроможність двійкового каналу з пам'яттю (вираз (5.34)) більше за пропускну спроможність 2СК без пам'яті (вираз (5.31)).

З формули (5.30) видно, що $C=0$ для m СК при $p = \frac{m-1}{m}$, $1-p = \frac{1}{m}$. Це саме й відповідає випадку «обриву каналу зв'язку», оскільки кожний із вхідних символів з рівною ймовірністю переходить у будь-який з вихідних.

Спостерігаючи вихідні символи, не можна надати перевагу жодному із вхідних символів, а це й відповідає поняттю обриву каналу, коли передача інформації по ньому виявляється зовсім марною, оскільки той же результат може бути отриманий при випадковому вгадуванні вхідних символів у точці прийому.

На рис. 5.3 показана залежність пропускної спроможності 2СК без пам'яті від імовірності помилки символу p у каналі

зв'язку. Як і слід було сподіватися, пропускна спроможність дорівнює нулю при $p=1/2$. Трохи несподіваним, на перший погляд, може здатися те, що $C=1$ також і при $p=1$.

Однак у дійсності випадок, коли $p=1$, – це аж ніяк не стан каналу з найбільшими завадами, а стан з так називаною *зворотною роботою*, коли всі нулі переходять в одиниці, а одиниці – у нулі.

Однак оскільки ця властивість каналу передбачається заздалегідь відомою (бо відомо, що $p=1$), то можемо здійснювати декодування за правилом $0 \rightarrow 1, 1 \rightarrow 0$. Тоді всі вхідні символи будуть прийматися абсолютно правильно, і тому цілком природно, що пропускна спроможність такого каналу дорівнює максимальній величині.

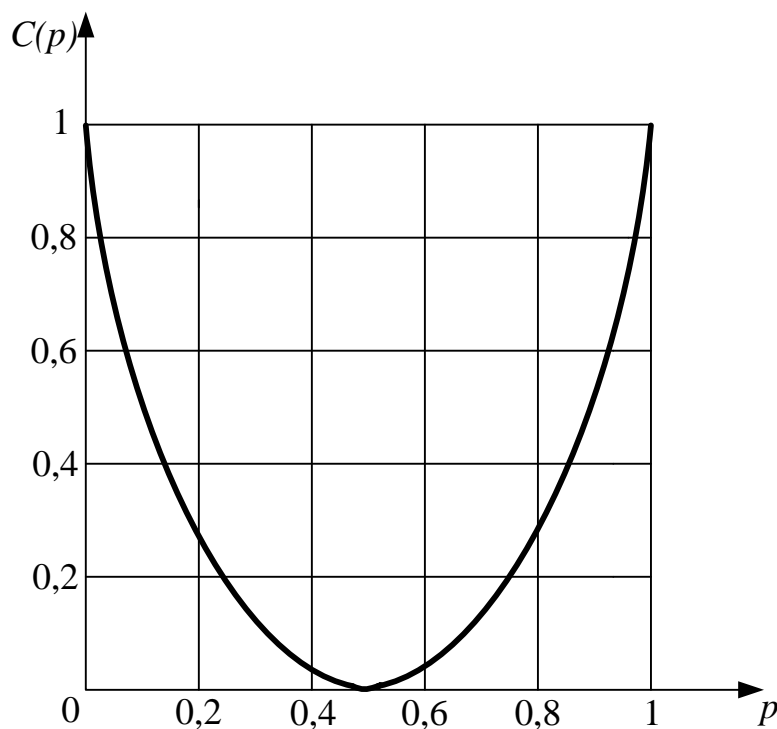


Рис. 5.3. Залежність пропускної спроможності 2СК без пам'яті від імовірності помилки символу

Зазначимо, що ця ситуація відрізняється від «зворотної роботи», яка описана у гл. 3, й є наслідком квазікогерентного прийому ФМ сигналів. Справа в тому, що стрибки фази там виникають у випадкові моменти часу, приводячи до ділянок то

правильної, то "зворотної роботи". Оскільки моменти стрибків є випадковими, то неможливо скоригувати їх зворотним декодуванням. Модель такого каналу аж ніяк не еквівалентною 2СК, що має ймовірність помилки $p=1$.

Підкреслимо, що формула (5.31) має найбільш простий вигляд при виборі двійкової основи логарифма, коли пропускна спроможність каналу зв'язку вимірюється у двійкових одиницях на символ або у бітах на символ.

Можна визначити пропускну спроможність каналу C' в одиницю часу як

$$C' = v_k C,$$

де v_k – швидкість передачі символів по каналу зв'язку (число символів в 1 с).

Якщо при визначенні використаний двійковий логарифм, то C' буде вимірятися в бітах за секунду.

Аналогічно до умовної ентропії можна ввести поняття середньої умовної взаємної інформації. Обмежуючись для простоти випадком каналу без пам'яті, одержуємо вираз

$$I(X, Y|Z) = \sum_{x \in X} \sum_{y \in Y} \sum_{z \in Z} P(x, y, z) \log \frac{P(x, y|z)}{P(x|z)P(y|z)}.$$

Однак на відміну від умовної ентропії, для якої завжди справедлива нерівність (5.18), умовна взаємна інформація $I(Z, X/Y)$ може бути меншою, більшою або дорівнювати безумовній взаємній інформації $I(X, Y)$.

Крім середньої умовної взаємної інформації можна визначити також середню взаємну інформацію між парою X, Y й Z :

$$I(X, Y; Z) = \sum_{x \in X} \sum_{y \in Y} \sum_{z \in Z} P(x, y, z) \log \frac{P(x, y, z)}{P(x, y)P(z)}.$$

Між даними величинами існують такі співвідношення [44]:

$$I(X, Y; Z) = I(X, Z) + I(Y, Z|X) = I(Y, Z) + I(X, Z|Y). \quad (5.35)$$

Для того щоб наочно пояснити різницю між величинами $I(X,Z;Y)$ і $I(X,Z|Y)$, які, на перший погляд, можуть здатися однаковими, розглянемо таку модель каналу зв'язку:

$$x = z \oplus y \oplus e,$$

де \oplus – означає підсумовування по *mod 2*;

x , y , e – деякі двійкові випадкові величини, причому наочно z – вхід каналу, x – вихід каналу, e – завада в каналі, y – додаткова перетворювальна послідовність на передачі.

Тоді якщо y не залежить від z , то $I(Z,Y) = 0$ й відповідно до виразу (5.35) $I(Z,X;Y) = I(Z,X|Y) = 1 - H(E)$, де $H(E)$ – ентропія завади E . Якщо ж z й y взаємно залежні, то $I(Z,Y) > 0$ й тому $I(Z,X;Y) > I(Z,X|Y)$. В окремому випадку, якщо $x=y$, то $I(Z,X|Y) = 0$, а $I(Z,X;Y) = 1$.

Теоретико-інформаційні поняття, наведені в даній главі, можна розглядати у двох аспектах: як розвиток математичного апарату, що примикає до теорії ймовірностей, і як характеристику, що пояснює процес передачі інформації з каналів зв'язку з завадами.

Середня інформативність джерела може бути кількісно оцінена його ентропією, а кожному джерелу, з'єднаному з каналом зв'язку, можна приписати деяке число, що виражає кількість інформації, переданої по цьому каналу. Більш того, кожному каналу відповідає деяка гранична кількість інформації, яка називається його пропускною спроможністю, більше за яку даний канал зв'язку не може передати ні від якого з джерел повідомлень.

Ця пропускна спроможність максимальна за відсутності завад і дорівнює нулю при втраті статистичного зв'язку між входом і виходом каналу. Здавалося б, ми одержали відомості, що дозволяють нам значно просунутися вперед у напрямку розуміння процесів передачі повідомлень по каналах з завадами.

Однак це ще лише деяке «передбачення результатів». Дійсно, як ми можемо використати значення ентропії джерела для чого-небудь більшого, ніж тавтологія, що стверджує, ніби він генерує середню інформацію, яка дорівнює ентропії? Як

перетворити наше визначення пропускної спроможності каналу на більш реальні в практичному відношенні характеристики процесу передачі інформації – надійність і швидкість передачі символів джерела?

Багато це чи мало, що пропускна спроможність одного каналу більше за іншу на 5 %, на 50 %, у 10 разів? Важливий факт говорить про те, що обрив каналу ($C=0$) виникає за відсутності статистичного зв'язку входу й виходу, тобто легко виходить і без використання інформаційного апарату.

Зазначимо, що для доведення й наочного трактування теорем кодування зовсім не потрібно інтуїтивного розуміння ентропії, умовної ентропії, кількості інформації й т. п. Цілком достатнім є розглядати даний розділ саме як розвиток деякого математичного апарату, що доводить властивості цих специфічних функцій імовірнісних розподілів.

5.4. Теореми кодування Шеннона для дискретного каналу зв'язку

Розглянемо спочатку теорему кодування в каналі без завад, який, очевидно, є окремим випадком каналу з завадами. Найкраще кодування полягає в тому, щоб при заданому каналі зв'язку, який визначається обсягом алфавіту m і швидкістю передачі каналних символів v_K , а також при заданому джерелі інформації, обумовленому імовірнісним розподілом $p(a)$ послідовностей, складених із символів, що належать алфавіту джерела A , забезпечити найбільшу можливу швидкість передачі символів джерела $v_{дж}$.

Зазначимо, що змістовність проблеми кодування в каналі без завад визначається не тільки розбіжністю обсягів алфавітів джерела й каналу, але також і тією обставиною, що якщо джерело має пам'ять або (і) його символи не рівноімовірні, то ці особливості можна істотно використати для збільшення швидкості передачі інформації від даного джерела.

Конструктивні методи такого кодування будуть розглянуті в наступній главі, а зараз розглянемо асимптотичні результати, тобто кодування при необмежених послідовностях символів джерела й каналу зв'язку, що порівнюються один з одним. Для

цього нам знадобиться одна теорема, що є деякою версією закону великих чисел, відомого в теорії ймовірностей.

5.4.1. Теорема про властивість асимптотичної рівноімовірності (ВАР)

Для будь-яких заданих як завгодно малих позитивних чисел ε і δ можна знайти таке число n (яке залежить від ε , δ і властивостей джерела), що з імовірністю, більшою, ніж $1-\delta$, джерело повідомлень видає послідовність $a^{[n]}$ довжини n , яка має ймовірність $P(a^{[n]})$, що задовольняє нерівність [24]

$$\left| \frac{1}{n} \log \frac{1}{P(a^{[n]})} - H(A) \right| < \varepsilon, \quad (5.36)$$

де $H(A)$ – ентропія даного джерела.

Доведення теореми для стаціонарних джерел без пам'яті має наступний вигляд. Відповідно до закону великих чисел при досить великому n частота n_i/n події, що полягає в появі символу a_i у послідовності довжини n , наближається до ймовірності $P(a_i)$, тобто

$$n_i/n \rightarrow P(a_i). \quad (5.37)$$

З іншого боку, якщо деяка послідовність $a^{[n]}$ містить n_0 символів a_0 , n_1 символів a_1, \dots, n_{K-1} символів a_{K-1} , то ймовірність її появи дорівнює

$$P(a^{[n]}) = P^{n_0}(a_0) \cdot P^{n_1}(a_1) \cdot \dots \cdot P^{n_{K-1}}(a_{K-1}).$$

Тоді маємо

$$\begin{aligned} \frac{1}{n} \log \frac{1}{P(a^{[n]})} &= \frac{n_0}{n} \log \frac{1}{P(a_0)} + \frac{n_1}{n} \log \frac{1}{P(a_1)} + \\ &+ \dots + \frac{n_{K-1}}{n} \log \frac{1}{P(a_{K-1})}. \end{aligned} \quad (5.38)$$

Підставляючи вираз (5.37) в вираз (5.38), одержуємо, що величина $\frac{1}{n} \log \frac{1}{P(a^{[n]})}$ буде прагнути за ймовірністю до виразу

$$-\sum_{i=0}^{K-1} P(a_i) \log P(a_i) = H(A).$$

Власне кажучи теорема ВАР означає, що при досить великій величині n всі послідовності, які видаються джерелом, розбиваються на дві групи, що називаються відповідно *типовими* й *нетиповими*.

Перші послідовності приблизно рівноімовірні, і кількість їх приблизно дорівнює

$$N_{min} = \frac{1}{P(a^{[n]})} \approx 2^{nH(A)}. \quad (5.39)$$

Що ж стосується нетипових послідовностей, то вони можуть мати різні ймовірності, але ймовірність появи хоча б однієї з них наближається до нуля при $n \rightarrow \infty$. Зазначимо, що число типових послідовностей певної довжини виявляється значно менше від загального можливого числа послідовностей такої самої довжини.

Наприклад, загальне число послідовностей букв російської мови довжини $n=10$ дорівнює $32^{10} \approx 10^{15}$, у той час як число типових послідовностей такої самої довжини, якщо прийняти ентропію російської мови 1,5 біт/симв., виявляється лише $2^{10 \cdot 1,5} \approx 33000$, що приблизно відповідає обсягу типового словника.

Теорема ВАР виявляється справедливою й для пар послідовностей $x^{[n]}$, $v^{[n]}$, щодо яких стверджується, що при досить великих n з імовірністю $1-\delta$ джерело видає пари послідовностей, для яких їхні умовні ймовірності будуть задовольняти нерівності

$$\left| \frac{1}{n} \log \frac{1}{P(x^{[n]}|y^{[n]})} - H(X|Y) \right| < \varepsilon, \quad (5.40)$$

$$\left| \frac{1}{n} \log \frac{1}{P(y^{[n]}|x^{[n]})} - H(Y|X) \right| < \varepsilon, \quad (5.41)$$

де ε – як завгодно мала величина.

5.4.2. Теорема 1. Про кодування джерела повідомлень

Існує спосіб кодування, при якому середня довжина послідовності каналних символів \bar{n} , що доводиться на один символ джерела повідомлень,

$$\bar{n} = \frac{V_{\kappa}}{V_{\partial}} = \frac{H(A)}{\log m} + \varepsilon, \quad \varepsilon > 0. \quad (5.42)$$

Не існує способу кодування, при якому \bar{n} менше, ніж $\frac{H(A)}{\log m}$.

Прикладний зміст цієї теореми полягає в тому, що при найкращому кодуванні в каналі без завад ми можемо передавати повідомлення джерела зі швидкістю $V_{\partial ж}$, як завгодно близької до величини

$$V_{\partial ж} = V_{\kappa} \frac{\log m}{H(A)}, \quad (5.43)$$

і неможливо передавати повідомлення зі швидкістю $v_{дж}$ більшою, ніж у виразі (5.43). Очевидно, що швидкість передачі виявляється тим більшою, ніж менше ентропія джерела або чим більше його надмірність, що відповідає нашій інтуїції підказує можливість більш швидкої передачі повідомлень за рахунок усунення надмірності, що втримується в них. Такий метод кодування називається також *кодуванням джерел* повідомлень або іноді –

статистичним або ощадливим кодуванням (стиском) повідомлень.

Оскільки теорема 1 дає умову лише для середньої довжини блоку каналних символів, то очевидно, що в окремі моменти часу ці довжини можуть виявитися значно більшими від середньої довжини, а це потребує для джерел з фіксованою швидкістю використання спеціального буферного пристрою для поглинання затримки повідомлень, які поступають.

Можна показати, що з імовірністю одиниця накопичувач будь-якої кінцевої ємності рано чи пізно буде переповнений, тобто відбудеться втрата інформації. Отже, зазначене вище кодування є звичайно нерівномірним. Тому для джерел повідомлень із фіксованою швидкістю має сенс така видозміна теореми 1, в якій використовується рівномірне кодування.

5.4.3. Теорема 2. Про кодування в каналі без завад

Існує спосіб кодування й декодування в каналі без завад, при якому для довжини n послідовності каналних символів, що доводяться на один символ джерела, буде виконуватися співвідношення (5.42), причому ймовірність помилки не перебільшує будь-якої як завгодно малої величини $\delta > 0$.

Зазначимо, що тут під імовірністю помилки розуміється імовірність того, що послідовність символів, яка видана одержувачу, буде відрізнятися від відповідної їй послідовності символів, переданої джерелом повідомлень. Незважаючи на те, що ми маємо справу з каналом без завад, помилки в прийнятому повідомленні з'являються внаслідок спеціального способу кодування.

З виразу (5.39) випливає, що число типових послідовностей джерела довжиною n_d буде асимптотично (при великих n_d) дорівнювати $2^{n_d H(A)}$. Умовимося кодувати каналними символами тільки ці типові послідовності, а будь-яку нетипову послідовність передавати за допомогою однієї й тієї самої послідовності, що, мабуть, при декодуванні на прийомі й буде призводити до помилок.

Однак оскільки при $n_D \rightarrow \infty$ ймовірність появи нетипової послідовності буде прагнути до нуля, то це означає, що ймовірність помилки може бути зроблена як завгодно малою величиною.

Всі типові послідовності довжини n_D будемо кодувати послідовностями довжини n_K , складеними з каналних символів. Для забезпечення однозначності декодування (відсутності помилок у цьому випадку) необхідне виконання умови $2^{n_D H(A)} = m^{n_K} - 1$, звідки при великих значеннях n_D і n_K одержуємо вираз

$$\bar{n} = \frac{n_K}{n_D} \approx \frac{H(A)}{\log_2 m}.$$

Інтерпретація цієї теореми для заданих швидкостей джерела й каналу призводить також до співвідношення (5.43). Різниця полягає лише в тім, що можлива помилка з деякою малою ймовірністю δ , але кодування виявляється рівномірним, що не призводить до переповнення буферної пам'яті кодера.

Перейдемо до розгляду теорем кодування Шеннона в каналі з завадами.

Нехай повідомлення деякого джерела інформації передаються по каналу з завадами, заданому вхідним X і вихідним Y алфавітами, умовним розподілом $P(y|x)$ і швидкістю передачі ν_k . Будемо здійснювати кодування, зіставляючи з різними послідовностями символів джерела різні послідовності (комбінації) символів каналу довжини n . Назвемо останні дозволеними кодовими комбінаціями блокового коду, позначивши їх через $x_i^{[n]}$, $i = 1, 2, \dots, M$, де M – повне число таких комбінацій.

Якби в каналі не було завад, то прийняті комбінації збігалися б у точності з дозволеними кодовими комбінаціями. У каналі ж з завадами прийнята комбінація може стати з деякою ймовірністю кожною із m^n послідовностей довжини n , складених з вихідних каналних символів. Розіб'ємо всю множину таких

послідовностей Y^n на M непересічних підмножин $Y_1^n, Y_2^n, \dots, Y_M^n$ й установимо наступне правило декодування.

Якщо прийнята послідовність $y^{[n]} \in Y_i^n$, то приймається рішення про те, що передавалася кодова комбінація $x_i^{[n]}$. Таке подання будемо називати також вирішальною схемою. У цьому випадку можна визначити ймовірність правильного прийому (або правильного декодування) як імовірність $P\{y^{[n]} \in Y_i^n | x_i^{[n]}\}$ того, що при передачі комбінації $x_i^{[n]}$ прийнята комбінація $y^{[n]}$ попадає у відповідну підмножину Y_i^n . Коротко будемо записувати цю ймовірність як $P(Y_i^n | x_i^{[n]})$.

5.4.4. Теорема кодування в дискретному каналі з завадами

Якщо канал має пропускну спроможність C біт/символів і задані будь-які числа $\delta > 0$, $H < C$, то завжди знайдеться таке ціле число n_0 , що при будь-якому $n \geq n_0$ існує блоковий код $x_1^{[n]}, x_2^{[n]}, \dots, x_M^{[n]}$ довжиною n , що складається з $M = 2^{nH}$ комбінацій, і вирішальна схема $Y_1^n, Y_2^n, \dots, Y_M^n$, які забезпечують виконання нерівностей

$$P(Y_i^n | x_i^{[n]}) \geq 1 - \delta, \quad i = 1, 2, \dots, M. \quad (5.44)$$

Якщо $H > C$, то нерівність (5.44) при довільному δ не виконується, яким би не було велике n_0 . Дана теорема справедлива для досить широкого класу (хоча й не для всіх) каналів з пам'яттю, але ми її доведемо лише для каналу без пам'яті й для співпадаючих вхідних і вихідних алфавітів каналу, оскільки в протилежному випадку доведення помітно ускладнюється [20].

Доведення. Доведемо спочатку пряму частину теореми, тобто факт, що умова виразу (5.44) виконується при $M = 2^{nH}$, де $nH < C$. Нехай пропускну спроможність каналу досягається при

деякому розподілі ймовірностей вхідних символів $P(x_0) = \tilde{P}_0, \dots, P(x_{m-1}) = \tilde{P}_{m-1}$.

Будемо тоді називати послідовність довжиною n , складену із вхідних символів каналу, ε -типовою, якщо нерівність (5.36) (з очевидною заміною $a^{[n]}$ на $x^{[n]}$ й $H(A)$ на $H(X)$) виконується для неї при вхідному розподілі $\tilde{P}_0, \dots, \tilde{P}_{m-1}$.

Аналогічно будемо називати послідовність вихідних символів ε -умовною типовою послідовністю для ε -типової вхідної послідовності, якщо для неї виконується нерівність (5.41), у якій вхідний розподіл задається як $\tilde{P}_0, \dots, \tilde{P}_{m-1}$, а перехідні ймовірності визначаються каналом зв'язку. Припустимо, що вже побудовано блоковий код $x_1^{[n]}, x_2^{[n]}, \dots, x_M^{[n]}$ і вирішальну схему $Y_1^n, Y_2^n, \dots, Y_M^n$, що задовольняють такі умови:

1. Всі комбінації коду є ε -типовими послідовностями.
2. Множини $Y_i^n, i=1, 2, \dots, M$ містять всі ε -умовні типові послідовності для кожного з $x_i^{[n]}$, які не входять в об'єднання попередніх множин, тобто в $Y_1^n \cup Y_2^n \dots \cup Y_{i-1}^n$.

3. Імовірності правильного декодування при виборі даного коду й даної вирішальної схеми обмежені знизу нерівностями $P(Y_i^n | x_i^{[n]}) \geq 1 - \delta, i=1, 2, \dots, M$.

4. Обраний код не може бути розширений шляхом приєднання до нього ще однієї комбінації $x_{M+1}^{[n]}$ й множини Y_{M+1}^n , для яких виконуються умови 1-3.

При досить великому n число комбінацій такого коду, що задовольняє умови 1-4, виявляється більше за 2^{nH} , де $H^{nH} < C$ (див. рис. 5.4).

Зазначимо насамперед, що існує такий код і вирішальна схема хоча б для $M=1$. Дійсно, для цього досить вибрати одну ε -типову послідовність $x_i^{[n]}$ і таке n , щоб нерівність (5.44) виконувалася з імовірністю, більшою за $1 - \delta$.

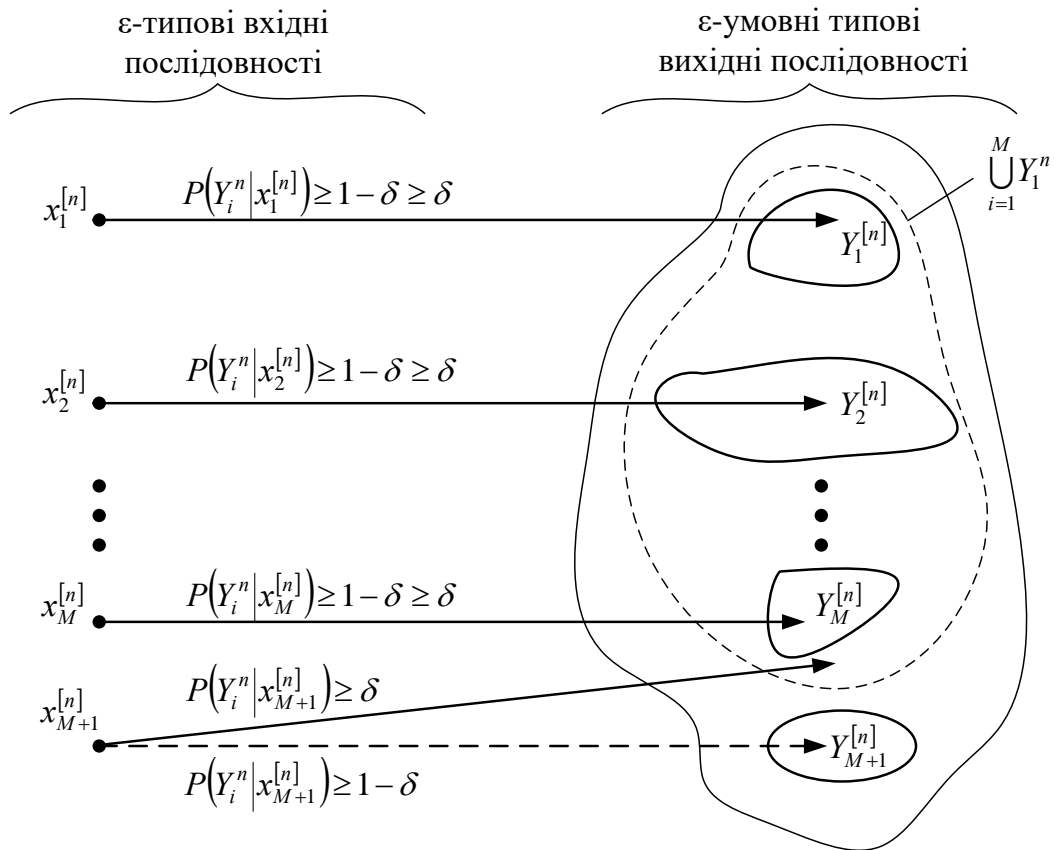


Рис. 5.4. Зв'язок між ϵ -типovими вхідними та ϵ -умовними типovими вихідними послідовностями

1. Для будь-якої ϵ -типovої (відносно $\tilde{P}_0, \dots, \tilde{P}_{m-1}$) вхідної послідовності $x^{[n]}$:

$$P\left(\bigcup_{i=1}^M Y_i^n | x^{[n]}\right) \geq \delta \quad \text{при} \quad P(x^{[n]}) \geq a.$$

$$2. P\left(\bigcup_{i=1}^M Y_i^n\right) \geq a\delta.$$

$$3. P(y - \epsilon\text{-умовна типова вихідна послідовність}) \leq 2^{-n(H(Y)-\epsilon)}.$$

$$4. \left| \bigcup_{i=1}^M Y_i^n \right| \geq a\delta 2^{n(H(Y)-\epsilon)}.$$

$$5. |Y_i^n| \geq 2^{n(H(Y|X)+\varepsilon)}.$$

$$6. M \geq \frac{a\delta 2^{n(H(Y)-\varepsilon)}}{2^{n(H(Y|X)+\varepsilon)}} = a\delta 2^{n(H(Y)-H(Y|X)-\varepsilon)}.$$

Нехай $x_{M-1}^{[n]}$ – будь-яка ε -типова послідовність, що не входить у код, тоді ймовірність того, що при передачі $x_{M-1}^{[n]}$ на виході з'явиться ε -типова умовна послідовність $Y^{[n]}$, що попадає в $Y_1^n \cup Y_2^n \dots \cup Y_M^n$, буде меншою за δ .

Дійсно, якщо це не так, то $x_{M-1}^{[n]}$ можна приєднати до коду й вибрати Y_{M-1}^n як множину вихідних ε -умовних типових послідовностей $x_{M-1}^{[n]}$, що не ввійшли в $Y_1^n \cup Y_2^n \dots \cup Y_M^n$. У цьому випадку $P(Y_{M-1}^n | x_{M-1}^{[n]}) \geq 1 - \delta$, тобто розширений код задовольняє умови 1-3, що суперечить умові 4.

Якщо ж передається кодова комбінація $x_i^{[n]}$, то з умов 1-3 видно, що ймовірність появи на виході ε -типової умовної послідовності, яка попадає в $Y_1^n \cup Y_2^n \dots \cup Y_M^n$ буде не меншою ніж $1 - \delta$ і, таким чином, при досить більших n (малих δ) також не менш ніж δ . Отже, яка б ε -типова послідовність відносного вхідного розподілу $\tilde{P}_0, \dots, \tilde{P}_{m-1}$ не передавалася, ймовірність того, що на виході каналу з'явиться послідовність, що належить $Y_1^n \cup Y_2^n \dots \cup Y_M^n$, буде не менше δ .

Відповідно до теореми ВАР зі збільшенням довжин блоків n ймовірність появи ε -типової вхідної послідовності наближається до одиниці й, таким чином, починаючи з якогось $n \geq n_0$ стає не менше постійної $a > 0$. Тому при досить великих n безумовна ймовірність появи на виході послідовності, що попадає в $Y_1^n \cup Y_2^n \dots \cup Y_M^n$, буде не менше за $a\delta$.

Для вихідних послідовностей буде також справедлива теорема ВАР, відповідно до якої при досить великих n

імовірність появи певної типової вихідної послідовності виявляється не більшою за $2^{n(H(Y))}$.

Множина $Y_1^n \cup Y_2^n \dots \cup Y_M^n$ складається саме з ε -типових вихідних послідовностей, тому число елементів $\left| \bigcup_{i=1}^M Y_i^{[n]} \right|$ не менше ніж $\frac{a\delta}{2^{n(H(Y))}} = a\delta 2^{n(H(Y))}$.

З іншого боку, число $\left| Y_i^{[n]} \right|$ ε -умовних типових вихідних послідовностей при заданій ε -типовій вхідній послідовності не перевершує $2^{n(H(Y|X))}$. Звідси випливає, що число M повинне задовольняти нерівність $M > \frac{a\delta 2^{n(H(Y))}}{2^{n(H(Y|X))}} = a\delta \cdot 2^{n(H(Y)-H(Y|X)-2)}$.

Оскільки вхідний розподіл $\tilde{P}_0, \dots, \tilde{P}_{m-1}$ реалізує пропускну спроможність каналу, то $H(Y) - H(Y|X) = C$ й $M > a\delta 2^{n(C-2)} = 2^{n\left(C-2 - \frac{\log_2 a\delta}{n}\right)}$. З огляду на те, що a й δ фіксовано, а ε може бути обране довільно малим, одержимо ствердження прямої теореми. Доведемо зворотну теорему. Для цього розглянемо код $x_1^{[n]}, x_2^{[n]}, \dots, x_M^{[n]}$ і вирішальну схему $Y_1^n, Y_2^n, \dots, Y_M^n$, для яких

$$P\left(Y_j^n \mid x_i^{[n]}\right) = \begin{cases} \frac{p_{n\delta}}{M-1}, & i \neq j, \\ q = 1 - p_{n\delta}, & i = j, \end{cases} \quad (5.45)$$

де $p_{n\delta}$ – імовірність помилки декодування блоку символів. (Якщо умова (5.45) не виконується, то завжди можна змінити вирішальну схему так, щоб зменшити максимальну ймовірність помилки й виконати цю умову для деякої величини $p'_{n\delta}$).

Тепер код разом з вирішальною схемою можна розглядати як розширений m СК із імовірностями переходів (вираз (5.45)) і обсягами вхідного й вихідного алфавітів, що дорівнюють M . Тоді

кількість інформації, переданої по такому каналу, не може бути більшою за його пропускну спроможність,

$$\log M + q \log q + p_{n0} \log \frac{p_{n0}}{M-1}.$$

Якщо цю кількість інформації віднести до одного символу n , то вона, мабуть, не зможе перевершити пропускну спроможність C вихідного каналу, тому що будь-які перетворення не збільшують кількості інформації (див. властивість 8 кількості інформації).

Тому одержуємо нерівність

$$\frac{1}{n} (\log M + q \log q + p_{n0} \log \frac{p_{n0}}{M-1}) < C. \quad (5.46)$$

Припустимо, що $H > C$, тобто $M = 2^{nH} > 2^{nC}$, тоді, підставляючи це значення M у вираз (5.46), одержуємо нерівність

$$\frac{1}{n} q \log q + \frac{p_{n0}}{n} \log p_{n0} - \frac{p_{n0}}{n} \log(2^{nH} - 1) < C - H < 0. \quad (5.47)$$

Звідси видно, що при $n \rightarrow \infty$ ліва частина виразу (5.47) завжди наближається до нуля, якщо $p_{n0} \rightarrow 0$. Однак це неможливо, оскільки права частина обмежена негативною постійною. Отже, якщо $M = 2^{nH}$, де $H > C$, то досягнення як завгодно малої помилки декодування ($p_{n0} \rightarrow 0$) шляхом збільшення довжин блоків $n \rightarrow \infty$ при будь-якому виборі коду й *вирішальної схеми* виявляється неможливим.

5.4.5. Інтерпретація теореми кодування

Розглянемо інтерпретацію теореми кодування. Нехай є дискретне джерело з обсягом алфавіту K і розподілом імовірностей його символів, що ми не знаємо або не хочемо використовувати при кодуванні повідомлень джерела. Тоді,

порівнюючи з кожною кодовою комбінацією довжини \tilde{n} послідовність символів джерела довжини n , що задовольняють очевидну умову взаємної однозначності кодування й декодування, запишемо

$$K^{\tilde{n}} = M = 2^{nH}$$

і за умови відсутності затримок у часі $\tilde{n}T_{дж} = nT_K$, де $T_{дж}$ й T_K – тривалості символів джерела й каналу відповідно, одержуємо

$$v_{дж} = \frac{1}{T_{дж}} = \frac{\tilde{n}}{n} v_K = \frac{v_K H}{\log_2 K}.$$

Як виходить з доведеної теореми кодування, $p_{nd} \rightarrow 0$ при $n \rightarrow \infty$, коли $H < C$, і тому одержуємо умову на максимально можливу швидкість передачі символів у такому вигляді:

$$v_{дж} < \frac{v_K C}{\log_2 K} = \frac{C'}{\log_2 K}. \quad (5.48)$$

В окремому випадку двійкового джерела одержуємо з виразу (5.48) швидкість передачі символів, біт/с,

$$v_{дж} < C'.$$

Припустимо, що буде враховуватися статистика джерела повідомлень, тобто виробляється додатково й кодування джерела повідомлень. Тоді, як виходить з теореми 1 про кодування джерела, середня довжина послідовності каналних символів, що припадає на один символ джерела повідомлень, буде визначатися співвідношенням (5.42).

Припустимо, що в нашому випадку символами каналу є кодові комбінації довжини n , число яких дорівнює $M = 2^{nH} = m$. Тоді у виразі (5.42) необхідно покласти $\log_2 m = nH$. Умова відсутності затримок у часі набуде в цьому випадку такого вигляду (знехтуючи величиною ε у виразі (5.42)):

$$T_{дж} = \bar{n}nT_k = \frac{H(A)n}{nH}T_k = \frac{H(A)}{H}T_k.$$

Оскільки для забезпечення $p_{nd} \rightarrow 0$ при $n \rightarrow \infty$ згідно з доведеною теоремою з кодуванням у каналі з завадами необхідно забезпечити виконання нерівності $H < C$, то з попереднього виразу знаходимо

$$\frac{H(A)}{T_{дж}} = H'(A) < \frac{C}{T_k} = C'. \quad (5.49)$$

Нерівність (5.49) становить *основну теорему Шеннона*, що формулюється :

Якщо продуктивність джерела $H'(A)$ менше за пропускну спроможність C' в одиницю часу дискретного каналу з завадами, то при будь-якому $\delta > 0$ існує спосіб кодування й декодування джерела й каналу, при якому повідомлення передаються одержувачеві з імовірністю помилки меншою за δ і в середньому без зростаючих затримок у часі. Якщо $H'(A) > C'$, то такого способу кодування не існує.

Зазначимо, що хоча нерівність (5.49) і є більш сильною за нерівність (5.48), але це не означає, що завжди варто використовувати як каналне кодування, так і кодування джерела. Останнє для досягнення помітного ефекту може виявитися нереалізовано складним.

Важливою особливістю теорем кодування Шеннона є та обставина, що вони носять характер *теорем існування* й майже нічого не говорять про практичні способи реалізації процедур кодування й декодування. (Цим питанням буде спеціально присвячена наступна глава.)

Дотепер у даній главі ми всюди розглядали так званий дискретний канал зв'язку з завадами. У дійсності таких каналів не існує й можна говорити лише про моделі відображення безперервного каналу зв'язку на дискретний. Власне кажучи це означає, що ми в рамках розгляду даного дискретного каналу фіксуємо певний спосіб модуляції й демодуляції.

Якщо нам додатково відомо, який саме обраний спосіб модуляції й демодуляції, то може бути відома й імовірність помилки $p(h^2)$ як функція параметра $h^2 = \frac{P_c T_k}{N_0}$. Припущення про

те, що при фіксованих значеннях P_c й N_0 ми можемо змінювати тривалість канальних символів T_k , що еквівалентно зміні канальної швидкості передачі v_k , відкриває нові можливості для оптимізації системи зв'язку.

Так можна порушити питання про оптимізацію величини v_k з метою забезпечення найбільшої пропускної спроможності каналу зв'язку в одиницю часу C' , що відповідно до теореми кодування забезпечить найбільшу швидкість передачі v_k повідомлень джерела при як завгодно високій правильності прийому.

Це завдання для 2СК зводиться до знаходження максимуму функції:

$$\max_{v_k} C'(v_k) = \max_{v_k} \{v_k [1 + p(v_k) \log p(v_k) + (1 - p(v_k)) \times \log(1 - p(v_k))]\}, \quad (5.50)$$

де $p(v_k) = p\left(\frac{P_c}{h^2 N_0}\right)$, оскільки $h^2 = \frac{P_c T_k}{N_0} = \frac{P_c}{v_k N_0}$.

Приведемо вираз (5.50) до більш зручного вигляду для знаходження максимуму

$$\max_{v_k} C'(v_k) = \frac{P_c}{N_0} \max_{h^2} \left\{ \frac{1}{h^2} \left(1 + p(h^2) \log p(h^2) + (1 - p(h^2)) \log(1 - p(h^2)) \right) \right\}. \quad (5.51)$$

У роботі [47] показано, що при оптимальному когерентному прийомі двійкових сигналів у детермінованому каналі із БГШ імовірність помилки як функція h^2 визначається рівнянням

$$p = \Phi(\sqrt{\alpha h^2}), \quad (5.52)$$

$$\text{де } \alpha = \begin{cases} 2, & \text{для протилежних сигналів;} \\ 1, & \text{для ортогональних сигналів рівної енергії;} \\ 1/2, & \text{для системи сигналів з пасивною паузою;} \end{cases}$$

$\Phi(x)$ – інтеграл імовірності.

При оптимальному некогерентному прийомі в тім самому каналі вираз (5.52) має вигляд

$$p = \frac{1}{2} e^{-\beta h^2}, \quad (5.53)$$

$$\text{де } \beta = \begin{cases} 1, & \text{для ВФМ сигналів;} \\ 1/2, & \text{для сигналів рівної енергії, ортогональних у підсиленому сенсі;} \\ 1/4, & \text{для системи сигналів з пасивною паузою.} \end{cases}$$

Підставляючи вирази (5.52) і (5.53) у вираз (5.51) і знаходячи екстремум функції звичайними аналітичними й чисельними методами [28], одержуємо, що для когерентного прийому максимум досягається при $\nu_k \rightarrow \infty$ і дорівнює $0,46 \frac{\alpha P_c}{N_0}$.

При некогерентному прийомі максимум C' досягається при $T_k = \frac{1}{\nu_k} \approx 1,551 \frac{N_0}{\alpha P_c}$ і дорівнює $0,33\beta \frac{P_c}{N_0}$.

Таким чином, максимум при когерентному прийомі двійкових сигналів досягається при необмеженій смузі пропускання приймача F , оскільки $F \rightarrow \infty$ при $\nu_k \rightarrow \infty$, у той час як при некогерентному прийомі максимум C' реалізується при обмеженій смузі частот каналу зв'язку. (Однак, що при будь-якій швидкості передачі ν_k і тих самих значеннях P_c й N_0 когерентний прийом завжди забезпечить більше значення C' , ніж некогерентний.)

Допустимість зміни тривалостей каналних символів при фіксованому способі модуляції й демодуляції призводить до можливості незвичайної інтерпретації теореми кодування.

Припустимо, що інформаційна швидкість передачі двійкового джерела $\nu_{дж}$, біт/с, задана так само, як P_c й N_0 . Виникає питання, чи можна за цих умов забезпечити $p_{нд} \rightarrow 0$ при якому-небудь, нехай як завгодно складному кодуванні (тобто при $n \rightarrow \infty$) і найкращому виборі тривалості канального символу?

Як було показано раніше, максимальне значення C' при оптимальному когерентному прийомі двійкових сигналів досягається при $\nu_k \rightarrow \infty$ й дорівнює $0,46 \frac{\alpha P_c}{N_0}$. Оскільки

відповідно до теореми кодування $\nu_{дж} < C'$, то $T_{дж} > \frac{1}{C'}$ й, отже, маємо

$$h^2 = \frac{P_c T_{дж}}{N_0} > \frac{P_c}{N_0 C'}, \quad (5.54)$$

де h^2 – відношення сигнал-шум при заданій інформаційній швидкості передачі.

Підставляючи у вираз (5.54) максимальне значення $C' = 0,46 \frac{\alpha P_c}{N_0}$, одержуємо, що $h^2 > \frac{2,17}{\alpha}$, а тоді за формулою (5.52)

максимально припустима ймовірність помилки (для заданої інформаційної швидкості передачі), при якій кодування ще може забезпечити як завгодно високу правильність прийому, дорівнює

$$p = \Phi(\sqrt{2,17}) \approx 0,072. \quad (5.55)$$

Якщо ж виявилось, що для необхідної інформаційної швидкості передачі ймовірність помилки дорівнює або більше за цю величину, то будь-яке кодування з метою підвищення завадостійкості виявляється марним. (Для підвищення надійності необхідно або змінити вид модуляції, або поліпшувати параметри каналу P_c й N_0 .)

Виконуючи аналогічні викладення для оптимального некогерентного прийому двійкових ВФМ сигналів, можна

показати, що мінімальна величина h^2 дорівнює $3/\beta$, а ймовірність помилки 0,025.

5.4.6. Енергетичний вигравш кодування

У реальних системах зв'язку звичайно задаються обмеженими значеннями (p_3) імовірності помилки, не вимагаючи виконання умови $p_3 \rightarrow 0$. Тому при заданій інформаційній швидкості передачі $\nu_{дж}$ й фіксованому параметрі N_0 завжди можна забезпечити необхідну правильність прийому за допомогою вибору необхідного значення потужності сигналу P_c .

Використання кодування може дозволити знизити необхідне значення потужності сигналу до деякої величини $P_{ск}$ при збереженні тих самих значень N_0 й $\nu_{дж}$. Виражене в децибелах відношення цих двох потужностей називається енергетичним вигравшем кодування (ЕВК) у порівнянні з використанням якогонебудь методу модуляції-демодуляції без кодування, тобто

$$\eta = 10 \lg \frac{P_c}{P_{ск}}, \quad (5.56)$$

де P_c – потужність сигналу в системі зв'язку без кодування;

$P_{ск}$ – потужність сигналу в тій самій системі зв'язку, але з кодуванням.

Тут ми оцінюємо ЕВК при необмеженій складності кодування, тобто при $n \rightarrow \infty$, тоді як у наступній главі ми будемо оцінювати його для кодів з кінцевими довжинами блоків. Очевидно, що ЕВК в останньому випадку завжди буде менше, ніж при $n \rightarrow \infty$. Тому тут ми одержимо найкращі можливі результати.

Визначимо швидкість блокового коду $x_1^{[n]}, x_2^{[n]}, \dots, x_M^{[n]}$ як відношення логарифма числа дозволених кодових комбінацій M до логарифма будь-якого числа комбінацій довжини n , утворених m -ковим кодом

$$R = \frac{\log_2 M}{\log_2 m^n}.$$

Для забезпечення тієї самої заданої інформаційної швидкості передачі v_k тривалості канальних символів за відсутності кодування T_k і за наявності кодування T_{kk} будуть зв'язані співвідношенням

$$T_{kk} = RT_k.$$

Тому для розрахунку ЕВК при використанні коду із заданою швидкістю (у тому числі й в асимптотичному випадку, коли $R = \lim_n \frac{\log_2 M}{\log_2 m^n}$) співвідношення (5.56) може бути подане в такому вигляді:

$$\eta = 10 \lg \left(\frac{h_c^2}{h_{ck}^2} R \right), \quad (5.57)$$

де h_c^2 – відношення сигнал-шум у каналі зв'язку, що забезпечує необхідну величину p_3 імовірності помилкового прийому символу без кодування;

h_{ck}^2 – відношення сигнал-шум у каналі зв'язку, що забезпечує необхідну величину p_3 при використанні даного коду.

Величина h_c^2 в цьому співвідношенні може бути знайдена як рішення рівняння $p_3 = p(h_c^2)$, де $p(h_c^2)$ – імовірність помилки для заданого способу модуляції-демодуляції й моделі каналу зв'язку як функції параметра h_c^2 . Величина h_{ck}^2 відповідно до теореми кодування Шеннона може бути знайдена як рішення рівняння

$$R = \frac{C(h_{ck}^2)}{\log_2 m},$$

де $C(h_{ck}^2)$ – пропускна спроможність дискретного каналу зв'язку, представлена як функція параметра h_{ck}^2 .

В окремому випадку для 2СК одержуємо рівняння

$$R = 1 + p(h_{ск}^2) \log_2 p(h_{ск}^2) + (1 - p(h_{ск}^2)) \log_2 (1 - p(h_{ск}^2)). \quad (5.58)$$

Якщо швидкість коду не задана, то ми можемо її оптимізувати, домагаючись максимізації η у виразі (5.57) при виконанні умови (5.58).

Це у свою чергу реалізується при максимізації за R величини

$$\frac{R}{h_{ск}^2} = \frac{C(h_{ск}^2)}{h_{ск}^2}.$$

Як було показано раніше, ця величина при оптимальному когерентному прийомі досягає максимуму, коли $\nu_k \rightarrow \infty$, а отже, при $R \rightarrow 0$ дорівнює $0,46\alpha$.

При оптимальному некогерентному прийомі ця величина максимізується при $T_{opt} = 1,551 \frac{N_0}{h_{ск}^2}$, що відповідає оптимальному

значенню $h_{ск\ opt}^2 = 1,551/\beta$ й "оптимальній" імовірності помилки символу в каналі зв'язку $p_{opt} = 0,5 e^{-\beta h_{ск\ opt}^2} = 0,5 e^{-1,551} \approx 0,106$.

При оптимальному значенні ймовірності помилки пропускна спроможність 2СК приблизно дорівнює $0,51$, що й збігається з оптимальною швидкістю коду. У цьому випадку, як було показано раніше, максимальне значення $R/h_{ск}^2$ дорівнює $0,33\beta$.

Так, якщо задатися величиною $p_3 = 10^{-6}$, то ЕВК при когерентному прийомі й оптимальному виборі швидкості коду ($R = 0$) складе $10,2$ дБ, а при оптимальному некогерентному прийомі ЕВК при оптимальному виборі швидкості буде дорівнювати приблизно $6,3$ дБ.

Знаючи величини потенційних вигравів при оптимальному кодуванні, ми можемо зробити попередній висновок про те, чи доцільно намагатися застосовувати регулярне кодування, тобто "обмінювати" ці децибели на ускладнення устаткування або програмного забезпечення. Зазначимо, що вибір фіксованої

швидкості коду R дозволяє контролювати розширення смуги частот сигналів у порівнянні з некодованою передачею повідомлень.

5.5. Потенційні можливості безперервних каналів зв'язку при передачі дискретних повідомлень

5.5.1. Кодування й декодування в безперервному каналі зв'язку

Переходячи від дискретного каналу зв'язку до безперервного, ми фактично відмовляємося від деяких обмежень, а точніше кажучи – від вибору фіксованих способів модуляції й демодуляції, які визначають умовні ймовірності переходів $p(y|x)$ у дискретному каналі зв'язку з завадами.

Тому ми вправі очікувати збільшення пропускної спроможності безперервного каналу в порівнянні з будь-яким відображенням його дискретним каналом. Однак і при завданні безперервного каналу також повинні існувати певні обмеження, які накладаються на вхідні і вихідні безперервні сигнали.

Визначимо формально безперервний канал зв'язку з завадами як пару просторів $S^T = \{s(t)\}$ й $Z^T = \{z(t)\}$ відповідно вхідних $s(t)$, $t \in (0, T)$, вихідних $z(t)$, $t \in (0, T)$, безперервних сигналів і заданого на них умовного імовірнісного розподілу $P(z/s)$, $z \in Z^T$, $s \in S^T$. Простим, але досить важливим окремим випадком безперервного каналу зв'язку є безперервний канал з адитивним шумом, для якого простір вихідних сигналів задається так

$$z(t) = s(t) + n(t), \quad (5.59)$$

де $n(t)$ – випадковий сигнал (*адитивна завада*), що не залежить від вхідного сигналу $s(t)$.

Для повного опису такого каналу досить задати обмеження на припустиму множину вхідних сигналів S^T й імовірнісні розподіли $P(n)$ для адитивної завади. Існують різні способи завдання обмежень на вхідні сигнали.

Найчастіше для цього використовуються такі поняття, як *середня* або *пікова потужності сигналів*, смуга частот, форма спектра сигналів. У цьому пункті ми так само, як й у попередніх пунктах цієї глави, вивчаємо дискретні джерела повідомлень, і тому під *кодуванням* у безперервному каналі будемо розуміти зіставлення послідовностей, складених із символів джерела повідомлень, з безперервними сигналами тривалості T , що належать простору S^T , а під *декодуванням* у безперервному каналі – зіставлення будь-якого безперервного сигналу тривалості T , що належить простору Z^T , з послідовностями символів джерела.

Будемо множину таких обраних безперервних сигналів $s_1(t), \dots, s_M(t)$ називати T -кодом, а розподіл Z^T на M підмножин $z_1^T, z_2^T, \dots, z_M^T$, що відповідає декодуванню, – *вирішальною схемою*. Очевидно, що завдання T -коду, вирішальній схеми й каналу зв'язку буде повністю визначати ймовірність помилкового декодування p_{nd} . У цьому випадку процедури кодування й декодування називають також модуляцією й демодуляцією.

Так само, як й у випадку дискретного каналу, завдання кодування й декодування полягає: по-перше, в узгодженні різних алфавітів джерела й каналу; по-друге, у забезпеченні максимально можливої швидкості передачі повідомлень джерела при одночасному досягненні як завгодно високої правильності прийому, яка розуміється як ймовірність помилкового декодування p_{nd} .

5.5.2. Пропускна спроможність безперервного каналу зв'язку

5.5.2.1. Взаємна інформація пари відліків для адитивної гаусівської завади

Розглянемо спочатку канал зв'язку із сигналами, що мають дискретний час і безперервну амплітуду на вході й виході.

Припустимо, що кожен такий «імпульс» передається незалежно від всіх інших, тобто в каналі відсутня пам'ять і задана умовна щільність ймовірності $w(z/s)$, де $z \in Z$, $s \in S$, а Z й S –

простори припустимих значень амплітуд відповідно вхідних і вихідних сигналів.

Нехай відома також щільність імовірності $w(s)$, $s \in S$ вхідних амплітуд. Тоді такий безперервний канал можна перетворити до дискретного, здійснюючи квантування множин S й Z з інтервалами Δs й Δz відповідно. У цьому випадку вхідний розподіл імовірностей буде мати вигляд

$$P(s_i) = P\{s_i \leq S \leq s_i + \Delta s\} \approx w(s_i)\Delta s, \quad i = 0, 1, \dots, \quad (5.60)$$

а умовні ймовірності переходів вхідних символів у вихідні

$$P(z_j/s_i) = P\{z_j \leq Z \leq z_j + \Delta z | s_i\} \approx w(z_j | s_i)\Delta z. \quad (5.61)$$

Спільна ймовірність появи вхідних s_i і вихідних z_j символів у такому дискретизованому каналі буде дорівнювати

$$P(s_i, z_j) = P\{s_i \leq S \leq s_i + \Delta s, z_j \leq Z \leq z_j + \Delta z\} \approx w(s_i, z_j)\Delta s\Delta z. \quad (5.62)$$

Для розрахунку кількості інформації $I_{\Delta s, \Delta z}(S, Z)$, переданої по такому каналу, скористаємося отриманим раніше виразом (5.21), що перетворимо так:

$$\begin{aligned} I_{\Delta s, \Delta z}(S, Z) &= \sum_i \sum_j P(s_i, z_j) \log \frac{P(s_i, z_j)}{P(s_i)P(z_j)} \approx \\ &\approx \sum_i \sum_j w(s_i, z_j)\Delta s\Delta z \log \frac{w(s_i, z_j)}{w(s_i)w(z_j)}, \end{aligned} \quad (5.63)$$

причому підсумовування у виразі (5.63) поширюється на всі $s_i \in S$, $z_j \in Z$. Спрямувавши Δs й Δz у вираз (5.63) до нуля, перейдемо від дискретного за амплітудою каналу до каналу з безперервною амплітудою. Тоді кількість інформації, переданої по такому каналу, буде

$$\begin{aligned}
I(S,Z) &= \lim_{\substack{\Delta s \rightarrow \infty \\ \Delta z \rightarrow \infty}} I_{\Delta s, \Delta z}(S,Z) = \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(s,z) \log \left(\frac{w(s,z)}{w(s)w(z)} \right) ds dz = \\
&= \int_{-\infty}^{\infty} w(s) \log \frac{1}{w(s)} ds - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(s,z) \log \frac{1}{w(s|z)} ds dz.
\end{aligned} \tag{5.64}$$

Для кількості інформації $I(S,Z)$, переданої по такому каналу з безперервною амплітудою, будуть справедливі такі властивості:

1. $I(S,Z) \geq 0$, причому $I(S,Z)=0$ тоді і тільки тоді, коли вхід і вихід каналу статистично незалежні, тобто $w(s,z)=w(s)w(z)$. (Ця властивість випливає з виразу (5.64) і властивостей 2 й 3 кількості інформації для дискретного каналу.)

2. $I(S,Z)=I(Z,S)$, що дозволяє назвати $I(S,Z)$ взаємною інформацією пари випадкових величин. (Випливає з виразу (5.64).)

3. $I(S,Z)=\infty$, якщо завади в каналі відсутні, тобто коли $z=s$.

Дійсно, за відсутності завад $I(S,Z) = \lim_{\Delta s \rightarrow 0} H_{\Delta s}(S)$,

де $H_{\Delta s}(S)$ – ентропія джерела, отриманого після квантування вхідних сигналів з інтервалом Δs , тобто

$$\begin{aligned}
H_{\Delta s}(S) &= \sum_i w(s_i) \Delta s \log \frac{1}{w(s_i) \Delta s} = \\
&= \sum_i w(s_i) \log \frac{1}{w(s_i)} \Delta s + \log \frac{1}{\Delta s} \sum_i w(s_i) \Delta s.
\end{aligned} \tag{5.65}$$

Спрямовуючи Δs до нуля, одержуємо, що перший член у правій частині виразу (5.65) наближається до кінцевої величини

$$h(S) = \int_{-\infty}^{+\infty} w(s) \log \frac{1}{w(s)} ds, \tag{5.66}$$

а другий завжди наближається до нескінченності незалежно від виду щільності розподілу ймовірностей $w(s)$ вхідного сигналу.

Наочний сенс цієї властивості полягає в тому, що канал зв'язку, який дозволяє як завгодно точно передавати нескінченну множину значень амплітуди, передає нескінченну кількість інформації. Дійсно, яким би не було великим значення ентропії $H(A)$ деякого дискретного джерела повідомлень, ці повідомлення завжди можна закодувати безперервним сигналом у вигляді одного імпульсу й *абсолютно точно* передати по безперервному каналу зв'язку, якщо в ньому повністю відсутні завади.

5.5.2.2. Умовна диференціальна ентропія

Величина $h(S)$, яка обумовлена виразом (5.66), називається *диференціальною ентропією*. Вона збігається з першим доданком у виразі (5.64). Другий доданок називається *умовною диференціальною ентропією* й позначається $h(S/Z)$. Тому кількість інформації, яка передана по безперервному каналу, дорівнює різниці цих диференціальних ентропій:

$$I(S,Z)=h(S)-h(S/Z)=h(Z)-h(Z/S), \quad (5.67)$$

причому остання рівність у виразі (5.67) впливає з властивості 2 кількості інформації, де $h(Z)$ і $h(Z/S)$ – диференціальна ентропія виходу каналу й диференціальна умовна ентропія виходу при відомому вході відповідно.

Розглянемо окремий випадок безперервного за амплітудою каналу з дискретним часом – канал з адитивним безперервним шумом, для якого

$$z=s+n, \quad (5.68)$$

де s – реалізація випадкової величини S ;

n – реалізація випадкової величини N , що не залежить від S .

Використовуючи вираз (5.67), легко показати, що для розглянутого каналу

$$I(S,Z)=h(Z)-h(N), \quad (5.69)$$

де $h(N)$ – диференціальна ентропія адитивної завади.

Знайдемо диференціальну ентропію гаусівського адитивного шуму з нульовим середнім і дисперсією $\sigma^2 = P_{uu}$ із виразу

$$\begin{aligned} h(N) &= \int w(n) \log \left(\sqrt{2\pi P_{uu}} e^{\frac{n^2}{2P_{uu}}} \right) dn = \\ &= \log(\sqrt{2\pi P_{uu}}) \int w(n) dn + \frac{\log e}{2P_{uu}} \int n^2 w(n) dn = \quad (5.70) \\ &= \log(\sqrt{2\pi P_{uu}}) + \frac{\log e}{2} = \log(\sqrt{2\pi e P_{uu}}). \end{aligned}$$

Диференціальна ентропія довільної випадкової величини N із заданою дисперсією σ^2 не може перевершувати диференціальної ентропії гаусівської випадкової величини з тією самою дисперсією, тобто

$$h(N) = \int_N w(n) \log \frac{1}{w(n)} dn \leq \log \sqrt{2\pi e P_{uu}}. \quad (5.71)$$

Для доказу розглянемо інтеграл

$$\int_N w(n) \log \left(\sqrt{2\pi \sigma^2} e^{\frac{n^2}{2\sigma^2}} \right) dn = \log(\sqrt{2\pi \sigma^2}).$$

Отже,

$$\begin{aligned} h(N) - \log(\sqrt{2\pi \sigma^2}) &= \int_N w(n) \left[\log \frac{1}{w(n)} - \log \left(\sqrt{2\pi \sigma^2} e^{\frac{n^2}{2\sigma^2}} \right) \right] dn = \\ &= \int_N w(n) \log \left(\frac{e^{\frac{n^2}{2\sigma^2}}}{w(n) \sqrt{2\pi \sigma^2}} \right) dn. \quad (5.72) \end{aligned}$$

Використаємо нерівність $\ln x \leq x - 1$, причому точна рівність буде досягтися тільки при $x=1$. Звідси, переходячи до двійкових логарифмів, одержуємо $\log_2 x \leq \log_2(x - 1)$.

Тому

$$\log_2 \left(\frac{e^{-\frac{n^2}{2\sigma^2}}}{w(n)\sqrt{2\pi\sigma^2}} \right) \leq \log_2 e \cdot \left(\frac{e^{-\frac{n^2}{2\sigma^2}}}{w(n)\sqrt{2\pi\sigma^2}} - 1 \right), \quad (5.73)$$

причому рівність має місце, якщо тільки

$$w(n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{n^2}{2\sigma^2}}.$$

Підставляючи вираз (5.73) у вираз (5.72), одержуємо

$$\begin{aligned} h(N) - \log(\sqrt{2\pi e\sigma^2}) &\leq \\ &\leq \left[\int_{-\infty}^{+\infty} w(n) \frac{e^{-\frac{n^2}{2\sigma^2}}}{w(n)\sqrt{2\pi\sigma^2}} dn - \int_{-\infty}^{+\infty} w(n) dn \right] \log_2 e = \\ &= \left[\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-\frac{n^2}{2\sigma^2}} dn - \int_{-\infty}^{+\infty} w(n) dn \right] \log_2 e = 0, \end{aligned}$$

звідки й випливає вираз (5.71).

5.5.2.3. Пропускна спроможність безперервного каналу зв'язку

Раніше вже зазначалося, що на простір припустимих вхідних сигналів повинні бути накладені деякі обмеження.

Розглянемо найбільш важливий окремий випадок, коли обмежена середня потужність (дисперсія) сигналу P_c . Тоді для каналів без пам'яті з дискретним часом це обмеження вимагає використання тільки таких вхідних щільностей імовірностей $w(s)$, для яких виконується умова

$$\int_{-\infty}^{+\infty} w(s)(s - \bar{S})^2 ds \leq P_c, \quad (5.74)$$

$$\text{де } \bar{S} = \int_{-\infty}^{+\infty} sw(s) ds.$$

Будемо називати пропускною спроможністю C такого каналу максимальне значення кількості інформації $I(S, Z)$ по всіх щільностях імовірності $w(s)$, що задовольняють вираз (5.74), тобто

$$C = \max_{w(s) \in W} \{I(S, Z)\}, \quad (5.75)$$

де W – множина щільностей імовірності, що задовольняють вираз (5.74).

Розрахуємо пропускну спроможність для каналу з дискретним часом і адитивним БГШ потужності P_w .

Використовуючи вирази (5.69) і (5.70), одержуємо

$$C = \max_{w(s) \in W} \left\{ h(z) - \log \sqrt{2\pi e P_w} \right\}. \quad (5.76)$$

Для даного каналу дисперсія виходу дорівнює сумі дисперсій входу й адитивного шуму, тобто

$$D\{z\} = D\{s\} + D\{n\} = P_c + P_w.$$

Тому відповідно до доведеної вище властивості одержуємо, що диференціальна ентропія виходу $h(Z)$ буде максимальна для гаусівського розподілу Z , а отже, і для гаусівського розподілу S і вона дорівнює правій частини виразу (5.71), якщо P_w замінити на

$\sigma_z^2 = P_c + P_u$. Підставляючи це значення у вираз (5.76), знаходимо, що

$$C = \log \sqrt{2\pi(P_c + P_u)} - \log \sqrt{2\pi P_u} = \frac{1}{2} \log \left(1 + \frac{P_c}{P_u} \right). \quad (5.77)$$

Вираз (5.77) дає значення пропускної спроможності каналу, що має при двійковому логарифмі розмірність біт/відлік (імпульс). Якщо швидкість видачі відліків (імпульсів) у секунду дорівнює ν_K , то пропускна спроможність в одиницю часу буде визначатися співвідношенням

$$C' = \nu_K C = \frac{\nu_K}{2} \log \left(1 + \frac{P_c}{P_u} \right). \quad (5.78)$$

Отриманий вираз дозволяє легко перейти до пропускної спроможності безперервного каналу з безперервним часом, у якого вхідні сигнали мають обмежену смугу частот F й обмежену середню потужність P_c . Крім того, будемо припускати, що завадою в ньому є "квазібілий" шум зі спектральною щільністю N_0 у смузі частот F , тобто гаусівський шум з рівномірним спектром і середньою потужністю $P_u = N_0 F$.

Оскільки корисні сигнали обмежені смугою частот F , то на виході такого каналу зв'язку можна поставити ідеальний фільтр, що пропускає тільки частоти в цій смузі, не втративши при цьому ніякої інформації. За теоремою відліків Котельникова [29], сигнали на вході й виході такого каналу будуть повністю визначатися відліковими значеннями в точках $t_k = k\Delta$, де $\Delta = 1/2F$. Отже, вся інформація, яка передається по такому каналу, буде міститися в цих відлікових значеннях.

Оскільки енергетичний спектр завади на виході рівномірний у смузі частот F , то відліки завади виявляються статистично незалежними й завдання зводиться до розрахунку пропускної спроможності безперервного каналу без пам'яті з дискретним

часом. Використовуючи отримане для цього співвідношення (5.78) при $\nu_k = \frac{1}{\Delta} = 2F$, знаходимо

$$C' = F \log \left(1 + \frac{P_c}{P_{ш}} \right) = F \log \left(1 + \frac{P_c}{N_0 F} \right). \quad (5.79)$$

Співвідношення (5.79) відоме як *формула Шеннона* для пропускної спроможності безперервного гаусівського каналу з обмеженою смугою частот й обмеженою середньою потужністю сигналу.

Проведемо аналіз формули Шеннона (5.79). Якщо $F = \text{const}$, а $P_c/P_{ш}$ зростає, то, як видно із цього співвідношення, C' буде також зростати, але її зростання виявляється досить повільним, тому що воно підкоряється логарифмічному закону. Тому якщо, наприклад, при смузі частот 100 Гц і відношенні $P_c/P_{ш} = 10^3 = 30$ дБ пропускна спроможність $C' \approx 1000$ біт/с, а її потрібно збільшити приблизно у два рази при збереженні колишньої смуги частот 100 Гц, то цього можна досягти, лише збільшивши відношення сигнал/шум до $10^6 = 60$ дБ.

Розглянемо тепер залежність пропускної спроможності каналу C' від смуги частот F при фіксованих значеннях P_c й N_0 . Графік залежності нормованої пропускної спроможності C'/P_c від $\frac{FN_0}{P_c}$ показаний на рис. 5.5. $C'(F)$, як впливає з виразу (5.79)

– це монотонно зростаюча функція, що при $F \rightarrow \infty$ асимптотично наближається до величини, біт/с,

$$\begin{aligned} C'_\infty &= \lim_{F \rightarrow \infty} C'(F) = \log_2 e \cdot \lim_{F \rightarrow \infty} F \cdot \ln \left(1 + \frac{P_c}{N_0 F} \right) = \\ &= \frac{P_c}{N_0} \log_2 e \approx \frac{P_c}{N_0} \cdot 1,44, \end{aligned} \quad (5.80)$$

яка може бути названа пропускною спроможністю безперервного каналу зв'язку з необмеженою смугою частот при адитивній заваді у вигляді БГШ.

Таким чином, хоча зі зростанням смуги пропускання можливості безперервного каналу з передачі інформації збільшуються, однак у смузі пропускання не укладені необмежені можливості збільшення C' . (Уже при виборі смуги $F = 3P_c/N_0$ реалізується приблизно 86 % від C'_∞ , (див. рис. 5.5).)

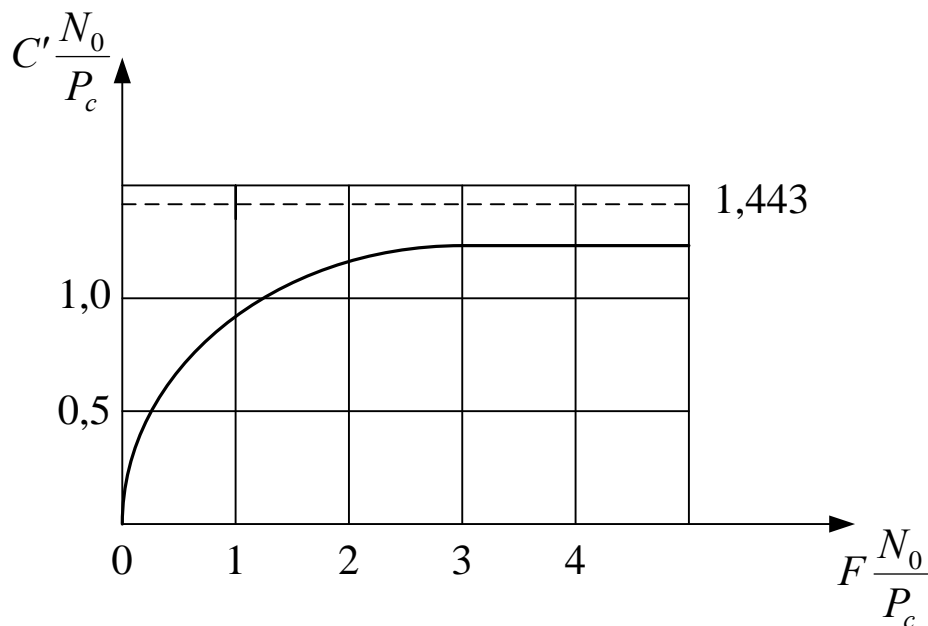


Рис. 5.5. Залежність нормованої пропускної спроможності безперервного каналу із БГШ від смуги пропускання

Хоча визначення поняття пропускної спроможності для безперервного каналу зв'язку й дозволяє судити про його можливості з передачі інформації, але для того щоб визначити ці можливості в більш конкретних термінах правильності й швидкості передачі, необхідно сформулювати теорему кодування Шеннона для безперервного каналу зв'язку. Однак перед цим визначимо ще одне поняття.

5.5.2.4. ϵ -ентропія безперервного джерела повідомлень

Іноді вводиться міра інформативності (непередбачуваності) безперервного джерела $s \in S$, яка називається ϵ -ентропією. Епсілон-ентропія $H_\epsilon(S)$ визначається як *мінімальна кількість*

інформації, що міститься в $Z(t)=S(t)+E(t)$ щодо сигналу $S(t)$, при якому $Z(t)$ і $S(t)$ еквівалентні. Еквівалентність приймається як близькість у середньоквадратичному сенсі:

$\overline{(Z(t)-S(t))^2} = \overline{E^2} \leq \varepsilon_0^2$ – припустиме значення середнього квадрата шуму спостереження.

Отже, за визначенням,

$$H_\varepsilon(S) = \min[h(S) - h(S|Z)] = h(S) - \max h(S|Z),$$

де мінімум береться за всіма умовними розподілами $w(s/z)$, для яких $\overline{E^2} \leq \varepsilon_0^2$. Оскільки $S(t)=Z(t)-E(t)$, то умовна диференціальна ентропія $h(S|Z)$ при заданому сигналі $z(t)$ повністю визначається шумом відтворення $E(t)$.

Якщо шум відтворення має фіксовану дисперсію $\sigma_\varepsilon^2 = \overline{E^2(t)}$, то диференціальна ентропія $h(E)$ максимальна, як було показано вище, при гаусівському розподілі й дорівнює $h(E) = \log \sqrt{2\pi e \sigma_\varepsilon^2}$.

Якщо джерело $S(t)$ є гаусівським, то при заданій дисперсії σ_s^2 його диференціальна ентропія $h(S) = \log \sqrt{2\pi e \sigma_s^2}$. Таким чином,

$$H(S) = \log \sqrt{2\pi e \sigma_s^2} - \log \sqrt{2\pi e \sigma^2} = 0,5 \log \frac{\sigma_s^2}{\sigma^2}.$$

Величина $\frac{\sigma_s^2}{\sigma^2} = \frac{P_s}{P}$ характеризує мінімальне відношення

сигнал-шум, при якому сигнал $S(t)$ і процес $Z(t)$ ще еквівалентні («схожі»). Позначимо це відношення ρ_0 , тоді $H(S) = 0,5 \log \rho_0$. Можна ввести поняття ε -продуктивності безперервного джерела

$$H'(\varepsilon) = \nu_{дж} H(\varepsilon),$$

де $\nu_{дж}$ – число відліків в одиницю часу.

Для ε -продуктивності безперервного гаусівського джерела безперервного часу без пам'яті

$$H'(\varepsilon) = F_c \lg \rho_0,$$

де F_c – смуга частот сигналу, у межах якої СЦП процесу $S(t)$ вважається рівномірною.

5.6. Теорема про кодування в безперервному каналі з завадами

Нехай є деякий безперервний канал зв'язку, для якого в п.п. 5.5.1 визначені поняття T -коду й вирішальної схеми. Тоді можна сформулювати таку теорему кодування для безперервного каналу, що є аналогом теореми кодування для дискретного каналу з завадами.

Теорема про кодування в безперервному каналі з завадами.

Якщо безперервний канал має пропускну спроможність C' і задані будь-які числа $\delta > 0$ і $H' < C'$, то завжди знайдеться таке T_0 , що при будь-якому $T > T_0$ існує T -код, що складається із $M = 2^{TH'}$ сигналів, і вирішальна схема, які забезпечують виконання нерівностей

$$P(z_i^T | s_i(t)) \geq 1 - \delta, \quad i = 1, 2, \dots, M, \quad z_i^T = s_i(t) + n(t). \quad (5.81)$$

Якщо $H' > C'$, то нерівність (5.81) не виконується, яким би не було велике значення T_0 .

Інтерпретація даної теореми мало відрізняється від інтерпретації відповідної теореми для дискретного каналу. Дійсно, якщо ми маємо деяке двійкове джерело інформації, то блоки довжини n даного джерела можна погодити з T -кодом без затримок у часі при виконанні таких умов:

$$2^n = 2^{TH'}, \quad nT_{дж} = T, \quad (5.82)$$

де $T_{дж}$ – тривалість символів джерела.

Перетворюємо вираз (5.82) та одержуємо необхідні й достатні умови зменшення помилки до нуля при кодуванні в безперервному каналі зв'язку у виразі

$$v_{дж} = H' < C'. \quad (5.83)$$

Бачимо, що ця умова відрізняється від умови (5.48), яка отримана для кодування в дискретному каналі, тільки тим, що пропускна спроможність дискретного каналу заміняється на пропускну спроможність безперервного каналу.

Оскільки, як вже зазначалося раніше, безперервний канал завжди має більшу пропускну спроможність, ніж будь-який відображуючий його дискретний канал, то кодування в безперервному каналі забезпечує завжди більшу інформаційну швидкість передачі, ніж у дискретному. Це властивість цілком очевидна, тому що кодування й декодування в безперервному каналі є більш загальними процедурами, чим у дискретному.

Безперервний канал зв'язку з нескінченною смугою пропускання, для якого пропускна спроможність визначається виразом (5.80), є достатньою умовою для зменшення ймовірності помилки до нуля при використанні ортогональних сигналів [39]. Тому необхідна й достатня умова забезпечення високої надійності передачі в такому каналі в дійсності буде мати вигляд

$$v < 1,44 P_c / N_0. \quad (5.84)$$

Для безперервного каналу залишається справедливою також й основна теорема Шеннона, якщо в ній розуміти під C' пропускну спроможність безперервного каналу.

Для каналу із пропускнуною спроможністю C' , на вхід якого підключено безперервне джерело з ε -продуктивністю $H'_\varepsilon(S)$, К. Шеннон довів таку теорему [50]: *якщо при заданому критерії еквівалентності повідомлень джерела ε^2_0 його ε -продуктивність $H'_\varepsilon(S)$ менше за пропускну спроможність каналу C' , то існує спосіб кодування й декодування (перетворення повідомлення на каналний сигнал і назад — каналного сигналу на повідомлення),*

при якому неточність відтворення як завжди близька до ε^2_0 .
 При $H'_\varepsilon(S) > C'$ такого способу не існує.

Для гаусівського безперервного каналу першу частину теорему можна записати як

$$H'_\varepsilon(S) < F_\kappa \log(1 + \rho_\kappa),$$

де ρ_κ – відношення сигнал-завада в каналі, $\rho_\kappa = \frac{P_c}{P_{ш}}$.

При гаусівському джерелі умова «неспотвореної передачі» у гаусівському каналі можна записати як

$$F_s \log \rho_0 < F_\kappa \log(1 + \rho_\kappa)$$

або при $\rho_\kappa \gg 1$

$$F_s \log \rho_0 < F_\kappa \log \rho_\kappa. \quad (5.85)$$

Перемножуючи ліву й праву частину виразу (5.85) на T , одержуємо нерівність

$$V_{is} < V_{i\kappa}, \quad (5.86)$$

де V_{is} – інформаційний об'єм сигналу, $V_{is} = TF_s \lg \rho_0$;

$V_{i\kappa}$ – інформаційний об'єм каналу, $V_{i\kappa} = TF_\kappa \lg \rho_\kappa$.

Нерівність (5.86) збігається з умовою неспотвореної передачі, вираженої в термінах фізичного обсягу сигналу й каналу [27].

Формула Шеннона (5.79) може бути використана для оцінки потенційних можливостей безперервного каналу зв'язку не тільки щодо його енергетики, але й займаного ним спектра.

Дійсно, використаємо два найважливіших показники системи зв'язку: енергетичний параметр $h_e^2 = P_c / (N_0 R_{дж})$ і частотну ефективність $\gamma = R_{дж} / F$.

Відповідно до теореми Шеннона, приймаючи при оптимальному узгодженні дискретного джерела й безперервного гаусівського каналу $R_{дж} = C'$, з виразу (5.79), маємо

$$\gamma = \frac{C'}{F} = \log_2 \left[1 + \frac{P_c C'}{N_0 F C'} \right] = \log_2 [1 + h_e^2 \gamma]$$

$$\text{або } h_e^2 = \frac{2^\gamma - 1}{\gamma} . \quad (5.87)$$

На рис. 5.6 побудована залежність енергетичної ефективності $\beta = \frac{1}{h_e^2}$, виражена в децибелах, як функція від γ ($\beta\gamma$ -номограма), що відповідає виразу (5.87).

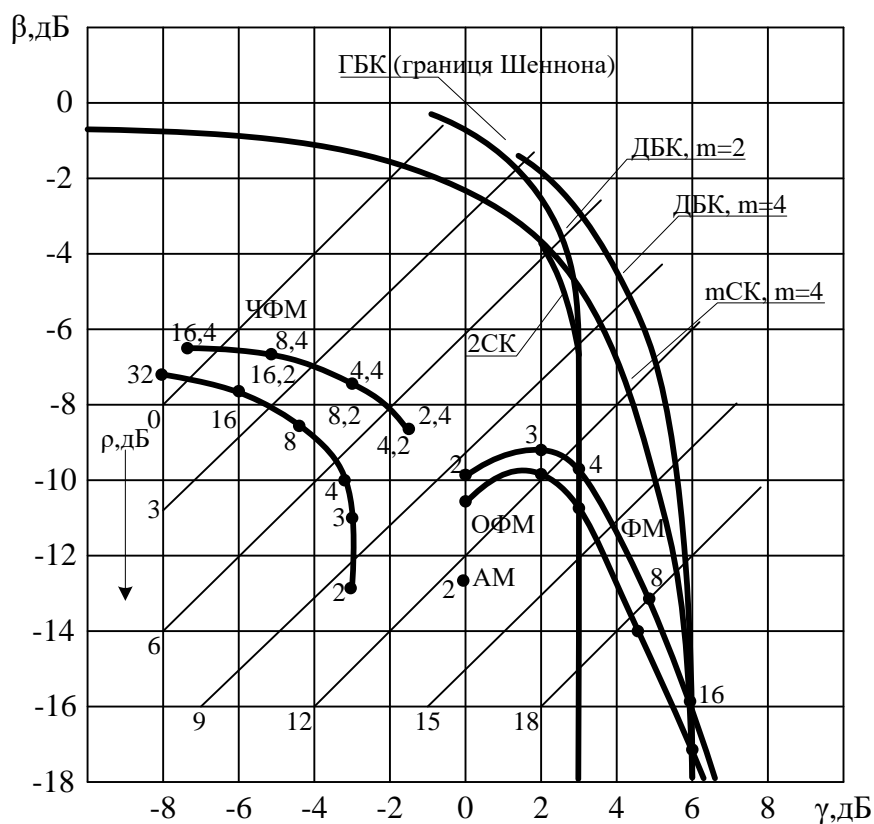


Рис. 5.6. Криві енергетичної та частотної ефективності цифрових систем зв'язку

Неможливо побудувати систему зв'язку, яка мала б пари чисел $(\beta = \frac{1}{h_e^2}, \gamma)$, що лежить вище зазначеної кривої на рис. 5.6.

При $\gamma = 1$ одержуємо $\beta = \frac{1}{h_e^2} = 1$ (0 дБ), а при $\gamma \rightarrow 0$, тобто коли

на смугу частот каналу не накладається ніяких обмежень, $h_e^2 \rightarrow \ln 2 \approx -1,6$ дБ, $\beta \rightarrow 1/\ln 2$. У той же час при $\gamma \rightarrow \infty$, як видно з виразу (5.87),

$$h_e^2 \sim \frac{2^\gamma}{\gamma} \sim e^{\gamma \ln 2 - \ln \gamma},$$

тобто мінімально необхідна бітова енергія експоненціально зростає при зростанні спектральної ефективності.

Якщо говорити про реальні системи зв'язку, то, використовуючи, наприклад, m -кові ортогональні сигнали (ЧМ) для $m = 8, 16, 32$, можна одержати при ймовірності помилки на біт $p_b = 10^{-5}$ $\beta\gamma$ -номограми, що відповідають точкам, зазначеним на рис. 5.6 ($\gamma < 0$ дБ). Щоб у цій області наблизитися до граничної кривої, необхідно використати завадостійке кодування, що описано в 6 главі.

Що ж стосується області $\beta\gamma$ -номограм при $\gamma > 0$ і поблизу граничної кривої, то вона виявляється досить важко досяжною для реальних систем зв'язку. Навіть для того щоб одержати $\gamma \geq (2-4)$ і $h_e^2 \leq (10-12)$ [дБ], доводиться використовувати комбіновані методи багаторазової фазової й амплітудної модуляції. Подальше ж наближення до граничної кривої в даній області можливе тільки при сполученні кодування й модуляції, тобто в так званих системах зв'язку з кодованою модуляцією.

5.7. Потенційні можливості каналів з багатьма користувачами

У попередньому пункті даної глави розглядалася модель системи зв'язку, у якій є одне джерело й один одержувач, а передача повідомлень здійснюється по єдиному каналу зв'язку.

Останнім часом значна увага приділяється так називаним мережам зв'язку, які реалізують обмін інформацією між більшою кількістю абонентів, що з'єднані між собою за допомогою безлічі розділених або загальних каналів зв'язку.

Загальна структура такої системи з багатьма користувачами показана на рис. 5.7. Тут є множина джерел повідомлень ДжП₁, ДжП₂, ..., Джп; множина передавальних пристроїв ПП₁, ..., ПП_к; множина прийомних пристроїв ПрП₁, ..., Прп, множина одержувачів О₁, ..., О_р і мережа каналів, що з'єднують передавальні й приймальні пристрої. У загальному випадку кожному джерелу доступна деяка підмножина передавальних пристроїв, а одержувачам повідомлень доступні виходи деякої підмножини прийомних пристроїв.

Основне завдання передачі повідомлень у такій системі зв'язку залишається тим самим – необхідно максимізувати швидкості передачі повідомлень при як завгодно високій правильності прийому. Однак на відміну від випадку з одним джерелом й одержувачем тут виникає множина припустимих швидкостей передачі між певними джерелами й одержувачами. Така постановка завдання призводить до поняття *області припустимих швидкостей* передачі, що призводять до надійної передачі повідомлень.

Найбільш частими випадками такої системи зв'язку є канал із *множинним доступом*, коли той самий канал зв'язку використовується для передачі повідомлень від декількох джерел до того самого одержувача, і так званий *широкомовний канал*, коли те саме повідомлення передається декільком одержувачам.

У роботі [24] розглянуто узагальнення теоретико-інформаційних понять для даної системи зв'язку й доведені теореми кодування, що визначають області досяжних швидкостей передачі при як завгодно малих імовірностях помилок.

Існує й інша постановка завдання, коли один з одержувачів є *недружнім* (інакше кажучи, він відповідає каналу підслуховування). Тоді завдання кодування полягає в мінімізації кількості інформації, що надходить до *недружнього одержувача*, при максимізації швидкості передачі інформації для *дружнього одержувача*.

У наш час достатньо розроблені й конструктивні методи кодування й декодування для каналів з багатьма користувачами, однак виклад цього матеріалу виходить за рамки даного навчального посібника.

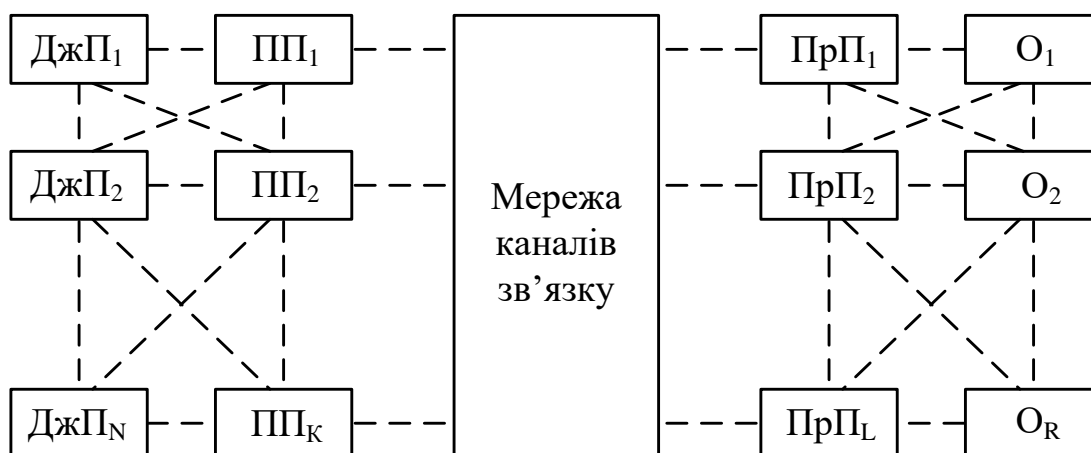


Рис. 5.7. Структура системи зв'язку з багатьма користувачами

Висновки до глави 5

1. Повідомлення джерела з надмірністю можуть бути «стислі», тобто закодовані більш ощадливо за допомогою узгодження алфавітів джерела й каналу зв'язку. Граничні можливості стиску визначаються такою характеристикою джерела повідомлень, як ентропія, що залежить від статистичних властивостей повідомлень.

2. Потенційні можливості за швидкістю для передачі достовірної інформації по каналах зв'язку з завадами визначаються пропускнуою спроможністю каналів зв'язку. Ця характеристика залежить від розподілів завад і параметрів каналів зв'язку.

3. Пропускна спроможність двійкового симетричного каналу без пам'яті залежить тільки від імовірності помилки символу p у каналі зв'язку. Вона максимальна при $p = 0$ або $p = 1$ і дорівнює нулю при $p = 1/2$, коли настає обрив каналу зв'язку.

4. Пропускна спроможність безперервного гаусівського каналу зв'язку залежить від відношення сигнал-шум і смуги частот каналу. При необмеженому розширенні смуги частот каналу пропускна спроможність наближається до скінченної величини.

Контрольні питання та задачі

5.1. Якщо в m -ковій системі зв'язку з ортогональними сигналами середня потужність сигналів і спектральна щільність шуму постійні, то що потрібно зробити, щоб забезпечити наближення до нуля ймовірності помилки при збереженні постійної інформаційної швидкості передачі?

5.2. Що таке кодування й декодування в дискретному й безперервному каналах зв'язку? Що в них загального й чим вони відрізняються один від одного?

5.3. Які види модуляції-демодуляції призводять до моделі двійкового симетричного каналу, а які до моделі несиметричного каналу?

5.4. Що зміниться, якщо при визначенні всіх інформаційних понять змінити основу логарифма? Чи вплине це на основні результати, які дає теорія інформації?

5.5. Властивості ентропії, умовної ентропії, кількості інформації можуть бути пояснені інтуїтивно. Однак якби це було не так, то чи збереглися б основні результати теорії інформації?

5.6. У скільки разів можна стиснути російський текст, переданий заголовними буквами ($K = 32$), якщо вважати, що ентропія джерела, що видає цей текст, $H(A) = 1,5$ біт/символ.

5.7. Чи можна за допомогою взаємної інформації $I(X, Y)$ вимірювати ступінь залежності випадкових величин X і Y ? Якщо це так, то чим це краще оцінки ступеня залежності за допомогою коефіцієнта кореляції?

5.8. Ентропія дискретного джерела на вході каналу $H(X) = 5$ біт/символ, втрати в дискретному каналі без пам'яті $H(X|Y) = 0,2$ біт/символ. Знайдіть ентропію шуму в каналі, якщо ентропія символів на виході каналу $H(Y) = 5,5$ біт/символ.

5.9. Поясніть наочно сенс того факту, що при тому самому значенні $p_c = p$ пропускна спроможність 2СК без пам'яті буде завжди менше, ніж пропускна спроможність двійкового за входом каналу зі стиранням.

5.10. Доведіть, що для двійкового за входом каналу зі стиранням з імовірностями переходів вираз (5.2) пропускна спроможність (біт/симв) $C = (1 - p_c)$.

5.11. Яке призначення кодування в каналі без завад?

5.12. Яке практичне значення має теорема кодування в каналі з завадами? Чи можна, використовуючи доведення цієї теореми, будувати реальні схеми кодування й декодування?

5.13. У скільки разів зменшується диференціальна ентропія рівномірно розподіленої випадкової величини (з нульовим МО) у порівнянні з випадком гаусівського розподілу, якщо фіксована дисперсія випадкової величини?

5.14. Чи є гаусівська завада найгіршою (і якщо так, то в якому розумінні) серед всіх завад однакової потужності?

5.15. На графіку рис. 5.4 показано, що при розширенні смуги частот каналу зв'язку пропускна спроможність завжди залишається менше, ніж C' . Чи має тоді практичний зміст розширювати смугу частот ліній зв'язку (наприклад, кабельних) понад таке значення F_0 , при якому $C'(F_0) \approx C'$?

ГЛАВА 6

ОСНОВИ ТЕОРІЇ ЗАВАДОСТІЙКОСТІ СИСТЕМ ПЕРЕДАЧІ ДИСКРЕТНИХ ПОВІДОМЛЕНЬ

6.1. Завдання синтезу оптимальних демодуляторів

Системи й методи передачі дискретних повідомлень займають основне місце в сучасних системах електричного зв'язку, що пояснюється не тільки наявністю безлічі джерел дискретних повідомлень, таких як дані з виходу ЕОМ, телеграфні джерела й ін., але й передачею повідомлень безперервних джерел (мова, музика, телебачення, факсимільна передача й ін.) більш завадостійкими дискретними цифровими методами.

На виході модулятора в ізохронних системах передачі відбувається перетворення цифрового сигналу $b_u(t)$ на послідовність $u(t)$ елементарних сигналів несучої $f(t - kT, b_k)$ заданої форми

$$u(t) = \sum_{k=0}^{\infty} f(t - kT, b_k), \quad (6.1)$$

де T - тактовий інтервал передачі в каналі, що визначає технічну швидкість передачі $V = 1/T$, яка вимірюється в Бодах;

b_k - безліч елементів коду.

Елементарний сигнал $f(t - kT, b_k)$ може формуватися по-різному. Так, у залежності від того, чи виконуються модуляція й кодування послідовно один за одним або спільно, він може мати різні частотно-часові властивості. Найбільш просту форму приймає вираз (6.1) при лінійній (багаторівневій) модуляції:

$$u(t) = \sum_{k=0}^{\infty} b_k f(t - kT, b_k). \quad (6.2)$$

Для стиску сигналу $f(t - kT, b_k)$ у часі та одночасно зменшення пік-фактора сигналу $u(t)$ (відношення пікової й

середньої потужності) можна в якості $f(t - kT, b_k)$ обрати прямокутні радіоімпульси тривалістю T . Але тоді ширина смуги частот сигналу теоретично не обмежена. Якщо ж формувати сигнали $f(t - kT, b_k)$ з рівномірним спектром у гранично вузькій смузі частот, то зростає пік-фактор сигналу $u(t)$ і, крім того, сигнали $f(t - kT, b_k)$ теоретично не обмежені в часі. У цьому випадку елементарні сигнали, що відповідають різним символам b_k , накладаються один на одного, тобто має місце міжсимвольна інтерференція (МСІ) вже на передачі. Однак ця обставина не перешкоджає безпомилковому виділенню інформації b_k з сигналу $u(t)$, якщо виконуються властивості сигналу $f(t - kT, b_k)$.

На практиці є розумний компроміс між двома зазначеними граничними подіями формування елементарних сигналів $f(t - kT, b_k)$.

На вхід демодулятора при передачі сигналу (вираз (6.2)) надходить коливання (сигнал - шум)

$$z(t) = \sum_{k=0}^{\infty} S(t - kT, b_k) + n(t), \quad (6.3)$$

де $S(t - kT, b_k)$ - відгук каналу на елементарний сигнал $f(t - kT, b_k)$;

$n(t)$ - адитивний шум у каналі зв'язку.

При лінійній (багаторівневій) модуляції маємо

$$z(t) = \sum_{k=0}^{\infty} b_k S(t - kT) + n(t), \quad (6.4)$$

де $S(t)$ - відгук каналу на несучу $f(t)$.

Для спрощення подальшого аналізу будемо вважати, що елементарний сигнал на передачі $f(t - kT, b_k)$ локалізований у межах тактового інтервалу.

Тоді в каналі з відносною пам'яттю Q реакція $S(t - kT, b_k)$ існує на інтервалі тривалістю $T_a = (1 + Q)T$.

Спочатку будемо аналізувати демодулятор для каналу без пам'яті ($Q = 0$, $T_a = T$), а далі будемо обговорювати демодулятор у каналі із міжсимвольною інтерференцією.

Позначимо аналізований елемент колювання (сигнал + шум) на вході демодулятора в каналі без МСІ:

$$z(t) = S(t, b_i) + n(t),$$

де $S(t, b_i)$ — сигнал, що відповідає символу b_i , $i \in \overline{0, m-1}$.

На виході демодулятора виникає дискретний сигнал, тобто послідовність кодових символів. Найчастіше елемент тривалістю T безперервного сигналу перетворюється демодулятором на один кодовий символ (поелементний прийом). Якби цей кодовий символ завжди збігався з переданим (який поступив на вхід модулятора), то зв'язок був би безпомилковим. Але, як відомо, завади призводять до неможливості з абсолютною імовірністю відновити за прийнятим сигналом переданий кодовий символ.

Кожен демодулятор описується законом, за яким безперервний сигнал, який надійшов на його вхід, перетворюється на кодовий символ. Цей закон називається *правилом вирішення*, а реалізуюча його схема — *вирішальною*. Демодулятори з різними правилами вирішення будуть видавати взагалі різні рішення, з яких одні правильні, а інших - помилкові.

Сучасна статистична теорія зв'язку дозволяє відшукати найкращу операцію обробки вхідного сигналу $z(t)$, що забезпечує максимальну якість оцінки b_i . Будемо вважати, що властивості джерела повідомлення й кодера відомі. Крім того, задано модулятор, а також відомо яка реалізація елемента сигналу відповідає тому або іншому кодовому символу, а також задана математична модель безперервного каналу. Потрібно визначити, який повинен бути оптимальний демодулятор (правило вирішення), щоб забезпечити найкращу якість прийому. Таке завдання було вперше поставлене й вирішено (для гаусівського неспотворюючого каналу із БГШ) В.О. Котельниковим [29].

У цій постановці якість оцінювалася ймовірністю правильного прийому символу. Максимум цієї ймовірності при

заданому виді модуляції В.О. Котельников назвав потенційною завадостійкістю, а демодулятор, що забезпечує цей максимум, - ідеальним приймачем. Із цього визначення слідує, що в жодному реальному демодуляторі середня ймовірність правильного прийому символу не може бути більше, ніж в ідеальному приймачі.

На перший погляд, принцип оцінки якості прийому ймовірністю правильного прийому символу здається цілком природним і навіть єдино можливим. Далі буде показано, що це не завжди так і що існують й інші критерії якості, які можуть бути застосовані в тих або інших окремих випадках.

Ознайомимося докладніше із статистичним підходом до завдання прийому дискретних повідомлень на фоні шумів. Припустимо, що при передачі дискретних повідомлень, закодованих кодом із основою m , у місці прийому очікуються сигнали $S_i(t), t \in [0, T]$, що відповідають кодовим символам b_i ($i = 0, 1, 2, 3, \dots, m - 1$). На відрізку часу $[0, T]$ на вхід прийомного пристрою надходить коливання $z(t)$, що внаслідок перекручувань і завад у каналі не збігається в точності з жодним із елементів сигналу на передачі $u_i(t)$. Отже, у цьому випадку прийомний пристрій повинен вибрати одну з m можливих взаємовиключних (альтернативних) гіпотез: передавався кодовий символ b_0 , тобто очікується сигнал $S_0(t)$; передавався кодовий символ b_1 , тобто очікується сигнал $S_1(t)$; передавався кодовий символ b_{m-1} , тобто очікується сигнал $S_{m-1}(t)$. Для двійкової системи ($m = 2$) прийомний пристрій обирає одну з двох альтернативних гіпотез про передачу символу 1 або 0.

Сукупність всіх можливих реалізацій $z(t)$ можна інтерпретувати точками в просторі Z прийнятих фінітних сигналів. Він є безкінечномірним простором Гілберта або, з деякими (прийнятними для практики) застереженнями, багатомірним простором Евкліда. Графічно зображення реалізації прийнятих сигналів $S_i(t)$ і завад $n(t)$ (тривалістю T) приймають точками на площині (рис. 6.1) або відповідними векторами на площині, що відкладають від початку координат 0. Якщо правило вирішення обране, то це означає, що кожній точці простору

прийнятих коливань (кінцю вектора $z = s + n$) приписується одна з m гіпотез, тобто певний переданий кодовий символ b_i . Простір прийнятих сигналів виявиться при цьому розбитим на m непересічних областей \hat{B}_i , кожна з яких відповідає прийняттю певної гіпотези. У такому трактуванні різні прийомні пристрої відрізняються один від одного способом розподілу простору сигналів на області \hat{B}_i , тобто правилом прийняття рішення. Можливий розподіл простору коливань, що приймаються на непересічні області, схематично показаний на рис. 6.1.

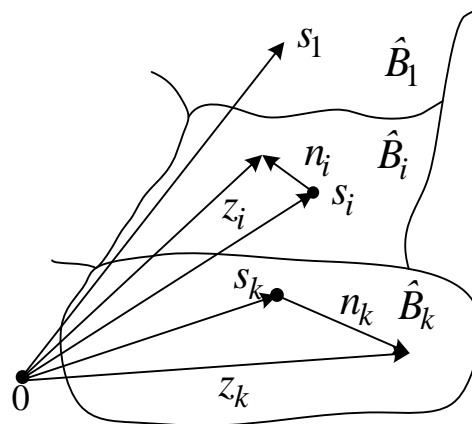


Рис. 6.1. Розподіл простору коливань, що приймаються на непересічні області

У двійковій системі простір Z розбивають на дві непересічні області \hat{B}_0 і \hat{B}_1 . Припустимо, на відрізку $[0, T]$ приймається коливання

$$z(t) = S_i(t) + n(t),$$

де $S_i(t)$ — корисний сигнал у місці прийому, що пройшов канал зв'язку;

$n(t)$ - реалізація адитивної завади.

Якщо завади відсутні, можливі значення $z(t)$ зображуються точками $S_i(t)$ ($i = 0, 1, 2, 3, \dots, m - 1$). За наявності завади й передачі сигналу з номером i кінець вектора прийнятого коливання z відхиляється від точки. На рис. 6.1 це показано для

сигналів $S_i(t)$ і $S_k(t)$. Область \hat{B}_i містить точку $S_i(t)$. У тих випадках, коли завада не виводить точку z за межі області \hat{B}_i , рішення виявляється правильним при сигналі $S_i(t)$. В іншому випадку виникає помилка. Очевидно, змінюючи границі між областями, можна впливати на ймовірність помилкового прийому окремих переданих символів. Наприклад, якщо в розподілу, показаному на рис. 6.1, розширити область \hat{B}_i за рахунок області \hat{B}_k , то зменшиться ймовірність помилкового прийому символу \hat{b}_k замість переданого символу b_i . Однак, у цьому випадку зростає ймовірність помилкового прийому \hat{b}_i при переданому b_k . Очевидно, завжди існує таке розташування областей, що у певному розумінні краще за будь-яке інше.

Якщо задано критерій якості, то найкращий розподіл простору прийнятих сигналів (оптимальна вирішальна схема прийомного пристрою) досягається методами теорії статистичних рішень.

6.2. Критерії якості й правила прийому дискретних повідомлень

Розглянемо спочатку широко розповсюджений критерій Котельникова або критерій ідеального спостерігача, відповідно до якого якість демодулятора оцінюють безумовною (середньою) імовірністю правильного прийому символу. Будемо спочатку припускати, що простір переданих і прийнятих сигналів є кінцевомірним евклідовим. Це може бути, наприклад, простір фінітних сигналів, що представляють кінцевою тригонометричною сумою. Надалі це обмеження буде знято.

У n -мірному просторі випадковий сигнал $z(t)$ характеризується n -мірною щільністю ймовірностей вектора z : $w(z)$. Її можна розглядати як щільність імовірності коефіцієнтів розкладання $z(t)$ за будь-яким ортонормованим базисом. Якщо передається деякий символ $b_i(t)$, тобто приймається сигнал $S_i(t)$, то можна визначити умовну n -мірну щільність імовірності

$w(z|b_i) = w(z|S_i)$ - функцію правдоподібності i -тої гіпотези ($i = 0, m - 1$).

На вхід демодулятора протягом відрізка $[0, T]$ приходять деякий елемент сигналу $z(t)$. Припустимо, що демодулятор приймає при цьому рішення, що передано символ b_i , тобто видає оцінку \hat{b}_i . Ймовірність того, що це рішення правильне, вочевидь, дорівнює умовній ймовірності $P(b_i|z(t))$ того, що дійсно передавався символ b_i , за умови приходу реалізації елемента сигналу $z(t)$. Її називають *апостеріорною ймовірністю символу b_i* (тобто ймовірністю, визначеною після досліду, що полягає в спостереженні й аналізі сигналу $z(t)$).

Очевидно, що ймовірність правильного прийому буде дорівнювати

$$P_{\text{і}} = \int_z w(z) \sum_{i=0}^{m-1} P(b_i, \hat{b}_i | z)$$

та досягає максимуму, якщо для кожної конкретної реалізації $z(t)$ сума

$$\sum_{i=0}^{m-1} P(b_i, \hat{b}_i | z) = \sum_{i=0}^{m-1} P(b_i | z) P(\hat{b}_i | z)$$

максимальна. Тут $P(b_i | z)$ - апостеріорна ймовірність передачі b_i , а $P(\hat{b}_i | z)$ у сукупності утворюють вирішальне правило

демодулятора. Оскільки $\sum_{i=0}^{m-1} P(\hat{b}_i | z) = 1$, то найкращим

вирішальним правилом буде таке, при якому $P(\hat{b}_i | z) = 1$ для i , відповідного максимальній $P(b_i | z)$.

Інакше кажучи, критерій ідеального спостерігача забезпечується вирішальною схемою, побудованою за правилом максимуму апостеріорної ймовірності — рішення \hat{b}_i приймається в тому випадку, якщо виконується система з $m - 1$ нерівностей:

$$P(b_i|z) > P(b_j|z), \quad j \in \overline{0, m-1}, j \neq i. \quad (6.5)$$

Для скорочення запишемо це правило в такій формі:

$$\hat{i} = \text{Arg max}_i \{P(b_i|z)\}, \quad (6.6)$$

де $\text{Arg max}_i \{A_i\}$ – значення i , при якому A_i максимально.

Для двійкової системи сигналів згадане правило зводиться до перевірки нерівності

$$P(1|z) > P(0|z). \quad (6.7)$$

При виконанні нерівності (6.7) реєструється символ 1, в іншому випадку – 0.

Відповідно до відомої формули Байєса

$$P(b_i|z) = \frac{P(b_i)w(z|b_i)}{w(z)}, \quad (6.8)$$

де $P(b_i)$ — апіорна ймовірність передачі символу b_i (тобто та ймовірність, що має місце до спостереження й аналізу і обумовлена статистикою джерела повідомлення й правилом кодування).

Підставивши вираз (6.8) у вираз (6.5) і з огляду на те, що $w(z)$ - безумовна щільність імовірності, яка не є функцією i , можна записати правило рішення за критерієм ідеального спостерігача в такій формі:

$$P(b_i)w(z|b_i) > P(b_j)w(z|b_j), \quad j \in \overline{0, 1, \dots, m-1}, j \neq i,$$

або скорочено

$$\hat{i} = \text{Arg max}_i \{P(b_i)w(z|b_i)\}. \quad (6.9)$$

Приймач, що реалізує алгоритм (6.9), називають приймачем Котельникова. Для двійкової системи правило (6.9) зводиться до перевірки нерівності

$$P(1)w(z|1) > P(0)w(z|0), \quad (6.10)$$

при виконанні якої реєструється символ 1, а при невиконанні символ 0.

Для побудови вирішальної схеми за правилом (6.9) необхідно знати апіорні ймовірності символів $P(b_i)$, а також властивості модулятора й каналу, визначальні умовні щільності $w(z|b_i)$ - функції правдоподібності.

Правило (6.9) можна записати інакше. Рішення про те, що передавався символ b_i , має прийматися, якщо для всіх $j \neq i$ виконуються $m - 1$ нерівностей

$$\frac{w(z|b_i)}{w(z|b_j)} > \frac{P(b_j)}{P(b_i)}. \quad (6.11)$$

Відношення в лівій частині цієї нерівності називається *відношенням правдоподібності* двох гіпотез: про те, що передавався символ b_i , і про те, що передавався символ b_j . Його позначають $A_{i,j}$.

У випадку, коли всі m символів передаються рівноймовірно, тобто $P(b_i) = 1/m$, правило (6.11) спрощується:

$$A_{i,j} > 1, \quad i \in \overline{0, m-1}, \quad j \neq i. \quad (6.12)$$

Іноді вводять у розгляд, крім m гіпотез про передачу символів $b_i (i = \overline{0, m-1})$, ще додаткову ("шумову") гіпотезу про те, що жоден сигнал не передавався, тобто $z(t) = n(t)$ — тільки завада. Відношення правдоподібності $\frac{w(z|b_i)}{w(z|b_{uu})} = A_{iu}$ звичайно позначають просто A_i . Тоді правило (6.12) можна записати так:

$$A_i > A_j \text{ при } j \neq i, \quad (6.13)$$

або скорочено $\hat{i} = \underset{i}{\text{Arg max}} [A_i]$. Таке правило максимуму правдоподібності реалізує критерій ідеального спостерігача за тієї умови, що всі символи передаються рівномірно.

Для двійкової системи правило (6.13) зводиться до перевірки нерівності

$$A_1 > A_0. \quad (6.14)$$

Як зазначалося, критерій ідеального спостерігача не є єдиним "розумним" критерієм оптимальності вирішальної схеми. Справа в тому, що в багатьох випадках різні помилки призводять до різних наслідків. Наприклад, у системі автоматичної пожежної сигналізації небезпечніше не виявити сигнал про пожежу, ніж оголосити "фіктивну тривогу", коли в дійсності пожежі немає.

Облік наслідків помилок різного роду (пов'язаних з передачею різних символів) призводить до узагальнення критерію ідеального спостерігача, відомого за назвою критерію мінімального середнього ризику (або байєсівського критерію). Уведемо деякі поняття.

Якщо при передачі символу b_i прийнято символ \hat{b}_j , то при $j \neq i$ має місце помилка. Щоб урахувати нерівноцінність різних помилок, будемо з кожною парою символів b_i і \hat{b}_j зв'язувати деяку чисельну величину, яка названа "втратою", позначивши її L_{ij} . Величина "втрати" залежить, таким чином, від того, який символ \hat{b}_j прийнятий замість переданого b_i . Правильному прийому при цьому звичайно приписується нульова "втрата". Значення L_{ij} визначаються в кожному конкретному випадку важливістю правильного прийому даного елемента сигналу й величиною небезпеки різних помилок.

Оскільки при передачі символу b_i символи \hat{b}_j з'являються з певними ймовірностями як реалізації деякої дискретної випадкової величини, можливо говорити про умовне

математичне очікування величини "втрати" при передачі конкретного символу b_i . Назвемо це умовне математичне очікування *умовним ризиком*:

$$R_i = \sum_{j=0}^{m-1} P(\hat{b}_j | b_i) L_{i,j} = \sum_{j=0}^{m-1} L_{ij} \int_{\hat{B}_j} w(z | b_i) dz. \quad (6.15)$$

Інтеграл у виразі (6.15) береться за областю \hat{B}_j вирішальної схеми й є ймовірністю того, що сигнал $z(t)$ потрапив у цю область, якщо передавався символ b_i . Зробимо середнім умовний ризик R_i за всіма символами b_i і одержимо величину, названу *середнім ризиком*:

$$R_{\bar{n}\bar{a}\bar{d}} = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} P(b_i) L_{ij} \int_{\hat{B}_j} w(z | b_i) dz. \quad (6.16)$$

Критерій мінімального середнього ризику полягає в тім, що оптимальною вважається вирішальна схема, яка забезпечує найменше значення середнього ризику $R_{сер}$. Приймач, що працює за таким критерієм, називають *байєсівським*.

З виразу (6.16) видно, що при використанні цього критерію потрібно, крім апіорних ймовірностей $p(b_i)$ передачі окремих символів, знати і величини втрат L_{ij} . Зазначимо, що якщо вважати всі помилки рівноцінними ($L_{ij} = const$ при $j \neq i$ й $L_{ij} = 0$), то критерій мінімального середнього ризику збігається із критерієм ідеального спостерігача, а байєсівський приймач збігається з ідеальним приймачем Котельникова. У загальному ж випадку в оптимальному байєсівському приймачі частіше будуть виникати помилки, пов'язані з малими втратами, і рідше - з великими втратами.

Ситуація, у якій практично неможливо визначити апіорну ймовірність передачі окремих елементарних повідомлень, а наслідки помилок різного роду неоднакові, типова для

радіолокації, коли приймач, аналізуючи прийняте коливання $z(t)$ (відбитий сигнал плюс завада), повинен визначити, є в даному напрямку й на даній відстані об'єкт спостереження (ціль) чи ні. Як правило, апіорна ймовірність наявності відбитого від цілі сигналу (передачі 1) заздалегідь не відома. Наслідки двох родів помилок фіктивної тривоги (приймач фіксує, що ціль існує, у той час як у дійсності її немає) і пропускання цілі (приймач відзначає відсутність цілі, у той час як фактично вона є) - нерівноцінні.

У цій й інших подібних ситуаціях найчастіше користуються критерієм прийому, відомим за назвою критерію Неймана - Пірсона. Суть його полягає в тім, що вирішальна схема вважається оптимальною, якщо при заданій ймовірності фіктивної тривоги $p_{фт}$ забезпечується мінімальна ймовірність пропускання цілі $p_{при}$. Уведемо в розгляд функції правдоподібності гіпотези про відсутність цілі $w[z|0]$ і про наявність цілі $w[z|1]$.

Вочевидь, можливо різними способами розбити простір прийнятих коливань $z(t)$ на дві області: \hat{B}_0 (область рішення про відсутність цілі) і \hat{B}_1 (про наявність цілі) — так, щоб ймовірність фіктивної тривоги

$$p_{\delta\delta} = \int_{\hat{B}_1} w(z|0) dz \quad (6.17)$$

дорівнювала заданій величині. Оскільки в локації символ 0 (відсутність цілі) передається паузою, то $w[z|0]$ - це щільність розподілу завади. Отже, ймовірність фіктивної тривоги визначається ймовірнісними характеристиками завади й вибором області \hat{B}_1 . Але від вибору цієї області залежить і ймовірність правильного виявлення цілі:

$$p_{\delta\delta} = \int_{\hat{B}_1} w(z|1) dz = 1 - p_{\delta\delta} = 1 - \int_{\hat{B}_0} w(z|1) dz, \quad (6.18)$$

де $p_{прц}$ - ймовірність пропускання цілі.

Інтеграл у виразах (6.17), (6.18) і в аналогічних інших формулах, узяті за векторною змінною, вочевидь, багаторазові.

Максимізація виразу (6.18) при заданій величині виразу (6.17) досягається, якщо рішення про наявність цілі приймається при виконанні нерівності

$$A_{1,0} = \frac{w(z|1)}{w(z|0)} > \lambda, \quad (6.19)$$

де λ — граничний, пороговий рівень, обумовлений заданою ймовірністю фіктивної тривоги $P_{фт}$.

Існують й інші критерії якості прийому, що не потребують знання апріорних ймовірностей символів [44].

У техніці зв'язку переважно застосовують правило максимальної правдоподібності (вирази (6.13), (6.14)). У тому випадку, коли всі символи передаються рівноймовірно, правило максимальної правдоподібності реалізує критерій ідеального спостерігача. Однак дуже часто це правило рішення застосовують і при невідомих або відомих, але не однакових апріорних ймовірностях символів. Звичайно, воно не забезпечує в цих випадках максимуму ймовірності правильного прийому.

Змінивши вирішальну схему на схему, побудовану за правилом максимальної апостеріорної ймовірності (вираз (6.7)), що реалізує критерій ідеального спостерігача, можна було б зменшити ймовірність помилок. При цьому, мабуть, довелося б скоротити області прийому малої ймовірності і розширити області високої ймовірності символів. У результаті рідко передані символи приймалися б менш надійно, ніж часто передані. Але рідкі символи несуть більше інформації, ніж часті. Тому перехід від правила максимальної правдоподібності до правила максимальної апостеріорної ймовірності, хоча й зменшує безумовну ймовірність помилки, може призвести до збільшення втрати інформації при демодуляції. Легко показати, що правило максимальної правдоподібності реалізує критерій мінімуму

середнього ризику (вираз (6.16)), якщо прийняти $L_{ij} = 0$ при $i = j$ й $L_{ij} = 1/p(b_i)$ при $i \neq j$.

Внаслідок сказаного будемо надалі користуватися, якщо не обговорено зворотне, правилом максимальної правдоподібності й вирішальну схему, яка реалізує правило (6.13), називати оптимальною.

6.3. Оптимальні алгоритми прийому при повністю відомих сигналах (когерентний прийом)

Припустимо, що всі спотворення в каналі строго детерміновані й випадковим є тільки гаусівський адитивний шум $n(t)$, який спочатку припускаємо є "білим", зі спектральною щільністю N_0 . Це означає, що при передачі символу b_i ($i = 0, 1, \dots, m-1$) прийняте коливання можна описати моделлю

$$z(t) = S_i(t) + n(t), \quad 0 \leq t \leq T, \quad (6.20)$$

де усі $S_i(t) = \gamma f(t - \tau - kT, b_k^{(i)}) = \gamma u_i(t - \tau)$ відомі (γ – відомий коефіцієнт передачі каналу). Невідома лише реалізація завади та позиція (індекс i) дійсно переданого сигналу, який й повинна визначити вирішальна схема.

Будемо також уважати, що всі S_i є фінітними сигналами, тривалість яких T . Це має місце, якщо передані сигнали $u_i(t)$ є фінітними й мають однакову тривалість (система синхронна), а в каналі немає ні багатопроменевого поширення, ні лінійних спотворень, що викликають збільшення тривалості сигналу (або вони скоректовані).

У подальшому будемо вважати, що в системі забезпечена надійна тактова синхронізація, тобто границі тактового інтервалу, на якому приходить сигнал $S_i(t)$, відомі точно. Момент початку послілки $S_i(t)$ приймемо за нуль.

Визначимо в цих умовах алгоритм роботи оптимального (тобто заснованого на правилі максимальної правдоподібності)

демодулятора, що аналізує сигнал на тактовому інтервалі $0...T$. Для цього необхідно знайти відносини правдоподібності для всіх m можливих сигналів щодо нульової гіпотези ($z(t) = n(t)$).

Завдання ускладнюється тим, що ширина спектра сигналу нескінченна (оскільки він фінітний), а тому простір сигналів безкінцевомірний $L_2(t)$. Для таких сигналів (або безкіцевомірних векторів) не існують щільності ймовірностей. Однак існують n -мірні щільності ймовірностей для будь-яких n перетинів сигналу.

Замінімо спочатку "білий" шум "квазібілим", що має таку саму однобічну спектральну щільність потужності N_0 , але тільки в деякій смузі частот $\Delta F = n/2T$, де $n \gg 1$. Розглянемо додаткову гіпотезу, тобто будемо вважати, що $z(t)$ — стаціонарний шум з нульовим МО. Візьмемо на тактовому інтервалі n рівновіддалених перетинів через $\Delta t = 1/2\Delta F = T/n$. Відліки z_1, \dots, z_n у цих перетинах для "квазібілого" гаусівського шуму незалежні відповідно до виразу (2.89). Тому n -мірна щільність імовірності для взятих відліків

$$w[z_1, z_2, \dots, z_n, t_1, t_2, \dots, t_n | z(t) = n(t)] = \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{k=1}^n z^2(t_k)\right\}, \quad (6.21)$$

де σ^2 - дисперсія (потужність) "квазібілого" шуму, $\sigma^2 = N_0 F$.

При гіпотезі, що передавався символ b_i , згідно з виразом (6.20) $n(t) = z(t) - S_i(t)$. Отже, умовна n -мірна щільність імовірності перетинів $z(t)$ визначиться такою самою формулою, як й вираз (6.21), якщо $z(t_k)$ замінити різницею $z(t_k) - S_i(t_k)$, що представляє при цій гіпотезі шум:

$$w[z_1, z_2, \dots, z_n, t_1, t_2, \dots, t_n | b_i] = \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{k=1}^n [z(t_k) - S_i(t_k)]^2\right\}.$$

Відношення правдоподібності для сигналу S_i (відносно додаткової гіпотези), обчислене для n перетинів,

$$\begin{aligned} \Lambda_i^{[n]} &= \frac{w[z_1, z_2, \dots, z_n, t_1, t_2, \dots, t_n | b_i]}{w[z_1, z_2, \dots, z_n, t_1, t_2, \dots, t_n | u]} = \\ &= \exp \left\{ -\frac{1}{2\sigma^2} \sum_{k=1}^n [z(t_k) - S_i(t_k)]^2 + \frac{1}{2\sigma^2} \sum_{k=1}^n z^2(t_k) \right\}. \end{aligned}$$

Замінімо дисперсію σ^2 її виразом: $\sigma^2 = N_0 F = N_0 / 2\Delta t$.

Тоді маємо

$$\Lambda_i^{[n]} = \exp \left\{ -\frac{1}{N_0} \sum_{k=1}^n [z(t_k) - S_i(t_k)]^2 \Delta t + \frac{1}{N_0} \sum_{k=1}^n z^2(t_k) \Delta t \right\}.$$

За правилом максимуму правдоподібності у випадку "квазібілого" шуму вирішальна схема повинна вибирати значення i , що забезпечує максимум $\Lambda_i^{[n]}$. Замість $\Lambda_i^{[n]}$ можна відшукувати максимум його логарифма:

$$\ln \Lambda_i^{[n]} = -\frac{1}{N_0} \sum_{k=1}^n [z(t_k) - S_i(t_k)]^2 \Delta t + \frac{1}{N_0} \sum_{k=1}^n z^2(t_k) \Delta t. \quad (6.22)$$

Повернемося до вихідного завдання для "білого" шуму. Для цього будемо розширювати смугу ΔF , тоді число перетинів n прагне до нескінченності, а Δt — до нуля. Суми у виразі (6.22) перетворюються на інтеграли, і після розкриття квадрата в першому доданку правило рішення (вибору оцінки \hat{b}_i) можна написати так:

$$\hat{b}_i = \text{Arg max}_i \left\{ \frac{2}{N_0} \int_0^T z(t) S_i(t) dt - \frac{1}{N_0} \int_0^T S_i^2(t) dt \right\}. \quad (6.23)$$

Правило прийому (вираз (6.23)) зводиться до перевірки системи нерівностей

$$\int_0^T z(t)S_i(t)dt - 0,5E_i > \int_0^T z(t)S_j(t)dt - 0,5E_j, \quad j \neq i, \quad (6.24)$$

де E_i - енергія очікуваного сигналу $S_i(t)$, $E_i = \int_0^T S_i^2(t)dt$.

Вираз (6.24) визначає ті операції (*алгоритм прийому*), які має виконувати оптимальний приймач над вхідним коливанням.

Для двійкової системи алгоритм (6.24) зводиться до перевірки однієї нерівності

$$\int_0^T z(t)S_1(t)dt - 0,5E_1 > \int_0^T z(t)S_0(t)dt - 0,5E_0. \quad (6.25)$$

При виконанні нерівності (6.25) реєструється символ "1", в іншому випадку "0".

Пристрій, що безпосередньо обчислює скалярний добуток (або *кореляційний інтеграл*)

$$(z, S_i) = \int_0^T z(t)S_i(t)dt, \quad (6.26)$$

називають *активним фільтром* або *корелятором*, тому приймач, що реалізує алгоритм (6.24), називають *кореляційним*.

На рис. 6.2 показана структурна схема прийомного пристрою, що працює відповідно до виразу (6.25). Блоки "х" — перемножники; Γ_0, Γ_1 — генератори опорних сигналів $S_0(t), S_1(t)$; \int — інтегратори; "—" пристрої, що віднімають; ВП - вирішальний пристрій, що визначає в моменти часу, кратні T (при замиканні ключа - К), номер i -тої гілки з максимальним сигналом ($i = 0, 1$). При $m > 2$ у схемі рис. 6.2 та інших нижчеподаних

схемах росте відповідно число гілок обробки сигналу, що попадають на ВП.

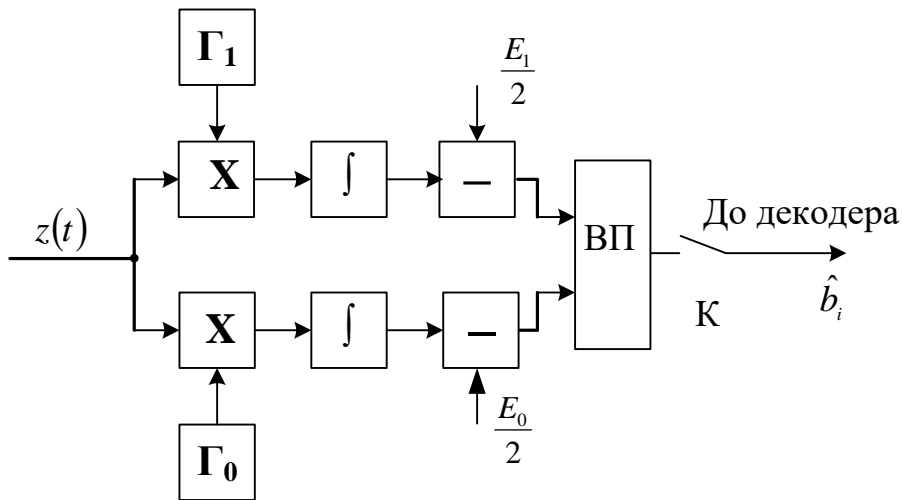


Рис. 6.2. Оптимальний демодулятор при точно відомих сигналах, який побудовано за кореляційною схемою

Якщо сигнали $u_i(t)$ обрані таким чином, що усі їх реалізації (а отже, і всі реалізації $S_i(t)$) мають однакові енергії ($E_i = \text{const}$), алгоритм прийому вираз (6.24) (і відповідно його реалізація) спрощується (відпадає необхідність у пристроях, що віднімають) і набуває такого вигляду:

$$\int_0^T z(t)S_i(t)dt > \int_0^T z(t)S_j(t)dt, \quad j \neq i, \quad (6.27)$$

або

$$\hat{i} = \text{Arg} \max_i \left\{ \int_0^T z(t)S_i(t)dt \right\}, \quad i = 0, 1, 2, \dots, m-1.$$

Із виразу (6.27) видно, що правило рішення не зміниться, якщо сигнал $z(t)$, який надходить на вхід демодулятора, помножити на будь-яке число. Тому система, у якій всі реалізації сигналу мають рівну енергію, відрізняється тим, що оптимальний алгоритм прийому в ній не вимагає знання "масштабу" сигналу, що приймається, або, інакше кажучи, знання коефіцієнта передачі

в каналі. Ця особливість обумовила надзвичайне поширення систем сигналів з рівною енергією, що важливо для каналів із завмираннями, у яких коефіцієнт передачі флукує.

Для двійкової системи нерівність (6.25) можна подати в більш простому вигляді:

$$\int_0^T z(t)S_{\Delta}(t)dt > \lambda, \quad (6.28)$$

де $S_{\Delta}(t)$ - різницевий сигнал, $S_{\Delta}(t) = S_1(t) - S_0(t)$;

λ - пороговий рівень, $\lambda = 0,5(E_1 - E_0)$.

Для реалізації виразу (6.28) у схемі рис. 6.2 потрібна лише одна гілка.

Для того щоб надати алгоритму оптимального прийому (6.24) наочного геометричного змісту, додамо в обох частинах

нерівності однакове значення $0,5 \int_0^T z^2(t)dt$. Тоді алгоритм (6.24)

матиме такий вигляд:

$$-0,5 \int_0^T [z(t) - S_i(t)]^2 dt > -0,5 \int_0^T [z(t) - S_j(t)]^2 dt, \quad j \neq i.$$

Перемножуючи ліву й праву частину нерівності на -2 (знак нерівності змінюється на зворотний) одержуємо алгоритм еквівалентний прийому (6.24):

$$\int_0^T [z(t) - S_i(t)]^2 dt > \int_0^T [z(t) - S_j(t)]^2 dt, \quad j \neq i. \quad (6.29)$$

Зазначимо, що саме у такому вигляді вперше одержав алгоритм оптимального прийому В.О. Котельников [29].

На рис. 6.3 для $m = 2$ наведена структурна схема приймального пристрою, що працює відповідно до алгоритму (6.29).

Схема складається з таких блоків: $(-)$ - пристрої, що віднімають; Γ_0, Γ_i - генератори опорних сигналів $S_0(t), S_i(t)$; $(\cdot)^2$ - квадратори; \int - інтегратори; ВП - вирішальний пристрій, що визначає у моменти часу, кратні T (при замиканні ключа К), номер гілки з мінімальним сигналом.

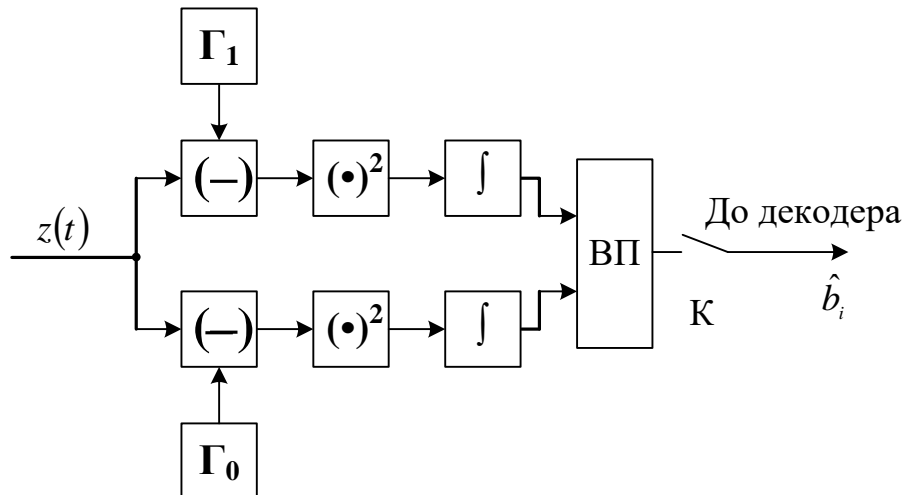


Рис. 6.3. Структурна схема оптимального приймального пристрою при точно відомих сигналах, що містять квадратори

У просторі Гілберта $\sqrt{\left(\int_0^T [z(t) - S_i(t)]^2 dt\right)}$ визначає норму різниці векторів z й S_i або відстань між ними. Тому алгоритм (6.29) можна записати як

$$\|z - S_i\| < \|z - S_j\|, \quad j \neq i$$

й додати йому просту геометричну інтерпретацію: оптимальний демодулятор повинен реєструвати той із сигналів $S_i(t)$ (відповідний символу b_i), що "ближче" до прийнятого колювання $z(t)$. Як приклад на рис. 6.4 показано оптимальний розподіл двовимірного простору прийнятих сигналів $z(t)$ при передачі сигналів $S_1(t)$ і $S_0(t)$. Області ухвалення рішення на користь символів 0 та 1 розташовані по обидва боки від прямої 0-0',

перпендикулярної відрізка, що з'єднує точки сигналів і ділить його навпіл.

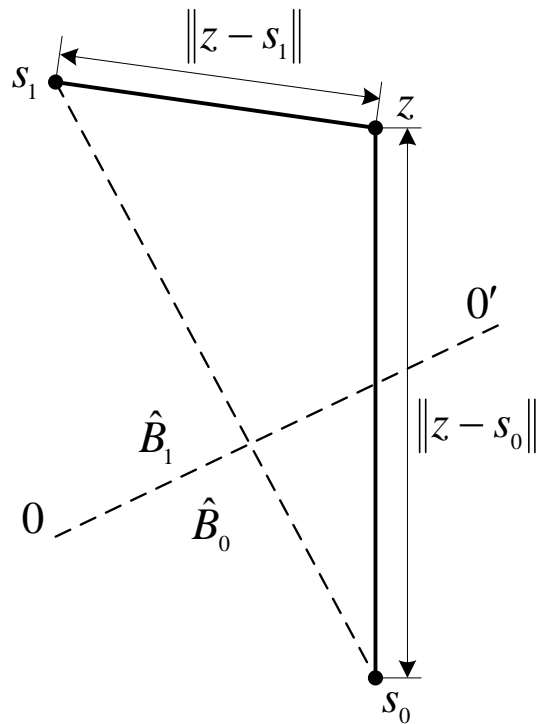


Рис. 6.4. Оптимальний розподіл простору прийнятих коливань при двійковому коді й точно відомих сигналах

Розглянемо випадок, коли гаусівський шум у каналі не "білий" і не "квазібілий", а *пофарбований*, тобто має нерівномірну щільність потужності $G(f)$ у смузі спектра сигналу. Пропустимо на вхід демодулятора суму сигналу й шуму через фільтр із передатною функцією $K(j2\pi f)$, таку, щоб у смузі спектра сигналу добуток $G(f)|K(j2\pi f)|^2$ був постійною величиною N . Із всіх можливих фільтрів з АЧХ, що задовольняють цю умову й розрізняються тільки фазочастотною характеристикою, можна вибрати мінімально фазовий, у якого $\varphi(\omega)$ пов'язана з натуральним логарифмом АЧХ $\ln K(\omega)$ парою перетворень Гілберта [4]:

$$\varphi(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\ln K(\xi)}{\omega - \xi} d\xi,$$

$$\ln K(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\varphi(\xi)}{\xi - \omega} d\xi.$$

Очевидно, що на виході фільтра шум виявиться "квазібілим": $G_{вих}(f) = N$. Тому такий фільтр називається фільтром, що *віблює*.

Сигнал $S_i(t)$ після проходження через фільтр, що *віблює*, перетвориться на деякий інший сигнал, позначений $S'_i(t)$. Вид його можна визначити, знаючи $S_i(t)$ і $K(j2\pi f)$. Якщо тепер подати коливання з виходу фільтра, що *обіляє*, на демодулятор, який є оптимальним для прийому сигналів $S'_i(t)$ ($i = \overline{0, m-1}$), то одержимо схему (рис. 6.5), що, мабуть, є оптимальною для сигналів $S_i(t)$ при пофарбованому шумі.

Слід звернути увагу на те, що в схемах рис. 6.3 та рис. 6.4 опорні сигнали повинні мати ті самі початкові фази, що й очікувані сигнали, які надходять або, інакше кажучи, повинні бути. Ця вимога звичайно ускладнює реалізацію демодулятора й вимагає введення в нього, крім зазначених на рисунках блоків, додаткових пристроїв, призначених для регулювання фаз опорних сигналів.

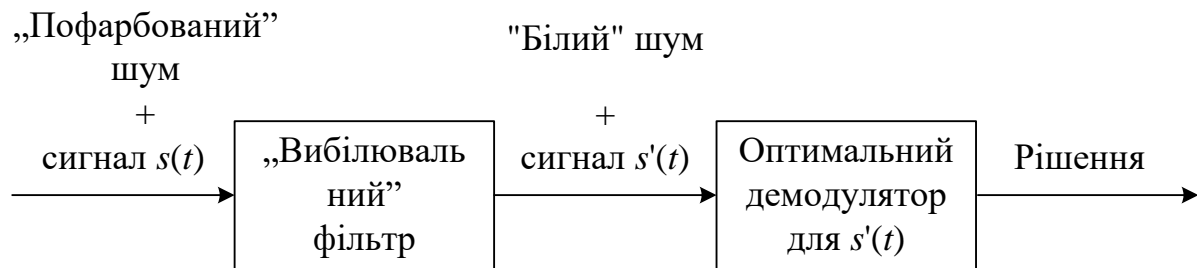


Рис. 6.5. Оптимальний демодулятор з віблювальним фільтром при гаусівському "пофарбованому" шумі

Всі методи прийому, для реалізації яких необхідне точне апріорне знання початкових фаз сигналів, що поступають, називають *когерентними*. У тих випадках, коли відомості про початкові фази очікуваних сигналів витягують із самого прийнятого сигналу (наприклад, якщо фаза флукує, але настільки повільно, що може бути завбачена за попередніми

елементами сигналу), прийом називають *квазікогерентним*. Якщо ж відомості про початкові фази сигналів, що приходять, відсутні або з будь-яких міркувань їх не використовують, то прийом називають *некогерентним*.

6.4. Прийом сигналів на узгоджений фільтр

Скалярний добуток виразу (6.26) можна обчислити не тільки за допомогою активного фільтра (корелятора), описаного в попередньому пункті, але й за допомогою пасивного лінійного фільтра з постійними параметрами. Якщо на вхід фільтра подати прийнятий сигнал $z(t)$, то напруга на виході фільтра в момент часу $t = T$:

$$y(T) = \int_0^T z(\tau)g(T - \tau)d\tau,$$

де $g(\tau)$ - імпульсна характеристика фільтра.

Оберемо її такою, щоб у момент $t = T$ одержати значення $y(T)$, що дорівнює скалярному добутку виразу (6.26). Легко побачити, що це буде виконано при наступному узгодженні $g(T - \tau) = S_i(\tau)$ або

$$g(t) = S_i(T - \tau). \quad (6.30)$$

У більш загальному випадку *узгодженим фільтром* для сигналу $S(t)$ називають *лінійний пасивний фільтр* з постійними параметрами та ІХ

$$g(t) = aS_i(t_0 - t), \quad (6.31)$$

де a, t - постійні.

Функція $g(t)$ є дзеркальним відображенням $S(t)$ щодо осі, проведеної через точку $t_0/2$ (рис. 6.6). Для фізичної реалізованості фільтра необхідно й досить, щоб $g(t) = 0$ при $t < 0$. Зокрема для фінітного сигналу $S(t)$, що надходить на вхід фільтра

в момент $t=0$ та закінчується в момент T , умова фізичної реалізованості узгодженого фільтра свідомо виконується, як видно з рис. 6.6, якщо постійна t_0 (момент відліку) задовольняє умову

$$t_0 - T \geq 0 \text{ або } t_0 \geq T. \quad (6.32)$$

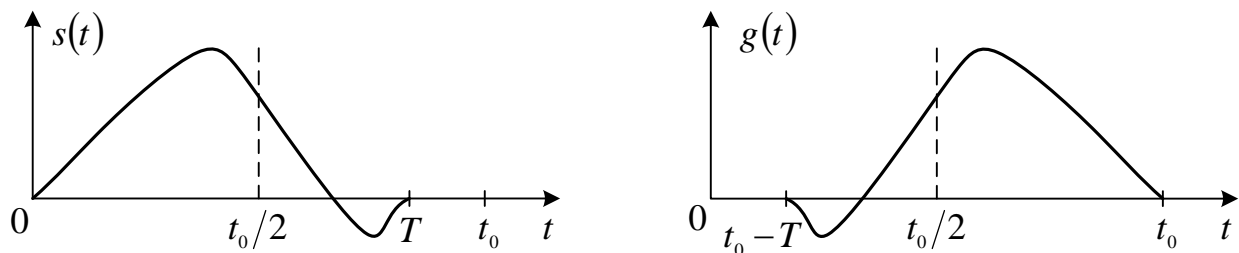


Рис. 6.6. Сигнал $S(t)$ та імпульсна характеристика $g(t)$ лінійного фільтра, узгодженого з цим сигналом

Передаточна функція (частотна характеристика) узгодженого фільтра з IX (з виразу (6.31)) визначається перетворенням Фур'є виразом [17].

$$\begin{aligned} K_{y\phi}(j\omega) &= \int_{-\infty}^{\infty} g(t)e^{-j\omega t} dt = a \int_{-\infty}^{\infty} S(t_0 - t)e^{-j\omega t} dt = \\ &= a \int_{-\infty}^{\infty} S(\tau)e^{-j\omega(t_0 - \tau)} dt = a\dot{S}(j\omega)e^{-j\omega t_0} \end{aligned} \quad (6.33)$$

де $\dot{S}(j\omega)$ - функція, комплексно-сполучена зі спектральною щільністю сигналу $S(t)$.

Отже, з точністю до коефіцієнта a АЧХ узгодженого фільтра визначається амплітудним спектром сигналу $S(t)$ (тобто фільтр краще передає ті частоти, які дають більший внесок в енергію сигналу), а його фазочастотна характеристика (без обліку доданка $-\omega t_0$, обумовленого затримкою t_0) зворотна за знаком фазовому спектру сигналу $S(t)$. Завдяки цьому в момент t_0 всі складові спектра прийнятого сигналу складаються у фазі й дають максимальний відгук.

Згідно з виразом (6.30) у момент часу T напруга на виході узгодженого фільтра пропорційна сигналу на виході інтегратора активного фільтра в схемі рис. 6.2. Тому демодулятор, що реалізує алгоритм (6.25), може бути виконаний і на базі узгоджених фільтрів. Структурна схема такого демодулятора для двійкової системи показана на рис. 6.7, де УФ — фільтр, узгоджений із сигналом $S_i(t)$.

Відгук узгодженого фільтра на фінітний сигнал тривалістю T , поданий до входу в момент часу $t = 0$, існує лише на фінітному інтервалі довжиною $2T$. Дійсно, якщо на вхід фільтра поданий сигнал, з яким він узгоджений, то сигнальна складова на виході узгодженого фільтра має вигляд

$$y_c(t) = \int_0^t S(t-\tau)g(\tau)d\tau = a \int_0^t S(t-\tau)S(t_0-\tau)d\tau = aB_s(t_0-t), \quad (6.34)$$

де $B_s(t_0-t)$ - ФК сигналу $S(t)$ при аргументі t_0-t .

Для фінітного сигналу вона визначена на інтервалі $(0, 2T)$ і має максимум у точці $t = t_0 = T$. Підкреслимо, що форми корисного сигналу на вході й виході узгодженого фільтра, як правило, істотно відрізняються один від одного. Завданням узгодженого фільтра є не відновлення форми сигналу, перекрученої шумом, а одержання одного відліку, за яким можна судити про присутність або відсутність на вході фільтра сигналу відомої форми.

Доведемо, що довільний лінійний стаціонарний фільтр із імпульсною характеристикою $g(t)$ дає на своєму виході в момент часу $t_0 \geq T$ значення відношення пікової потужності сигналу до дисперсії шуму:

$$\rho_{\text{пик}}(t_0) = \frac{y_c^2(t_0)}{y_n^2(t)} \leq 2h^2, \quad (6.35)$$

де $2h^2$ - значення $\rho_{nik}(t_0)$ для узгодженого фільтра ($h^2 = E/N_0$ - відношення енергії сигналу тривалості T до спектральної щільності шуму на позитивних частотах).

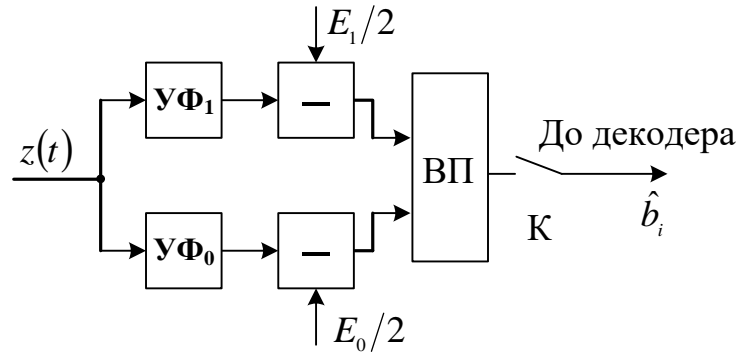


Рис. 6.7. Оптимальний демодулятор на основі узгодженого фільтра

Дійсно, нехай на вхід фільтра в момент $t = 0$ при нульових початкових умовах подається суміш $z(t) = s(t) + N(t)$. Тоді сигнальна та шумова складові на виході УФ мають вигляд

$$y_c(t_0) = \int_0^{t_0} S(\tau) g(t_0 - \tau) d\tau,$$

$$y_n(t_0) = \int_0^{t_0} n(\tau) g(t_0 - \tau) d\tau.$$

Якщо шум $n(\tau)$ центрований, стаціонарний й δ -корельований, то $\overline{y_n(t_0)} = 0$, тобто

$$\sigma_n^2 = \overline{y_n^2(t_0)} = \int_0^{t_0} \int_0^{t_0} \overline{n(\tau_1)n(\tau_2)} g(t_0 - \tau_1)g(t_0 - \tau_2) d\tau_1 d\tau_2 =$$

$$= \frac{N_0}{2} \int_0^{t_0} \int_0^{t_0} \delta(\tau_1 - \tau_2) g(t_0 - \tau_1)g(t_0 - \tau_2) d\tau_1 d\tau_2.$$

Враховуючи фільтруючу властивість δ -функції, маємо

$$\sigma_n^2 = \frac{N_0}{2} \int_0^{t_0} g^2(t_0 - \tau) d\tau = \frac{N_0}{2} E_g,$$

$$\text{де } E_g = \int_0^{t_0} g^2(t) dt.$$

Тоді отримуємо

$$\rho_{\text{ник}}(t_0) = \frac{2y_{c(t_0)}^2}{N_0 E_g} = \frac{2 \left| \int_{-\infty}^{\infty} S(\tau) g(t_0 - \tau) d\tau \right|^2}{N_0 E_g} \leq 2h^2. \quad (6.36)$$

Відповідно до нерівності Буняковського-Шварца маємо

$$\left| \int_{-\infty}^{\infty} S(\tau) g(t_0 - \tau) d\tau \right|^2 \leq \int_{-\infty}^{\infty} S(\tau) d\tau \int_{-\infty}^{\infty} g^2(t_0 - \tau) d\tau, \quad (6.37)$$

причому рівність має місце лише при виконанні умови (6.31), тобто для узгодженого фільтра. З урахуванням виразів (6.36) і (6.37) одержимо в результаті вираз (6.35).

Розглянемо можливості реалізації узгоджених фільтрів. Узгоджений фільтр для *фінитного* сигналу довільного виду $S(t)$ можна у принципі побудувати на основі довгої лінії, що не спотворює і забезпечує затримку сигналу на час T , з нескінченною щільністю відводів. Практично можна брати відводи в дискретних точках з рознесенням $\Delta = 0,5/\Delta F$, де ΔF — ефективна ширина спектра сигналу.

Дійсно, за допомогою схеми, показаної на рис. 6.8, можна із заданою точністю синтезувати будь-який сигнал $S(t)$, представлений усіченим рядом Котельникова [17],

$$S(t) = \sum_{k=0}^n a_k \frac{\sin[2\pi F(t - k\Delta)]}{2\pi F(t - k\Delta)},$$

де $a_k = S(k\Delta)$;

$$\Delta = \frac{1}{(2F)}.$$

Такий сигнал можна одержати на виході ідеального ФНЧ зі смугою пропускання ΔF , подаючи на його вхід послідовність δ -імпульсів з вагами $a_k' = a_k \Delta$ через інтервали часу Δ . Це здійснюється з відомим наближенням у схемі рис. 6.8. Якщо на вхід лінії в початковий момент подається один короткий імпульс, що апроксимує δ -функцію, то з відводів знімаються такі самі імпульси, рознесені на інтервали Δ , які, пройшовши через блоки, що зважують, a_k , надходять по черзі на вхід ФНЧ. Блоки, що зважують, містять атенюатори або підсилювачі з коефіцієнтом підсилення $|a_k'|$, а також при негативних a_k' - інвертори.

Схема на рис. 6.8 являє собою лінійний фільтр, називаний *трансверсальним*, з імпульсною реакцією $S(t)$. Якщо вхідний імпульс подати не на вхід A , а на вихід B , то буде синтезовано сигнал, що є дзеркальним відображенням $S(t)$. Тому та сама схема із входом у точці B виявляється фільтром, узгодженим з $S(t)$.

Існують й інші способи реалізації фільтра, точно або приблизно узгодженого із сигналом заданої форми.

Так, варіант фільтра, узгодженого із прямокутним радіоімпульсом $S(t) = a \cos(\omega_0 t + \varphi_0)$, заданим на інтервалі $(0, T)$, причому $\omega_0 T = 2\pi n$, n - ціле, показаний на рис. 6.9. Він складається з ідеального коливального контуру без втрат, настроєного на частоту ω_0 , і фазообертача, що зрушує фазу коливаний контуру на $-\varphi_0$. Лінія затримки на T у схемі й інвертор забезпечують згасання коливаний фільтра поза інтервалом $(0, T)$, імпульсна ж реакція фільтра на цьому інтервалі з урахуванням фазообертача $g(t) = a \cos(\omega_0(t - T) - \varphi_0) = a \cos(\omega_0(T - t) + \varphi_0) = a S(T - t)$, що й забезпечує узгодження виразу (6.31).

Схема з узгодженими фільтрами, на перший погляд, здається простішою за схему з активними фільтрами, оскільки в ній немає опорних генераторів і не виникає проблеми забезпечення їхньої когерентності (узгодження за фазою із сигналом, що надходить). Однак й у схемі з узгодженими

фільтрами є свої практичні труднощі. У цьому можна переконатися, порівнявши, наприклад, епюри напруг (без обліку завад у каналі) на виході фільтра (рис. 6.10, б), погодженого із прямокутним радіоімпульсом (рис. 6.10, а), і на виході інтегратора активного фільтра (рис. 6.10, в). Зазначимо, що всюди, за винятком точки $t = T$, напруги на виходах обох фільтрів відрізняються одна від одної.

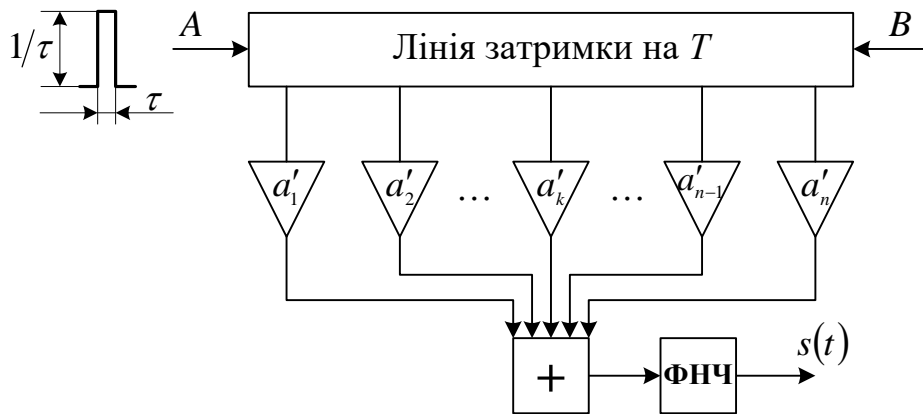


Рис. 6.8. Реалізація фільтра, узгодженого з довільним безперервним сигналом на основі лінії затримки з відводами та блоками зважування

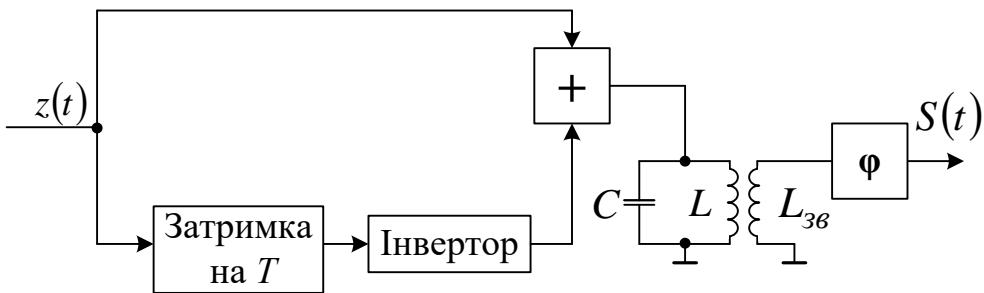


Рис. 6.9. Реалізація фільтра, узгодженого з прямокутним радіоімпульсом

З рис. 6.10 видно, що припустима неточність у часі Δt зняття відліку максимуму сигналу на виході активного фільтра значно більше, ніж при знятті відліку максимуму сигналу на виході узгодженого фільтра. При активному фільтрі досить зажадати,

щоб неточність узяття відліку була мала в порівнянні з тактовим інтервалом, а при узгодженому фільтрі - у порівнянні з періодом високочастотного заповнення радіоімпульсу (так званий *когерентний відлік*). Труднощі забезпечення когерентного відліку в узгодженому фільтрі цілком порівняна із труднощами реалізації *когерентних опорних генераторів* в активному фільтрі.

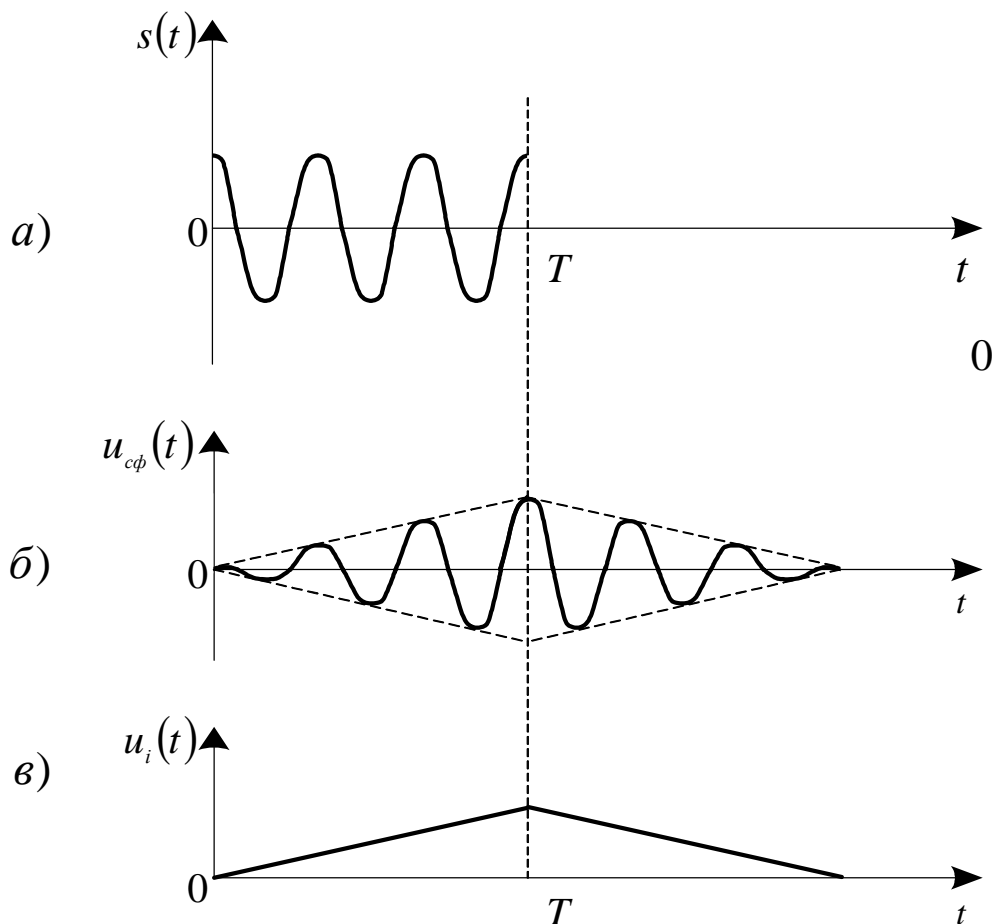


Рис. 6.10. Сигнали на виході узгодженого фільтра й кореляційної схеми при поданні на вхід прямокутного радіоімпульсу:
 а) імпульс на вході; б) імпульс на виході узгодженого фільтра;
 в) напруга на виході інтегратора активного фільтра

Зазначимо одну важливу обставину, яка властива узгодженому фільтру. Нехай на його вхід надходить коливання $z(t)$, а фільтр погоджений з фінітним сигналом $S(t)$. Тоді сигнал на його виході в момент часу t запишемо як

$$y_s(t) = \int_0^t z(\tau) S(\tau - (t - t_0)) d\tau. \quad (6.38)$$

Спектральна щільність, за Фур'є, цього сигналу дорівнює

$$\dot{S}_{y_s}(f) = S_s^*(f) \dot{S}_z(f) e^{-j\omega t_0}.$$

На виході фільтра, узгодженого зі сполученим сигналом $\widehat{S}(t)$, коливання $z(t)$ дає відгук

$$y_{\widehat{s}}(t) = \int_0^t z(\tau) \widehat{S}(\tau - (t - t_0)) d\tau. \quad (6.39)$$

Спектральна щільність, за Фур'є, цього сигналу дорівнює

$$\dot{S}_{y_{\widehat{s}}}(f) = j \operatorname{sign}(f) \dot{S}_s(f) \dot{S}_z(f) e^{-j\omega t_0}.$$

Таким чином, сигнали виразів (6.38) та (6.39) спряжені, за Гілбертом, з точністю до знаку. Обвідна на виході фільтра, узгодженого з сигналом $S(t)$ (або $\widehat{S}(t)$), описується формулою

$$V_s(t) = \sqrt{y_s^2(t) + y_{\widehat{s}}^2(t)}. \quad (6.40)$$

Якщо сигнал $S(t)$ отримує фазовий зсув θ , то маємо такі вирази:

$$\begin{aligned} y_{s,\theta}(t) &= y_s(t) \cos \theta - y_{\widehat{s}}(t) \sin \theta; \\ y_{\widehat{s},\theta}(t) &= y_{\widehat{s}}(t) \cos \theta + y_s(t) \sin \theta. \end{aligned}$$

Оскільки обвідна сигналу $S(t)$ від θ не залежить, то вираз буде мати такий вигляд:

$$V_{s,\theta}(t) = \sqrt{y_{s,\theta}^2(t) + y_{\hat{s},\theta}^2(t)} = \sqrt{y_s^2(t) + y_{\hat{s}}^2(t)} = V_s(t). \quad (6.41)$$

Якщо коливання $z(t)$ ортогональне $S(t)$ на інтервалі $(0, T)$, але не є ортогональним на цьому інтервалі $\hat{S}(t)$, то $y_s(T) = 0$, $\hat{y}_s(T) = 0$ та $V(T) \neq 0$. Обвідна на виході фільтра дорівнює нулю, якщо $z(t)$ та $S(t)$ ортогональні у посиленому розумінні.

6.5. Завадостійкість оптимального когерентного прийому

Визначимо потенційну завадостійкість для двійкової системи з адитивним БГШ у каналі, коли при прийомі точно відомі обидва очікуваних сигнали $S(t)$ і $S_0(t)$, вирішуючи, що апіорні ймовірності цих сигналів однакові. Сигнал $z(t)$, що надходить, є випадковим, тому що, по-перше, заздалегідь не відома реалізація переданого сигналу; по-друге, він містить випадкову заваду $N(t)$.

У цьому випадку згідно з виразом (6.25) алгоритм оптимального прийому буде мати вигляд

$$\int_0^T z(t)[S_1(t) - S_0(t)] dt > 0,5(E_1 - E_0), \quad (6.42)$$

де $E_i = \int_0^T S_i(t) dt$, $i = 0, 1$.

При виконанні нерівності (6.42) оптимальний приймач реєструє символ 1, що відповідає сигналу $S_1(t)$, в іншому випадку - символ 0, що відповідає сигналу $S_0(t)$. Якщо дійсно передається символ 1, то $z(t) = S_1(t) + N(t)$. При цьому ймовірність помилки визначається ймовірністю того, що нерівність (6.42) не виконана, тобто ймовірністю виконання зворотної нерівності

$$\int_0^T S_1(t)[S_1(t) - S_0(t)] dt + \int_0^T N(t)[S_1(t) - S_0(t)] dt < 0.5 \int_0^T [S_1^2(t) - S_0^2(t)] dt,$$

яку легко привести до такого вигляду:

$$\int_0^T N(t)[S_1(t) - S_0(t)] dt < -0.5 \int_0^T [S_1(t) - S_0(t)]^2 dt. \quad (6.43)$$

Аналогічне співвідношення набуває вигляду, якщо припустити, що передається символ 0. Отже, в обох випадках імовірності помилки $p(0|1) = p(1|0) = p$ і сформований модемом двійковий дискретний канал є симетричним.

Запишемо вираз (6.43) як

$$\xi < -0.5E, \quad (6.44)$$

де $\xi = \int_0^T N(t)S_{\Delta}(t)dt;$

$$E = \int_0^T S_{\Delta}(t)dt;$$

$$S_{\Delta}(t) = S_1(t) - S_0(t).$$

Якщо $N(t)$ - нормальний стаціонарний "білий" шум з нульовим середнім й однобічною спектральною щільністю потужності N_0 , то ξ - нормально розподілена величина (оскільки вона визначається лінійною операцією над нормальним випадковим процесом). Її математичне очікування

$$\overline{\xi} = \int_0^T \overline{N}(t)S_{\Delta}(t)dt = 0,$$

а дисперсія

$$\begin{aligned}
D(\xi) &= \overline{\xi^2} = \int_0^T \int_0^T \overline{N(t_1)N(t_2)} S_{\Delta}(t_1)S_{\Delta}(t_2) dt_1 dt_2 = \\
&= \frac{N_0}{2} \int_0^T \int_0^T \delta(t_1 - t_2) S_{\Delta}(t_1)S_{\Delta}(t_2) dt_1 dt_2 = \frac{N_0}{2} \int_0^T S_{\Delta}^2(t) dt = \quad (6.44, a) \\
&= 0,5N_0E_0.
\end{aligned}$$

Тому імовірність виконання нерівності (6.44), тобто імовірність помилки, буде визначатися з виразу

$$\begin{aligned}
p &= \int_{-\infty}^{-0,5E_e} W_1(\xi) d\xi = \frac{1}{\sqrt{2\pi D(\xi)}} \int_{-\infty}^{-0,5T_e} \exp\left[\frac{-\xi^2}{2D(\xi)}\right] d\xi = \\
&= \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{-\frac{r^2}{2}} dt, \quad (6.45)
\end{aligned}$$

де здійснено заміну змінної $t = -\frac{\xi}{\sqrt{D(\xi)}}$ та введено позначення

$$\alpha = \frac{0,5E_e}{\sqrt{D(\xi)}} = \sqrt{\frac{E_e}{2N_0}}. \quad (6.46)$$

Функція $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{r^2}{2}} dt$ табульована та названа

додатковою функцією помилок (див. дод. 1). Через Q -функцію можна вираз (6.45) записати як

$$p = Q\left(\sqrt{\frac{E_e}{2N_0}}\right). \quad (6.47)$$

При заданій інтенсивності завади N_0 потенційна заводостійкість двійкової системи залежить лише від так названої *еквівалентної енергії сигналів*

$$E_e = \int_0^T (S_1(t) - S_0(t))^2 dt = d_{S_1 - S_0}^2, \quad (6.48)$$

яка дорівнює квадрату відстані між сигнальними точками в просторі Гілберта. Завадостійкість вище (ймовірність помилки менше) тієї системи, у якій більша еквівалентна енергія використовуваних сигналів незалежно від форми використовуваних сигналів. Останні, зокрема, можуть бути як простими (відірзками синусоїди з малою базою), так і складними (шумоподібними, з великою базою).

На рис. 6.11 у двовимірному просторі показані точки сигналів для двійкової системи: рис. 6.11, а - АМн при $S_0(t) = 0$,

$\int_0^T S_1^2(t) dt = E$; рис. 6.11, б - ЧМн із ортогональними сигналами

$\int_0^T S_1(t)S_0(t) dt = 0$, $\int_0^T S_1^2(t) dt = \int_0^T S_0^2(t) dt = E$; рис. 6.11, в - ФМн з

протилежними сигналами $S_1(t) = -S_0(t)$.

З рис. 6.11 видно, що в порівнянні із двійковою АМн для двійкової ЧМн еквівалентна енергія сигналу $E_e = \|S_1 - S_0\|^2$ в 2 рази більше, а для двійкової ФМн - у 4 рази більше.

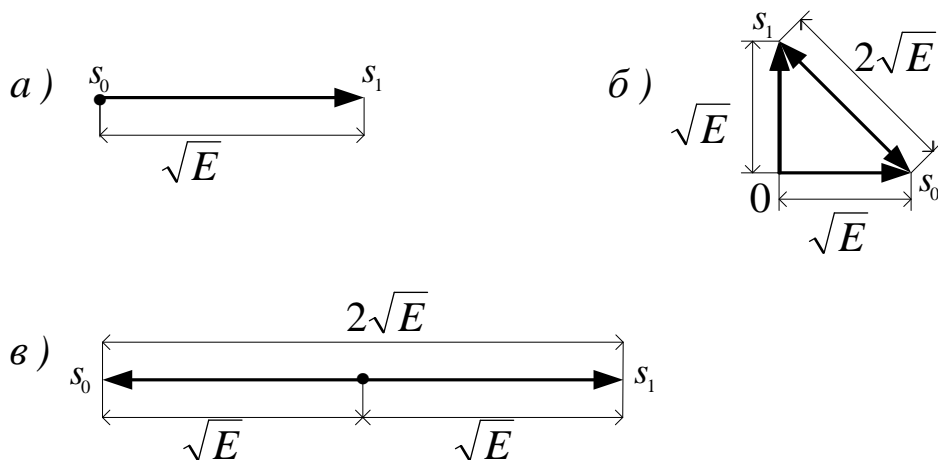


Рис. 6.11. Еквівалентні енергії двійкових систем АМн, ЧМн, ФМн

Співвідношення (6.48) дозволяє здійснювати оптимальний вибір сигналів $S_1(t)$ та $S_0(t)$ або відповідно $u_1(t)$ і $u_0(t)$, що забезпечують максимально можливу завадостійкість при заданій енергії сигналів E . Справді, для такої оптимальної системи величина E_e повинна бути максимальною за умови, що

$$E_1 = \int_0^T S_1^2(t) dt > E; \quad E_2 = \int_0^T S_0^2(t) dt \leq E. \quad (6.49)$$

Можна написати $E_e = 2E_1 + 2E_0 - \int_0^T (S_1(t) + S_0(t))^2 dt$. Для одержання максимуму цього виразу потрібно зробити E_1 й E_0 можливо більшими, а інтеграл у правій частині – якомога меншим. Максимально можливі значення E_1 й E_0 будуть, якщо з огляду на умови (6.49) дотримано рівняння

$$E_1 = E_0 = E. \quad (6.50)$$

Інтеграл $\int_0^T (S_1(t) + S_0(t))^2 dt$ приймає тільки ненегативні значення, тому його мінімум дорівнює нулю й досягається за умови $S_1(t) = -S_0(t)$, що не суперечить умові (6.50). Таким чином, у двійковому каналі з постійними параметрами й адитивним БГШ оптимальною виявляється система із *протилежними сигналами*. Цій умові задовольняють, наприклад, двополярні імпульси, сигнали двійкової фазової модуляції (ФМр), якщо різниця фаз сигналів $\Delta\varphi = \pi$ і т. п.

Для всіх таких систем $E_e = 4E$ й ймовірність помилки дорівнює

$$p = Q\left(\sqrt{\frac{2E}{N_0}}\right) = Q(\sqrt{2h}), \quad (6.51)$$

де h^2 - відношення енергії сигналу на вході демодулятора до спектральної щільності потужності флуктуаційної завади, $h^2 = \frac{E}{N_0}$.

Для системи з ортогональними сигналами рівної енергії (наприклад, при відомих умовах для системи двійкової ЧМн), коли $\int_0^T S_1(t)S_0(t) dt$, $E_e = 2E$, мінімальна ймовірність помилки дорівнює

$$p = Q(h). \quad (6.52)$$

Порівнюючи вирази (6.51) і (6.52), дійдемо висновку, що перехід від системи з ортогональними сигналами до системи з оптимальними (протилежними) сигналами дозволяє в розглянутому каналі забезпечити незмінну якість зв'язку (ймовірність помилки) при зниженні середньої потужності передавача в 2 рази, тобто дає енергетичний виграш в 2 рази (або на 3 дБ). Цей висновок зроблено також з рис. 6.11.

У двійковій системі з пасивною паузою, розуміючи $S_0(t) = 0$ і $\int_0^T S_1^2(t) dt = E$, одержуємо для мінімальної ймовірності помилки

$$p = Q\left(\frac{h}{\sqrt{2}}\right). \quad (6.53)$$

Звідси видно, що при переході від системи АМн до системи ЧМн енергетичний виграш за максимальною потужністю дорівнює 2, а при переході до системи ФМ - 4. Якщо ж порівняння проводити не за піковою, а за середньою потужністю, то перехід від АМн до ЧМн не дає енергетичного виграшу, оскільки при ЧМн середня потужність дорівнює максимальній, а при АМн - удвічі менше за максимальну, якщо сигнали $S_1(t)$ та $S_0(t)$ передаються з однаковою ймовірністю).

Система ФМн, як й інші системи із протилежними сигналами, забезпечує максимальну для двійкової системи потенційну завадостійкість. Однак реалізація демодулятора для когерентного прийому ФМн - сигналів має певні труднощі. При

побудові демодулятора з активним фільтром виникає проблема підтримки рівності фаз опорного генератора й сигналу, що приходить. Якщо намагатися будувати його на основі узгодженого фільтра (див. рис. 6.7), то виникає не менш важке завдання узяття когерентного відліку.

У практичних схемах опорний сигнал $S_I(t)$ формується із прийнятого коливання. Для цього необхідно за прийнятим сигналом відновити немодульований гармонійний сигнал $\cos(\omega_0 t + \varphi_0)$.

Завдання виділення опорного сигналу особливо ускладнюється при ФМн, тому що якщо елементи $S_I(t)$ і $S_Q(t) = S_0(t)$ передаються рівноймовірно, то спектр сигналу ФМн взагалі не містить складової на частоті ω_0 . Для його одержання доводиться використовувати нелінійні пристрої зняття модуляції. Це досягається різними схемами, наприклад, схемою, запропонованою О.О. Пістолькорсом. Схема містить множник частоти на 2, вихідний сигнал якого через вузькосмуговий фільтр, який настроєно на частоту $2\omega_0$, надходить на дільник частоти на 2. Якщо сигнал на вході множника записати як $u_1 = \cos(\omega_0 t + k\pi + \varphi_0)$, $k = 0$ або $k = 1$, то сигнал на виході множника $u_2 = \cos(2\omega_0 t + 2\varphi_0)$, а сигнал на виході дільника $u_3 = \cos(\omega_0 t + \varphi_0)$.

Однак всі схеми формування опорного сигналу такі, що внаслідок різних неконтрольованих факторів можливі випадкові зміни знака опорного сигналу. Це, зокрема, стосується й дільника частоти на 2 у схемі О.О. Пістолькорса, оскільки ця операція неоднозначна - фаза вихідного сигналу дільника може прийняти кожне із двох значень: $(2\varphi_0 + 2\pi) / 2 = \varphi_0 + \pi$. Це означає, що символи, які реєструються на виході приймача, навіть за відсутності адитивної завади в каналі $[z(t) = S_i(t)]$ після випадкового перескоку фази опорного сигналу інвертуються (нулі будуть записані як 1, а 1 - як 0). Це буде тривати до наступного перескоку фази опорного сигналу. Виникає так називане явище *зворотної роботи*, внаслідок якого практичне впровадження системи із двійковою фазовою модуляцією виявилось ускладненим.

Ефективний метод усунення цього явища був знайдений шляхом переходу до *відносних методів модуляції*, запропонованим М.Т. Петровичем [48]. Вони зводяться до модуляції інформаційного параметра переданої послідовності елементів сигналу щодо того самого параметра попередньої послідовності. При *відносній фазовій маніпуляції* (ВФМн) повідомлення втримується не в абсолютному значенні фази елементів сигналу, а в різниці фаз двох сусідніх елементів, при цьому символ 1 передається повторенням тієї реалізації сигналу, яка і мала місце в якості попереднього елементу, а символ 0 - реалізацією зі зворотною фазою, або навпаки.

Сигнали ВФМн можуть прийматися різними методами. Розглянемо квазікогерентний прийом сигналів ВФМн (названий *методом порівняння полярностей*). Систему ВФМн можна розглядати як звичайну систему з фазовою модуляцією (ФМн), але зі спеціальним перекодуванням символів. Це означає, що оптимальний прийом сигналів ВФМн можна здійснити з перекодуванням прийнятих символів. Перекодування виконується порівнянням полярностей напруг на виході інтегратора для двох сусідніх елементів, для чого потрібна затримка вихідних символів у комірці пам'яті (КП) на час T .

Така схема демодулятора зображена на рис. 6.12 (без пристрою підстроювання фази опорного генератора Γ , що може бути виконано, наприклад, за схемою Пістолькорса). Оскільки ВФМн - система з рівною енергією окремих позицій, то граничний рівень у демодуляторі нульовий і вирішальний пристрій перетворюється на дискримінатор полярності (ДП). Полярності сусідніх елементів порівнюються в схемі порівняння полярностей (СПП). Символ 1 реєструється на виході приймача, наприклад, при збігу полярностей двох сусідніх послідовностей, символ 0 - якщо ці полярності різні, або навпаки. При такому методі прийому перескок фази опорного сигналу (за відсутності завади у каналі) викликає помилку тільки в одному символі. Наступні ж символи реєструються правильно, тобто явище "зворотної роботи" не виникає.

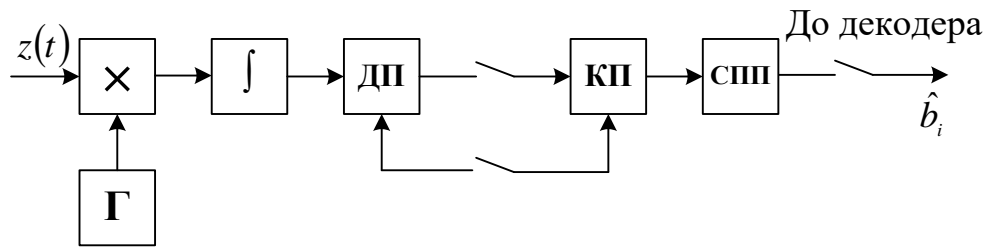


Рис. 6.12. Схема оптимального прийому сигналів ВФМн методом порівняння полярностей (когерентний прийом)

Визначимо ймовірність помилки в системі ВФМн при обліку флуктуаційної завади в каналі при когерентному прийомі. Очевидно, що помилкова реєстрація символу при прийомі методом порівняння полярностей можлива в результаті однієї із двох несумісних подій;

а) знак даного елемента прийнято помилково, а знак попереднього - правильно;

б) знак даного елемента прийнято правильно, а знак попереднього - помилково.

Кожне із цих подій має ймовірність $p_{\text{фм}}(1 - p_{\text{фм}})$. Таким чином, одержуємо

$$p_{\text{ВФМ}} = 2p_{\text{фм}}(1 - p_{\text{фм}}). \quad (6.54)$$

В умовах експлуатації, коли потрібно $p_{\text{фм}} \ll 1$, маємо

$$p_{\text{ВФМ}} = 2p_{\text{фм}} = 2Q[\sqrt{2h}].$$

Таким чином, "платою" за усунення зворотної роботи є подвоєння ймовірності помилки, обумовленої шумом у каналі.

Очевидно, що при розглянутому методі прийому сигналів ВФМн дискретний канал, що утворюється, є марковським. Ймовірність помилки в ньому залежить від того, правильно чи помилково прийняті попередні символи. Переважна більшість помилок групується по дві.

Для недвійкових систем ($m > 2$) знаходження ймовірності помилкового прийому p_m у загальному випадку ускладнюється, тому що тепер доводиться аналізувати сукупність із $(m - 1)$ нерівностей (6.24).

Ймовірність помилки в m -ній системі при передачі символу b_i визначається ймовірністю об'єднання подій формулою

$$P_m(\text{ном}|b_i) = P \left(\bigcup_{\substack{j=0 \\ i \neq j}}^{m-1} (A_j \cap A_i) \right).$$

При оптимальному когерентному прийомі в каналі із БГШ маємо

$$A_j = \int_0^T (N(t) + S_i(t)) S_j(t) dt - 0,5E_j.$$

Для ортогональної системи сигналів з рівною енергією (система виявляється також *еквідистантною* — всі сигнальні точки перебувають на однаковій відстані одна від одної) ймовірність помилки p_m (однакова при передачі будь-якого символу) виражається таким інтегралом [28]:

$$p_m = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} [1 - Q(u)]^{m-1} \exp \left[-0,5(u - \sqrt{2h})^2 \right] du. \quad (6.55)$$

Згідно з виразом (6.55) ймовірність помилки монотонно падає зі зростанням енергетичного параметра h^2 . При $m = 2$ із виразу (6.55) випливає вираз (6.52).

У детермінованому неспотворювальному каналі із БГШ оптимальною (яка забезпечує мінімальну ймовірність помилки при заданому значенні h^2) виявляється *еквідистантна система*, сигнальні точки якої утворюють *багатомірний симплекс* (вони лежать на однаковій відстані одна від одної в $(m - 1)$ -мірній *гіперсфері*). Ймовірність помилки для цієї оптимальної системи сигналів можна в області малих помилок також визначити формулою (6.55).

Будемо враховувати, що ймовірність об'єднання подій (у загальному випадку сумісних) дорівнює сумі ймовірностей окремих подій мінус імовірності сполучення подій, тоді можна

одержати просту формулу (адитивну верхню границю для ймовірності помилки або нерівність Буля)

$$P_m(\text{ном}|b_i) \leq \sum_{\substack{j=0 \\ j \neq i}}^{m-1} P[A_j > A_i], \quad (6.56)$$

де $P[A_j > A_i]$ - імовірність помилки у двійковій системі із сигналами $S_i(t)$ і $S_j(t)$.

Для систем рівноймовірних ортогональних сигналів рівної енергії канал симетричний і можна оцінити ймовірність помилки простою нерівністю

$$P_m \leq (m-1)Q(\sqrt{h^2}). \quad (6.57)$$

Як випливає з виразу (6.57), імовірність помилки залежить від додаткової функції помилок.

6.6. Прийом сигналів з невизначеною фазою (некогерентний прийом)

Багато каналів можна описати моделлю із флюктуючою фазою. Природно, якщо фаза (або який-небудь інший параметр, що визначає апіорну інформацію, необхідну для когерентного прийому) прийнятого сигналу флюктує настільки повільно, що шляхом вимірювання (оцінювання) її можна досить надійно передбачити, оптимальний прийом загалом реалізується так само, як при точно відомому сигналі (з додаванням блоків оцінювання). Така ситуація характерна для багатьох каналів провідного і рідше радіозв'язку. Однак нерідко фаза флюктує досить швидко, і точну її оцінку одержати не вдається. Крім того, оцінка фази вимагає іноді застосування складних пристроїв. Тому навіть у тих випадках, коли принципово можна оцінити початкову фазу сигналу, що надходить, часом від цього відмовляються й використовують алгоритм, побудований з припущенням, що початкова фаза сигналу, яка надходить, не відома й може

приймати будь-яке значення на інтервалі $(0, 2\pi)$. Такий метод прийому є *некогерентним*. Алгоритм оптимального некогерентного прийому вперше отриманий Л.М. Фінком [48].

Для виведення правила оптимального некогерентного прийому будемо виходити з логарифма відношення правдоподібності $\ln \Lambda_i$ для сигналу $S_i(t)$, яке при точно відомій початковій фазі виражається формулою (6.22). Використовуючи подання сигналу із формули (6.20), формулу для $\ln \Lambda_i$ можна (після наближення у виразі (6.22) $\Delta t \rightarrow 0$) записати так:

$$\begin{aligned} \ln \Lambda_i = & \frac{2\gamma}{N_0} \int_0^T z(t) [\cos \theta u_i(t) - \sin \theta \hat{u}_i(t)] dt - \\ & - \frac{\gamma^2}{N_0} \int_0^T [\cos \theta u_i(t) - \sin \theta \hat{u}_i(t)]^2 dt, \end{aligned} \quad (6.58)$$

де θ - випадковий зсув в каналі.

Тут $\ln \Lambda_i$ є випадковою величиною, що приймає різні значення при різних θ . Правило максимуму правдоподібності в такій ситуації полягає у виборі такого рішення, для якого математичне очікування Λ_i , - буде найбільшим. Таке правило оцінки символу \hat{b}_i , якщо зсув фази θ рівномірно розподілено на інтервалі $(0 - 2\pi)$, коротко записується так:

$$\begin{aligned} \hat{i} = & \text{Arg max}_i \{ \overline{\Lambda}_i(\theta) \} = \text{Arg max}_i \left\{ \int_0^{2\pi} \Lambda_i(\theta) w(\theta) d\theta \right\} = \\ = & \text{Arg max}_i \left\{ \frac{1}{2\pi} \int_0^{2\pi} \Lambda_i(\theta) d\theta \right\}, \end{aligned}$$

де $w(\theta)$ - щільність розподілу ймовірності θ , $w(\theta) = \frac{1}{2\pi}$ при $0 \leq \theta \leq 2\pi$.

При знаходженні $\bar{\Lambda}_i(\theta)$ помітимо, що другий інтеграл у правій частині виразу (6.58) від θ не залежить і дорівнює енергії $E_{\Pi i}$ сигналу $u_i(t)$ на вході каналу (на передатчику). Це зрозуміло з того, що підінтегральною функцією є квадрат сигналу $u_i(t)$, зсунутого по фазі на θ , що, як відомо, не впливає на його енергію. Таким чином, враховуємо, що $\frac{\gamma^2 E_{\Pi i}}{N_0} = h_i^2$, одержуємо

$$\bar{\Lambda}_i(\theta) = \exp(-h_i^2) \frac{1}{2\pi} \int_0^{2\pi} \exp\left\{\frac{2}{N_0} [\cos\theta y_i - \sin\theta \hat{y}_i]\right\} d\theta,$$

де введено позначення

$$y_i = \gamma \int_0^T z(t) u_i(t) dt; \quad \hat{y}_i = \gamma \int_0^T z(t) \hat{u}_i(t) dt.$$

Позначивши

$$V_i = \sqrt{y_i^2 + \hat{y}_i^2} \quad \text{та} \quad \theta_i = \arctg\left(\frac{\hat{y}_i}{y_i}\right), \quad (6.59)$$

можна одержати

$$\begin{aligned} \bar{\Lambda}_i(\theta) &= \exp(-h_i^2) \frac{1}{2\pi} \int_0^{2\pi} \exp\left\{\frac{2}{N_0} V_i [\cos(\theta + \theta_i)]\right\} d\theta = \\ &= \exp(-h_i^2) \frac{1}{2\pi} I_0\left(\frac{2V_i}{N_0}\right), \end{aligned} \quad (6.60)$$

де $I_0(x)$ - модифікована функція Бесселя нульового порядку,

$$I_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp[x \cos(\varphi + \theta)] d\varphi.$$

Замість того щоб порівнювати відношення правдоподібності, можна порівняти їх логарифми, що призведе до такого правила (алгоритму):

$$\hat{i} = \underset{i}{\text{Arg max}} \left\{ \ln I_0 \left(\frac{2V_i}{N_0} \right) - h_i^2 \right\}. \quad (6.61)$$

Для двійкової системи сигналів правило оптимального некогерентного прийому буде виражено нерівністю

$$\ln I_0 \left(\frac{2V_1}{N_0} \right) - h_1^2 \geq \ln I_0 \left(\frac{2V_0}{N_0} \right) - h_0^2. \quad (6.62)$$

При виконанні цієї нерівності реєструється 1, в іншому випадку - 0.

Величини y_i й \hat{y}_i можна одержати в момент відліку T на виході активного фільтра з опорними сигналами, що дорівнюють відповідно $u_i(t)$ і $\hat{u}_i(t)$. З урахуванням сказаного зрозуміло побудову на основі активних фільтрів схеми (рис. 6.13), яка названа квадратурною і реалізує алгоритм (6.62).

На цій схемі блоки Γ_0, Γ_1 - відповідно генератори опорних сигналів $u_0(t), u_1(t)$; $\varphi_{\pi/2}$ - фазообертач всіх сигнальних компонентів на $\pi/2$ (перетворювач Гілберта); БВМ - блок визначення модуля вектора $V_i = |y_i + j\hat{y}_i|$ за ортогональними компонентами; НП — нелінійні безінерційні пристрої з характеристикою

$$u_{\text{вих}} = \ln I_i \left(\frac{2}{N_0} u_{\text{вх}} \right).$$

Підкреслимо, що величини V_i не залежать від початкової фази сигналів $u_i(t)$ і, як видно з виразу (6.59), пропорційні, обвідної (у моменти відліку, кратні T) на виході фільтра, узгодженого із сигналом $u_i(t)$.

Таким чином, алгоритм (6.62) можна реалізувати й на базі узгоджених фільтрів, як показано на рис. 6.14. Ідеальний детектор виділяє обвідну напруги на виході узгодженого фільтра.

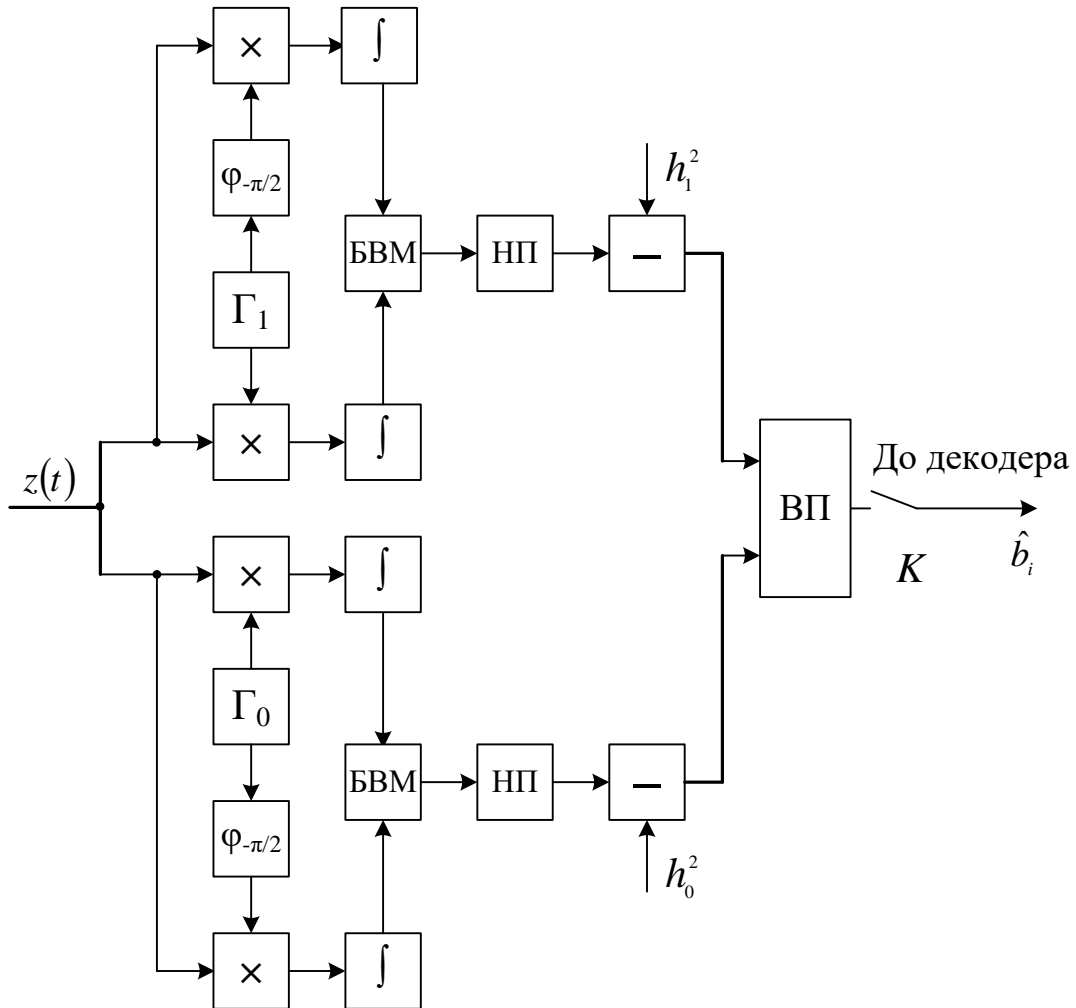


Рис. 6.13. Квадратурна схема реалізації оптимального прийому дискретних повідомлень при невизначеній фазі сигналу

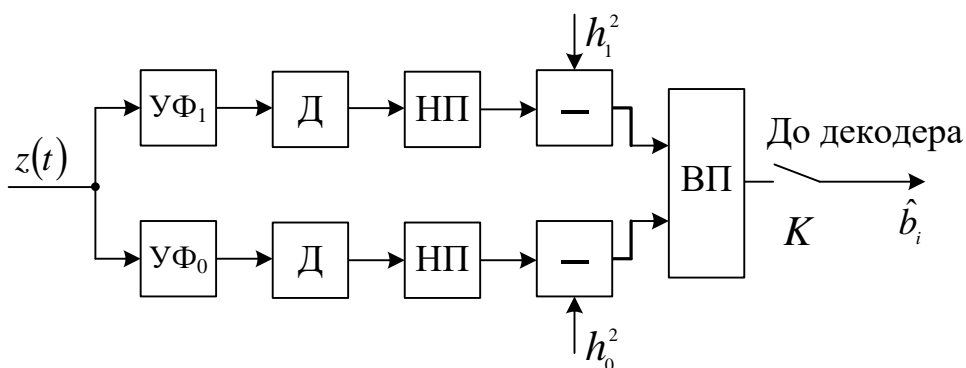


Рис. 6.14. Схема реалізації оптимального прийому дискретних повідомлень на базі узгоджених фільтрів при невизначеній фазі сигналу

Для двійкової системи з пасивною паузою, припускаючи, що символ 0 передається сигналом $u_0(t) = 0$, правило (6.61) можна записати як

$$V_1 > \lambda, \quad (6.63)$$

де пороговий рівень

$$\lambda = \frac{N_0}{2} f(h^2),$$

а функція $x = f(y)$ зворотна функції $y = \ln I_0(x)$. При виконанні нерівності (6.63) (перевищення V_1 над порогом) реєструється символ 1, в іншому випадку - символ 0.

Алгоритм (6.61) і відповідно його реалізація істотно спрощуються для системи сигналів з рівною енергією ($E_i = E$). Для них з урахуванням монотонного характеру функції $\ln I_0(x)$ алгоритм оптимального некогерентного прийому можна записати так:

$$V_i > V_j, j \neq i, \quad \hat{i} = \text{Arg max}_i \{V_i\}, i = \overline{0, m-1}. \quad (6.64)$$

Для двійкової системи правило (6.64) зводиться до перевірки однієї нерівності

$$V_1 > V_0. \quad (6.65)$$

При її виконанні реєструється символ 1, в іншому випадку – символ 0.

При реалізації алгоритму (6.65) у схемах рис. 6.13 й 6.14 не потрібні блоки НП та блоки вирахування. Більш того, алгоритм за цієї умови інваріантний щодо коефіцієнта передачі каналу й

спектральної щільності шуму N_0 , оскільки V_i не залежить від N_0 , зі зміною γ всі значення V_i змінюються пропорційно, що не впливає на вираз (6.63). Саме це є основною перевагою систем з рівною енергією сигналів, що визначило їх широке застосування.

Дослідження ймовірності помилок у каналі з невизначеною фазою й адитивним гаусівським шумом при поелементному прийомі показало [25], що мінімальну ймовірність помилки забезпечує система сигналів з рівною енергією, що задовольняє умовам ортогональності в посиленому змісті:

$$\int_0^T u_i(t)u_j(t)dt = 0, \quad i \neq j;$$

$$\int_0^T u_i(t)\hat{u}_j(t)dt = 0, \quad \text{при } i, j = 0, 1, \dots, m-1. \quad (6.66)$$

Це можна пояснити тим, що за відсутності шуму й передачі сигналу $u_i(t)$, що задовольняє умови (6.66), праві частини (V_j) в алгоритмі (6.64) приймають мінімально можливі значення (дорівнюють нулю), у той час як $V_i = E$ (формули (6.67) і (6.68)). Якщо шум стаціонарний і гаусівський, то найбільш ймовірними є малі відхилення V_j й V_i від зазначених значень. А це означає, що для системи сигналів вираз (6.66) нерівності (6.64) при передачі сигналу $u_i(t)$ виконуються майже завжди, що призводить до мінімізації ймовірності помилки.

Визначимо ймовірність помилки при прийомі за алгоритмом (6.65) двійкових сигналів, що задовольняють умови (6.66). Якщо передається символ 1, то з урахуванням виразів (6.20), (6.58) і (6.59) маємо

$$V_1 = \sqrt{(\xi_1 + E \cos \theta)^2 + (\hat{\xi}_1 + E \sin \theta)^2}, \quad (6.67)$$

$$V_0 = \sqrt{\xi_0^2 + \hat{\xi}_0^2}, \quad (6.68)$$

де

$$\xi_j = \gamma \int_0^T N(t) u_j(t) dt; \quad \hat{\xi}_j = \gamma \int_0^T N(t) \hat{u}_j(t) dt; \quad j = 0, 1. \quad (6.69)$$

Міркуючи так само, як при виведенні виразів (6.44) і (6.44 а), легко переконатися, що величини $\xi_j, \hat{\xi}_j$ розподілені нормально з нульовим середнім значенням і дисперсією $0,5N_0E$.

Легко також показати, що коефіцієнти кореляції $\overline{\xi_1 \hat{\xi}_0}$ й $\overline{\hat{\xi}_1 \xi_0}$ при системі сигналів, ортогональній в посиленому значенні, дорівнюють нулю. Некорельованість гаусівських величин забезпечує і їх незалежність. Як наслідок, випадкові величини V_1 й V_0 незалежні, причому для V_0 маємо розподіл Релея

$$w(V_0) = \frac{2V_0}{N_0E} \exp\left(-\frac{V_0^2}{N_0E}\right), \quad V_0 \geq 0, \quad (6.70)$$

а також узагальний розподіл Релея (Райса)

$$w(V_1) = \frac{2V_1}{N_0E} \exp\left(-\frac{V_1^2 + E^2}{N_0E}\right) I_0\left(\frac{2V_1E}{N_0E}\right), \quad V_1 \geq 0. \quad (6.71)$$

Імовірність прийому символу 0 при передачі символу 1 визначається ймовірністю виконання нерівності

$$\begin{aligned}
P(\hat{0}|1) &= \int_0^\infty w(V_1) \int_{V_1}^\infty w(V_0) dV_0 dV_1 = \\
&= \int_0^\infty w(V_1) \int_{V_1}^\infty \frac{2V_0}{N_0 E} \exp\left(-\frac{V_0^2}{N_0 E}\right) dV_0 dV_1 = \\
&= \int_0^\infty w(V_1) \exp\left(-\frac{V_1^2}{N_0 E}\right) dV_1 = \\
&= 0,5 \exp\left(-\frac{E}{2N_0}\right) \int_0^\infty \frac{4V_1}{N_0 E} \exp\left[-\frac{4V_1^2 + E^2}{2N_0 E}\right] I_0\left(\frac{2V_1}{N_0}\right) dV_1.
\end{aligned} \tag{6.72}$$

В отриманій формулі інтеграл легко обчислити за допомогою теорії ймовірностей. Якщо замінити змінну, $2V_1 = x$, то з'ясується, що це інтеграл від щільності ймовірностей випадкової величини $A \geq 0$, що має узагальнений розподіл Релея, з параметрами $a = E$ та $\sigma^2 = N_0 E$. Оскільки інтеграл береться по всій області визначення A , то він дорівнює 1. Остаточно

$$P(0|1) = 0,5 \exp(-0,5h^2), \tag{6.73}$$

де, як і раніше, h^2 - відношення енергії елемента прийнятого сигналу до спектральної щільності потужності шуму.

З урахуванням симетрії зрозуміло, що такою самою буде ймовірність прийому символу 1 при передачі 0. Тому ймовірність помилки не залежить від переданого символу. Вона однакова для всіх двійкових систем, ортогональних у посиленому значенні (при однакових енергіях сигналу), і визначається виразом (6.73). Зокрема ця формула справедлива для систем ЧМн, систем з часовою маніпуляцією й будь-яких інших, для яких виконується вираз (6.66).

На рис. 6.15 показана залежність $p(h^2)$ згідно з виразом (6.73) (крива 2). Там же для порівняння наведена крива, що характеризує потенційну завадостійкість тієї самої системи при когерентному прийомі й визначена виразом (6.53) (крива 1).

Величина h^2 для зручності виражена в децибелах, а ймовірності помилок відкладені в логарифмічному масштабі.

Порівняння кривих показує, що для розглянутої системи сигналів (ортогональної в посиленому розумінні з рівною енергією) апріорне знання фази й когерентний прийом дають лише дуже невеликий енергетичний вигравш у порівнянні з випадком некогерентного прийому. Цей вигравш тим менше, чим нижче припустима ймовірність помилки.

Систему ФМн так само, як й інші системи із протилежними сигналами, що відрізняються зсувом фаз на π , при некогерентному прийомі застосовувати не можна, тому що при невідомій початковій фазі такі сигнали нерозрізнені. Однак, якщо зсуви фази в каналі змінюється досить повільно, то різниці фаз між сусідніми елементами практично зберігаються й можуть бути вимірювані у прийомному пристрої.

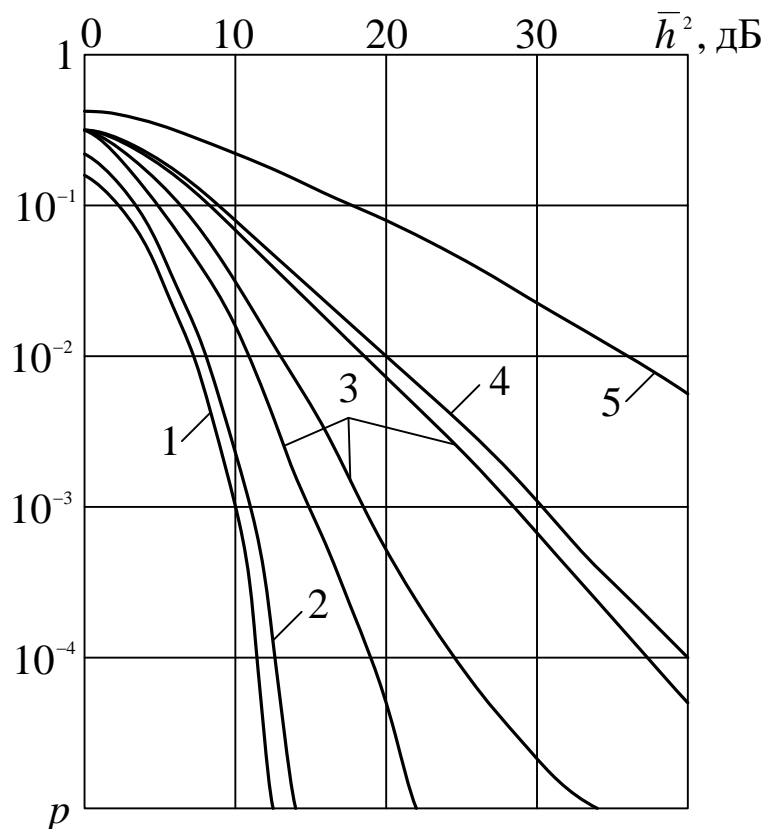


Рис. 6.15. Залежність ймовірності похибки у двійковій системі, ортогональній у посиленому розумінні, з активною паузою (наприклад, ЧМн) при оптимальному прийомі та різноманітних параметрах каналу: 1) канал з незмінними параметрами (когерентний прийом); 2) канал

з невизначеною фазою завмирання (некогерентний прийом); 3) узагальнені релеєвські завмирання; 4) релеєвські завмирання; 5) односторонні нормальні завмирання

Оскільки ВФМн інформаційний параметр сигналу визначається двома сусідніми елементами $(n-1)$ - M інтервалу $(-T,0)$ та на n -му інтервалі $(0,T)$, то оптимальний алгоритм можна записати як

$$\hat{i} = \underset{i}{\text{Arg max}} \left\{ \sqrt{\left[\int_{-T}^T z(t) u_i(t) dt \right]^2 + \left[\int_{-T}^T z(t) \hat{u}_i(t) dt \right]^2} \right\}, \quad i = 0,1. \quad (6.74)$$

Сигнал $S(t)$, що приходить, на двох тактових інтервалах при ВФМн можна уявити у залежності від символу, який передається n -тим елементом, так:

$$S_1(t) = a \cos(\omega_0 t + \psi), \quad -T < t < T \text{ (при передачі символу 1);}$$

$$S_0(t) = \begin{cases} a \cos(\omega_0 t + \psi), & -T < t < 0, \\ -a \cos(\omega_0 t + \psi), & 0 < t < T \text{ (при передачі символу 0),} \end{cases} \quad (6.75)$$

де ψ - випадкова початкова фаза, невідома при прийомі, яка залежить від символу, що передається $(n-2)$ -тим елементом.

Формули (6.75) являють собою двійкову систему сигналів з рівною енергією, ортогональну у підсиленому значенні на інтервалі тривалістю $2T$, а не T . Тому ймовірність помилки при прийомі сигналів ВФМн за алгоритмом (6.74) визначається на основі виразу (6.73), але з урахуванням того, що енергія сигналу на інтервалі $(-T;T)$ дорівнює $2E$. Тоді одержуємо

$$P = 0,5 \exp(-h^2), \quad (6.76)$$

де h^2 - відношення енергії сигналу на інтервалі тривалістю T до спектральної щільності потужності шуму.

Як слід очікувати, імовірність помилки виразу (6.76) більше, ніж обчислена для когерентного прийому двійкової ВФМн (вираз (6.54)), однак розходження між ними дуже мале.

Для схемної реалізації алгоритм (6.74) можна спростити. Для цього підставимо систему сигналів виразу (6.75) у вираз (6.74) і після скорочення однакових доданків приведемо алгоритм прийому до вигляду

$$X_a X_b + Y_a Y_b > 0, \quad (6.77)$$

$$\left. \begin{aligned} X_a &= \int_T^0 z(t) \cos(\omega_0 t + \psi) dt; & X_b &= \int_0^T z(t) \cos(\omega_0 t + \psi) dt; \\ Y_a &= \int_T^0 z(t) \sin(\omega_0 t + \psi) dt; & Y_b &= \int_0^T z(t) \sin(\omega_0 t + \psi) dt. \end{aligned} \right\} (6.78)$$

Фаза ψ може бути хоча й випадковою, але постійною на інтервалі $(-T; T)$, можна легко показати, що ліва частина виразу (6.77) інваріантна до значення цієї фази.

На рис. 6.16 показана кореляційна схема, що реалізує алгоритм прийому (6.77) на основі активних фільтрів. Величини X_a, X_b, Y_a, Y_b отримуються шляхом інтегрування добутку елемента прийнятого коливання на опорні сигнали $\cos(\omega_0 t + \psi)$ і $\sin(\omega_0 t + \psi)$ на інтервалі тривалістю T .

У моменти часу, кратні T , величини X_b, Y_b знімаються безпосередньо з інтегратора, а X_a, Y_a - з виходу ланцюга затримки на час T . На рис. 6.16 не показані ланцюги, що здійснюють скидання інтегратора до кінця інтервалу інтегрування та уведення накопиченого на ньому результату в перемножник і ланцюг затримки на час T .

Некогерентний прийом ВФМн можна реалізувати також у схемі з узгодженим фільтром і лінією затримки (рис. 6.17). Сигнал надходить на фільтр УФ, узгоджений з елементом

сигналу $u_1(t) = \cos(\omega_0 t + \psi)$ тривалістю T . Відгук фільтра надходить на два входи перемножника, на один з них безпосередньо, а на інший — через лінію затримки (ЛЗ), що забезпечує затримку на час T . Таким чином, поблизу моменту відліку на перемножник надходять напруги, що відповідають двом сусіднім елементам сигналу: тому, що надійшов, та попередньому, що пройшов через лінію затримки. Можна показати, що перша з цих напруг виражається формулою $u_1(t) = X_a \cos \omega_0 t + Y_a \sin \omega_0 t$, а друга $u_2(t) = X_b \cos \omega_0 t + Y_b \sin \omega_0 t$. Після їх перемножування й фільтрації результату у ФНЧ одержуємо напругу $X_a X_b + Y_a Y_b$, що у ВП порівнюється з нульовим порогом, тобто реалізується алгоритм (6.77). Описану схему називають *схемою порівняння фаз*.

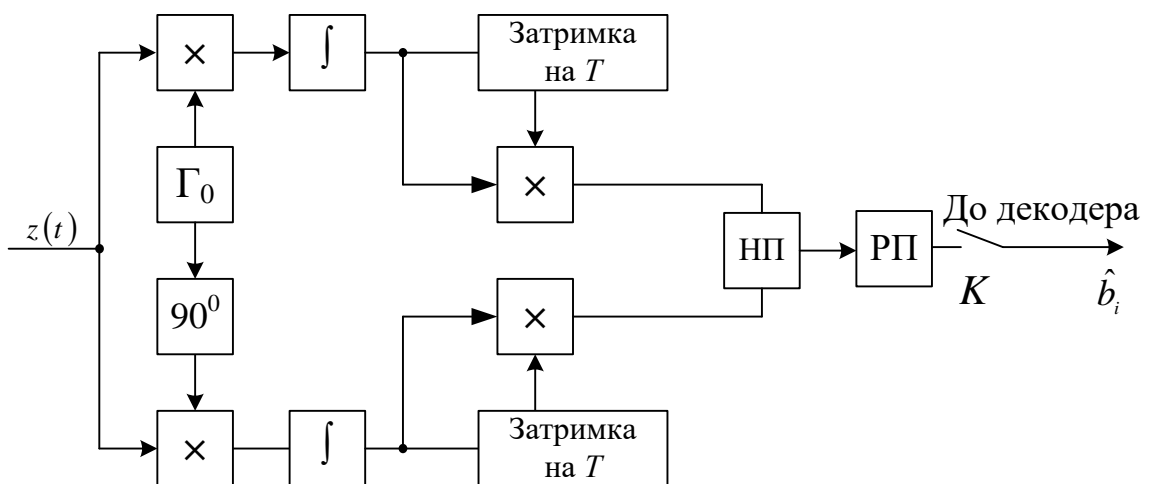


Рис. 6.16. Схема оптимального некогерентного прийому сигналів ВФМ на базі активних фільтрів

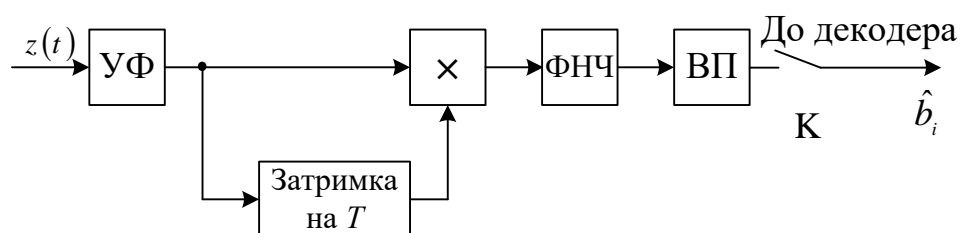


Рис. 6.17. Схема оптимального некогерентного прийому з узгодженим фільтром і лінією затримки для сигналів ВФМн

Зупинимося на деяких схемах *неоптимального прийому* при невизначеній фазі сигналу, які широко використовуються у сучасній апаратурі зв'язку. При прийомі сигналів двійкової АМн поширена схема наведена на рис. 6.18, яка складається з амплітудного детектора (Д) і фільтра нижніх частот (ФНЧ), який виділяє обвідну $r(t)$ прийнятого коливання, що пройшло вхідний вибірний блок — смуговий фільтр (СФ) з ефективною смугою пропускання ΔF , достатньою для одержання всіх найбільш істотних компонентів сигналу. Обвідна з виходу ФНЧ у певні моменти часу (наприклад, у середині посилки) порівнюється в ВП з деяким граничним рівнем λ . При виконанні нерівності

$$r > \lambda \quad (6.79)$$

реєструється символ 1, в іншому випадку - 0. Порівнюючи нерівність (6.79) з алгоритмом (6.63) і їх схемні реалізації, можна побачити, що схема на рис. 6.18 відрізняється від оптимальної некогерентної схеми використанням СФ і післядетекторного ФНЧ замість одного узгодженого фільтра до детектора.

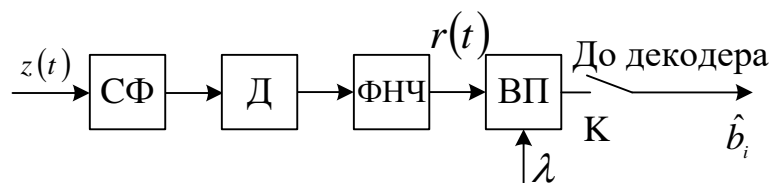


Рис. 6.18. Схема неоптимального прийому сигналів АМн методом порівняння обвідної з пороговим рівнем

При прийомі сигналів двійкової ЧМн поширена схема, наведена на рис. 6.19, де $СФ_1$ і $СФ_0$ — роздільні смугові фільтри, що пропускають без істотних перекошувань відповідно сигнали $S_0(t), S_1(t)$, Д - амплітудний детектор. Різницевий сигнал двох детекторів піддається фільтрації у ФНЧ, а результат для вибору рішення порівнюється з нульовим порогом.

Аналіз такої системи призводить до таких результатів. Імовірність помилок у схемі на рис. 6.19 більша, ніж при

оптимальному некогерентному прийомі, причому її зростання обумовлюється двома основними факторами:

- а) зменшенням відношення потужності сигналу до потужності шуму в порівнянні з узгодженим фільтром;
- б) міжсимвольними завадами, створюваними перехідними процесами у фільтрах (залишковими власними коливаннями, що виникли в результаті впливу попередніх елементів сигналу).

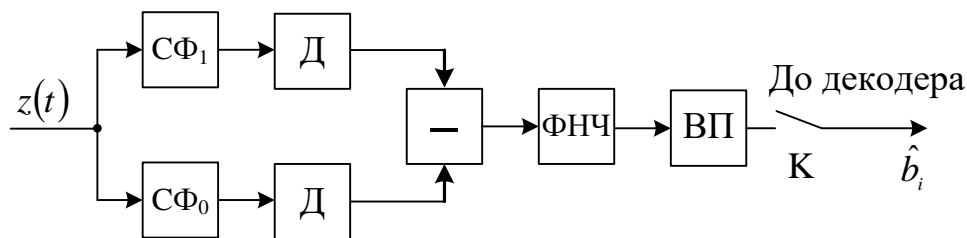


Рис. 6.19. Схема неоптимального некогерентного прийому сигналів ЧМн з роздільними смуговими фільтрами

Як зазначалося в п. 6.4 перший із цих факторів викликає найменшу втрату завадостійкості, якщо смуга пропускання смугового фільтра $\Delta f = 1,37/T$. Однак при такій смузі пропускання досить істотні похибки вносяться за рахунок другого фактора - міжсимвольної інтерференції. Тому найменша ймовірність помилок у схемі зі смуговими роздільними фільтрами за відсутності ФНЧ досягається при більш широкій смузі пропускання - приблизно $\Delta f = 3/T$. Можна показати, що для одержання однакової ймовірності помилок у схемі зі смуговими роздільними фільтрами потрібно в Δf T раз (у даному прикладі в 3 рази) більша потужність сигналу, ніж у схемі оптимального некогерентного прийому, що й визначає енергетичний програш при заміні узгоджених фільтрів смуговими.

Схему приймачів з *неоптимальною фільтрацією до й після детектора* широко використовують на практиці в тих випадках, коли частотна стабільність недостатня, тобто неточність частоти сигналу $|\delta F| > 1/T$ і, отже, реалізація оптимального прийому з узгодженою фільтрацією фактично неможлива. Це має місце, наприклад, внаслідок ефекту Доплера при зв'язку з об'єктами, що

рухаються, або при використанні супутника, що рухається, для ретрансляції при більших нестабільностях частот автогенераторів тощо.

Якщо смуги пропускання вхідних фільтрів Δf у схемах рис. 6.18 й 6.19 задовольняють умову $\Delta f > 2|\delta F|$, то сигнал залишиться в смузі пропускання фільтра при будь-яких флуктуаціях частот. При цьому значення $\Delta f T$ може виявитися значно більше 1, і якби не було фільтрації сигналу після детектора, енергетичний програш (у порівнянні з оптимальним прийомом при стабільній частоті сигналу) був би досить істотним. Однак є можливість значно зменшити цей програш, застосувавши фільтрацію напруги з виходу детектора. При $2|\delta F|T \gg 1$ смуговий фільтр майже не спотворює обвідну вхідного сигналу, тому за відсутності завад напруга на виході детектора в схемі на рис. 6.18 являє собою однополярні імпульси, а на виході блока вирішування схеми на рис. 6.19 — двополярні. При невеликому рівні шуму на вході детектора умови на його виході приблизно такі самі, як при прийомі прямокутних імпульсів на фоні "білого" гаусівського шуму. Тому необхідно включити після детектора фільтр, узгоджений із прямокутним імпульсом. На практиці часто застосовують і неузгоджений послідовний фільтр нижніх частот, як це показано в схемах на рис. 6.18 й 6.19.

6.7. Прийом дискретних повідомлень в умовах флуктуації фаз й амплітуд сигналів

У більшій частині радіоканалів, а також у деяких інших каналах флуктує не тільки початкова фаза, але й амплітуда очікуваних сигналів $S_i(t)$ (коефіцієнт γ). При відносно швидких (у порівнянні із тривалістю посилки T) завмираннях сигналу не можна судити за результатами прийому попередніх елементів про значення амплітуд й фаз наступних елементів. Знайдемо оптимальний алгоритм прийому при цих умовах.

Нехай канал описується моделлю, тобто є однопроменевим гаусівським із загальними завмираннями. Алгоритм оптимального прийому в умовах флуктуації як фази, так й

амплітуди сигналу можна (на основі правила максимальної правдоподібності) одержати при обчислюванні математичного очікування (вираз (6.60)) за γ

$$\bar{A}_i = \int_0 A_i(\gamma) w(\gamma) d\gamma \quad (6.80)$$

і порівнюючи між собою відносини правдоподібності A_i з різними індексами i . Однак для систем з рівною енергією сигналів результат легко вказати й без додаткових пояснень - він визначається співвідношенням (6.64). Це очевидно, оскільки вираз (6.64) є алгоритмом прийому при невизначеній фазі, не залежить від амплітуди (коефіцієнта γ), отже, цей алгоритм залишається оптимальним при будь-якій амплітуді (при будь-якому законі розподілу амплітуд). При цьому завадостійкість прийому істотно залежить від розподілу γ .

Визначимо, наприклад, ймовірність помилки для двійкової системи, ортогональної в посиленому значенні, з однаковою енергією сигналів за умови, що завмирання в каналі повільні, тобто коли γ можна вважати незмінним протягом елемента сигналу T і маломінливим від посилки до посилки.

Якщо умовну ймовірність помилки при деякому фіксованому значенні γ позначити $p(\gamma)$, то безумовна ймовірність помилки при повільних завмираннях дорівнюватиме

$$p = \int_0^{\infty} p(\gamma) w(\gamma) d\gamma. \quad (6.81)$$

У цьому випадку умовна ймовірність помилки визначається формулою (6.73), у якій величина $h^2 = \gamma^2 E_{II} / N_0$ пропорційна γ^2 , E_{II} - енергія сигналу на вході каналу (на передачі).

Нехай, наприклад, γ має розподіл Релея, який можна подати як [38]

$$w(\gamma) = \frac{2\gamma}{\gamma^2} \exp\left(-\frac{\gamma^2}{\gamma^2}\right). \quad (6.82)$$

Підставивши це у вираз (6.81) і позначивши $\gamma^2 \sqrt{\gamma^2} = x^2$, з урахуванням виразів (6.73) і (6.82) знайдемо ймовірність помилки для двійкових ортогональних у посиленому значенні сигналів при релеєвських завмираннях у такому вигляді:

$$\begin{aligned} p &= \int_0^{\infty} 0,5 \exp\left[-0,5\gamma^2 \frac{\overline{h^2}}{\gamma^2}\right] \frac{2\gamma}{\gamma^2} \exp\left(-\frac{\gamma^2}{\gamma^2}\right) d\gamma = \\ &= \int_0^{\infty} x \exp\left[-x^2\left(1 + 0,5\overline{h^2}\right)\right] dx = \frac{1}{2 + \overline{h^2}}, \end{aligned} \quad (6.83)$$

де $\overline{h^2} = \overline{\gamma^2} \frac{E}{N_0}$.

Ця залежність представлена на рис. 6.15 кривою 4.

Аналогічно визначається ймовірність помилок і при інших законах завмирань. Так, якщо розподіл ймовірностей - узагальнене релеєвське, то

$$p = \frac{1 + q^2}{2 + 2q^2 + \overline{h^2}} \exp\left[-\frac{\overline{h^2} q^2}{2 + 2q^2 + \overline{h^2}}\right], \quad (6.84)$$

де q^2 — відношення потужностей регулярної й флуктуючої складових.

На рис. 6.15 (крива 3) показана ця залежність при $q = 1, 5$ й 10 . Легко перевірити, що при $q \rightarrow \infty$ (відсутність завмирань) формула (6.84) переходить у формулу (6.73), а при $q^2 = 0$ (відсутність регулярної складової) — у формулу (6.83).

Приведемо ще результат для випадку, коли випадкова величина γ розподілена за одностороннім гаусівським законом (це "найгірший" радіоканал у рамках загальної гаусівської моделі [47]):

$$w(\gamma) = \frac{2}{\sqrt{2\pi\gamma^2}} \exp\left[-\frac{\gamma^2}{2\gamma^2}\right], \quad \gamma \geq 0.$$

При цьому

$$p = \frac{1}{2\sqrt{1+h^2}}. \quad (6.85)$$

Для загального гаусівського каналу формули для ймовірності помилок можна знайти в роботі [28].

Зазначимо, що всі отримані для двійкових систем вирази ймовірностей помилок при h^2 (або $\overline{h^2}$), що прагне до нуля, приймають значення 0,5. Це й слід було б очікувати, тому що при $p = 0,5$ по двійковому каналі жодна інформація не передається. При $h^2 \rightarrow \infty$ ймовірність помилок прагне до нуля. Це означає, що у всіх розглянутих каналах можна одержати як завгодно малу ймовірність помилки, збільшуючи потужність сигналу. Однак ступінь цього збільшення різний для різних каналів. Порівняння кривих на рис. 6.15 показує, що при завмираннях сигналу завадостійкість систем зв'язку значно нижче, ніж у каналі без завмирань за тієї самої середньої потужності передавача. Для підтримки заданої якості зв'язку в цих умовах доводиться мати певний запас потужності передавача.

Ймовірність помилок при прийомі дискретних повідомлень можна істотно зменшити за допомогою *рознесеного прийому*, сутність якого полягає в тім, що демодулятор ухвалює рішення щодо переданого символу не за одним, а по двома або більше сигналами, що несє ту саму інформацію. Рознесений прийом є одним з основних способів підвищення завадостійкості зв'язку при завмираннях сигналу.

У радіозв'язку застосовуються різні способи рознесеного прийому: за часом (він зводиться до повторення сигналу кілька разів на передачі до накопичення на прийомі); за частотою (сигнал дублюється по частотних каналах); прийом сигналу на різні антени, рознесені в просторі; поляризаційне рознесення (прийом на антени, що розташовані в одному місці, але приймають електромагнітні хвилі різної поляризації); рознесення

по окремих променях у багатопробеневому каналі. В останньому випадку промені розділяються або чітко спрямованими антенами за кутом приходу в горизонтальній чи вертикальній площині (такий спосіб застосовується головним чином на УКХ), або за часом приходу (часу запізнювання). Можна показати, що повний поділ променів за часом приходу вимагає застосування шумоподібних сигналів.

З перерахованих методів у радіозв'язку найпоширенішим є *прийом на рознесені в просторі антени*, на другому місці - *рознесений прийом за частотою*, використовуваний у каналах. Інші методи рознесення також застосовуються, але значно рідше. У каналах без завмирань рознесений прийом підвищує правильність, якщо є можливість скласти сигнали, що прийшли по n гілках когерентно, тобто звести наявні між ними різниці фаз до скільки завгодно малого значення.

При когерентному додаванні однакових сигналів сумарний сигнал буде мати в n раз більшу "амплітуду", тобто в n^2 раз більшу потужність, ніж окремий сигнал. При цьому завади, які звичайно в різних гілках незалежні, складаються некогерентно, так що потужність сумарної завади буде тільки в n раз більше за потужність завади в одній гілці, у результаті відношення потужності сигналу до потужності завади збільшується в n раз.

Можна показати [30], що якщо по n гілках приймаються сигнали з різними потужностями, а завади, що присутні у них, мають різну інтенсивність, то найкращі результати виходять при когерентному додаванні сигналів, помножених на вагові коефіцієнти, пропорційні

$$h_r = \sqrt{\frac{E_r}{N_{0,r}}},$$

де E_r - енергія елемента сигналу, що приходить;

$N_{0,r}$ — спектральна щільність шуму в n -тій гілці ($r = 1, \dots, n$).

При цьому в сумарному сигналі відношення h^2_{Σ} енергії сигналу до спектральної щільності шуму

$$h^2_{\Sigma} = \sum_{r=1}^n h_r^2. \quad (6.86)$$

У каналах із завмираннями є й інший, більш ефективний механізм підвищення правильності при рознесеному прийомі. Він ґрунтується на тім, що при одиночному прийомі помилки виникають головним чином тоді, коли рівень сигналу впаде нижче деякого граничного значення, а при рознесеному прийомі - коли рівень сигналу виявиться нижче за граничний у всіх гілках.

Якщо завмирання в галузях слабо корельовано, то ймовірність одночасного падіння рівнів сигналу у всіх гілках дуже мала.

Існують різні способи *комбінування (додавання)* сигналів окремих гілок при рознесеному прийомі. Відзначимо найбільш простий, досить ефективний і широко розповсюджений спосіб автовибору гілки з найбільш сильним сигналом (див. рис. 6.20). У цій схемі постійно виміряється коефіцієнт передачі каналу (або потужності прийнятого сигналу) по окремих гілках (приймачах), а до демодулятора підключається приймач із найбільш сильним сигналом.

На рис. 6.21 показана залежність імовірності помилок при некогерентному рознесеному прийомі. Тут n - число гілок рознесення.

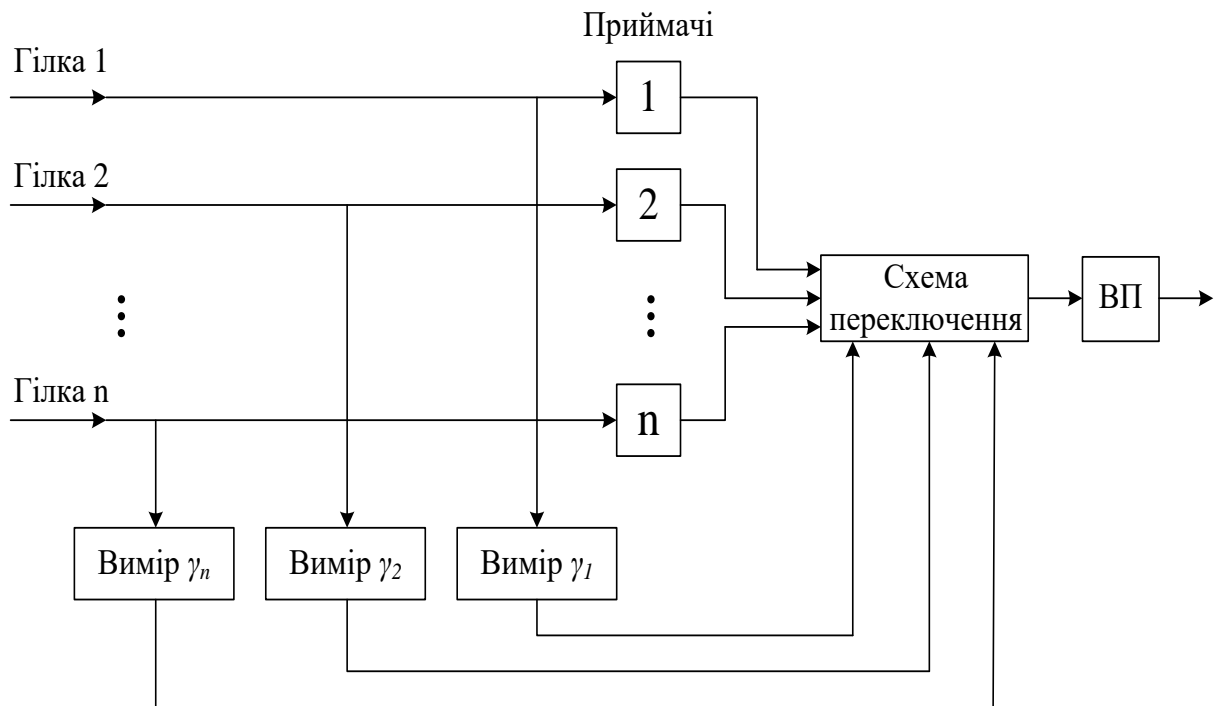


Рис. 6.20. Структурна схема рознесеного прийому при автовиборі гілки з найбільш сильним сигналом

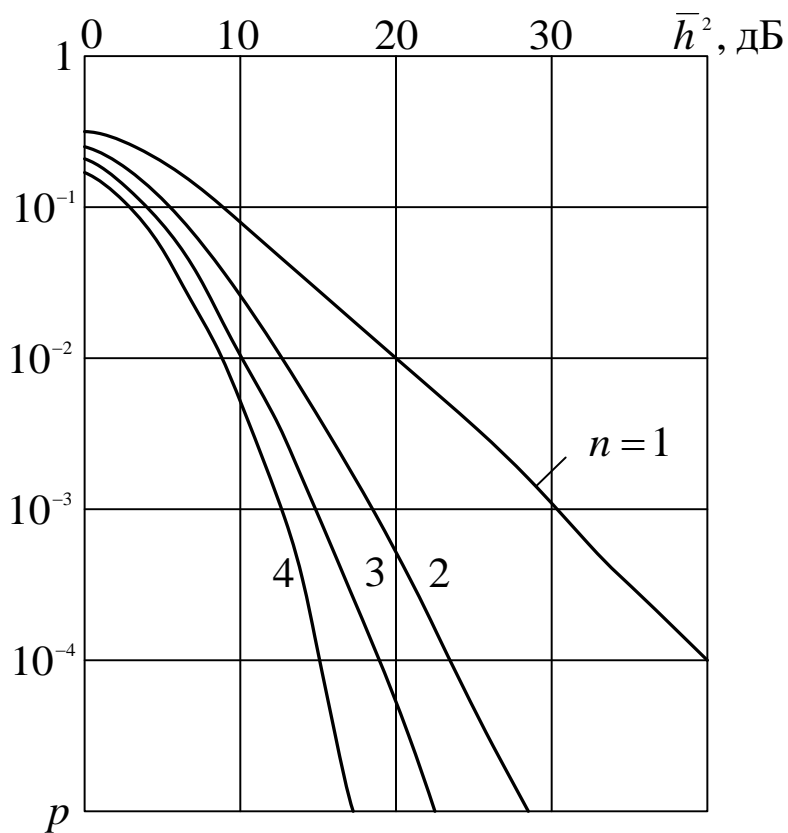


Рис. 6.21. Залежність імовірності помилки при рознесеному прийомі з релеєвськими завмираннями від h^2 (двійкова система з активною паузою, ортогональна у підсиленому значенні при автовиборі гілки з найбільш сильним сигналом), n -число гілок рознесення

6.8. Прийом дискретних повідомлень у каналах із зосередженими за спектром й імпульсними завадами

Завадостійкість дискретних систем зв'язку визначалася з урахуванням неминучого адитивного флуктуаційного або іншого гаусівського шуму. Однак на практиці доводиться враховувати й дію у каналі інших адитивних завад, породжуваних зовнішніми джерелами, і насамперед віднесених до класу зосереджених за спектром ("гармонійних") та імпульсних.

Зазначимо, що якщо на вхід прийомного пристрою надходить велике число слабо корельованих завад від різних джерел порівнянної потужності, то їх добуток відповідно до центральної граничної теореми являє процес, близький до гаусівського. Додаючись до флуктуаційного шуму апаратури, він може істотно збільшити спектральну щільність гаусівської завади, що вимагає відповідного збільшення потужності сигналу. Однак нерідко серед безлічі малопотужних завад на вхід приймача надходять окремі потужні імпульси або зосереджені завади, у такому каналі, якщо не вживати спеціальних заходів, прийом дискретних повідомлень супроводжується додатковими помилками й зв'язок може бути повністю порушено.

Всі заходи щодо захисту від зовнішніх завад можна розбити на три групи. До першої відносять ті, які спрямовані на заглушування завад у місці їх виникнення, зокрема екранування джерел промислових завад, застосування іскрогасних конденсаторів, зниження рівня побічних випромінювань радіопередавачів і т. п. Ці заходи регулюються спеціальними законодавчими актами й стандартами.

Друга група - це заходи, мета яких перешкодити проникненню завад на вхід демодулятора. Для цього в системах провідного зв'язку схрещують проводи повітряних ліній,

удосконалюють конструкцію кабелів для зменшення взаємних впливів і т. д. У радіозв'язку для цього встановлюється раціональний розподіл частот між окремими службами й каналами з урахуванням розміщення передавачів і приймачів та умов поширення радіохвиль. При вживанні заходів щодо перших двох груп важливу роль відіграють міжнародні органи, що виробляють припустимі норми й контролюють їх дотримання, - міжнародні консультативні комітети з телеграфії й телефонії (МККГТ) і з радіозв'язку (МККР).

Третя група заходів охоплює вибір ансамблю сигналів і побудову прийомного пристрою з метою попередити влучення зовнішніх завад безпосередньо у вирішальний пристрій (демодулятор) і мінімізувати ймовірність викликаних ними помилок, якщо вони таки проникнуть у нього. Питання, пов'язані із взаємним впливом різних систем передачі повідомлень (зв'язку) один на одного, за останні роки інтенсивно вивчаються теорією електромагнітної сумісності (ЕМС).

Зосереджені завади спостерігаються майже винятково в радіоканалах. Захист демодулятора від їх влучення здійснюється звичайно лінійними ланцюгами спеціальних блоків (вхідні вибірні ланцюги, перетворювачі частоти, резонансні й смугові підсилювачі й т. п.) різних прийомних пристроїв. Здатність послабити зосереджену заваду на вході вирішальної схеми приймача визначає його вибірковість. *Частотна вибірковість* забезпечується тим, що до подачі сигналу на вхід демодулятора він фільтрується в згаданих вище лінійних ланцюгах, смуга пропускання яких достатня для того, щоб сигнал пройшов без істотних перекручувань, а зосереджені завади, що лежать поза смугою пропускання, при цьому заглушуються. Крім частотної вибіркової, широко використовують також *просторову вибірковість*, засновану на застосуванні вузьконаправлених прийомних антен. Важливо зазначити, що вплив зосереджених завад зростає при збільшенні нелінійності вхідних каскадів приймача, оскільки виникаючі при цьому комбінаційні частоти (навіть якщо завада на вході приймача безпосередньо й не потрапила в смугу пропускання) можуть виявитися в корисній області частот. Питання захисту радіоприймача від зосереджених

завад складають основний зміст курсу радіоприймальних пристроїв.

Для зменшення ймовірності влучення зосередженої завади в смугу частот спектра сигналу бажано використовувати якомога більше вузькосмугові сигнали. Саме тому протягом багатьох десятиліть для передачі дискретних повідомлень по радіоканалах застосовувалися тільки прості вузькосмугові сигнали (АМ, ЧМ, ВФМ), елементи яких є відрізками синусоїди. Однак за останні 30-35 років намітився й інший підхід, пов'язаний з істотним розширенням спектра сигналу ускладненням його форми або з використанням шумоподібних сигналів. Шумоподібні сигнали дозволяють успішно передавати повідомлення в багатопробієвих каналах. Але, як це ні парадоксально, застосування шумоподібних сигналів виявилось корисним й для захисту від вузькосмугових зосереджених завад.

Справа в тому, що якщо спектр вузькосмугового сигналу перебивається потужною зосередженою завадою, то практично не вдається уникнути виникнення помилок. Якщо ж така завада виявиться в смузі широкосмугового сигналу, то в принципі існує можливість "вирізати" її режекторним фільтром (або іншими способами) і за частиною спектра, що залишилася, широкосмугового сигналу відновити передану інформацію. Тому, хоча ймовірність влучення зосередженої завади в спектр широкосмугового сигналу більше, ніж у спектр вузькосмугового, ймовірність помилок, створюваних такою завадою, при широкосмуговому сигналі (і раціонально побудованому приймачі) може виявитися значно меншою.

Найпростішим способом побудови широкосмугового сигналу для захисту від зосереджених завад є об'єднання декількох вузькосмугових сигналів, що передають однакову інформацію, на суміжних смугах частот зі здійсненням частотно-рознесеного прийому. При цьому схема автовибору будується так, що до вирішального пристрою підключаються тільки гілки, не уражені зосередженими завадами. Більш складні системи часто будуються з використанням "блока захисту від зосереджених завад", що являє собою ряд паралельно включених вузькосмугових фільтрів із суміжними смугами пропускання, розрахованих так, що разом вони пропускають без істотних

перекручувань увесь широкопasmовий сигнал. Цей блок включається на вхід демодулятора й управляє пристроєм, що аналізує напругу на виході кожного фільтра й замикає ти з них, у яких виявляються потужні зосереджені завади.

Для захисту від імпульсних завад запропоновані різні способи, найбільш ефективні з яких засновані на амплітудному обмеженні вхідного сигналу до його фільтрації або на миттєвому запиранні приймача на час дії завади. О.М. Щукін показав, що, застосовуючи обмежник у широкопasmовому тракті приймача й пропускаючи обмежений сигнал через вузькопasmовий фільтр, можна при належному виборі смуг пропускання загасити імпульсні завади без помітного погіршення завадостійкості щодо зосереджених і флуктуаційних завад. Така система одержала назву ШОВ (широкопasmовий фільтр, обмежник, вузькопasmовий фільтр) [22]. У сучасних пристроях роль вузькопasmового фільтра виконують звичайно узгоджені фільтри демодулятора.

Нехай вхідний сигнал приймача подається на двосторонній амплітудний обмежник. Якщо рівень обмеження U_0 обрано дещо вище за напругу корисного сигналу, то за відсутності імпульсної завади схема приймача залишається лінійною. Якщо ж з'явиться імпульсна завада з рівнем більшим, ніж U_0 , вона буде обмежена. Таким чином, імпульсна завада тривалістю τ_i із як завгодно великою амплітудою на вході трансформується в імпульс із площею $U_0\tau_i$. Амплітуда цього імпульсу приблизно дорівнює амплітуді сигналу, а спектр його сильно відрізняється від спектра сигналу. Тому після проходження через вузькопasmовий (або узгоджений) фільтр переважна частина енергії імпульсної завади відфільтровується й вона не викликає помилок.

Однак у реальних умовах рівень U_0 досягається й зосередженою завадою, а через нелінійний елемент у схемі (обмежника) утворюються комбінаційні частоти зосередженої завади, які надалі важко відфільтрувати. Установлення обмежника після вузькопasmового фільтра, що усуває вплив зосередженої завади, неефективне, тому що на виході такого фільтра тривалість імпульсної завади зростає й умова $\tau_i \ll T$ не може бути виконаною.

Метод миттєвого запирання приймача на час дії імпульсної завади також не позбавлений недоліків. По-перше, під час запирання й відмикання виникають перехідні процеси, що спотворюють роботу демодулятора; по-друге, сумарне вхідне коливання (сигнал плюс зосереджена й флуктуаційна завади) виявляється при цьому промодульованим імпульсом запирання, через що з'являються додаткові частотні складові, які можуть потрапити в смугу сигналу.

Можна відзначити частотно-часову дуальність між гармонійною й імпульсною завадами (спектральні характеристики зосередженої за спектром завади нагадують часові характеристики імпульсної, і навпаки). Ця обставина пояснює, чому заходи боротьби з імпульсною й зосередженою завадами у прийомному пристрої взаємно протилежні. Згадані вище шумоподібні сигнали можна з успіхом використовувати й для боротьби з імпульсними завадами внаслідок їх розходження за формою. Фільтр приймача, узгоджений з таким шумоподібним сигналом, перетворить сигнал на короткий інтенсивний (залежно від енергії сигналу) імпульс, тривалість якого обернено пропорційна смузі частот сигналу. Імпульсна ж завада перетворюється цим фільтром на коливання малої інтенсивності, що має характер шуму, який слабо маскує сигнал.

За останні роки запропоновані схеми захисту від зосереджених й імпульсних завад, засновані на оцінюванні цих завад і вирахуванні сигналу оцінки із прийнятого коливання (компенсаційні методи). При цьому такі схеми в умовах зміни характеристик завад стають *адаптивними*. Ефективним заходом захисту від зосереджених й імпульсних завад є рознесення прийому одночасно за частотою й часом. З гілок частотного рознесення варто вибирати ті, у яких менше (чи немає зовсім) зосереджених перешкод, а з гілок рознесення у часі - ті, де немає імпульсної завади. Досить ефективні також методи захисту від різних завад, засновані на завадостійкому кодуванні.

6.9. Завадостійкість прийому дискретних повідомлень в оптичному діапазоні хвиль

Узагальнена структурна схема *оптичного приймача* наведена на рис. 6.22. *Фотодетектор* (зазвичай це *фотодіод*) перетворює прийнятий *оптичний сигнал* на електричний струм, пропорційний потужності оптичного сигналу. Наступний за фотодіодом підсилювач підсилює отриманий струмовий сигнал і перетворює його на напругу. Остання поступає на демодулятор, у якому приймаються рішення (оцінки) про передачу 1 або 0. В оптичному приймачі є різні джерела шумів, що визначають завадостійкість передачі.

Насамперед, це *шум фотодіода* — квантовий шум, породжений випадковим квантовим процесом детектування оптичного сигналу. Цей шум можна назвати *дробовим*. Далі варто враховувати *теплові (флуктуаційні) шуми* в оптичному приймачі й *шум підсилювача*. Розглянемо спочатку вплив на завадостійкість дробового (квантового) шуму. Згідно з п. 4.6 ймовірність того, що на інтервалі аналізу T число створених фотодетектором пар носіїв заряду дорівнює K ,

$$p_K(T) = \frac{N^K}{K!} e^{-N}. \quad (6.87)$$

Зазначимо, що у випадкової величини, розподіленої по Пуассону (вираз (6.87)), математичне очікування дорівнює дисперсії $\bar{K} = \sigma_k^2 = N$. В ідеальній системі зв'язку зміна числа генерованих пар носіїв заряду (дробовий шум) - єдине джерело шуму. У такій системі оптична енергія приймається, а носії заряду генеруються лише тоді, коли передається 1. Якщо приймач досить чутливий, щоб виявити єдину електронно-діркову пару, створену світлом, то поріг може бути встановлений на цьому рівні. Немає помилки при передачі 0, оскільки не приймається ніяка енергія й не генерується ніякий сигнал. Тільки коли надходить на фотоприймач оптична енергія, що відповідає 1, він взагалі не генерує які-небудь носії заряду, замість 1 фіксується 0, тобто має місце помилка.

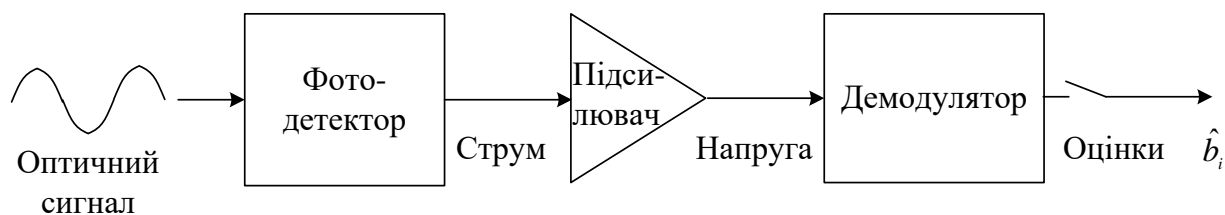


Рис. 6.22. Узагальнена структурна схема оптичного приймача

З урахуванням рівної ймовірності передачі 1 й 0 в ідеальній системі середня ймовірність помилки

$$p = 0,5 \left[\frac{N^0 e^{-N}}{0!} + 0 \right] = 0,5 e^{-N}. \quad (6.88)$$

Для одержання $p < 10^{-9}$ (сучасні вимоги для цифрових систем передачі) треба, як випливає з виразу (6.88), $N > 20$. Тоді одержуємо оцінку для мінімальної середньої потужності на вході фотоприймача

$$P_{np} = 0,5 E_{np} R_i = 0,5 R_i \frac{NE_\phi}{\eta} > 10 E_\phi \frac{R_i}{\eta}. \quad (6.89)$$

Значення $10 E_\phi \frac{R_i}{\eta}$ характеризує абсолютну квантову межу детектованості. При $\eta = 1$ та $\lambda = 0,9 \mu\text{м}$ одержуємо

$$\frac{R_i}{\eta} = h \frac{c}{\lambda} = 2,2 * 10^{-19} \text{Вт} * \text{с} \text{ та } P_{np} > 2,2 \text{ пВт}/(\text{Мбіт}/\text{с}). \quad (6.90)$$

Раніше було розглянуто завадостійкість оптичного приймача при обліку тільки дробового шуму. Розглянемо інший крайній випадок, коли дробовим шумом можна нехтувати у порівнянні з тепловим шумом і шумом підсилювача в приймачі. Звичайно думають, що випадкові флуктуації напруги й струми підкоряються гаусівському розподілу. Для аналізу зручніше віднести флуктуації, що спостерігаються на вході вирішального

пристрою демодулятора (рис. 6.22), до еквівалентного числа пар носіїв заряду, які варто було б створити у фотодіоді, щоб одержати той самий результат на вході ВП демодулятора. У випадку теплового шуму й шуму підсилювача цим дійсним числом доводиться приписувати гаусівський розподіл щодо середнього N . Таким чином, ймовірність того, що на вході ВП загальна напруга в точці відліку (сигнал + шум $y_c(t_0) + y_n(t_0)$) відповідає k парам носіїв заряду, створеним у фотодіоді, дорівнює в гаусівському наближенні

$$P(K|N) = \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp\left[-\frac{(K-N)^2}{2\sigma_u^2}\right], \quad (6.91)$$

де σ_u^2 - дисперсія шумового відліку $y_n(t_0)$.

Як і раніше припускаємо, що середнє число генерованих пар носіїв дорівнює N при прийомі 1 і нулю при прийомі 0. Оптимальний поріг для ухвалення рішення в ВП для системи з пасивною паузою (див. вираз (6.25)) дорівнює $y_n(t_0)/2$, що відповідає $N/2$ парам носіїв заряду.

Середня ймовірність помилки

$$P = 0,5[P(0|1) + P(1|0)] = 0,5\left[\sum_{k=0}^{N/2} P(K|N) + \sum_{k=N/2} P(K|0)\right]. \quad (6.92)$$

Для подальшого розрахунку треба у вираз (6.92) підставити гаусівський розподіл (вираз (6.91)). При $N \gg 1$ легше виконати обчислення, якщо замість доданку у виразі (6.92) виконати інтегрування

$$P = 0,5 \frac{1}{\sqrt{2\pi\sigma_u^2}} \left[\int_0^{N/2} \exp\left[-\frac{(K-N)^2}{2\sigma_u^2}\right] dK + \right. \\ \left. + \int_{N/2} \exp\left[-\frac{K^2}{2\sigma_u^2}\right] dK \right]. \quad (6.93)$$

Оскільки при виконанні симетрії два інтеграли рівні між собою, то

$$P = \frac{1}{\sqrt{2\pi\sigma_u^2}} \int_{N/2}^{\infty} \exp\left(-\frac{K^2}{2\sigma_u^2}\right) dK = \frac{1}{\sqrt{2\pi}} \int_{N/2}^{\infty} e^{-\frac{t^2}{2}} dt = \quad (6.94)$$

$$= Q\left(\frac{N}{2\sigma_u}\right).$$

З таблиць функції $Q(x)$ зрозуміло, що для $p < 10^{-9}$ необхідно мати $\frac{N}{2\sigma_u} > 6$. Це означає, що необхідне еквівалентне значення $N_e > 12\sigma_u$, що істотно більше (при реальних значеннях σ_u), ніж квантова межа детектування $N > 20$.

Облік спільної дії дробового шуму, теплових шумів і шумів підсилювача призводить до того [22], що необхідний рівень прийнятої потужності в оптичному приймачі на два порядки вище, ніж межа.

6.10. Порівняння завадостійкості систем передачі дискретних повідомлень

Об'єктивне порівняння різних систем передачі дискретних повідомлень при оптимальному прийомі й заданій моделі безперервного каналу можливе, якщо скористатися поняттям еквівалентної ймовірності помилки p_e , уведеним Л.М. Фінком [48]. Під p_e розуміють ймовірність помилки в симетричному двійковому каналі без пам'яті з постійними параметрами, у якому система із примітивним кодуванням виявляється еквівалентною розглянутій системі (передає рівну кількість інформації I , біт, і забезпечує однакову ймовірність правильного прийому блоку символів, що несуть цю інформацію в заданому каналі).

Із визначення зрозуміло, що ймовірність правильного прийому блоку символів

$$P_{\text{прав бл}} = (1 - p_e)^I = 1 - P_{\text{нм бл}} \text{ або } 1 - p_e = \left(1 - P_{\text{нм бл}}\right)^{\frac{1}{I}},$$

де $P_{\text{нм бл}}$ - імовірність помилково прийнятого блоку символів.

У умовах $P_{\text{нм бл}} \ll 1$, використовуючи формулу бінома Ньютона $\left(1 - P_{\text{нм бл}}\right)^{\frac{1}{I}} \approx 1 - \frac{P_{\text{нм бл}}}{I}$, одержуємо

$$p_e \approx 1 - \frac{P_{\text{нм бл}}}{I} \approx p_b. \quad (6.95)$$

Таким чином, p_e приблизно визначає у будь-якому каналі ймовірність помилки на біт. За інших рівних вимог та система краща, для якої p_e менше.

Введемо у розгляд енергетичний параметр системи

$$h_e^2 = \frac{P_c}{N_0 R_i} = \frac{E}{N_0 R \log_2 m} = \frac{h^2}{R \log_2 m}, \quad (6.96)$$

де P_c , N_0 - середня потужність сигналу на вході приймача й спектральна щільність БГШ у каналі;

$$R_i - \text{швидкість передачі інформації, } R_i = \frac{\log_2 m R}{T} \left(\frac{\text{біт}}{\text{с}}\right);$$

T - тривалість каналної послілки (тактовий інтервал передачі);

R - швидкість коду - відношення числа інформаційних елементів кодової комбінації до загального числа, $R = \frac{k}{n}$.

Таким чином, параметр h_e^2 визначає нормовану енергію $h^2 = \frac{E}{T}$ на 1 біт інформації (бітову енергію). Для системи i -тої системи маємо

$$h_{ei}^2 = \frac{h_i^2}{\log_2 m_i R_i} \text{ або } h_i^2 = h_{ei}^2 \log_2 m_i R_i, \quad (6.97)$$

де h_i^2 - енергетичний параметр, $h_i^2 = \frac{P_{ci} T_i}{N_0}$.

Для i -тої системи $p_{nm\ бл}$ є функцією від h_{ei}^2 , тобто $p_{nm\ бл i} = f_i(h_{ei}^2)$. Тоді можливо записати $p_{ei}^2 = \frac{f_i[h_{ei}^2]}{I_i}$.

Визначимо виграш i -тої системи над j -тою за еквівалентною ймовірністю помилки (B_e):

$$B_e = \frac{p_{ej}}{p_{ei}} = \frac{f_j[h_{e,j}^2] I_i}{f_i[h_{e,i}^2] I_j} \text{ при } h_{ei}^2 = h_e^2 = const. \quad (6.98)$$

Можна також визначити енергетичний виграш ($E_{B\ i/j}$) i -тої системи над j -тою за формулою

$$E_{B\ i/j} = 10 \lg \left(\frac{h_{ej}^2}{h_{ei}^2} \right) = 10 \lg \frac{f_j^{-1}(p_{ej} I_j)}{f_i^{-1}(p_{ei} I_i)} \text{ при } p_{ei} = p_{e=const}. \quad (6.99)$$

Крім енергетичної характеристики h_{ei}^2 i -та система визначається також частотною ефективністю $\gamma_i = \frac{R_{ii}}{F_i}$ (F_i - смуга частот системи).

Всебічне дослідження залежності енергетичної ефективності різних систем передачі дискретних повідомлень

$\beta_i = \frac{1}{h_{ei}^2}$ від γ_i [20].

Енергетичний виграш за рахунок кодування обговорюється в гл. 5 при оптимальному кодуванні й у гл. 7 при реальному завадостійкому кодуванні.

Висновки до глави 6

1. При ізохронній передачі дискретних повідомлень функція повідомлень перетвориться модулятором на ізохронну послідовність елементарних сигналів. Вони займають різну частотно-часову область, але переносять повідомлення з незмінним тактовим інтервалом, що визначає технічну швидкість передачі.

2. При використанні у якості переносників повідомлень на тактовому інтервалі прямокутних радіоімпульсів забезпечується граничний стиск сигналу в часі, але його спектр необмежений. Навпаки, при використанні переносників з обмеженим рівномірним спектром час елементарної послідовності не обмежено (міжсимвольна інтерференція на передачі), однак зберігається властивість звітності, що дозволяє безпомилково виділяти інформаційну послідовність. Ця властивість губиться в каналі з пам'яттю (з розсіюванням).

3. У системах зв'язку як критерій оптимального прийому повідомлень використовують головним чином критерій максимуму середньої ймовірності правильного прийому символу. Це критерій ідеального спостерігача. Оптимальний приймач забезпечує потенційну завадостійкість для заданої системи сигналів.

4. При рівноймовірній передачі різних символів критерій ідеального спостерігача реалізується правилом максимальної правдоподібності.

5. Оптимальний приймач при точно відомому сигналі є когерентним і реалізується кореляційною схемою або схемою з узгодженими фільтрами.

6. Ймовірність помилки при оптимальному когерентному прийомі двійкових сигналів у каналі зі стаціонарним гаусівським "білим" шумом залежить тільки від еквівалентної енергії сигналів (квадрата відстані між сигнальними точками) і спектральної щільності шуму. Мінімальну ймовірність помилки забезпечує двійкова система із протилежними сигналами.

7. У каналах із МСІ (пам'яттю) використовуються адаптивні

коректори каналу, однак граничну завадостійкість забезпечує оптимальний прийом.

8. У каналі із МСІ великий практичний інтерес викликає субоптимальний алгоритм поелементного прийому (зі зворотним зв'язком за рішенням) Кловського-Ніколаєва й алгоритм Вітербі.

9. Оптимальний приймач при невизначеній (випадковій) фазі сигналу є некогерентним і реалізується квадратурною схемою або схемою з узгодженими фільтрами й детектором.

10. При невизначеній фазі оптимальний приймач істотно спрощується для систем з однаковою енергією сигналів. Мінімальну ймовірність помилки забезпечує система сигналів, ортогональних у посиленому значенні.

11. Для системи з однаковою енергією сигналів схема оптимального некогерентного прийому не залежить від амплітуди сигналу й, отже, не змінюється при випадкових змінах амплітуд сигналу (завмираннях у каналі). Однак середня ймовірність помилки істотно залежить від закону розподілу амплітуд сигналу.

12. Ефективним засобом підвищення завадостійкості систем зв'язку (у тому числі при наявності в каналі зосереджених за спектром й імпульсних завад) є рознесений прийом.

13. В ідеальній оптико-волоконній лінії зв'язку з модуляцією потужності оптичного несучого коливання єдиним джерелом шуму є квантовий (дробовий) шум фотодетектора. Для забезпечення ймовірності помилки 10^{-9} при рівноймовірній передачі 1 і 0 потрібна гранична потужність (абсолютна квантова межа детектованості) на вході фотодетектора на один біт інформації в секунду $P_{nop} = 10h \frac{\lambda R_i}{c}$ Вт/(Мбіт/с). При $\lambda=0,9$ мкм та $P_{nop} = 2,2$ пВт/(Мбіт/с).

14. Облік теплових шумів приймача й шумів підсилювача після фотодетектора призводить до підвищення граничної потужності, обумовленої квантовою межею, на два порядки.

15. Еквівалентна ймовірність помилки, тобто ймовірність помилкового прийому символу в еквівалентному 2СК без пам'яті, з постійними параметрами й примітивним кодом, при якому забезпечується передача тієї самої кількості двійкових одиниць

інформації й та сама ймовірність помилкового прийому блоку символів, що й у розглянутій системі (із заданим способом модуляції й кодування, моделі каналу й т. п.), є зручною характеристикою порівняння різних систем передачі дискретних повідомлень між собою. При заданій p_e можна знайти енергетичний виграш системи, порівнюючи їх енергетичні параметри
$$h_e^2 = \frac{P_c}{N_0 R_i}.$$

Контрольні питання та задачі

6.1. Переносником повідомлень служить сигнал з рівномірним амплітудним спектром в області позитивних частот в інтервалі $(f_0 - \Delta f/2, f_0 + \Delta f/2)$, $\Delta f \ll f_0$. Написати аналітичний вираз для сигналу на виході модулятора при ізохронній послідовній передачі кодових символів з тактовим інтервалом T і частотною модуляцією несучої.

6.2. Кодова послідовність модулює ізохронно несучу із завдання 6.1 за фазою. Написати аналітичний вираз для сигналу на виході модулятора.

6.3. Переносником повідомлень на кожному тактовому інтервалі служать прямокутні радіоімпульси із середньою частотою спектра f_0 . Написати аналітичний вираз для сигналу на виході модулятора при лінійній амплітудній (багаторівневій) модуляції.

6.4. Написати аналітичний вираз для сигналу на вході демодулятора при передачі сигналу із завдання 6.3 при відомій імпульсній характеристиці лінійного каналу.

6.5. При заданій реалізації прийнятого колювання апостеріорні ймовірності переданих символів 1 і 0 дорівнюють $p(1|z)=0,6$ і $p(0|z)=0,4$. Який символ зареєструє приймач, оптимальний за критерієм ідеального спостерігача?

6.6. Яке правило прийому переважно застосовують у техніці зв'язку й чому?

6.7. Показати, що при нерівноймовірній передачі символів $p(b_i)$ алгоритм оптимального прийому за критерієм ідеального

спостерігача при детермінованих сигналах $S_i(t)$ ($0 \leq t \leq T$) і адитивному стаціонарному БГШ у каналі з одnobічною спектральною щільністю N_0 можна записати як

$$\hat{i} = \text{Arg max}_i \left\{ \int_0^T z(t) S_i(t) dt - 0,5 [E_i + N_0 \ln P(b_i)] \right\}.$$

6.8. Які основні блоки містить кореляційна схема оптимального когерентного прийому в каналі з адитивним стаціонарним БГШ?

6.9. Що розуміють під узгодженим фільтром? У який момент часу на виході УФ забезпечується максимальне відношення сигнал-шум і чому воно дорівнює?

6.10. Знайти передатну функцію фільтра, узгодженого із прямокутним відеоімпульсом, і на її основі побудувати структурну схему фільтра.

6.11. Фільтр узгоджений з пачкою з N періодично переданих (з періодом T) імпульсів. Перший імпульс має спектральну щільність $S_1(j\omega)$. Знайти передатну функцію УФ і на її основі побудувати структурну схему фільтра.

6.12. Показати, що АЧХ фільтра з імпульсною характеристикою гаусівської форми $g(t) = Ae^{r^2}$, $-\infty < t < +\infty$ не задовольняє вимоги Пейлі – Вінера.

6.13. При оптимальному когерентному прийомі двійкових сигналів кореляційною схемою генератори опорних сигналів допустили фазовий зсув на $\Delta\varphi$. Знайти величину зсуву, при якому енергетичний програш стосовно схеми з ідеальною фазовою когерентністю не перевищує 1,1 (підвищення потужності передавача становить 10 %).

6.14. Фільтр узгоджений з біполярним сигналом коду Баркера (1,1,1,-1,1). Зобразити сигнал на виході фільтра.

6.15. Чим пояснити практичні переваги систем зв'язку з однаковою енергією сигналів при когерентному й некогерентному прийомі?

6.16. Знайдіть енергетичний програш двійкової системи, ортогональний у посиленому значенні, у каналі з релеєвськими завмираннями відносно каналу без завмирань при ймовірності

помилки $p = 10^{-3}$.

6.17. У чому зміст рознесеного прийому сигналів й які види рознесення ви знаєте?

6.18. Які методи боротьби із зосередженими (за спектром) і імпульсними (зосередженими в часі) завадами ви знаєте?

6.19. Зв'язок може вважатися якісним, якщо в смузі частот корисного сигналу немає зосередженої завади або як тільки вона з'явиться, амплітуда корисного сигналу перевищує амплітуду завади. Знайти вираз для ймовірності якісного зв'язку, якщо амплітуди сигналу й завади розподілені за Релеєм із середнім квадратом $\overline{\gamma_c^2}$ і $\overline{\gamma_n^2}$, а ймовірність появи зосередженої завади $P_{зз} = 0,2$. При яких співвідношеннях параметрів $\overline{\gamma_c^2}$ і $\overline{\gamma_n^2}$ ймовірність якісного зв'язку не менше 0,8?

ГЛАВА 7

ЗАВАДОСТІЙКЕ КОДУВАННЯ

7.1. Основні поняття і визначення

При передачі інформації по каналах зв'язку виникають помилки унаслідок завад і спотворень сигналів. Для їх виявлення і виправлення використовуються завадостійкі (коректуючі) коди.

Завадостійкими називають коди, що дозволяють виявляти і (або) виправляти помилки в прийнятому повідомленні. Спроможність коду виявляти і виправляти помилки заснована на введенні надмірності в кодоване повідомлення. Надмірні символи формуються за певними правилами і називаються перевірочними або контрольними. Збільшення числа таких символів у кодовій

комбінації підвищує виявляючу і виправляючу спроможності коду, але призводить до зниження швидкості передачі інформації.

Спрощена схема системи передачі інформації при завадостійкому кодуванні наведена на рис. 7.1.

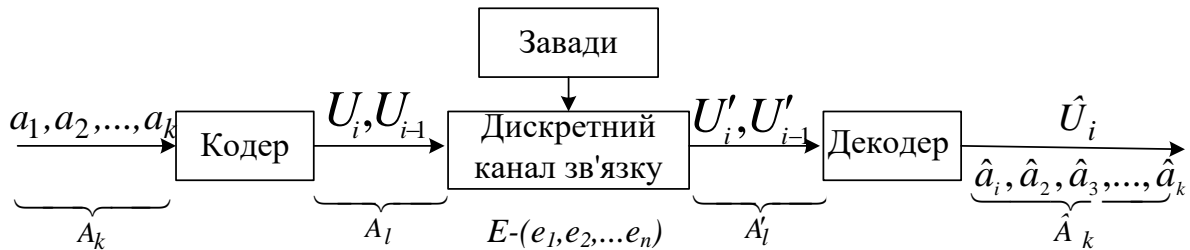


Рис. 7.1. Спрощена схема системи передачі інформації

У загальному випадку під кодуванням розуміється заміна послідовності символів повідомлення від дискретного джерела інформації з алфавітом A_k послідовністю символів кодуючого пристрою (кодера) з алфавітом A_l . Розрізняють два види кодування: блокове і безперервне.

При блоковому кодуванні послідовність символів повідомлення розділяється на блоки з k символів, які перетворюються на блоки з n символів коду ($n > k$). Символи повідомлення джерела називають інформаційними.

Послідовність з n символів на виході кодера називається кодовою комбінацією або кодовим словом. Сукупність кодових слів утворює (n, k) код. Коди, комбінації яких містять однакове число символів, називаються рівномірними. Їх застосування, на відміну від нерівномірних кодів, спрощує схеми кодерів і декодерів.

Якщо при кодуванні інформаційна послідовність не розділяється на блоки, то код називається безперервним. Такі коди також називають деревоподібними або ланцюговими. Блочні коди можна розглядати як окремий випадок безперервних.

При відомому розміщенні інформаційних і перевірочних символів у кодовій комбінації код називається систематичним. Зазвичай інформаційними є перші k символів кодової комбінації.

Безліч символів, з яких складається кодова комбінація, називається алфавітом коду, а число різних символів в алфавіті –

основою коду. Найбільшого поширення набули двійкові коди з основою коду, що дорівнює двом.

Завадостійкі коди, включаючи двійкові, поділяються на лінійні і нелінійні. Найбільш поширені лінійні, рівномірні, роздільні двійкові коди, кодові комбінації яких утворюють лінійний простір відносно операції порозрядного складання по модулю 2.

Складання по модулю 2 ($mod 2$) здійснюється за такими правилами: $0 \oplus 0 = 0$; $0 \oplus 1 = 1$; $1 \oplus 1 = 0$ або $1 = -1$ - це означає, що операції складання і віднімання по $mod 2$ співпадають.

Кодові комбінації таких кодів можуть бути подані у вигляді векторів в n -мірному лінійному просторі або багаточленів (поліномів) від формальної змінної x степеня $n-1$.

Подання кодових комбінацій як векторів в n -мірному просторі дозволяє визначити відстань між ними, звану відстанню Хеммінга. Під нею розуміється кількість символів, якими одна комбінація відрізняється від іншої [44].

Наприклад, між комбінаціями $U_1 = 110110$ і $U_2 = 011001$ відстань $d(U_1, U_2) = 5$, оскільки вони містять 5 неспівпадаючих символів. Відстань Хеммінга співпадає з вагою сумарної комбінації $(U_1 \oplus U_2) = d(U_1, U_2)$, під якою розуміється число її ненульових символів. Поняття ваги застосоване до будь-якої комбінації коду. Наприклад, комбінація U_1 має вагу $W(U_1) = 4$, оскільки містить 4 одиниці, а вага $W(U_2) = 3$.

Побудова завадостійких кодів заснована на використанні алгебраїчних структур (груп, кілець, полів), що визначають правила формування кодових комбінацій, виявлення і виправлення помилок у них.

Для передачі кодових комбінацій між кодером і декодером використовується дискретний канал зв'язку – сукупність технічних засобів, включаючи середовище розповсюдження, сигнали на вході і виході якого приймають кінцеве число значень (рис. 7.1). Найпростішою моделлю дискретного каналу є двійковий канал зв'язку з завадою, що адитивно взаємодіє з сигналом.

Завадою, або вектором помилки, в такому каналі називається послідовність з n символів $E = (e_1, e_2, \dots, e_n)$, яку треба порозрядно скласти по модулю 2 з переданою кодовою комбінацією U , щоб отримати прийняту: $U' = U \oplus E$. Символи вектора помилки $e_i = 1$ означають наявність помилки в прийнятій комбінації U' , а $e_i = 0$ - відсутність помилки. Число ненульових символів вектора помилки E називається вагою або кратністю помилки q . Кратність помилки є випадковою величиною, що приймає значення від 0 до n .

Якщо символи кодової комбінації в такому каналі спотворюються з однаковою імовірністю p_e , тобто $p(e_i = 1) = p_e, i = 1, 2, \dots, n$, то канал називається двійковим симетричним без пам'яті, а помилки – незалежними. У цьому випадку розподіл імовірності помилки q -тої кратності в кодових комбінаціях (n, k) коду є біноміальним:

$$P_n(q) = C_n^q p_e^q (1 - p_e)^{n-q}, \quad (7.1)$$

де C_n^q - число поєднань з n символів по q .

Наприклад, $p_e = 1 \cdot 10^{-4}$, а $n=5$. Тоді імовірність помилки першої, другої і п'ятої кратностей відповідно дорівнює

$$p_1 = C_5^1 p_e^1 (1 - p_e)^4 \approx 5 \cdot 10^{-4};$$

$$p_2 = C_5^2 p_e^2 (1 - p_e)^3 \approx 1 \cdot 10^{-7};$$

$$p_5 = C_5^5 p_e^5 (1 - p_e)^0 \approx 1 \cdot 10^{-20}.$$

З їх порівняння зрозуміло, що в каналах з незалежними помилками найбільш імовірні одиночні помилки. Цей результат обґрунтовує використання завадостійких кодів для виявлення і виправлення помилок малої кратності в каналах з незалежними помилками. Частіше зустрічаються канали, в яких помилки групуються в пакети.

Під пакетом помилок розуміється вектор помилки з n символів, з яких l підряд (від 2 до n) спотворені або починаються і закінчуються помилками, між якими можуть бути спотворені і неспотворені символи. Наприклад, вектори помилки $\mathring{A}_1 = 000111100\dots 0$; $\mathring{A}_2 = 001111000\dots 0$; $\mathring{A}_3 = 011110000\dots 0$ містять пакети помилок з 4 символів $l=4$. Під дією помилок комбінація на виході каналу зв'язку U' відрізняється від комбінації на вході каналу U . Для оцінки U' інформаційних символів за прийнятою комбінацією U' використовується декодуючий пристрій (декодер). При декодуванні вирішуються два завдання: оцінювання переданої кодової комбінації і формування інформаційних символів. Найскладнішим є перше завдання декодування. При рівній імовірності кодових комбінацій її оптимальне рішення забезпечує метод максимальної правдоподібності.

У разі двійкового симетричного каналу функція правдоподібності визначається імовірністю векторів помилки:

$$P(U'/U) = P(E_i) = p^{q_i} (1 - p_e)^{n - q_i}, \quad (7.2)$$

де q_i – вага вектора помилки $E_i = U \oplus U'$.

З формули (7.2) виходить, що $P(U'/U)$ максимальна, якщо q_i мінімальна. Тоді оцінкою U є кодова комбінація, спотворення якої для отримання прийнятої комбінації U' має мінімальну вагу:

$$\hat{U} = U' \oplus E_i. \quad (7.3)$$

Якщо таких векторів помилок E_i декілька, то найбільш імовірний з них визначається випадковим набором.

У цілому складність процедури декодування залежить від довжини кодованих інформаційних послідовностей, правил кодування, типу каналу зв'язку (односторонній, двосторонній) і характеру помилок у ньому (незалежні або пакети помилок).

У залізничних системах автоматики, телемеханіки і зв'язку найбільшого поширення набули двійкові лінійні коди. Тому далі розглядаються коди, що належать до класу двійкових лінійних кодів.

7.2. Принцип виявлення і виправлення помилок у завадостійких кодах

Для того щоб код виявляв і виправляв помилки, безліч всіх можливих комбінацій коду 2^n розділяється на дві непересічні підмножини: дозволених (кодових) і заборонених $2^n - 2^k$ комбінацій. Кодер використовує тільки дозволені комбінації. Заборонені комбінації не використовуються при передачі, а можуть утворюватися тільки під впливом завади.

Принцип виявлення помилок полягає в тому, що декодер визначає приналежність прийнятої комбінації коду. Якщо комбінація належить коду, то приймається рішення, що помилок немає. Якщо ж прийнята комбінація є забороненою, то приймається рішення про наявність помилки. Код може виявляти помилки, якщо число заборонених комбінацій $2^n - 2^k > 0$. Помилки не можна виявляти тільки в тих випадках, коли одна кодова комбінація перейшла в іншу під впливом завади. Імовірність виявлення помилок (n, k) коду визначається вибором дозволених кодових комбінацій і не залежить від властивостей дискретного каналу зв'язку.

Принцип виправлення помилок полягає в тому, що вся безліч комбінацій коду 2^n розділяється на 2^k непересічних підмножин, які називаються областями рішень. Кожна область

рішення відповідає одній з дозволених кодових комбінацій. Якщо прийнята комбінація потрапляє в одну з 2^k областей, то приймається рішення про передачу кодової комбінації, відповідної цій області. виправлення помилок відбуватиметься в тих випадках, коли кодове слово і прийнята заборонена комбінація належать до однієї області рішень. У кожному області рішень включають такі заборонені комбінації, при прийомі яких найбільш імовірною є одна з дозволених кодових комбінацій (n, k) коду, що відповідає принципу максимальної правдоподібності.

При передачі рівноймовірних кодових комбінацій (n, k) коду розподіл комбінацій за областями рішень визначається тільки характером помилок у каналі зв'язку. Наприклад, у двійковому симетричному каналі найбільш імовірні помилки малої кратності. Тому розбиття заборонених комбінацій між областями рішень повинне здійснюватися так, щоб перш за все виправлялися помилки малої кратності. Для цього кодові комбінації вибираються так, щоб при обмеженому числі заборонених комбінацій забезпечити можливість виправлення помилок малої кратності. Для однозначного декодування кожна із заборонених комбінацій повинна входити тільки в одну область рішень, тоді ці області не перекриваються.

Відмінності в комбінаціях (n, k) кодів як векторів лінійного n -мірного простору визначаються відстанню Хеммінга між ними $d(U_i, U_j)$, яка змінюється від 1 до n . Очевидно, що серед комбінацій (n, k) коду можна вибрати такі, відстань між якими, наприклад, не менше двох, і їх використовувати при кодуванні комбінацій і передавати в канал зв'язку. Такі комбінації є кодовими або дозволеними, а інші – забороненими. Мінімальна відстань між всіма кодовими комбінаціями називається мінімальною кодовою відстанню d_{min} і визначається як вага кодової комбінації U_i , найменше відмінна від нульової комбінації U_0 , що завжди належить коду $d_{min} = W(U_i, U_0)$. Мінімальну кодову відстань можна використовувати для оцінки виявляючої і виправляючої спроможності (n, k) кодів.

Для виявлення помилки в прийнятій комбінації достатньо, щоб її кратність q (або вага помилки) була менше від мінімальної кодової відстані d_{min} . У цьому випадку прийнята комбінація вважається забороненою. Тоді кратність помилки пов'язана із d_{min} співвідношенням

$$q_0 = d_{min} \leq 1. \quad (7.4)$$

Для виправлення помилок необхідно, щоб області рішень не перекривалися, і кожна із заборонених комбінацій входила тільки в одну область. Чим далі межа області від кодового слова, тим більшої кратності помилка може бути виправлена при передачі даної кодової комбінації. Допустимо, що межа областей двох сусідніх кодових слів стикаються, і відстань між словами d_{min} . Тоді помилки q -тої кратності, потрапляючи в одну з областей рішень, виправлятимуться. Це можливо при $(d_{min} - 1)/2$. Таким чином, максимальна кратність помилок, що виправляються, пов'язана із d_{min} співвідношенням

$$q_i \leq \left[\frac{d_{min} - 1}{2} \right], \quad (7.5)$$

де квадратні дужки означають цілу частину числа.

Виявляючу і виправляючу спроможність завадостійкого коду можна пояснити за допомогою рис. 7.2. На ньому зображені дозволені кодові послідовності U_{p1}, U_{p2} і U_{p3} , між якими розташовані заборонені U_{z1}, U_{z2}, U_{z3} і U_{z4} .

З рис. 7.2. і формул (7.4) і (7.5) випливає узагальнений вираз для мінімальної кодової відстані

$$d_{min} \geq q_0 + q_i + 1. \quad (7.6)$$

Коректуючий код дозволяє також виправляти деякі помилки кратності, більшої за q_i , якщо прийнята спотворена комбінація не виходить з області рішень для переданого кодового слова.

Як впливає з виразів (7.4) і (7.5), для підвищення кратності виявлених або виправлених помилок потрібно збільшувати мінімальну кодову відстань d_{min} , а отже, число надмірних символів коду. Тому одним із завдань теорії кодування є побудова кодів з мінімальним числом перевірочних (надмірних) символів при заданому d_{min} .

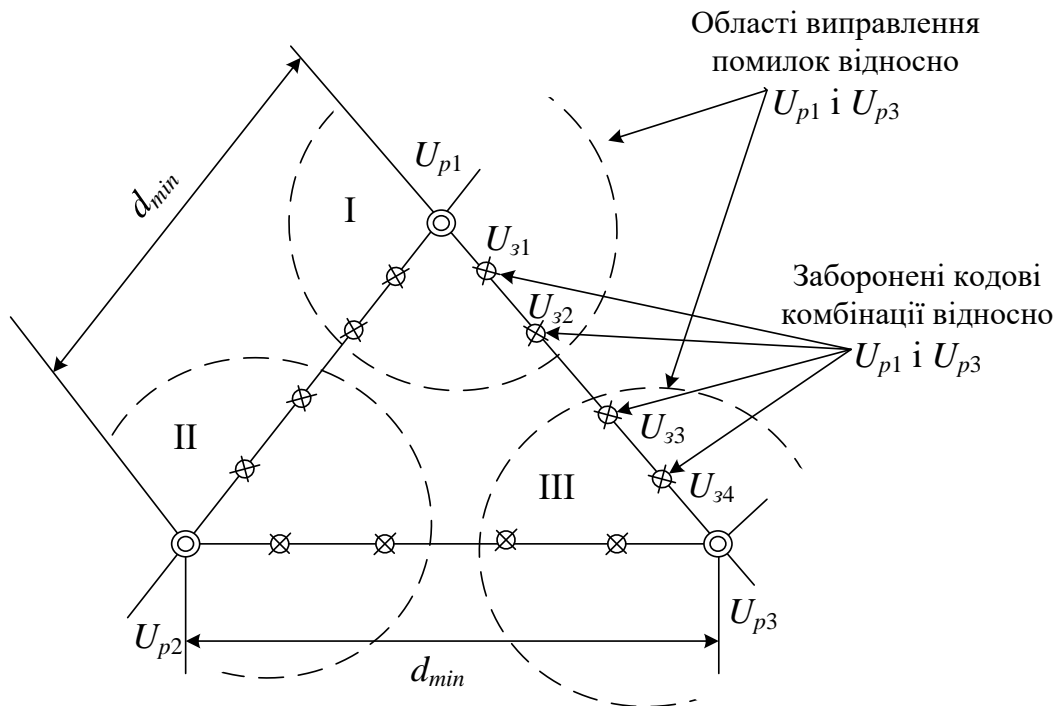


Рис. 7.2. Зв'язок мінімальної кодової відстані з числом виявлених і виправлених помилок

Загального рішення цієї задачі не існує, проте відомі граничні співвідношення для певних класів кодів. Так, наприклад, нижню межу для числа надмірних (контрольних) символів (n, k) кодів можна визначити, якщо порівняти число виправлених помилок, що дорівнюють числу заборонених комбінацій $2^n - 2^k$, до числа усіх помилок кратності від 1 до n будь-якої з $2^n - 2^k$ кодових комбінацій:

$$2^n - 2^k \geq 2^k \sum_{q=1}^{q_i} C_n^q$$

або

$$2^{n-k} - 1 \geq 2^k \sum_{q=1}^{q_i} C_n^q. \quad (7.7)$$

З виразу (7.7) випливає, що для виправлення помилок першої кратності параметри (n, k) коду зв'язані співвідношенням

$$2^{n-k} - 1 \geq n = k + r, \quad (7.8)$$

де r – число перевірочних (контрольних) символів (розрядів) у кодовій комбінації.

Враховуючи формулу (7.7), вираз (7.8) може бути поданий як

$$(n - k) \geq \left[\log_2 \sum_{q=0}^{\frac{d_{min}-1}{2}} C_n^q \right]. \quad (7.9)$$

Цей вираз називається межею Хеммінга для числа контрольних символів коду.

Якщо при деяких n і k нерівності (7.7)-(7.9) переходять у рівності (n, k) , код називається здійсненим або щільноупакованим. Такі коди виправляють всі помилки тільки кратністю до $q_i = 0,5[d_{min} - 1]$ і жодної помилки вищої кратності. Щільноупакованими коди називають тому, що всі заборонені комбінації потрапляють всередину областей рішень.

До досконалих належать три типи лінійних кодів: двійковий код непарної довжини з повторенням, коди Хеммінга і код Голя.

Код з повторенням $(n, 1)$ містить дві кодові комбінації, що містять один інформаційний символ і $n - 1$ перевірочних, які повторюють інформаційний:

$$a_1 = a_2 = a_3 = \dots = a_{n-1}.$$

Одна кодова комбінація цього коду складається тільки з нулів, інша – з одиниць. Мінімальна кодова відстань $d_{min} = n$ і

$$q_0 = n - 1, q_i = \left\lceil \frac{n-1}{2} \right\rceil.$$

Коди Хеммінга, запропоновані Р. Хеммінгом в 1948 р., належать до (n, k) кодів, виправляють всі одиночні помилки, мають мінімальну кодову відстань $d_{min} = 3$, $q_0 = 2$, $q_i = 1$. Коди Хеммінга утворюють такий ряд: (3,1), (7,4), (15,11), (31,26), (63,57), (127,120)... Отже, число перевірочних символів росте значно повільніше за число інформаційних. Це означає, що вигідно кодувати довші інформаційні послідовності. Для цих кодів досягається нижня межа числа перевірочних символів $n - k = \log_2(n + 1) = r$.

Код Голея (23,12) має $d_{min} = 7$, $q_0 = 3$ і є досконалим, оскільки $n - k = 23 - 12 = \log_2(C_{23}^0 + C_{23}^1 + C_{23}^2 + C_{23}^3)$ [12].

Значно ширше клас квазідосконалих кодів, які виправляють всі помилки кратності до q_i , частину помилок кратності $q_i + 1$ і жодної помилки вищої кратності.

7.3. Класифікація та основні характеристики завадостійких кодів

7.3.1. Класифікація завадостійких кодів

Відоме велике число кодів, систематизація і класифікація яких через їх численні ознаки є досить складною. Тому при розгляді класифікації, як правило, використовують структурні характеристики кодів (рис. 7.3) [44].

Перешкодостійкі коди поділяються на рівномірні і нерівномірні. Рівномірні коди – це коди, всі кодові комбінації яких містять постійну (однакову) кількість розрядів. Нерівномірні коди містять кодові комбінації з різним числом розрядів. З огляду на те, що нерівномірні коди через складність їх реалізації не знайшли широкого застосування, надалі їх розглядати не будемо.

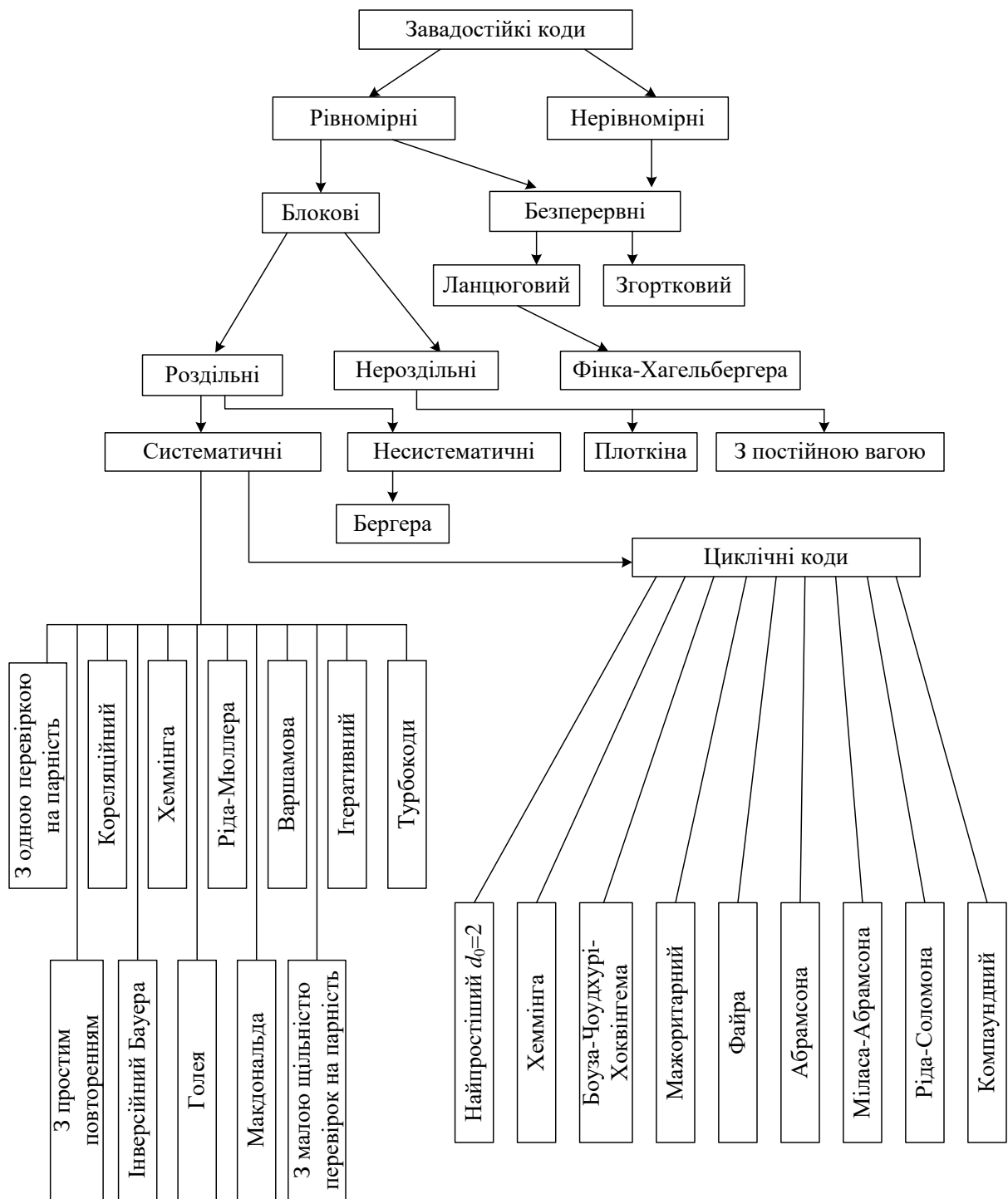


Рис. 7.3. Класифікація завдостійких кодів

Всі надмірні коди розділяються на два класи: безперервні і блокові.

У безперервних блоках процес кодування і декодування відбувається безперервно. Такі коди не отримали широкого розвитку. У блокових кодах кожному повідомленню відповідає своя кодова комбінація (блок) з n символів. Блоки кодуються окремо один від одного.

Надмірні коди, в яких певні розряди кодових комбінацій відводяться для інформаційних і перевірочних символів, називаються роздільними. Роздільні блокові коди позначаються символом (n, k) , де n - загальне число розрядів кодової комбінації; k - число розрядів, що відводяться для інформаційних символів, $k = (n - r)$; r - число перевірочних (контрольних) символів. Нероздільні коди не мають чіткого розділення кодової комбінації на інформаційні і перевірочні символи. До них належать, наприклад, коди з постійною вагою, коди Плоткіна та ін.

Роздільні блокові коди поділяються на несистематичні і систематичні. У несистематичних кодах перевірочні символи є сумами підблоків довжиною l , на які розбивається послідовність інформаційних символів. Такі коди можуть виявляти пачки помилок довжиною l . До них належать коди Бергера.

Систематичні коди складають найбільший клас роздільних кодів, являють найбільш розвинений напрям сучасної теорії кодування, що спирається на спеціальні розділи лінійної алгебри, векторного аналізу, полів Галуа, матричного числення і т. п.

У таких кодах перевірочні символи знаходяться в результаті виконання лінійних операцій над певними інформаційними символами.

Для двійкових кодів такі операції зводяться до вибору кожного перевірочного символу так, щоб його сума по модулю два з певними інформаційними символами дорівнювала нулю.

До систематичних кодів належать коди з перевіркою на парність (непарність), коди з повторенням, інверсний (код Бауера), кореляційний, коди Хеммінга, Голея, Ріда-Мюллера, Макдональда, Варшамова, коди з малою щільністю перевірок на парність, ітеративний код.

Різновидом систематичних кодів є циклічні коди. Окрім всіх властивостей систематичних кодів, циклічні коди (ЦК) мають таку властивість: якщо деяка дозволена кодова комбінація належить даному коду, то нова кодова комбінація, що отримується шляхом циклічної перестановки символів, також належить цьому коду.

До найбільш відомих циклічних кодів належать коди Хеммінга, Боуза-Чоудхурі-Хоквінгема, мажоритарні коди, коди Файра, Абрамсона, Міласа-Абрамсона, Ріда-Соломона, компаундні коди.

Останнім часом велика увага приділяється застосуванню згорткових кодів, які належать до класу безперервних кодів. Вони дозволяють ефективно виявляти і виправляти помилки в каналах із завмираннями сигналів.

7.3.2. Основні характеристики завадостійких кодів

Оцінка ефективності кодів проводиться за їх основними характеристиками, що виражають різні кількісні і якісні показники. Дані характеристики використовуються при виборі і обґрунтуванні кодів, призначених для передачі, зберігання і обробки інформації, наприклад довжина коду, основа коду, потужність коду, загальне число кодових комбінацій, число інформаційних і контрольних символів, надмірність коду, мінімальна кодова відстань, швидкість передачі коду, вага кодової комбінації, вагова характеристика коду, вірогідність помилкового декодування (при декодуванні з виправленням помилок), вірогідність невиявленої помилки (при декодуванні з виявленням помилок), оптимальність коду, коефіцієнт помилкових переходів і т. д.

Довжина коду n – число розрядів (символів), що складають кодову комбінацію.

Основа коду m – кількість значень ознак, використовуваних у кодових комбінаціях, що розрізняються між собою. Так, у двійковому коді $m=2$ (символи 0 і 1).

Потужність коду N_p – число дозволених кодових комбінацій, що дорівнює числу початкових інформаційних і

використовуваних для передачі повідомлень. Загальне число кодових комбінацій N_0 – число всіх можливих комбінацій, що дорівнює m^n (для двійкових кодів $N = 2^n$).

Число інформаційних символів k – кількість символів (розрядів) кодової комбінації, призначених для передачі власне повідомлення. Тоді $N_p = 2^k$.

Число перевірочних символів r – кількість символів (розрядів) кодової комбінації, необхідних для корекції помилок. Це число характеризує абсолютну надмірність коду.

У теорії кодування під надмірністю коду R розуміють відносну надмірність, що дорівнює відношенню числа перевірочних символів до довжини коду $R = r/n$.

У більш загальному випадку ця формула може бути подана як

$$R = 1 - \frac{\log_m N_p}{\log_m N_0}. \quad (7.10)$$

Швидкість передачі кодових комбінацій – це відношення числа інформаційних символів до довжини коду $R' = k/n$. Враховуючи, що $n = k + r$, маємо $R' = 1 - R$.

Вага кодової комбінації (коду) w – кількість одиниць (ненульових розрядів) у кодовій комбінації. Так, кодова комбінація 10111001 має вагу, що дорівнює 6.

Вагова характеристика коду $W(w)$ – число кодових комбінацій з вагою w . Наприклад, для коду, що містить кодові комбінації 10101, 0001, 10101, 11011, вагова характеристика $W(1) = 1$, $W(3) = 2$, $W(4) = 1$, тобто даний код складається з одного кодового слова вагою 1, двох слів вагою 3 і одного слова вагою 4.

Імовірність помилкового декодування при корекції помилок визначається виразом

$$P_{nd} \leq \sum_{q=[d_{min}/2]}^n C_n^q p^q (1-p)^{n-q}, \quad (7.11)$$

де p – імовірність одиночної помилки в каналі зв'язку;

q – кратність виникаючих помилок;

d_{min} – мінімальна кодова відстань даного коду;

C_n^q - число поєднань з n по q ;

$[d_{min}/2]$ позначає найбільшу цілу частину значення $d_{min}/2$.

Імовірність невиявленої помилки при виявленні помилок

$$P_{nn} \leq \sum_{q=d_{min}}^n C_n^q p^q (1-p)^{n-q}. \quad (7.12)$$

Оптимальність коду – властивість такого коду, який забезпечує найменшу імовірність невиявлення помилки серед всіх кодів тієї самої довжини n і надмірності r .

Коефіцієнт помилкових переходів

$$K_{nn}(d) = \frac{1}{N_p} \sum_{i=1}^N \frac{N_{pi}(d)}{C_n^d}, \quad (7.13)$$

де $N_{pi}(d)$ - число дозволених кодових комбінацій, які знаходяться від i -тої кодової комбінації на відстані d ;

C_n^d - число поєднань з n по d .

Даний коефіцієнт показує, яка частка помилок кратності d не виявляється.

7.4. Лінійні, блокові коди з виявленням і виправленням помилок

Лінійним (n, k) кодом називається сукупність 2^k кодових комбінацій із загального 2^n всіх можливих комбінацій, що задовольняє певне правило кодування. Кодові комбінації U_i систематичного блокового (n, k) коду записуються як

$$U = (a_1, a_2, a_3, \dots, a_k, a_{k+1}, a_{k+2}, \dots, a_n),$$

де a_1, \dots, a_k - число k інформаційних символів;

a_{k+1}, \dots, a_n - число (n, k) перевірочних (надмірних символів).

Числа k і n визначають швидкість передачі інформації кодом, що дорівнює k/n двійкових одиниць на один символ (де n - загальне число символів; r - перевірочні символи).

Завадостійкість (n, k) коду оцінюється за допомогою мінімальної кодової відстані d_{min} або числом виявлених і виправлених помилок $(q_v, q_{вин})$. Комбінації лінійного (n, k) коду можна розглядати як вектори в n -мірному просторі, координатами якого є двійкові символи. Для комбінацій лінійного (n, k) коду визначені операції складання $(U_i \oplus U_j)$, множення на скаляр α , тобто αU_j , що задовольняють умови комутативності, асоціативності і дистрибутивності.

7.4.1. Особливості кодування лінійних блокових кодів

Способами завдання лінійного (n, k) коду є різні формулювання правила кодування, формування кодової комбінації за відомими інформаційними символами. Лінійні (n, k) коди можуть бути задані за допомогою рівнянь формування і перевірок утворювальної G і перевірочної H матриць.

Формування перевірочних символів (n, k) коду полягає в складанні по модулю 2 певних інформаційних символів:

$$a_{k+j} = \sum_{i=1}^k a_i p_{ij}; \quad j = 1, \dots, n - k, \quad (7.14)$$

де p_{ij} - коефіцієнти, що приймають значення 0 або 1 залежно від участі інформаційного символу a_i у формуванні перевірочного a_{k+j} .

Рівнянь формування стільки, скільки перевірочних символів у кодовій комбінації (n, k) коду, тобто $(n - k)$. Якщо рівняння формування переписати як

$$\sum_{i=1}^k a_i p_{ij} \oplus a_{k+j} = 0, \quad (7.15)$$

то утворюються рівняння перевірок, використовувани при декодуванні прийнятої з каналу комбінації. Рівняння (7.15) називають узагальненою перевіркою на парність, оскільки воно показує, сума яких кодових комбінацій повинна бути парною, тобто дорівнювати 0 по модулю 2.

Лінійний (n, k) код задається вказівкою базису k -мірного підпростору, який містить k лінійно незалежних комбінацій (векторів) довжиною n . Лінійно незалежними називаються такі k комбінацій U_1, U_2, \dots, U_k , для яких $\sum_{i=1}^k \alpha_i U_i = 0$ тільки при всіх $\alpha_i = 0, i = 1, \dots, k$.

Як базисні зручно вибрати інформаційні комбінації, розташовані по осях координат k -мірного підпростору і які створюють одиничну підматрицю I . Сукупність k базисних комбінацій утворює утворювальну (породжувальну) матрицю G з k рядків і n стовпців, що складається з одиничної (інформаційної) підматриці I розмірністю $(k \times k)$ і перевіркової матриці P розмірністю $[(n - k) \times k]$, тобто

$$G = [I_k P] = \begin{bmatrix} 1000\dots 0 & P_{1,1} & P_{1,2} \dots P_{1,n-k} \\ 0100\dots 0 & P_{2,1} & P_{2,2} \dots P_{2,n-k} \\ & \vdots & \\ 0000\dots 1 & P_{k,1} & P_{k,2} \dots P_{k,n-k} \end{bmatrix} = \begin{bmatrix} U_{\sigma 1} \\ U_{\sigma 2} \\ \vdots \\ U_{\sigma k} \end{bmatrix}. \quad (7.16)$$

Значення P_{ij} в підматриці P вибираються так, щоб забезпечувалася задана мінімальна кодова відстань d_{min} коду (n, k) . Мінімальна кодова відстань d_{min} визначається мінімальною вагою рядка (або рядків) матриці G , тобто $d_{min} = W_{min}(U_{\sigma i})$ або мінімальним числом одиниць у ній.

Рядки матриці G є лінійно незалежними, що забезпечується одиничною підматрицею I , і дозволяють будь-яку кодову комбінацію (n, k) коду представити подати як

$$U = a_1 U_{\delta 1} + a_2 U_{\delta 2} + \dots + a_k U_{\delta k} = \sum_{i=1}^k a_i U_{\delta i} \quad (7.17)$$

або в матричній формі $U = AG$, де A – інформаційна послідовність символів a_1, a_2, \dots, a_k .

Наявність одиничної підматриці I в утворювальній матриці G є ознакою систематичного (n, k) коду. Це впливає з формули (7.14), оскільки перші k символів кодової комбінації співпадають з інформаційними символами.

Якщо матриця G не містить одиничну підматрицю I , то код є несистематичним. Це можливо, коли обираємо базисними інші лінійно незалежні комбінації k -мірного простору.

Перевірочна матриця H є символічним записом рівняння перевірок (7.15), в якій одиницями позначаються символи, що беруть участь у перевірках, а нулями - решта всіх символів комбінації (n, k) коду. Матриця H складається з $(n - k)$ рядків і n стовпців і для систематичного (n, k) коду має вигляд

$$H = [P^T I_{n-k}] = \begin{bmatrix} P_{11} & P_{21} & \dots & P_{k1} & 10\dots 0 \\ P_{12} & P_{22} & \dots & P_{k2} & 01\dots 0 \\ & \vdots & & & \vdots \\ P_{1n-k} & P_{2n-k} & \dots & P_{kn-k} & 00\dots 1 \end{bmatrix}. \quad (7.18)$$

Перевірочні і утворювальні матриці задовольняють рівняння $HG^T = 0$, якщо вони задають один і той самий (n, k) код. Для кодових комбінацій (n, k) коду виконується рівність, справедливість якої перевіряється при декодуванні,

$$HG^T = 0, \quad (7.19)$$

де індекс T означає транспонування матриці.

7.4.2. Кодування в кодї Хеммінга

Розглянемо завдання коду Хеммінга для виправлення однократних помилок на прикладі коду (7,4). Складемо три рівняння формування перевірочних символів коду (за числом перевірочних символів у кодовій комбінації):

$$\left. \begin{aligned} a_1 \oplus a_2 \oplus a_3 &= a_5 \\ a_1 \oplus a_2 \oplus a_4 &= a_6 \\ a_1 \oplus a_3 \oplus a_4 &= a_7 \end{aligned} \right\}. \quad (7.20)$$

Тут виконується складання по модулю 2.

З рівнянь формування легко випливають рівняння перевірок по модулю 2:

$$\left. \begin{aligned} a_1 \oplus a_2 \oplus a_3 \oplus a_5 &= 0 \\ a_1 \oplus a_2 \oplus a_4 \oplus a_6 &= 0 \\ a_1 \oplus a_3 \oplus a_4 \oplus a_7 &= 0 \end{aligned} \right\}. \quad (7.21)$$

Перевірочні і утворювальні матриці коду відповідно мають такий вигляд:

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}; \quad G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Як видно з матриці G , мінімальна кодова відстань $d_{min} = 3$, оскільки вага першого, другого та третього рядків підматриці P_{ij} однакова і дорівнює 3.

Кодери для лінійних (n, k) кодів можуть бути побудовані підставі співвідношення (7.20), що визначає спосіб формування кодової комбінації за заданими інформаційними символами.

Структурна схема кодера, що формує перевірочні символи відповідно до рівнянь (7.21) для систематичного коду Хеммінга (7,4), наведена на рис. 7.4.

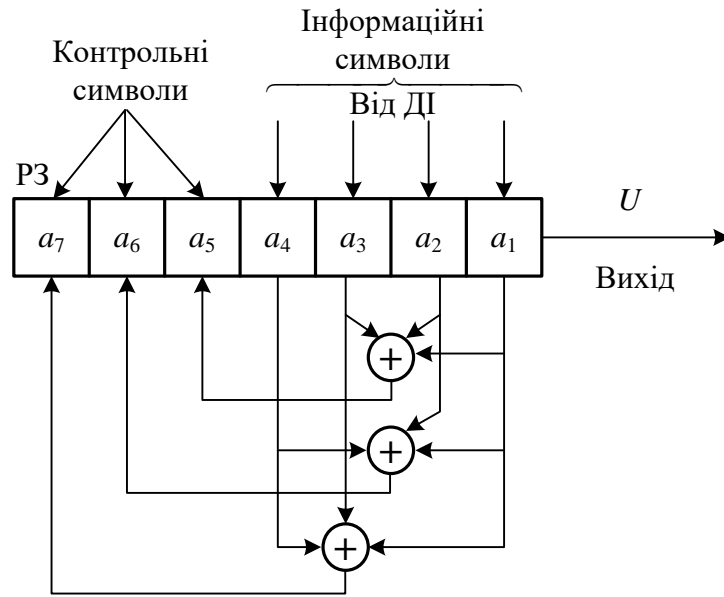


Рис. 7.4. Структурна схема кодера Хеммінга

Кодер містить семирозрядний регістр зсуву (PЗ) і три суматори по модулю 2. Чотири інформаційні символи $a_1 - a_4$ поступають від джерела інформації ДІ в PЗ. Складанням по модулю 2 певних інформаційних символів (див. вираз (7.20)) за один такт формуються перевірочні символи $a_5 - a_7$, які записуються в PЗ. Після чого сформована таким чином кодова комбінація U поступає з PЗ в канал зв'язку.

7.4.3. Особливості декодування лінійних блокових кодів

Декодування лінійних (n, k) кодів полягає у виявленні і(або) виправленні помилки в прийнятій комбінації U' . Для цього визначається синдром (ознака) прийнятої комбінації $U' = (a'_1, a'_2, a'_3, \dots, a'_n)$ вигляду

$$C = (c_1, c_2, c_3, \dots, c_{n-k}) = U'H^T. \quad (7.22)$$

Синдромом є вектор-рядок з $(n - k)$ символів $c_1, c_2, c_3, \dots, c_{n-k}$, які при підстановці формули (7.15) у вираз (7.22) можна подати як

$$C_j = \sum_{i=1}^k a'_i p_{ij} \oplus a_{k+j}, \quad j = 1, n-k.$$

Оскільки прийнята комбінація U' може відрізнитися від переданої кодової комбінації U на величину помилки $U' = U + E$, то синдром C визначається тільки вектором помилки E

$$C = U'H^T = (U \oplus E)H^T = UH^T \oplus EH^T = EH^T, \quad (7.23)$$

оскільки $UH^T = 0$ згідно з виразом (7.19).

Із співвідношення (7.19) випливає, що $C=0$, коли помилки в комбінації U' не було ($E=0$) або під впливом помилки ($E \neq 0$) одна дозволена кодова комбінація перейшла в іншу дозволена кодову комбінацію.

Якщо $C \neq 0$, у прийнятій комбінації є помилка, і для її виявлення необхідно обчислити синдром.

Виправлення помилок у прийнятій комбінації можливе, якщо встановлена взаємно однозначна відповідність між синдромом і конфігурацією помилки (однократна, двократна і т. д.). Таку відповідність можна встановити підстановкою вектора помилки $E = (e_1, e_2, \dots, e_n)$ у виразі (7.23), тоді

$$C = \sum_{i=1}^n e_i H_i^T, \quad (7.24)$$

де H_i^T - стовпець перевіркової матриці.

З формули (7.24) випливає, що при однократній помилці e_i вектор помилки E може мати вигляд $100\dots 0, 010\dots 0, \dots 00\dots 1$, тобто синдром співпадає з i -тим стовпцем матриці H або i -тим спотвореним символом прийнятої комбінації. Взаємна однозначна відповідність між синдромом C , а також помилкою e_i забезпечується, якщо стовпці перевіркової матриці різні і не нульові. У разі помилки другої кратності e_i, e_j (вектор помилки E має вигляд $101\dots 0, 011\dots 0, \dots 10\dots 1$) синдром формується складанням i -го і j -го стовпців матриці H , при цьому їх сума не

повинна співпадати з жодним із стовпців матриці. Така умова може бути забезпечена певним вибором матриці H .

За обчисленим синдромом $C = (c_1, c_2, \dots, c_n)$ здійснюється оцінка символів e_i вектора помилки E і оцінка символу переданої комбінації

$$\hat{a}_i = a'_i \oplus \hat{e}_i = a_i \oplus e_i \oplus \hat{e}_i.$$

При цьому \hat{a}_i правильна, якщо правильно оцінено символ завади $\hat{a}_i = a_i$.

7.4.4. Декодування в коді Хеммінга

Розглянемо процедуру декодування на прикладі коду Хеммінга (7,4). Схема декодера (рис. 7.5) містить регістр зсуву (РЗ), в який з каналу поступає комбінація U' . Три суматори по модулю 2 здійснюють три перевірки відповідно до співвідношень (7.21). Виходи суматорів підводяться на вхід схеми «АБО» для фіксації однократних і двократних помилок і на входи чотирьох дешифраторів помилок (схема L) за числом інформаційних символів комбінації.

Під час надходження всієї комбінації в РЗ її символи відповідно до співвідношень (7.21) подаються на три суматори для вироблення складових синдрому $C = (c_1, c_2, c_3)$. Якщо складові синдрому рівні, то згідно з рівняннями перевірок (7.21) помилки в прийнятій комбінації не було або даною перевіркою вона не виявлена. Якщо лише один з $c_i \neq 0$, то є помилка. Таким чином, синдромом помилки є сукупність результатів рівнянь перевірок, яка відповідає i -му спотвореному символу прийнятої комбінації. Оскільки стовпці матриці H різні (згідно з виразом(7.21)), то однократні помилки будуть виправлені. Для цього у схемі декодера встановлені дешифратори помилок (схеми L), налаштовані на комбінації, що відповідають першим чотирьом стовпцям матриці H .

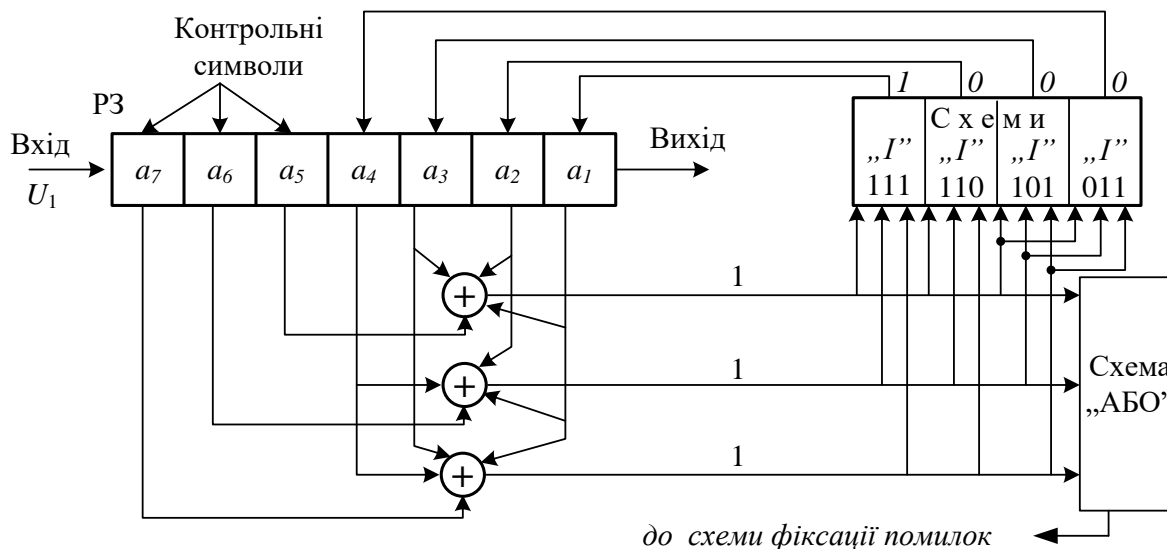


Рис. 7.5. Структурна схема декодера Хеммінга

Припустимо, що при перевірці комбінації на наявність помилки на виходах суматорів отримана комбінація вигляду 111. Тоді на виході схеми «АБО» з'являться одиниці, що означає наявність помилки, і на виході першого дешифратора також з'явиться 1, а значення a_1 в РЗ змінено на зворотне. Це означає, що для коду Хеммінга оцінка символу здійснюється за правилом $\hat{a}_i = a_i \oplus 1$.

Недоліком такої схеми декодера є швидке збільшення кількості дешифраторів помилок при зростанні числа інформаційних символів (n, k) коду, що обмежує їх застосування і вимагає використання лінійних кодів, які дозволяють зменшити об'єм обладнання декодера.

7.5. Циклічні коди

7.5.1. Способи завдання і кодування

Лінійний (n, k) код називається циклічним, якщо при циклічному зсуві символів кодової комбінації на один розряд управо (вліво) утворюються комбінації, що також належать коду.

Для алгебраїчного опису циклічних (n, k) кодів кожної комбінації $U = (a_1, a_2, a_3, \dots, a_n)$ ставиться у відповідність

багаточлен (поліном) $U(x) = (a_0x^0 + a_1x^1 + a_2x^2 + \dots + a_{n-1}x^{n-1})$ від формальної змінної x . Множення на x збільшує степінь багаточлена $U(x)$ і викликає подовження кодової комбінації на один символ. Щоб уникнути збільшення степеня $U(x)$, x^n замінюють на 1 ($x^n = 1$ або $x^n - 1 = 0$):

$$xU(x) = x(a_0x^0 + a_1x^1 + \dots + a_{n-1}x^{n-1}) = a_0x + a_1x^2 + \dots + a_{n-2}x^{n-1} + a_{n-1}x^n = a_{n-1} + a_0x + a_1x^2 + \dots + a_{n-2}x^{n-1}.$$

Для безлічі багаточленів, що відповідають n -символьним комбінаціям, вводяться операції складання, віднімання, множення, які співпадають із звичайними операціями над багаточленами. Складання коефіцієнтів багаточлена виконується по модулю 2, при цьому операції складання і віднімання співпадають. Множення багаточленів виконується по модулю $x^n - 1$, причому $x^n - 1 = x^n + 1$. Таким чином, для опису і побудови циклічних кодів використовується алгебра багаточленів по модулю $x^n + 1$, названим модулем за множенням.

При поданні кодових комбінацій багаточленами циклічним (n, k) кодом називається безліч багаточленів $U_i(x)$ степеня не більше $n - 1$, які діляться без залишку на деякий багаточлен $g(x)$ степеня $n - k$, названим утворюючим (породжучим).

Багаточлен $g(x)$ є утворювальним для циклічного (n, k) коду, якщо він є дільником багаточлена $x^n + 1$, тобто

$$x^n + 1 = g(x)h(x),$$

де $h(x)$ - багаточлен (поліном), степеня k , названий перевірочним багаточленом (n, k) коду, який так само, як і $g(x)$, задає циклічний (n, k) код. Таблиця поліномів, наведена у додатку 7.

Циклічно (n, k) код має мінімальну кодову відстань d_{min} , якщо вага комбінації, що відповідає утворювальному багаточлену

$g(x)$, не менше d_{min} . Це означає, що кількість доданків утворювального багаточлена $g(x)$ має бути не менше d_{min} .

Утворювальний багаточлен $g(x)$ циклічного коду, який виправляє однократні помилки, є неприводимим, тобто ділиться сам на себе і на 1. Для виправлення помилок більшої кратності $g(x)$ розкладається на множники $g_i(x)$, число яких дорівнює кратності помилки і кожний з них є неприводимим.

Припустимо, що потрібно побудувати циклічний (n, k) код $(7, 4)$, який виправляє всі однократні помилки. Утворювальний багаточлен $g(x)$ такого коду має степінь $n - k = 3$ і число доданків не менше трьох, оскільки мінімальна кодова відстань коду $d_{min} = 3$. У якості $g(x)$ виберемо багаточлен вигляду $x^3 + x + 1$. Переконаємося, що він є утворювальним для коду $(7, 4)$, для чого розділимо модуль за множенням $x^7 + 1$ на $x^3 + x + 1$:

$$\begin{array}{r|l} x^7+1 & x^3+x+1 \\ \hline x^7+x^5+x^4 & x^4+x^2+x+1 \\ \hline x^5+x^5+1 & \\ x^5+x^3+x^2 & \\ \hline x^4+x^5+x^3+1 & \\ x^4+x^3+x & \\ \hline x^3+x+1 & \\ x^3+x+1 & \\ \hline 0 & \end{array}$$

Багаточлен $x^4 + x^2 + x + 1$ є перевірочним $h(x)$. Оскільки ділення виконане без залишку, то багаточлен $x^3 + x + 1$ є утворювальним для коду $(7, 4)$. За багаточленами $g(x)$ і $h(x)$ можуть бути побудовані утворювальна (G) і перевірочна (H) матриці.

Утворювальна матриця G містить чотири рядки, утворені комбінаціями з семи символів, що відповідають багаточленам $g(x)$, $g(x)x$, $g(x)x^2$, $g(x)x^3$, і додатковими нулями:

$$U_1x = g(x) = 1 + x + x^3; U_1 = 1101000;$$

$$\begin{aligned}
U_2x &= xg(x) = x + x^2 + x^4; U_2 = 0110100; \\
U_3x &= x^2g(x) = x^2 + x^3 + x^5; U_3 = 0011010; \\
U_4x &= x^3g(x) = x^3 + x^4 + x^6; U_4 = 00011010;
\end{aligned}$$

$$G = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

Перевірочна матриця H містить три рядки, утворені комбінаціями, що відповідають багаточленам $h(x)$, $h(x)x$, $h(x)x^2$, і подається як

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

Порівняння отриманих матриць з матрицями (7.16) лінійного коду (7,4) показує, що вони відрізняються розташуванням стовпців, тому коди, які їм відповідають, є еквівалентними.

При заданому $g(x)$ кодові багаточлени $U_i(x)$ систематичного коду утворюються за правилом

$$U(x) = l(x)x^{n-k} + R(x), \quad (7.25)$$

де $l(x)$ - багаточлен, що відповідає інформаційній послідовності з k -символів;

$R(x)$ - залишок від ділення багаточлена $l(x)x^{n-k}$ на $g(x)$.

Розглянемо приклад для коду (7,4). Припустимо, що $l=1001$, їй відповідає $l(x) = 1 + x^3$. Багаточлен $l(x)x^{n-k} = (1 + x^3)x^3 = x^3 + x^6$ розділимо на $g(x) = 1 + x + x^3$:

$$\begin{array}{r} x^6+x^3 \\ x^6+x^4+x^3 \end{array} \left| \begin{array}{l} x^3+x+1 \\ x^3+x \end{array} \right. \\ \hline x^4 \\ x^4+x^2+x \\ \hline x^2+x$$

Звідси випливає, що залишок від ділення $R(x) = x^2 + x$, або в двійковому коді 110.

Кодовому багаточлену $U(x) = l(x)x^{n-k} + R(x) = x^6 + x^3 + x^2 + x$ відповідає кодова комбінація 1001110, в якій чотири перших символи - інформаційні (1001), три, що залишилися, - перевіірочні.

Схема кодера, що реалізує алгоритм кодування (7.25), будується за утворювальним багаточленом $g(x)$ і є схемою ділення на $g(x)$. Один з варіантів схеми показаний на рис. 7.6, а. Схема складається з $n-k$ -елементів реєстра зсуву на D -тригерах ($S_1 - S_3$) і суматорів по модулю 2, число яких дорівнює числу знаків складання в $g(x)$. Побудова схеми заснована на поданні $g(x) = 1 + x + x^3$ як $g(x) = g_0x^0 + g_1x^1 + g_2x^2 + g_3x^3$, де $g_0 = g_1 = g_3, g_2 = 0$.

Для коду (7,4) структурна схема кодера наведена на рис. 7.6, б. Роботу схеми зручно розглядати по тактах, оскільки D -тригери зазвичай використовуються як елементи реєстра зсуву. Для опису їх стану вводиться поняття затримки D на один такт.

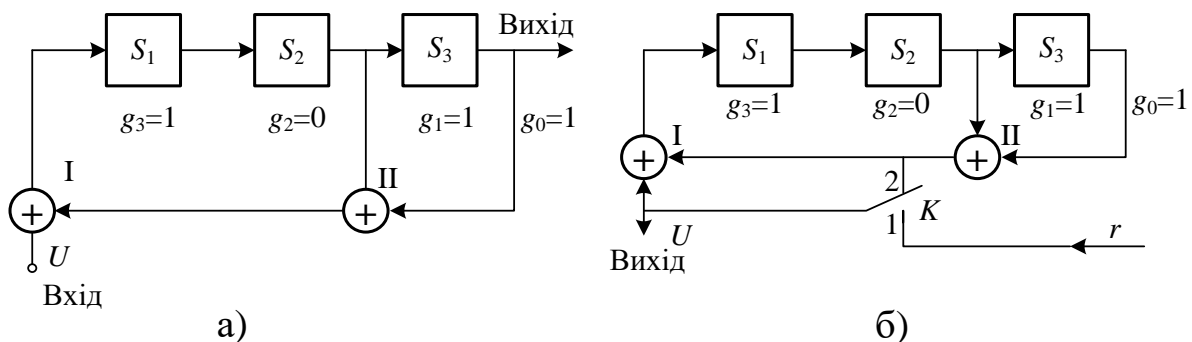


Рис. 7.6. Варіанти структурних схем кодерів циклічного коду

Початковий стан D -тригерів – нульовий. Схема працює таким чином. Перші чотири такти ключ K знаходиться в положенні 1, і інформаційна послідовність l з чотирьох символів поступає в канал і на вхід схеми, де відбувається її ділення на $g(x)$. Після 4-го такту ключ K переходить в положення 2, закорочуючи вхід суматора I . У результаті цього стан S_1 після 5-го такту залишається нульовим до кінця роботи схеми. Починаючи з 5-го такту в канал поступають три перевіірочні символи з виходу суматора II , які формуються складанням станів S_2 і S_3 . Робота схеми при подачі на її вхід послідовності $l = 1001$ або $l(x) = x^3 + 1$ пояснюється табл. 7.1. Кодер працює 7 тактів за числом символів у кодовій комбінації ($n = 7$). У перші 4 такти в канал поступають інформаційні символи ($k = 7$), а з 5-го по 7-й такти – перевіірочні ($n - k = 3$).

Таблиця 7.1

Стан елементів регістра зсуву у схемі кодера

Такт	Положення ключа K	l Стан елементів РЗ визначається такими співвідношеннями: $S_1 = D(l + S_2 + S_3);$ $S_2 = DS_1;$ $S_3 = DS_2$	Стан елементів РЗ				U
			S_1	S_2	S_3	$S_2 + S_3$	
1-й	1	1	0	0	0		1
2-й		0	1	0	0		0
3-й		0	0	1	0		0
4-й		1	1	0	1		1
5-й	2	0	0	1	0	1	1
6-й		0	0	0	1	1	1
7-й		0	0	0	0	0	0
8-й	-	0	0	0	0	0	

7.5.2. Декодування циклічних кодів

Декодування засноване на тому, що будь-який кодовий багаточлен $U_i(x)$ ділиться на утворювальний багаточлен $g(x)$ без залишку відповідно до формули (7.26). Якщо прийнята комбінація не належить циклічному коду, то відповідний їй багаточлен $U'(x) = U(x) + E(x)$ не ділиться на $g(x)$, тобто

$$\frac{U'(x)}{g(x)} = \frac{U(x) + E(x)}{g(x)} = Q(x) + \frac{E(x)}{g(x)} = Q(x) + C(x), \quad (7.26)$$

де $C(x)$ – залишок від ділення $U'(x)$ на $g(x)$;

$Q(x)$ – частка від ділення.

З формули (7.26) випливає, що багаточлен залишку $C(x)$ визначається тільки багаточленом помилки $E(x)$ і не залежить від кодового багаточлена $U(x)$. Коли $E(x)$ належить циклічному коду, тоді $E(x)$ ділиться на $g(x)$ і $C(x) = 0$. Якщо $E(x)$ не належить коду, то $C(x) \neq 0$. Це дозволяє за ненульовим залишком $C(x)$ від ділення $U'(x)$ на $g(x)$ виявити помилки.

Відповідні однократним помилкам багаточлени $C_i(x)$ отримують діленням на $g(x)$ багаточленів завади вигляду $E(x) = x^{n-i}$, де $i = \overline{1, n}$ – номер спотвореного символу комбінації (n, k) коду. Їм відповідають комбінації, які є стовпцями перевірконої матриці H лінійного коду і є відповідними синдромами S відповідно до формул (7.16) і (7.18).

При помилці q -тої кратності залишок $C(x)$ визначається сумою багаточленів залишків, що відповідають однократним помилкам, так само як для довільного лінійного коду синдром визначається складанням стовпців матриці H відповідно до співвідношень (7.16) і (7.18). Таким чином, залишок $C(x)$ є синдромом помилки в циклічних кодах.

Визначення синдрому діленням $U(x)$ на $g(x)$ вимагає менше обладнання, ніж його обчислення відповідно до формули (7.24). Таким чином, основним елементом у схемі декодера циклічного коду, як і в кодері систематичного коду, є схема

ділення, названа генератором синдромів (ГС). Узагальнена структурна схема декодера циклічного коду наведена на рис. 7.7.

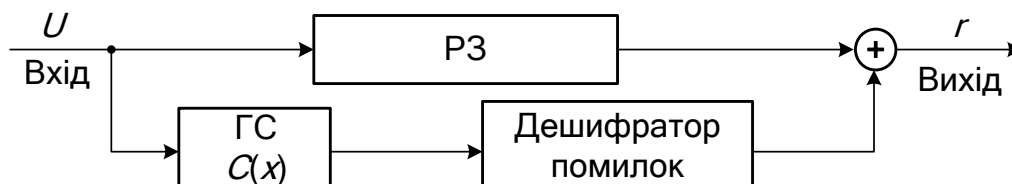


Рис. 7.7. Узагальнена структурна схема декодера циклічного коду

Складність декодера визначається дешифратором помилок і істотно залежить від кратності помилок, що виправляються. При однократній помилці дешифратор має один вихід і $(n - k)$ входів. У разі помилок другої кратності дешифратор має n виходів, об'єднаних схемою «АБО».

При виправленні помилок кратності q_{en} число виходів дешифратора визначається з виразу

$$N = \sum_{i=1}^{q_{en}} C_{n-1}^{i-1} = \sum_{i=1}^q \frac{(n-1)!}{(i-1)!(n-i)!}$$

і швидко зростає при збільшенні числа символів кодової комбінації n і кратності помилки q_i . Відомі різні схеми декодування циклічних кодів, але всі вони засновані на обчисленні синдромів помилки відповідно до співвідношень (7.16) і (7.18) і побудовані на схемах ділення $U(x)$ на $g(x)$.

Розглянемо один з варіантів на прикладі коду (7,4), що виправляє однократні помилки. Він заснований на обчисленні синдрому однократної помилки в старшому розряді комбінації (n, k) коду при діленні багаточлена помилки $E(x) = x^{n-1}$ на $g(x)$ і його використанні для виправлення помилки в будь-якому символі прийнятої комбінації. Обчислення цього синдрому здійснюється схемою ділення (рис. 7.8), на вхід якої подається вектор помилки $E=1000000$, що є багаточленом помилки

$E(x) = x^6$. Стан елементів $S_1 - S_3$ схеми ділення показаний в табл. 7.2.

Починаючи з 2-го по 8-й такти результати обчислення відповідає складовим синдромів помилки в першому (молодшому) розряді кодової комбінації [такт 2 (100)], у другому розряді [такт 3 (010)] і т. д. до 8-го такту, якому відповідає синдром $C=001$. У цьому легко переконатися, подавши на вхід схеми вектори помилки вигляду $E=0100000$, 0010000 і т. д. Якщо продовжити зсув складових синдромів після 8-го такту, то можна переконатися, що вони повторюються. На цій властивості циклічності складових синдромів (станів $S_1 - S_3$) засновано виправлення однократних помилок. Для цього як дешифратор помилки використовується схема I , що має три входи і один вихід і яка налаштована на синдром $C=001$.

Розглянемо алгоритм декодування циклічного коду, який полягає у наступному.

1. Комбінація U , що поступає з демодулятора, протягом 7-ми тактів записується в регістр зсуву і одночасно ділиться на утворювальний багаточлен $q(x)$ у схемі ГС. Якщо помилки в комбінації U не було або даною перевіркою вона не виявлена, то стан елементів $S_1 - S_3$ ГС на 8-му такті буде нульовим ($C=000$). За наявності помилки в комбінації U стан елементів ГС відмінний від нульового і визначає складові синдрому C' .

2. Починаючи з 8-го такту символи комбінації U «виштовхуються» з РЗ, а ГС здійснюють циклічний зсув складових синдрому C' до моменту встановлення елементами $S_1 - S_3$ стану 001, на яке настроєна схема I .

3. На виході схеми I з'являється 1, що подається на вхід суматора; на другий вхід суматора з РЗ поступає спотворений символ.

Цьому алгоритму відповідає схема декодера (рис. 7.8). Ключ $K1$ розмикається після 7-го такту, коли комбінація U повністю поступила в РЗ, а ключ $K2$ - після 11-го такту, коли чотири інформаційні символи вийдуть з РЗ. У загальному випадку схема працює $2n$ тактів. У даному випадку, враховуючи, що код систематичний, то достатньо 11 тактів роботи для виправлення помилок тільки в інформаційних символах.

Таблиця 7.2

Стан елементів $S_1 \div S_3$ схеми ділення

Такт	E	S_1	S_2	S_3
1-й	1	0	0	0
2-й	0	1	0	0
3-й	0	0	1	0
4-й	0	1	0	1
5-й	0	1	1	0
6-й	0	1	1	1
7-й	0	0	1	1
8-й	0	0	0	1
9-й	0	1	0	0
10-й	0	0	1	0
11-й	0	1	0	1

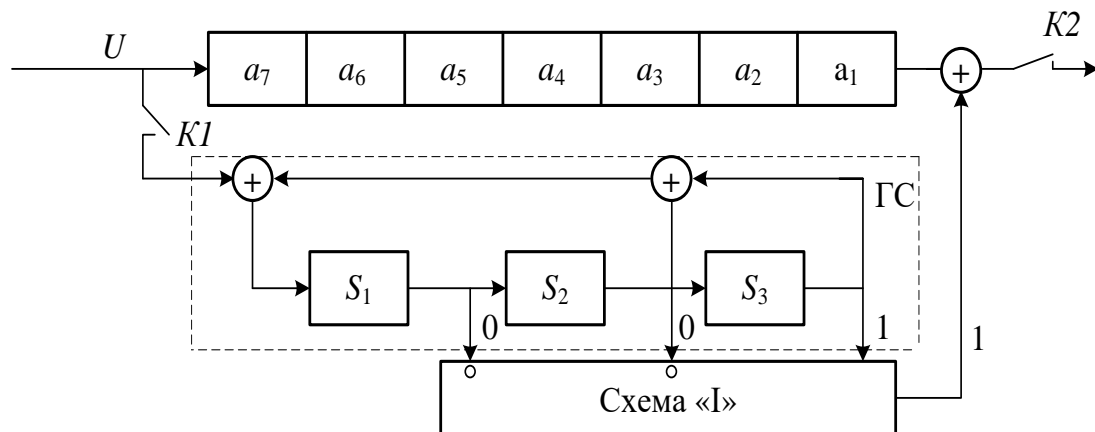


Рис. 7.8. Структурна схема декодера циклічного коду

У табл. 7.3 наведений стан елементів ГС під час появи з демодулятора спотвореної комбінації $U' = 1011110$ (неспотворена комбінація $U = 1001110$).

Як видно з табл. 7.3, на 10-му такті ГС приходить у стан 001, спрацьовує схема I, і 1 з її виходу поступає на суматор, на інший вхід якого поступає спотворений символ. У результаті складання відбувається виправлення помилки.

Таблиця 7.3

Стан елементів генератора синдромів помилок

Такт	E	S_1	S_2	S_3
1-й	1	0	0	0
2-й	0	1	0	0
3-й	1	0	1	0
4-й	1	0	0	1
5-й	1	0	0	0
6-й	1	1	0	0
7-й	0	1	1	0
8-й	0	1	1	1
9-й	0	0	1	1
10-й	0	0	0	1

7.5.3. Мажоритарне декодування циклічних кодів

Складність дешифратора помилок у схемі декодера (рис. 7.7) вимагає використовувати інші методи декодування циклічних кодів при великих n і q_i . Одним з них є мажоритарне декодування, яке застосоване до будь-яких лінійних кодів.

Принцип мажоритарного декодування полягає в тому, що для кожного символу кодової комбінації можуть бути отримані декілька оцінок. Рішення про переданий символ ухвалюється за більшістю однакових оцінок. Основою для отримання різних оцінок одного і того самого символу є рівняння перевірок або перевірна матриця H .

Розрізняють розділені (незалежні) і зв'язані (залежні) оцінки. Розділеними називають оцінки, в яких будь-який символ прийнятої комбінації бере участь не більше ніж в одному з рівняння оцінок, інакше оцінки є зв'язаними. Вони характеризуються коефіцієнтом зв'язаності α , що дорівнює максимальному числу оцінок, які містять один і той самий символ.

Загальна кількість оцінок для виправлення помилок кратності q_i

$$N = 2\alpha q_i + 1. \quad (7.27)$$

Наприклад, для коду (7,4), що виправляє однократні помилки, мінімальне число оцінок $N=3$, якщо коефіцієнт зв'язаності $\alpha=1$, тобто оцінки є розділеними. Для складання оцінок на прикладі коду (7,4) скористаємося перевірочним багаточленом $h(x)=1+x+x^2+x^4$ і запишемо матрицю H , в якій перший рядок відповідає перевірочному багаточлену, друга отримана двома циклічними зсувами першої, третя – двома зсувами другої:

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}.$$

Випишемо три оцінки \hat{a}_1 за рядками матриці H :

$$\hat{a}_1 = \hat{a}_2 \oplus \hat{a}_3 \oplus \hat{a}_5; \quad \hat{a}_1 = \hat{a}_3 \oplus \hat{a}_6 \oplus \hat{a}_7; \quad \hat{a}_1 = \hat{a}_4 \oplus \hat{a}_5 \oplus \hat{a}_6.$$

З рівнянь оцінок випливає, що вони є зв'язаними, оскільки символи a_3, a_5 і a_6 беруть участь в оцінках двічі, тобто $\alpha=2$. Тому для виправлення однократних помилок у даному випадку потрібно 5 оцінок згідно з виразом (7.27). Відсутні дві оцінки одержимо складанням трьох рядків матриці $H - \hat{a}_1 = \hat{a}_2 \oplus \hat{a}_4 \oplus \hat{a}_7$; тоді очевидна оцінка $\hat{a}_1 = \hat{a}_1$. Схема мажоритарного декодера в першому такті роботи наведена на рис. 7.9. Вхідний у схему мажоритарний елемент (МЕ), який складається з порогового елемента і схеми «І», дає на виході 1, якщо 1 подається на більшу частину входів, наприклад три з п'яти.

У наступний такт символи в регістрі зсовуються на 1, здійснюється оцінка \hat{a}_2 і т. д.

$$\hat{a}_2 = \hat{a}_2; \quad \hat{a}_2 = \hat{a}_3 \oplus \hat{a}_4 \oplus \hat{a}_6; \quad \hat{a}_2 = \hat{a}_3 \oplus \hat{a}_5 \oplus \hat{a}_1; \\ \hat{a}_2 = \hat{a}_4 \oplus \hat{a}_7 \oplus \hat{a}_1; \quad \hat{a}_2 = \hat{a}_5 \oplus \hat{a}_6 \oplus \hat{a}_7.$$

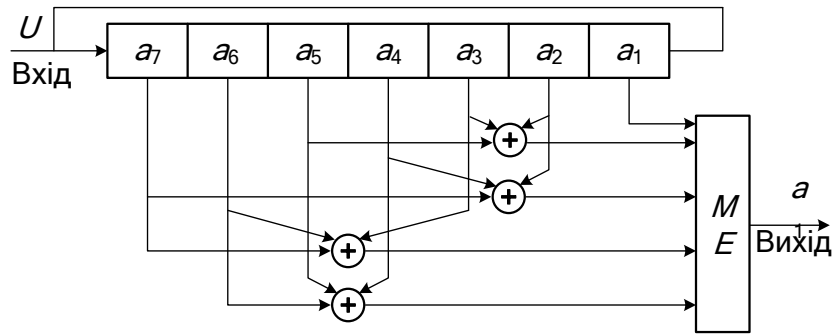


Рис. 7.9. Структурна схема мажоритарного декодера

Мажоритарний декодер вимагає менше обладнання, ніж циклічний, особливо для кодів з розділеними оцінками, що містять мале число символів у кожній оцінці.

7.6. Згорткові коди

7.6.1. Визначення, способи завдання і кодування

Згорткові коди належать до безперервних кодів. Ці коди мають велику надмірність, характеризуються простотою реалізації кодуючих пристроїв і найбільш придатні для виправлення пакетів помилок.

Згортковими коди називають тому, що кодове слово можна розглядати як відгук (згортку) кодера, який є лінійною системою, на вхідну інформаційну послідовність.

Згорткові коди поділяються на систематичні і несистематичні. У систематичних кодах за інформаційними символами ідуть перевірочні. Таке розділення символів дозволяє спростити формування оцінок інформаційних символів на приймальній станції. У систематичних згорткових кодах кодове слово складається з N блоків завдовжки n_0 з k інформаційними символами, що визначає швидкість коду k/n_0 . Перевірочний символ блоку залежить від інформаційних символів з блоків N . Це визначає основну відмінність згорткових кодів від блокових. Якщо $N=1$, згортковий код перетворюється на блоковий. Добуток $N n_0$ називають кодовим обмеженням.

Для отримання кодового слова систематичного згорткового коду потрібно за відомими інформаційними символами знайти перевірочні і розташувати їх на певних позиціях блоку n_0 .

Згортковий код може бути заданий декількома способами:

а) застосуванням $n_0 - 1$ рекурентних співвідношень, що визначають перевірочні символи за відомими інформаційними;

б) за допомогою утворювального багаточлена $g(x) = g_0 \oplus g_1x \oplus \dots \oplus g_mx^m$ при кодуванні однієї інформаційної послідовності $k=1$. Степінь $g(x)$ визначає число блоків, перевірочні символи яких залежать від інформаційних, тобто $N = m + 1$. При кодуванні k інформаційних послідовностей згортковий код задається за допомогою $k(n_0 - k)$ утворювальних багаточленів;

в) з використанням графа і кодових решіток, а також утворювальної і перевірочної матриць.

Розглянемо приклад найпростішого згорткового коду, в якому кожному інформаційному символу a_i відповідає один перевірочний b_i . Для такого коду $k = 1$, $n_0 = 2$, $N = 2$, швидкість коду $k/n_0 = 1/2$. Кодове слово є напівнескінченною послідовністю символів $a_0b_0, a_1b_1, a_2b_2, \dots, a_ib_i, \dots$

Перевірочний символ слова визначається через інформаційні лінійним рекурентним співвідношенням [39]

$$b_i = a_i \oplus a_{i-1}. \quad (7.28)$$

Породжуючий багаточлен коду $g(x) = g_0 \oplus g_1x = 1 + x$. При відомій інформаційній послідовності l і $g(x)$ визначаються перевірочні символи згорткового коду. Наприклад, інформаційна послідовність має вигляд $l=101100$, де $a_0 = 1, a_1 = 0, a_2 = 1, a_3 = 1, a_4 = 0, a_5 = 0$. Їй відповідає багаточлен $l(x) = 1 + x^2 + x^3 + \dots$. Тоді перевірочні символи утворюються таким чином:

$$b(x) = l(x)g(x) = l(x)(1 + x) = 1 + x + x^2 + x^4.$$

Цьому багаточлену відповідає послідовність $b=11101$, де $b_0 = 1, b_1 = 1, b_2 = 1, b_3 = 0, b_4 = 1$ - перевіірочні символи. Кодова послідовність $U=1101111001$. Граф згорткового коду показаний на рис. 7.10.

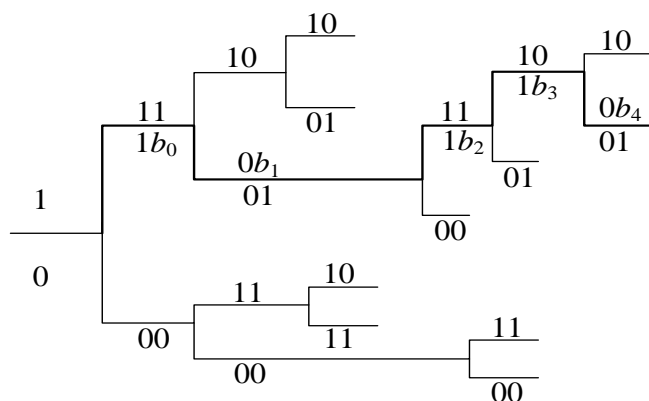


Рис. 7.10. Граф згорткового коду

Використовуючи рекурентне співвідношення (7.28), можна скласти систему перевірок для символів, що з'являються на виході кодера,

$$\left. \begin{aligned} a_0 \oplus b_0 &= 0 \\ a_0 \oplus a_1 \oplus b_1 &= 0 \\ a_1 \oplus a_2 \oplus b_2 &= 0 \\ a_2 \oplus a_3 \oplus b_3 &= 0 \end{aligned} \right\}.$$

Використовуючи перевірки як рядки перевіірочної матриці, складемо перевіірочну матрицю коду

$$H = \begin{pmatrix} 11000\dots \\ 1011000\dots \\ 001011\dots \\ 00001011\dots \\ \dots\dots\dots \end{pmatrix}.$$

Схема кодера такого коду наведена на рис. 7.6, а і може бути побудована за співвідношенням (7.23) або утворювальним багаточленом $g(x)$.

7.6.2. Декодування згорткових кодів

Для декодування комбінацій згорткових кодів розроблені алгебраїчні та імовірнісні методи. До алгебраїчних методів належать порогове (мажоритарне) декодування, до імовірнісних – послідовне декодування і декодування за максимумом правдоподібності (алгоритм Вітербі) [12].

Алгебраїчні методи засновані на виконанні рівнянь перевірок, що зв'язують інформаційні і перевірочні символи. Імовірнісні методи засновані на пошуку шляху за графом (кодовим деревом), найближчим до прийнятої послідовності.

При пороговому декодуванні для кожного блоку n_0 обчислюється синдром. Наприклад, для коду, розглянутого вище, відповідно до рекурентного співвідношення $b_i = a_i \oplus a_{i+1}$:

$$\begin{aligned}c_i &= b_i \oplus a_i \oplus a_{i-1} = e_i^b \oplus e_i^a \oplus e_{i-1}^a; \quad i=1,2,3,\dots; \\c_{i+1} &= b_{i+1} \oplus a_{i+1} \oplus a_i = e_{i+1}^b \oplus e_{i+1}^a \oplus e_i^a,\end{aligned}$$

де $e_i^b, e_i^a, e_{i+1}^a, e_{i+1}^b$ - компоненти вектора помилки i -го блоку на позиціях інформаційних і перевірочних символів.

Якщо $c_i=1$, то фіксується наявність помилки. У разі $c_i=0$ найімовірніше, що символи a_i, b_i, a_{i-1} неспотворені, проте при спотворенні двох символів з трьох $c_i=0$.

Правило оцінки компонента вектора помилки визначається співвідношенням

$$e_i^a = (c_i \oplus e_{i-1}^a) \oplus c_{i+1}. \quad (7.29)$$

Оцінка інформаційного символу здійснюється за правилом

$$\hat{a} = \hat{a}_i \oplus e_i^a. \quad (7.30)$$

Компонент вектора помилки e_{i-1}^a може бути оцінений неправильно, тоді його підстановка у вираз для c_i збільшує імовірність неправильного рішення при оцінці a_i , яка

використовується для оцінювання a_{i+1} і т. д., тобто виникає ефект розмноження помилок. Для підвищення достовірності декодованих символів необхідно аналізувати достатньо довгу послідовність синдромів $c_{i-1}, c_i, c_{i+1}, \dots, c_{i+k}$, що зазвичай виконується в алгоритмах імовірнісного декодування, які забезпечують кращу завадостійкість.

Структурна схема декодера, побудована відповідно до правил (7.29), (7.30) для згорткового коду $1/2, k=1, n_0=2$, наведена на рис. 7.11.

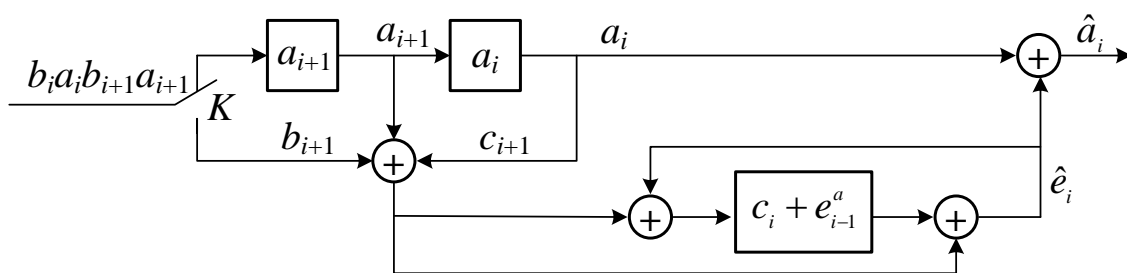


Рис. 7.11. Структурна схема декодера згорткового кода

При декодуванні комбінацій згорткових кодів з більшою надмірністю і більш великим кодовим обмеженням, ніж розглянутий код, застосовується послідовне декодування і декодування за максимумом правдоподібності. Відмінність між ними полягає в тому, що при послідовному декодуванні визначення правильного шляху за кодовим деревом на кожному кроці здійснюється з урахуванням всіх попередніх. Ухвалення рішення при декодуванні за максимумом правдоподібності (алгоритм Вітербі) здійснюється на кожному кроці з урахуванням попереднього. Мірою схожості прийнятої послідовності на передану є відстань Хеммінга.

7.7. Застосування завадостійких кодів у системах залізничної автоматики, телемеханіки і зв'язку

Завадостійкі коди в системах залізничної автоматики і телемеханіки використовуються переважно для виявлення помилок різної кратності.

У системах телемеханіки найширше використовуються коди з постійною вагою. Кодові комбінації таких кодів містять постійне число одиниць і нулів. Загальне число дозволених комбінацій визначається числом поєднань C_n^m , де n – число символів у кодовій комбінації, m – число одиниць у ній. Тому ці коди називають також кодами на одне поєднання. Коди мають мінімальну кодову відстань $d_{min}=2$ і виявляють однократні помилки.

Різновидом коду з постійною вагою, використовуваного, наприклад, у системі станційної кодової централізації (СКЦ), є розподільний код. Всі комбінації цього коду мають вагу, що дорівнює одиниці, тобто в будь-якій комбінації з n символів міститься тільки одна одиниця.

У системах телемеханіки розподільний код використовується для кодування номера команди, що міститься у виконавчій частині кодової комбінації. Для кодування адреси групи об'єктів у виборчій частині кодової комбінації застосовується код з великим числом одиниць. Наприклад, по каналу телекерування системи СКЦ передаються кодові комбінації, виконавча частина яких визначається кодом C_{10}^1 , а виборча частина - кодом C_9^3 .

В агрегатній системі диспетчерської централізації (АСДЦ) для виявлення помилок використовуються систематичні завадостійкі коди: код Бауера, що є різновидом коду з простим повторенням, і код з однією перевіркою на непарність (перевіркою за паритетом). Принцип кодування кодом Бауера полягає в повторенні інформаційної послідовності з k символів, якщо вона містить непарне число одиниць, і в інверсному повторенні інформаційних символів – при парному числі одиниць. Наприклад, інформаційним послідовностям 0111 і 0101 відповідають кодові комбінації $U_1=01110111$ і $U_2=001011010$. Очевидно, що надмірність коду Бауера дорівнює 0,5.

У комбінаціях коду з перевіркою на непарність використовується один перевірючий символ, при складанні якого з інформаційними символами по модулю 2 їх сума дорівнює одиниці.

Імовірність появи невиявленої помилки P_{nn} в коді Бауера для комбінації з n символів може бути визначена з виразу

$$P_{nn} = \left[C_{n/2}^2 p_e^2 (1 - p_e)^{0,5n-2} \right]^2 \approx \left(C_{n/2}^2 \right)^2 p_e^4,$$

де $C_{n/2}^2$ - число поєднань з $n/2$ елементів по 2;

p_e - імовірність помилкового прийому символу кодової комбінації.

При $n=8$ і $p_e=1 \cdot 10^{-4}$ імовірність появи невиявленої помилки $P_{nn}=3,6 \cdot 10^{-15}$.

Імовірність появи невиявленої помилки в коді з перевіркою на непарність знаходиться з формули

$$C_n^2 p_e^2 (1 - p_e)^{n-2} \approx C_n^2 p_e^2.$$

Якщо $p_e=1 \cdot 10^{-4}$ і $n=8$, то $P_{nn}=2,8 \cdot 10^{-7}$. У перспективних мікропроцесорних системах залізничної автоматики і телемеханіки для виявлення несправностей в окремих платах і вузлах використовується так званий сигнатурний аналіз, який засновано на принципі декодування циклічних кодів, і він полягає в тому, що тестова послідовність, яка поступає з виходу контрольованої плати або вузла, ділиться у схемі ділення на деякий утворювальний багаточлен $g(x)$. Залишок від ділення є ознакою несправності.

У вживаних на залізничному транспорті цифрових системах зв'язку і передачі даних найбільшого поширення набули циклічні коди для виявлення помилок різної кратності, звані кодами циклічної перевірки надмірності (CRC – Cyclic Redundancy Check), а в стільникових системах зв'язку з рухомими об'єктами також і згорткові коди для виправлення незалежних помилок у поєднанні з пристроями переміщення. Наприклад, у модемних протоколах корекції помилок і протоколах каналів передачі кадрів у мережах передачі даних використовується код CRC з утворювальним багаточленом $g(x) = x^{16} + x^{12} + x^5 + 1$. Код CRC-16 містить 16 перевірочних символів, має мінімальну кодову

відстань $d_{min}=4$ і дозволяє виявляти помилки до третьої кратності включно.

Імовірність виявлення помилок кодом CRC визначається за формулою

$$p_0 = 1 - \frac{1}{2^m},$$

де m – ступень утворювального багаточлена $g(x)$.

Так, код CRC-16 дозволяє виявити 99,9985 % помилок. Разом з ним у модемних протоколах V42 використовується код CRC з $d_{min}=15$ і утворювальним багаточленом

$$g(x) = x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1.$$

У мережах високошвидкісної передачі даних (АТМ мережах) використовуються укорочені циклічні коди БЧХ (Боуза-Чоудхурі-Хоквінгема). Наприклад, для виявлення помилок до третьої кратності і виправлення помилок першої кратності в заголовках осередків АТМ використовується код CRC з утворювальним багаточленом $g(x) = x^8 + x^2 + x + 1$. Для виявлення однократних і двократних помилок у протокольних блоках даних мереж АТМ використовується CRC код з утворювальним багаточленом $g(x) = x^4 + x + 1$ і $d_{min}=3$.

Принцип циклічної перевірки надмірності кодами CRC аналогічний до принципів кодування і декодування циклічних кодів і полягає в наступному. На станції, що передає, визначається залишок від ділення інформаційного багаточлена (інформаційної послідовності) на утворювальний багаточлен $g(x)$, який у вигляді перевірочних символів передається після інформаційних символів у загальному блоці даних. На приймальній станції виконується ділення прийнятої послідовності на $g(x)$, і в разі отримання залишку від ділення ухвалюється рішення про наявність помилки.

Для захисту від помилок у радіоканалах рухомого зв'язку стандарту GSM використовуються циклічні і згорткові коди у поєднанні з перемеженням.

Перемеження символів передаваної послідовності забезпечує перетворення пакетів помилок на незалежні помилки і полягає в тому, що кодова комбінація довжиною n , яка підлягає передачі по радіоканалу, записується по рядках у матрицю розміром $n_1 n_2$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n_1} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n_1} \\ & \vdots & & & \vdots \\ a_{n_2 1} & a_{n_2 2} & a_{n_2 3} & \dots & a_{n_2 n_1} \end{pmatrix},$$

потім зчитується по стовпцях у вигляді послідовності $a_{11} a_{21} a_{31} \dots a_{n_2 1} a_{12} a_{22} \dots a_{n_2 2} \dots$ і передається в канал. У результаті символи повідомлення $a_{11} a_{21} a_{31}$, які в початковій послідовності слідували один за одним, передаються в канал з інтервалом n_2 символів, який називається глибиною перемеження. Збільшення глибини перемеження підвищує ефективність згорткового кодування, але призводить до затримки передаваних символів.

На виході каналу відбувається зворотний перезапис прийнятої послідовності символів. Запис здійснюється по стовпцях, зчитування - по рядках. У відновленій послідовності пакетом помилок виявляються спотворені несуміжні, розкидані по всій довжині $n = n_1 = n_2$ символи, що спрощує їх виправлення згортковим кодом.

Для виправлення незалежних помилок у повношвидкісних каналах зв'язку систем GSM використовується згортковий код із швидкістю $1/2$ і $k = 5$. Декодування згорткового коду здійснюється за алгоритмом Вітербі.

Розглянутий приклад кодування–декодування повідомлень у стільникових системах зв'язку стандарту GSM характерний для всіх систем зв'язку з рухомими об'єктами і є одним з варіантів каскадного кодування–декодування.

Як впливає з наведених прикладів, у системах зв'язку і телемеханіки найбільшого застосування отримали завадостійкі

коди для виявлення і виправлення однократних помилок. Це обумовлено двома чинниками: недоліками, зв'язаними із застосуванням завадостійких кодів, і організацією каналів у системах зв'язку і телемеханіки. Недоліком застосування завадостійких кодів у системах передачі інформації є те, що методи завадостійкого кодування не адаптовані до заводової обстановки, що змінюється, тому надмірність (перевірочні символи) доводиться вводити, орієнтуючись на найсприятливіший випадок. У решті часу при низькому рівні завад швидкість передачі інформації по каналу нижче за можливу.

7.8. Способи підвищення правильності передачі інформації в каналах зв'язку з завадами

Принципи кодування і декодування завадостійких кодів полягають головним чином у припущенні передачі інформації по двійковому симетричному каналу зв'язку, що не завжди виконується.

Необхідність забезпечення високої правильності передачі інформації в умовах завад вимагає застосування складних завадостійких кодів і, як наслідок, складних і дорогих кодуючих і декодуючих пристроїв. Ці недоліки зводяться до мінімуму, якщо використовувати двосторонні (дуплексні) канали систем зв'язку і телемеханіки, коли повідомлення передаються в прямому і зворотному напрямках. У цьому випадку є можливість підвищити завадостійкість їх передачі завдяки використанню зворотного каналу: при виявленні помилок у прийнятій комбінації її передача повторюється. У деяких системах передачі даних однократні помилки виправляються, а помилки більшої кратності виявляються. Повторення відбувається до тих пір, поки не відбудеться передача без помилок.

Системи передачі із зворотним зв'язком мають властивість адаптації: вони автоматично змінюють швидкість передачі інформації при зміні рівня завад у каналах зв'язку.

Залежно від того, на якому боці системи передачі інформації ухвалюється рішення про повторення комбінації і спосіб використання зворотного зв'язку, розрізняють системи з

інформаційним зворотним зв'язком (ІЗЗ) і системи з вирішальним зворотним зв'язком (ВЗЗ). Структурні схеми таких систем наведені на рис. 7.12.

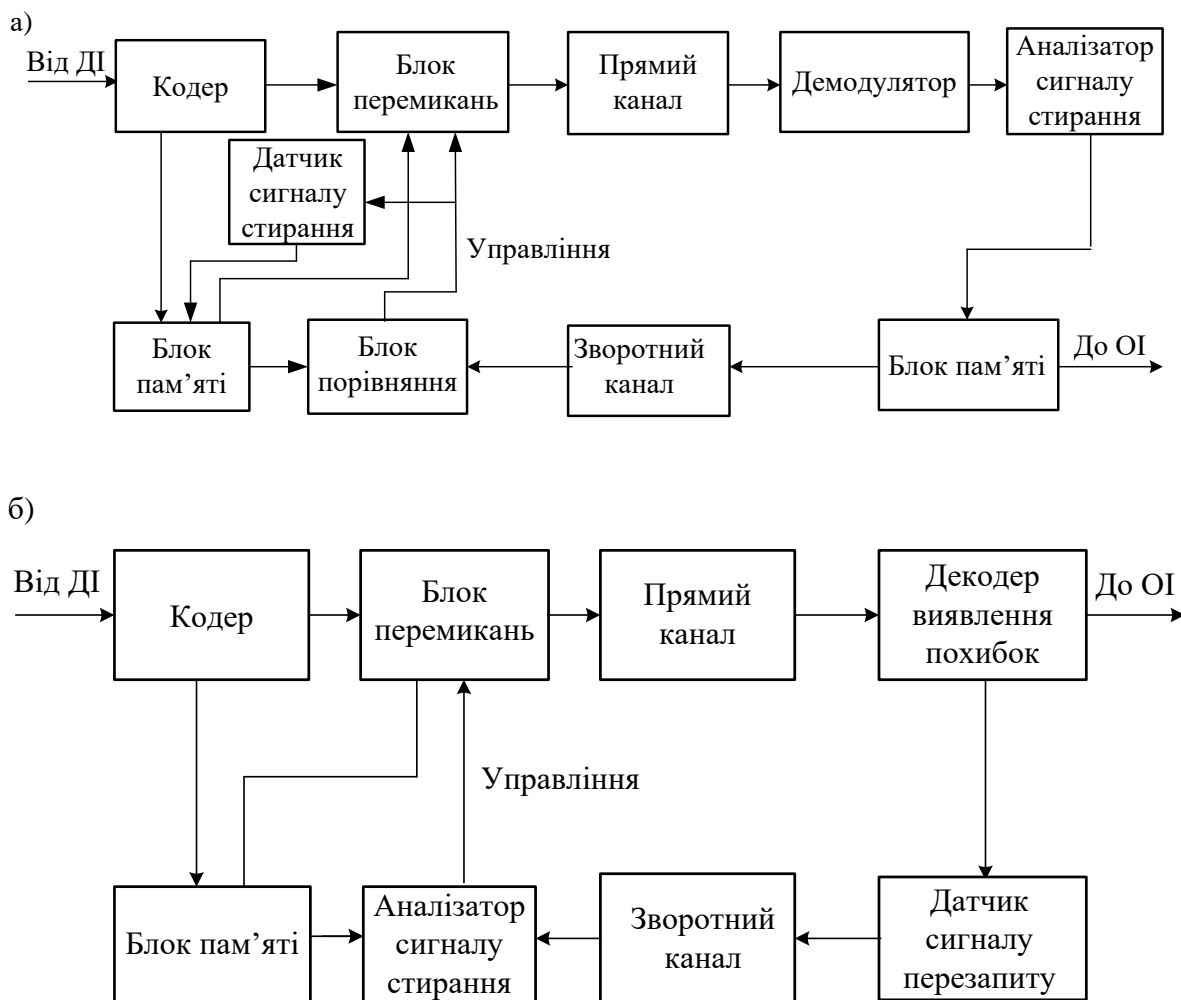


Рис. 7.12. Структурна схема систем із зворотним зв'язком:
а) інформаційна; б) вирішальна

У системі з інформаційним зворотним зв'язком на станції, що передає, повідомлення від джерела інформації (ДІ) кодується простим ненадмірним кодом, вводиться в блок пам'яті і передається по прямому каналу. На приймальній станції повідомлення записується в блок пам'яті і передається по зворотному каналу на станцію, що передає, де порівнюється з переданим повідомленням. При їх збігу передається наступне повідомлення, а прийнята інформація з блока пам'яті передається

до одержувача інформації (ОІ), інакше, передається сигнал стирання, і процес повторюється.

Найбільшого поширення на залізничному транспорті набули системи з вирішальним зворотним зв'язком, в яких повідомлення кодується надмірним кодом, що виявляє помилки, записуються в блок пам'яті і передаються в канал. Прийнята кодова комбінація декодується з виявленням помилок (рідше з виправленням однократних помилок). Якщо помилок не виявлено, то декодоване повідомлення поступає до одержувача (ОІ). При виявленні помилок комбінація стирається, і по зворотному каналу передається сигнал перезапиту, який є спеціальною кодовою комбінацією. На час її передачі припиняється передача інших повідомлень по зворотному каналу. Прийнятий сигнал викликає повторну передачу кодової комбінації, що зберігається в блоці пам'яті.

Висновки до глави 7

1. За своїм призначенням кодування поділяється на примітивне, економне і завадостійке. Примітивне кодування погоджує алфавіт джерела з алфавітом дискретного каналу, а в деяких випадках використовується для шифрування даних, які передаються, або забезпечення стійкої роботи системи синхронізації. Економне кодування дозволяє зменшити час передачі довгої послідовності повідомлень джерела по дискретному каналу. Завадостійке кодування дозволяє виявляти або виправляти помилки в прийнятих з каналу повідомленнях джерела.

2. При завадостійкому кодуванні виявлення і виправлення помилок можливе тому, що велика частина з 2^n двійкових комбінацій довжини n не використовується для передачі повідомлень джерела (заборонені комбінації). Поява забороненої комбінації на приймальному кінці однозначно свідчить про помилку в каналі.

3. Кодові (дозволені) комбінації повинні задовольняти деяку систему перевірок (яка задає код), що дозволяє відрізнити їх від заборонених комбінацій. Результатом обчислення перевірок для

прийнятої з каналу комбінації є синдром. Якщо синдром нульовий, то прийнята кодова комбінація (помилки немає). При виявленні помилок декодер відкидає ті прийняті комбінації, які мають ненульовий синдром. При виправленні помилок декодер за синдромом визначає положення помилкових символів у прийнятій комбінації та інвертує їх.

4. Найважливішими параметрами завадостійкого коду є надмірність і мінімальна відстань. Надмірність чисельно дорівнює відносній частці перевірочних символів у кодовій комбінації. При збільшенні надмірності розширюється смуга частот, яка зайнята каналним сигналом, або зменшується швидкість передачі інформації. Одночасно росте коректуюча спроможність коду (число помилок, що виявляються або виправляються). Мінімальна відстань коду d дорівнює найменшому числу помилок, яке може перетворити одну кодову комбінацію на іншу (і яке не буде виявлено кодом).

5. Для конструктивної реалізації процедури кодування можна використовувати лінійні блокові коди, зокрема циклічні коди. Для цих кодів можна оцінити величину мінімальної кодової відстані (а отже, і енергетичний виграш від застосування кодування), достатньо просто реалізувати пристрій кодування і декодування з виявленням помилок.

6. Для конструктивної реалізації процедури декодування з виправленням помилок, у тому числі і в каналах з групуванням помилок, можна використовувати підклас лінійних кодів, таких як БЧХ-коди або мажоритарні коди. Перспективним у таких каналах є використання каскадних кодів, де як зовнішні коди використовуються РС-коди.

7. При практичному використанні завадостійких кодів головним обмеженням є складність пристрою декодування, яка може бути виражена або числом логічних схем у декодері, або числом обчислювальних операцій, необхідних для декодування. Тому серед кодів, що забезпечують заданий виграш, слід вибирати ті, які допускають менш складну реалізацію, або, навпаки, при заданій складності декодування слід вибирати коди, що забезпечують найбільший виграш.

Контрольні питання та задачі

- 7.1. Які класи кодів (за призначенням) Ви знаєте?
- 7.2. Поясніть відмінність між рівномірним і нерівномірним кодуванням.
- 7.3. Поясніть відмінність між лінійним і нелінійним кодом.
- 7.4. Дайте визначення систематичного коду.
- 7.5. Що таке надмірність завадостійкого коду?
- 7.6. Що таке відносна швидкість завадостійкого коду?
- 7.7. Закодуйте кодом із загальною перевіркою на парність повідомлення дискретного джерела з об'ємом алфавіту $K=8$.
- 7.8. Закодуйте кодом Хеммінга повідомлення дискретного джерела з об'ємом алфавіту $K=16$.
- 7.9. Доведіть, що код із загальною перевіркою на парність і код Хеммінга є лійними.
- 7.10. Що таке відстань за Хеммінгом і вага кодової комбінації?
- 7.11. Чому дорівнює надмірність і відносна швидкість коду із загальною перевіркою на парність $(4, 3)$ і коду Хеммінга $(7, 4)$?
- 7.12. Що таке мінімальна відстань коду?
- 7.13. Як спрощується процедура відшукування мінімальної відстані для лінійного коду?
- 7.14. Як зв'язані мінімальна відстань коду, число виправлених та число виявлених помилок?
- 7.15. Чому дорівнюють мінімальна відстань, число виявлених та число виправлених помилок для коду Хеммінга $(7, 4)$ і коду із загальною перевіркою на парність $(4, 3)$?

ГЛАВА 8

ЗАВАДОДОСТІЙКІСТЬ ПЕРЕДАЧІ БЕЗПЕРЕРВНИХ ПОВІДОМЛЕНЬ

8.1. Критерії завадостійкості прийому безперервних повідомлень

Безперервні повідомлення того або іншого джерела (мова, відео, музика та ін.) можуть передаватися через лінію зв'язку безпосередньо або за допомогою модуляції. У першому випадку сигнал, що передається, пропорційний повідомленню, що передається, $u(t) = kb(t)$, де k - постійний множник. У другому - сигнал $u[t, b(t)]$ є деякою функцією повідомлення, що передається, $b(t)$. При ЧМ і ФМ, наприклад, ця функція нелінійна, а при АМ і балансній модуляції (БМ) - лінійна. Звідси відбуваються назви лінійні і нелінійні види модуляції.

Сигнал на вході приймача є сумарним коливанням корисного сигналу $s[t, b(t)]$ і шуму $N(t)$:

$$z(t) = s[t, b(t)] + N(t). \quad (8.1)$$

Припустимо, що шум $N(t)$ є стаціонарним гаусівським процесом з односторонньою спектральною щільністю потужності N_0 .

Задача приймача полягає в тому, щоб з вхідного коливання $z(t)$ отримати (відновити) вихідний первинний сигнал $\hat{b}(t)$, що найменше відрізняється, в розумінні деякого критерію, від переданого повідомлення $b(t)$. Відтворене з деякою похибкою повідомлення $\hat{b}(t)$ називають *оцінкою повідомлення (процесу)*.

Таким чином, задачу прийому безперервних повідомлень можна розглядати як задачу отримання оцінки процесу. В окремому випадку, коли $s(t, \lambda)$ є функцією деякого параметра λ , не залежного від часу, задача зводиться до оцінки параметра сигналу $s(t, \lambda)$.

При безпосередній передачі в неспотворюючому каналі (з коефіцієнтом передачі γ) одержуємо $s(t, b) = \gamma b(t)$, що відповідає лінійному зв'язку між сигналом і повідомленням. Тоді обчислення оцінки $\hat{b}(t)$ зводиться до лінійної фільтрації сигналу $s(t, b)$.

При передачі за допомогою модульованих сигналів $s(t, b)$ оцінка повідомлення $\hat{b}(t)$ в приймачі визначається (обчислюється) демодулятором. У багатьох випадках ця обробка зводиться до тих або інших методів фільтрації і може здійснюватися як до детектора, так і після нього. У звичайному приймачі (наприклад, радіомовному) додетекторна обробка проводиться за допомогою смугових підсилювачів, що забезпечують необхідну частотну вибірковість сигналів на вході детектора. Функцію ж післядетекторної обробки виконують підсилювачі низької частоти, які підсилюють і фільтрують низькочастотний первинний сигнал $\hat{b}(t)$, що поступає на вхід кінцевих пристроїв (гучномовця, телефону, магнітофона і т. д.).

Оптимальний демодулятор у загальному випадку є нелінійним пристроєм (нелінійним фільтром), що забезпечує найкраще (за заданим критерієм) виділення повідомлення $b(t)$ з прийнятого сигналу $Z(t)$. Теорія оптимального прийому безперервних повідомлень вперше була розроблена В.О. Котельниковим [29].

Мірою завадостійкості при передачі безперервних повідомлень може бути ступінь «відхилення» отриманої оцінки $\hat{b}(t)$ від переданого повідомлення $b(t)$. Зазвичай застосовується середньоквадратичне відхилення або середній квадрат помилки (середня потужність різниці):

$$\overline{E^2(t)} = \overline{|\hat{B}(t) - B(t)|^2}, \quad (8.2)$$

де усереднювання береться за всіма можливими реалізаціями $b(t)$ і $\hat{b}(t)$ з урахуванням щільності імовірності сумісного розподілення $w(b, \hat{b})$.

Різницю $\varepsilon(t) = \hat{b}(t) - b(t)$ можна розглядати як реалізацію завади $E(t)$ на виході приймача. Величина $\overline{E^2(t)} = P_\varepsilon$ - потужність завади на виході приймача. Потужність повідомлення, що передається, $B(t)$ вважається заданою $P_b = \overline{B^2(t)}$. Тоді можна визначити відношення сигнал-завада (ВСЗ) на виході приймача $\rho_{вих} = P_b / P_\varepsilon$. На вході приймача це відношення дорівнює $\rho_{вх} = P_c / P_{ш}$. У багатьох випадках як критерій завадостійкості приймають не середній квадрат помилки $\overline{E^2(t)}$, а ВСЗ, тобто $\rho_{вих}$. Відношення сигнал-завада в каналі, як вже зазначалося, може бути покращений приймачем. Це поліпшення залежить не тільки від способу прийому, але і від виду модуляції. Тому завадостійкість систем передачі безперервних повідомлень зручно оцінювати вирашем

$$g = \frac{\rho_{вих}}{\rho_{вх}} = \frac{P_b}{P_\varepsilon} \bigg/ \frac{P_c}{P_{ш}}. \quad (8.3)$$

При $g > 1$ відношення сигнал-завада при демодуляції поліпшується. У деяких випадках $g < 1$, це означає, що система модуляції дає не виграв, а програш. У різних системах сигнали можуть мати різну ширину спектра, і тому при порівняльній оцінці систем зв'язку слід визначати реальний або «узагальнений виграв системи» через відношення потужностей сигналу не до потужностей завади, а до їх середньої спектральної щільності:

$$g' = \frac{\rho'_{вих}}{\rho'_{вх}} = \frac{\rho_{вих}}{\alpha \rho_{вх}} = \frac{g}{\alpha}, \quad (8.4)$$

$$\text{де } \rho'_{вих} = \frac{P_b}{P_\varepsilon / F_c} = \rho_{вих} F_c;$$

$$\rho'_{вх} = \frac{P_c}{P_{ш} / F} = \frac{P_c}{N_0} = \rho_{вх} F;$$

$$\alpha = \Delta F / F_c;$$

ΔF - ширина спектра сигналу;

F_c - ширина спектра повідомлення.

Зазначимо, що в системі безпосередньої передачі, в якій сигнал пропорційний повідомленню, що передається, $s(t,b) = \gamma b(t)$, $F = F_c$ і $\rho_{вих} = \rho_{вх}$, виграш і узагальнений виграш однакові, $g = g' = 1$.

Це означає, що введений критерій виграшу зводиться до порівняння систем передачі модульованих сигналів з системою безпосередньої передачі. На практиці використовують і інші критерії правильності передачі безперервних повідомлень, наприклад, у телеметрії критерій розбірливості при передачі мовних повідомлень, критерій максимальної помилки $\delta = \left| \hat{b}(t) - b(t) \right|_{\max}$.

8.2. Оптимальна оцінка окремих параметрів сигналу

Задача оцінки параметрів сигналу виникає в багатьох практичних випадках. Так, часто потрібно визначити (виміряти) частоту або фазу сигналу, оцінити амплітуду сигналу при його передачі по каналу і т. д. У телеметрії і телеуправлінні необхідно оцінювати велике число параметрів, що характеризують стан об'єкта. Таким об'єктом може бути технологічний процес, нафтопровід, літальний апарат і т. п. У цьому випадку повідомленнями, що передаються, є контрольовані параметри - випадкові величини, не залежні від часу. Модуляція при цьому зводиться до встановлення деякого параметра λ сигналу $s(t)$ відповідно до переданого повідомлення, а демодуляція - до виявлення (оцінки) цього параметра з можливо більшою точністю.

У найпростішому випадку, коли оцінюють один параметр сигналу заданої форми, задачу ставлять таким чином. Хай коливання $Z(t)$, що приймається на інтервалі $(0, T)$, є адитивною сумішшю сигналу $s(t, \lambda)$, залежного від одного невідомого параметра λ , з шумом $N(t)$:

$$z(t) = s(t, \lambda) + N(t), \quad 0 \leq t \leq T. \quad (8.5)$$

Вважаємо, що параметр λ має постійне значення на інтервалі спостереження $(0, T)$ і відома апіорна щільність імовірності цього параметра $w(\lambda)$. Потрібно визначити оператора системи, що гарантує отримання якнайкращої оцінки параметра і розрахувати точність цієї оцінки.

Через шум в каналі і випадковий характер параметра λ точне вимірювання його неможливе. Можна лише вказати його наближену оцінку.

Якість оцінки параметра, названої точковою, зазвичай перевіряється на виконання трьох умов:

- *спроможності*, що полягає в тому, що оцінка $\hat{\lambda}$ зводиться за імовірністю до оцінюваного параметра λ при необмеженому збільшенні часу аналізу (або об'єму вибірки), тобто при $\zeta > 0$
 $\lim_{T \rightarrow \infty} P\{|\hat{\lambda} - \lambda| \geq \zeta\} = 0$. При цьому дисперсія помилки прагне до нуля:

$$\lim_{T \rightarrow \infty} D(E) = \lim_{T \rightarrow \infty} D(\lambda - \hat{\lambda}) = 0;$$

- *незмішуваності* оцінки, яка полягає в тому, що умовне МО помилки при всіх λ повинно дорівнювати нулю: $\bar{E} = \lambda - \hat{\lambda} = 0$ або $\bar{\hat{\lambda}} = \lambda$;

- *ефективності* оцінки, що полягає в тому, що дисперсія помилки $D(E)$ повинна при заданому часі аналізу або об'ємі вибірки бути мінімальною в класі всіх можливих оцінок:

$$D(E) = \overline{(\lambda - \hat{\lambda})^2} = \min .$$

Уся інформація про переданий параметр (повідомлення) λ після прийому коливання (вираз (8.5)) міститься в апостеріорному розподілі $w(\lambda|z)$, який згідно з формулою Байєса [47]

$$w(\lambda|z) = w(\lambda)w(z|\lambda)w(z). \quad (8.6)$$

На підставі аналізу апостеріорного розподілу (8.6) ухвалюється рішення про оцінку параметра - λ , що передається. При великих відносинах сигнал-шум апостеріорна щільність

імовірності має найбільший максимум в області дійсного значення параметра λ . Ця обставина вказує на те, що оцінкою доцільно взяти те значення $\hat{\lambda}$, яке обертає в максимум функцію $w(\lambda|z)$.

У багатьох практичних випадках апіорна щільність імовірності $w(\lambda)$ виявляється невідомою і її вважають рівномірною: $w(\lambda) = \text{const}$ на всьому можливому інтервалі L . При цьому координата максимуму апостеріорної імовірності співпадатиме з відповідною координатою умовного розподілу, який визначає функцію правдоподібності. У цьому випадку правило максимуму апостеріорної щільності імовірності переходить у правило максимальної правдоподібності. Тут оцінка параметра $\hat{\lambda}$ визначається з умови

$$-\frac{dw(z|\lambda)}{d\lambda} = 0. \quad (8.7)$$

Оцінку параметра, що отримується за цим критерієм, називають максимально правдоподібною. Рівняння правдоподібності (8.7) можна записати в еквівалентному вигляді

$$\frac{d[\ln w(z|\lambda)]}{d\lambda} = 0, \quad (8.8)$$

оскільки $\ln x$ - монотонна функція свого аргументу і, отже, коріння рівнянь (8.7) і (8.8) співпадають. Оцінка визначається тим коренем рівняння (8.8), який відповідає максимуму функції правдоподібності. Іншим вельми поширеним критерієм оцінки параметрів сигналу є оцінка за мінімумом середньоквадратичної помилки. При цьому критерій мінімізується за $\hat{\lambda}$:

$$\overline{E^2(\hat{\lambda})} = \int_{\Lambda} (\lambda - \hat{\lambda})^2 w(\lambda|z) d\lambda. \quad (8.9)$$

Оптимальна оцінка $\hat{\lambda}$ знаходиться з умови $\overline{\frac{dE^2(\hat{\lambda})}{d\hat{\lambda}}} = 0$.

Після диференціювання виразу (8.9) за $\hat{\lambda}$ з урахуванням того, що $\int_{\Lambda} w(\lambda|z) d\lambda = 1$, одержуємо $2\hat{\lambda} - 2 \int_{\Lambda} \lambda w(\lambda|z) d\lambda = 0$, звідки

$$\hat{\lambda} = \int_{\Lambda} \lambda w(\lambda|z) d\lambda, \quad (8.10)$$

тобто оптимальною оцінкою параметра є в даному випадку математичне очікування апостеріорного розподілу.

Критерій середньоквадратичної помилки є окремим випадком більш загального критерію, коли мінімізується математичне очікування деякої функції втрат $L(\lambda - \hat{\lambda})$, тобто

$$\overline{L(\lambda - \hat{\lambda})} = \int_{\Lambda} L(\lambda - \hat{\lambda}) w(\lambda|z) d\lambda. \quad (8.11)$$

Оцінку, що мінімізує цю величину, називають *байєсівською оцінкою*, а критерій, заснований на мінімізації (8.11), як і в дискретному випадку, - *критерієм середнього ризику*.

При $L(\lambda - \hat{\lambda}) = (\lambda - \hat{\lambda})^2$ критерій мінімуму середнього ризику (вираз (8.11)) переходить у критерій середньоквадратичної помилки виразу(8.9). В цьому випадку байєсівська оцінка визначається виразом (8.10).

Якщо $w(z|\lambda)$ симетрична відносно $\hat{\lambda}_{opt}$, що має місце при великому відношенні сигнал-шум, то критерій максимуму апостеріорної імовірності і максимуму функції правдоподібності співпадає з критерієм мінімуму середньоквадратичної помилки.

Якщо значення параметра λ постійне на інтервалі спостереження і прийняте колювання є адитивною сумішшю (вираз (8.5)) корисного сигналу $s(t, \lambda)$ і БГШ $N(t)$ зі спектральною щільністю $N_0/2$, то вектор z , яке визначає колювання, що приймається у функціональному просторі, є

випадковим гаусівським вектором, середнє значення якого $s(t, \lambda)$, а дисперсія співпадає з дисперсією шуму. Із роботи [20] випливає умова

$$w(z|\lambda) = A^* \exp\left(-\frac{1}{N_0} \int_0^T [z(t) - s(t, \lambda)]^2 dt\right), \quad (8.12)$$

де A^* - постійний коефіцієнт.

Тоді згідно з виразами (8.6) і (8.12) маємо

$$w(z|\lambda) = kw(\lambda)w(z|\lambda) = kw(\lambda) \exp\left(-\frac{1}{N_0} \int_0^T [z(t) - s(t, \lambda)]^2 dt\right), \quad (8.13)$$

де k - деяка постійна, яка може бути обчислена з умови нормування $\int_A w(\lambda|z) d\lambda = 1$.

Спростимо цей вираз, перетворивши показник експоненти,

$$w(z|\lambda) = kw(\lambda) \exp\left(-\frac{1}{N_0} \int_0^T z^2(t) dt\right) \exp\left(-\frac{1}{N_0} \int_0^T s^2(t, \lambda) dt\right) \times \\ \times \exp\left(\frac{2}{N_0} \int_0^T z(t)s(t, \lambda) dt\right).$$

У правій частині цієї рівності перший експоненціальний множник не залежить від λ , і його можна включити в постійну. Другий множник дорівнює $\exp(-E/N_0)$, де E - енергія сигналу. У тих випадках, коли параметр λ неенергетичний, тобто енергія сигналу не залежить від λ , цей член також можна включити в постійну k . За цієї умови останній вираз можна записати так:

$$w(z|\lambda) = kw(\lambda) \exp(q(\lambda)), \quad (8.14)$$

де

$$q(\lambda) = \frac{2}{N_0} \int_0^T z(t)s(t, \lambda) dt. \quad (8.15)$$

Неважко переконатися, що для сигналу $s(t)$, залежного від декількох параметрів, функція правдоподібності $w(z|\lambda)$ і апостеріорна щільність імовірності $w(z|\lambda)$ визначатимуться аналогічними виразами, в яких $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$.

Звідси випливає, що при відомій апріорній щільності імовірності визначення апостеріорної щільності імовірності зводиться до обчислення функції $q(\lambda)$. Ця функція з точністю до коефіцієнта дорівнює скалярному добутку сигналу, що прийшов, з очікуваним варіантом сигналу $s(t, \lambda)$. Її часто називають *кореляційним інтегралом*. Вона визначає ті істотні операції, які потрібно виконати над $z(t)$, щоб витягнути всю доступну інформацію про передане повідомлення λ . Оптимальний приймач максимальної правдоподібності відтворює те повідомлення λ , для якого функція $q(\lambda)$ максимальна.

Приклад. Знайти оптимальну оцінку коефіцієнта передачі каналу γ (амплітуду сигналу, що пройшов через канал), вважаючи, що сигнал, який приймається, $z(t) = \gamma s(t) + N(t)$, де $s(t)$ - точно відомо на прийомі, а $N(t)$ - БГШ із спектральною щільністю N_0 . Функція правдоподібності для цього випадку згідно з виразом (8.12)

$$w(z|\lambda) = A \exp \left(-\frac{1}{N_0} \int_0^T [z(t) - \gamma s(t)]^2 dt \right).$$

Рівняння (8.8) максимальної правдоподібності набуває вигляду

$$\frac{\partial [\ln w(z|\lambda)]}{\partial \gamma} = -\frac{2}{N_0} \int_0^T [z(t) - \hat{\gamma} s(t)] s(t) dt = 0$$

або після перетворень

$$-\frac{2}{N_0} \int_0^T z(t)s(t)dt + \frac{2\hat{\gamma}}{N_0} \int_0^T s^2(t)dt = 0.$$

Звідси оцінка дорівнює

$$\hat{\gamma} = \frac{1}{E} \int_0^T z(t)s(t)dt = \gamma + \frac{1}{E} \int_0^T N(t)s(t)dt. \quad (8.16)$$

Вимірник для отримання такої оцінки може бути реалізований фільтром, узгодженим з сигналом $s^2(t)/E$ або з еквівалентною кореляційною схемою. Визначимо якість знайденої оцінки. З виразу (8.16) видно, що при $T \rightarrow \infty$, тобто

$E \rightarrow \infty$, похибка оцінки $\gamma - \hat{\gamma} = \frac{1}{E} \int_0^T N(t)s(t)dt$ прагне до нуля,

тобто дана оцінка є спроможною, незміщеною (оскільки $\overline{N(t)} = 0$). Вона є також ефективною при гаусівському шумі $N(t)$. Дисперсія помилки (що має гаусівське розподілення) з урахуванням виразу (2.85) дорівнюватиме $D(E) = \frac{N_0}{2E}$.

8.3. Оптимальна демодуляція безперервних сигналів

Визначимо умови оптимального прийому безперервних повідомлень. Припустимо, що повідомлення є деяким стаціонарним процесом (первинним сигналом) з реалізацією $b(t)$. Він може безперервно змінюватися в часі і приймати будь-яку форму.

Для простоти аналізу вважатимемо, що функція $b(t)$ приймає значення від -1 до +1, реалізації повідомлення мають кінцеву тривалість T і їх спектр практично обмежений частотами від 0 до F_c . За цих умов функція $b(t)$ може бути розкладена за ортонормованим базисом $\{\psi_k(t)\}$ і подана як усічений ряд

$$b(t) = \sum_{k=1}^B \lambda_k \psi_k(t), \quad (8.17)$$

де λ_k - випадкові коефіцієнти, що визначають повідомлення, яке передається.

При розкладанні в тригонометричний ряд Фур'є λ_k пропорційні складовим спектра, а при розкладанні в ряд Котельникова - відліковим значенням $b(k\Delta)$ функції $b(t)$. Тут $\Delta = 1/(2F_c)$, $B = T/\Delta = 2TF_c$. Таким чином, при відомій системі базисних функцій $\{\psi_k(t)\}$ передача безперервних повідомлень $b(t)$ еквівалентна передачі B значень коефіцієнтів (параметрів) λ_k . Для передачі по каналу коливання $b(t)$ перетвориться на сигнал $s(t, b)$. Оскільки коливання виразу (8.17) визначається параметрами λ_k ($k=1, 2, \dots, B$), то і сигнал залежить від цих параметрів. Прийняте коливання з урахуванням накладення завади

$$z(t) = s(t, b) + n(t) = s(t, \lambda) + n(t). \quad (8.18)$$

Вплив завад призводить до того, що кожний параметр λ_k буде прийнятий з деякою похибкою $\Delta\lambda_k$. У результаті оцінка повідомлення дорівнює

$$\hat{b}(t) = \sum_{k=1}^B (\lambda_k + \Delta\lambda_k) \psi_k(t) = b(t) + \varepsilon(t), \quad (8.19)$$

де $\varepsilon(t)$ - похибка відтворення повідомлення (шум на виході приймача), $\varepsilon(t) = \Delta b = \sum_{k=1}^B \Delta\lambda_k \psi_k(t)$.

Таким чином, задача оптимального прийому безперервного повідомлення $b(t)$ зводиться до задачі сумісного оптимального прийому сукупності багатьох параметрів $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$. Ця задача є узагальненням розглянутої в п. 8.2 задачі оптимальної оцінки одного параметра.

Отже, за реалізацією $z(t)$ необхідно відновити передане повідомлення $b(t)$ з можливо більшою точністю, хоч би при слабких завадах. Для цього необхідно на основі аналізу прийнятого колювання $z(t)$ знайти максимум апостеріорного розподілу $w(b|z)$, який на основі формули Байєса може бути поданий як

$$w(b|z) = kw(b)w(z|b), \quad (8.20)$$

де k - постійний коефіцієнт.

Функція правдоподібності $w(z|b)$, що входить у вираз (8.20), відома (для даного гаусівського каналу - це гаусівське розподілення). Апріорне розподілення $w(b)$ залежить від вигляду і характеристик повідомлень, що передаються $b(t)$.

Вибір конкретної моделі апріорного розподілення $w(b)$ є не таким істотним, оскільки роль початкових, апріорних відомостей зменшується зі збільшенням об'єму спостережень. При великому об'ємі спостережень алгоритми обробки сигналів виходять асимптотично однаковими, тобто *малочутливими* до апріорного розподілення. Тому обмежимося розглядом моделі рівномірного розподілення $w(b) = \text{const}$. У цьому випадку розв'язання задачі спрощується, оскільки згідно з виразом (8.20) апостеріорне розподілення $w(b|z)$ повністю визначатиметься функцією правдоподібності $w(z|b)$, яка для гаусівського каналу визначається виразом, аналогічним до виразу (8.12):

$$w(z|b) = c \exp \left(- \frac{1}{N_0} \int_0^T [z(t) - s(t, b(t))]^2 dt \right).$$

Згідно з цим виразом максимуму функції правдоподібності $w(z|b)$, а отже, і функції $w(b|z)$ відповідає мінімум за $b(t)$ інтеграла

$$\min_{b(t)} \int_0^T [z(t) - s(t, b(t))]^2 dt. \quad (8.21)$$

Отже, оптимальний приймач повинен відтворювати повідомлення $\hat{b}(t)$, що відповідає, як і при передачі дискретних повідомлень, тому з можливих сигналів $s(t, \hat{b})$, який менше за інших відрізняється в середньоквадратичному розумінні від реалізації сигналу $z(t)$ на вході приймача.

За відсутності завад такий приймач відтворює повідомлення без спотворень (без помилок): $z(t) = s(t, b(t))$, $\hat{b}(t) = b(t)$ і $\overline{E^2(t)} = 0$, а за наявності завад помилка мінімальна.

Запишемо вираз (8.20) в іншому вигляді, подібному до виразу (8.15):

$$w(b|z) = kw(b) \exp[q(b)], \quad (8.22)$$

де

$$q(b) = \frac{2}{N_0} \int_0^T z(t) s(t, b(t)) dt. \quad (8.23)$$

Звідси випливає, що при відомій апіорній імовірності визначення апостеріорної імовірності зводиться до обчислення функції $q(b)$, тобто скалярного добутку прийнятого коливання $z(t)$ на передані (очікувані) сигнали $s(t, b(t))$.

У багатьох випадках для наближеного знаходження $q(b)$ доцільне застосування простих стежачих пристроїв. Розглянемо принципову можливість побудови таких пристроїв. Докладне обґрунтування на основі теорії нелінійної фільтрації наведено в п. 8.8.

При передачі безперервних повідомлень сигнал $s(t, b(t))$ не є повністю відомим. Проте зазвичай є деяка апіорна інформація про цей сигнал. Відомі, наприклад, несуча частота, вид модуляції, ширина спектра сигналу і т. п. Частина інформації можна отримати в результаті спостереження за прийнятою реалізацією сигналу $z(t)$ за попередній проміжок часу. У результаті є

можливість визначити оцінку сигналу $s(t, \hat{b}(t))$ і обчислити функцію $q(\hat{b})$ для цієї оцінки:

$$q(\hat{b}) = \frac{2}{N_0} \int_0^T z(t) s(t, \hat{b}(t)) dt. \quad (8.24)$$

Функцію $q(\hat{b})$ можна знайти за допомогою фільтра зі змінними параметрами (рис. 8.1) або схеми стежачого корелятора (СК) (рис. 8.2). Кожна з цих схем має основний інформаційний канал, на виході якого отримується оціночне значення $\hat{b}(t)$ повідомлення, що передається, і канал зворотного зв'язку, за допомогою якого в схемі на рис. 8.2 формується опорний сигнал $s(t, \hat{b}(t))$, а в схемі на рис. 8.1 за допомогою управляючого елемента (УЕ) проводяться зміни параметрів стежачого фільтра СФ так, щоб він був узгоджений з очікуваним сигналом $s(t, \hat{b}(t))$, що безперервно змінювався. У схемі на рис. 8.2 за допомогою УЕ змінюється модульований параметр несучого коливання, що формується генератором (G).

При частотній модуляції, наприклад, цим параметром буде частота, при часовій імпульсній модуляції - зсув імпульсів у часі і т. п. Фільтр нижніх частот ФНЧ в цій схемі виконує роль інтегратора на інтервалі спостереження T , який пов'язаний з максимальною частотою F_c в спектрі повідомлення, що передається співвідношенням $T = 1/(2F_c)$.

Розглянуті схеми є квазіоптимальними, оскільки отримувана оцінка $s(t, \hat{b}(t))$ не є якнайбільшою можливою. При різних видах модуляції принцип стежачого прийому залишається одним і тим самим. Вид модуляції визначає параметр, за яким повинне здійснюватися стеження. Інакше кажучи, оптимальний приймач повинен з найменшою помилкою стежити за передаваним випадковим коливанням $b(t)$. Схеми стежачого прийому дозволяють практично реалізувати завадостійкість, близьку до потенційної. При лінійній модуляції, коли $s(t, b(t)) = f(t)b(t)$, де $f(t)$ - відома функція (несуче коливання), оптимальний демодулятор можна реалізувати розімкненою схемою з синхронним детектором (рис. 8.3).

Перейдемо до визначення завадостійкості систем зв'язку при оптимальному прийомі. Зазначимо, що цю потенційну завадостійкість можна обчислити, не уточнюючи структури оптимального демодулятора. Для цього достатньо знати, що він видає рішення $\hat{b}(t)$, відповідне мінімуму виразу (8.21).

Перш ніж приступити до виведення формул, що визначають потенційну завадостійкість, нагадаємо основні принципи класифікації видів модуляції при передачі безперервних повідомлень. У загальному випадку модуляція полягає в тому, що безліч повідомлень (первинних сигналів) $B(t)$ перетворюється (відображається) на безліч вторинних сигналів $S(t) = S[t, B(t)]$. Цим записом підкреслюється, що значення сигналу S в деякий момент t визначається в загальному випадку всією поведінкою повідомлення $B(t)$ на всій осі часу.

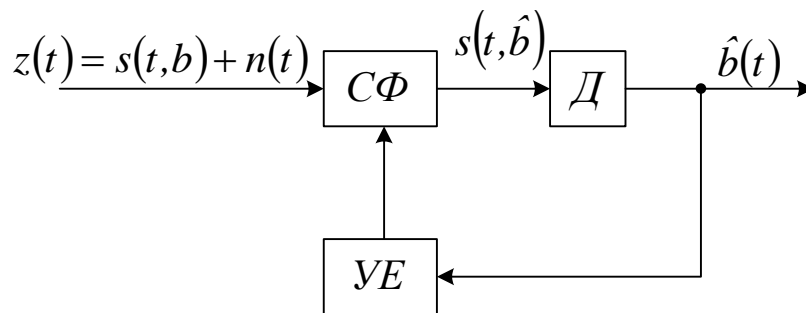


Рис. 8.1. Структурна схема демодулятора зі стежачим фільтром

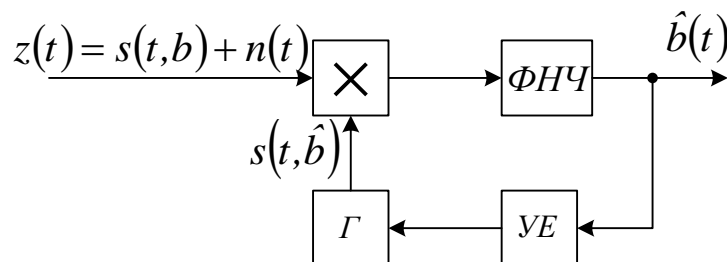


Рис. 8.2. Структурна схема стежачого кореляційного демодулятора

В окремому випадку, якщо сигнал $S(t)$ у будь-який момент t залежить не від усього ходу сигналу $B(t)$, а тільки від його

значення у момент t , то система модуляції називається *прямою*. У цьому випадку повідомлення $\hat{b}(t)$ входить безпосередньо у вираз сигналу $s(t)$. До прямих належить переважна більшість вживаних методів модуляції, наприклад АМ, БАМ і ФМ. Решта систем модуляції, в яких $S(t)$ залежить від загальної поведінки сигналу $B(t)$, називаються *непрямими*. Серед них особливий інтерес являють *інтегральні* системи, в яких $B(t)$ входить у вираз $S(t)$ під інтегралом.

Система модуляції називається *лінійною*, якщо $S(t)$ можна отримати із $B(t)$ за допомогою лінійних операцій. Лінійні системи можуть бути прямими (наприклад амплітудна АМ) і непрямими (наприклад односмугова ОМ).

Геометрично модуляцію можна розглядати як відображення простору B повідомлень у просторі сигналів S , а демодуляцію - як зворотне відображення. При демодуляції завада $n(t)$ на вході приймача відображається як похибка оцінки повідомлення (шум відтворення або шум на виході приймача) $\varepsilon(t)$.

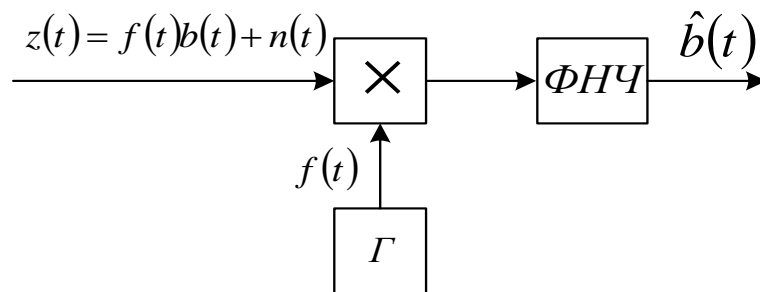


Рис. 8.3. Структурна схема демодулятора з синхронним детектором

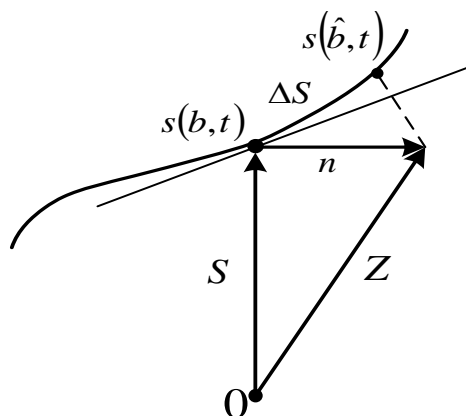


Рис. 8.4. Геометричне подання сигналу і шуму

Розглянемо прийом безперервного повідомлення на фоні БГШ зі спектральною щільністю N_0 . При досить слабкому шумі $n(t)$ похибка $\varepsilon(t) = \hat{b}(t) - b(t)$ (шум на виході приймача) є також гаусівським процесом зі спектральною щільністю $G_\varepsilon(f)$, яку і визначатимемо. Для цього зручно скористатися геометричним поданням. У просторі сигналів кожної реалізації сигналу $s(t, b(t))$ при різних $b(t)$ відповідає точка. Якщо $s(b)$ залежить безперервно від b (що має місце у всіх аналогових системах зв'язку), то всі ці точки утворюють деяку криву (рис. 8.4).

Прийнятий сигнал $z(t)$ є також точкою в просторі сигналів, що як правило, не лежить на кривій $s(b)$. Максимально правдоподібна оцінка $\hat{b}(t)$ відповідає тому сигналу $s(t, \hat{b}(t))$, який зображується на сигнальній кривій точкою, найближчою до точки z .

Позначимо $\Delta s = s(t, \hat{b}(t)) - s(t, b(t))$, де $b(t)$ - дійсно передане повідомлення. При малій заваді і, отже, малому відхиленні Δs відрізок між $s(b)$ і $s(\hat{b})$ можна апроксимувати прямою лінією, яка є дотичною до лінії сигналу в точці $s(b)$. Тоді Δs є проекцією вектора n на цю пряму. У цьому випадку справедливий вираз

$$\Delta s(t) \approx \Delta b \frac{\partial s}{\partial b} = n_1(t). \quad (8.25)$$

Тут $n_1(t)$ - складова (координата) шумового вектора в просторі сигналів, що являє собою низькочастотний гаусівський ергодичний процес з нульовим МО і зі спектральною щільністю N_0 в смузі частот від 0 до F_c . $\Delta b = \hat{b}(t) - b(t)$. Тоді з урахуванням роботи [19] в одиничній смузі частот маємо

$$\overbrace{n_1^2(t)} = N_0 = \overbrace{(\Delta b)^2} \left(\frac{\partial s(t)}{\partial b(t)} \right). \quad (8.26)$$

Оскільки процес $\Delta b(t)$ (компонента шуму на виході приймача) змінюється значно повільніше за процес $\frac{\partial s(t)}{\partial b(t)}$, то

$$\overline{\Delta b^2(t) \left(\frac{\partial s(t)}{\partial b(t)} \right)^2} = \overline{\Delta b^2(t)} \overline{\left(\frac{\partial s(t)}{\partial b(t)} \right)^2} = N_0. \quad (8.27)$$

З (8.27) випливає, що одностороння спектральна щільність потужності (СЩП) шуму на виході приймача дорівнює

$$G_{0,\varepsilon}(f) = N_0 / \overline{\left(\frac{\partial s}{\partial b} \right)^2}. \quad (8.28)$$

При лінійних системах модуляції $\overline{\left(\frac{\partial s}{\partial b} \right)^2}$ не залежить від частоти. Таким чином, при лінійних системах модуляції шум на виході приймача "квазібілий", тобто має рівномірний спектр у смузі частот F_c .

У разі інтегральних систем повідомлення $B(t)$ входить у вирази сигналу під знаком інтеграла:

$$S(B, t) = S \left[\int_0^t B(\tau) d\tau, t \right] = S(\psi(t), t), \text{ де } \psi(t) = \int_0^t B(\tau) d\tau.$$

Оскільки

$$\hat{B}(t) = B(t) + E(t),$$

то

$$\hat{\psi}(t) = \int_0^t B(t) dt + \int_0^t E(t) dt = \psi(t) + \xi(t).$$

Отже, СЩП шуму на виході приймача для інтегральних систем можна визначити як СЩП похідної $\xi'(t)$. На підставі відомої теореми про спектр похідної $G_{0\varepsilon}(f) = (2\pi f)^2 G_{0\xi}(f)$, де $G_{0\xi}(f)$ визначається за формулою (8.28), якщо в останню замість $\frac{\partial s}{\partial b}$ підставити $\frac{\partial s}{\partial \psi}$.

Таким чином, для інтегральних систем СЩП шуму на виході приймача дорівнює

$$G_{0\varepsilon}(f) = \frac{(2\pi f)^2 N_0}{(\partial s / \partial \psi)^2}, \quad (8.29)$$

тобто СЩП завади на виході приймача в інтегральних системах пропорційна квадрату частоти.

Всі ці результати справедливі для лінійної модуляції або при довільній модуляції для слабких завад, коли можна вважати $\Delta s \approx \Delta b \frac{\partial s}{\partial b}$. Вони характеризують так звані нормальні помилки.

Очевидно, потужність шуму на виході приймача в смузі частот від нуля до F_c буде $P_\varepsilon = \int_0^{F_c} G_{0\varepsilon}(f) df$.

З іншого боку, потужність P_b повідомлення на виході приймача, що дорівнює $\overline{B^2(t)}$, можна виразити через пік-фактор повідомлення $\Pi = \frac{|B(t)|_{max}}{\sqrt{\overline{B^2(t)}}}$. Вважаючи, що повідомлення нормоване і $|B(t)|_{max} = 1$, одержуємо

$$\overline{B^2(t)} = P_b = 1/\Pi^2. \quad (8.30)$$

Тоді відношення потужностей сигналу і шуму на виході приймача визначається з виразу

$$\rho_{вих} = \frac{P_b}{P_\varepsilon} = \frac{1}{\Pi^2 \int_0^{F_c} G_\varepsilon(f) df}, \quad (8.31)$$

а вираз для виграшу g і узагальненого виграшу g' відповідно до визначень (8.3) і (8.4) можна записати так:

$$g = \frac{\rho_{вих}}{\rho_{ex}} = \frac{N_0 F}{\Pi^2 \int_0^{F_c} G_{0,\varepsilon}(f) df}; \quad g' = \frac{g}{\alpha} = \frac{N_0 F_c}{\Pi^2 P_c \int_0^{F_c} G_{0,\varepsilon}(f) df}. \quad (8.32)$$

Для гармонійного сигналу $\Pi = \sqrt{2}$, а для телефонного повідомлення $\Pi \approx 3$.

Використовуючи теорему Шеннона [50], можна знайти максимальні можливі значення виграшу й узагальненого виграшу при заданих параметрах системи зв'язку. Розглянемо це питання для найбільш простого випадку, коли безперервне повідомлення є гаусівським процесом з рівномірним спектром у смузі частот F_c ("квазібілий" шум), а в каналі існує адитивна завада у вигляді "квазібілого" шуму в смузі F з односторонньою спектральною щільністю N_0 .

Згідно з теоремою Шеннона передача повідомлення із заданим значенням $P_b/P_\varepsilon = \rho_0$ можлива у разі, коли $H'_\varepsilon(B) < C'$. Тут $H'_\varepsilon(B)$ - епсилон-продуктивність джерела, яка в даному випадку дорівнює $F_c \log \rho_0$, а пропускна спроможність гаусівського каналу C' згідно [44] дорівнює

$$C' = F \log \left(1 + \frac{P_c}{N_0 F} \right) = F \log(1 + \rho_{ex}),$$

де F - смуга пропускання каналу і в загальному випадку $F \neq F_c$.

У гіпотетичній *ідеальній системі зв'язку*, в якій повністю використовується пропускна спроможність каналу і $\rho_{вих} = \rho_0$,

$$F_c \log \rho_{вих} = F \log(1 + \rho_{вх}). \quad (8.33)$$

У реальних системах зв'язку зазвичай вдається лише частково використовувати пропускну спроможність каналу. Назвемо *ефективністю* η системи зв'язку відношення епсилон-продуктивності джерела до пропускну спроможності каналу, при якій забезпечується задана правильність, тобто $\rho_{вих} = \rho_0$. Для такої реальної системи замість виразу (8.33) маємо

$$F_c \log \rho_{вих} = \eta F \log(1 + \rho_{вх}). \quad (8.34)$$

З виразів (8.33) і (8.34) видно, що при $\eta F > F_c$ можна забезпечити високу правильність (велике значення $\rho_{вих}$) при відносно малих $\rho_{вх}$, тобто отримати великий виграш g .

Таким чином, виграш досягається в результаті обміну ширини спектра на динамічний діапазон, про що йшлося в гл. 1. Великий виграш можна отримати тільки при великому відношенні $\alpha = \frac{F}{F_c}$. Зазначимо, що великий виграш може мати місце і при малій ефективності η , і навпаки. Отже, при оцінці різних систем зв'язку необхідно враховувати принаймні два показники: ефективність і завадостійкість. Сукупність цих двох показників складає достатньо повну характеристику системи.

Якнайкращою вважається система, яка забезпечує найбільшу завадостійкість при заданій ефективності або найбільшу ефективність при заданій завадостійкості.

Для ідеальної системи $\eta=1$ і з виразу (8.33) витікає

$$\rho_{вих} = (1 + \rho_{вх})^\alpha. \quad (8.35)$$

Звідси при $\rho_{вх} \gg 1$ одержуємо

$$g \approx \rho_{вх}^{\alpha-1}; \quad g' \approx \frac{1}{\alpha} \rho_{вх}^{\alpha-1}. \quad (8.36)$$

Таким чином, в ідеальній системі виграш g зростає зі збільшенням α за експоненціальним законом. Ніяка реальна система не може забезпечити при заданому α більш вищу завадостійкість, ніж ідеальна.

8.4. Завадостійкість систем передачі безперервних повідомлень при слабких завадах

Для передачі безперервних повідомлень по каналу зв'язку застосовують різні види модуляції. При синусоїдальному переноснику на практиці використовують амплітудну модуляцію (АМ), балансну модуляцію (БМ), односмугову модуляцію (ОМ), фазову модуляцію (ФМ) і частотну модуляцію (ЧМ). Визначимо потенційну і реальну завадостійкість цих систем.

Амплітудна модуляція формує сигнал вигляду $S(t, B) = U_0(1 + mB(t))\cos \omega_0 t$, де $m \leq 1$. Тоді можна записати

$$\begin{aligned} \frac{\partial s(t, b)}{\partial b} &= mU_0 \cos \omega_0 t; \\ \overbrace{\left(\frac{\partial s}{\partial b}\right)^2} &= \frac{1}{T} \int_0^T \left(\frac{\partial s}{\partial b}\right)^2 dt = \frac{1}{T} \int_0^T m^2 U_0^2 \cos^2 \omega_0 t dt = \\ &= \frac{1}{2T} m^2 U_0^2 \int_0^T [1 + \cos 2\omega_0 t] dt = \frac{1}{2} m^2 U_0^2. \end{aligned}$$

Згідно з виразом (8.28) СЩП шуму на виході приймача АМ дорівнює

$$G_{0,\varepsilon} = \frac{2N_0}{m^2 U_0^2}. \quad (8.37)$$

Потужність сигналу на вході приймача визначається з виразу

$$P_c = \overline{S^2(t, B)} = \frac{U_0^2}{T} \int_0^T [1 + 2mB(t)]^2 \cos^2 \omega_0 t dt =$$

$$= \frac{U_0^2}{2T} \int_0^T [1 + 2m\overline{B(t)} + m^2 \overline{B^2(t)}] [1 + \cos 2\omega_0 t] dt.$$

Оскільки повідомлення - центрований процес, то $\overline{B(t)} = 0$, а унаслідок прийнятого нормування $\overline{B^2(t)} = \frac{1}{\Pi^2}$, $P_c = \frac{U_0^2}{2} \left(1 + \frac{m^2}{\Pi^2}\right)$.

Тут і далі вважаємо, що $\omega_0 T \gg 1$, і тому значенням $\int_0^T \cos 2\omega_0 t dt = \frac{\sin 2\omega_0 T}{2\omega_0} \leq \frac{1}{2\omega_0}$ можна нехтувати в порівнянні з T .

На підставі виразу (8.32) з урахуванням того, що при АМ $F = 2F_c$, одержуємо такі вирази для виграшу й узагальненого виграшу при оптимальному прийомі:

$$g_{AM} = \frac{\alpha m^2}{m^2 + \Pi^2} = \frac{2m^2}{m^2 + \Pi^2}; g'_{AM} = \frac{m^2}{m^2 + \Pi^2}. \quad (8.38)$$

Граничне значення виграшу при АМ дорівнює 1. Воно досягається тоді, коли $m=1$ і $\Pi=1$. Практично завжди $m < 1$, а $\Pi > 1$, і тому $g_{AM} < 1$ і $g'_{AM} < 0,5$, тобто система АМ дає програш. Так, при передачі мови $\Pi \approx 3$, і тоді при $m = 0,7$ $g_{AM} = 0,103$ і $g'_{AM} = 0,051$. Малі значення виграшу при АМ обумовлені тим, що лише невелика частина потужності сигналу поміщена в бічних смугах, що несуть корисну інформацію. Отже, усунення несучої в сигналі АМ може призвести до збільшення виграшу, що і має місце при балансній і односмуговій модуляції (без несучої).

При балансній модуляції (АМ з пригніченою несучою) сигнал формується шляхом простого перемноження первинного сигналу $B(t)$ і коливання переносників $U_0 \cos \omega_0 t$:

$$S(t, B) = B(t)U_0 \cos \omega_0 t.$$

Середня потужність такого сигналу запишеться як

$$P_c = \frac{1}{2} U_0^2 \overline{B^2(t)} = \frac{1}{2} \frac{U_0^2}{\Pi^2}, \text{ а } \left(\frac{\partial s}{\partial b} \right)^2 = \frac{1}{2} U_0^2.$$

З урахуванням цих співвідношень маємо (при $\alpha=2$) $g_{\text{бм}} = 2, g'_{\text{бм}} = 1$. Звідси видно, що виграш при БМ не залежить від пік-фактора повідомлення.

Односмугова модуляція, як відомо, є просто перенесенням спектра первинного сигналу вгору на частоту ω_0 . Це лінійна операція, при якій не змінюється ширина спектра, а також співвідношення між потужностями складових. Тому при демодуляції вхідна завада перетвориться на вихідну таким самим чином, як і сигнал. З цих міркувань можна, не вдаючись до громіздких обчислень, записати $g_{\text{ом}} = g'_{\text{ом}} = 1$. Таким чином, узагальнений виграш при БМ і ОМ однаковий. Проте смуга частот, займана сигналом при ОМ, в 2 рази менше, ніж при БМ.

Системи АМ, БМ і ОМ - лінійні, тому отримані вище співвідношення для оптимального приймача справедливі як при слабких, так і при сильних завадах на вході. Ці співвідношення визначають граничну (потенційну) завадостійкість систем.

З'ясуємо, чи реалізується ця завадостійкість при звичайних способах прийому, використовуваних у реальних приймачах. При «лінійному» детектуванні граничні значення $g'_{\text{АМ}} = 1$ досягаються лише при великому відношенні сигнал-шум. Можливо показати, що при синхронному детектуванні АМ коливання потенційна завадостійкість реалізується при довільному відношенні сигнал-шум у каналі. Можна показати, що при детектуванні сигналів БМ і ОМ синхронним детектором потенційна завадостійкість також реалізується при будь-якому рівні шуму на вході приймача.

Фазова модуляція також належить до прямих систем модуляції, проте до нелінійних. При фазовій модуляції сигнал можна записати як $S(t, B) = U_0 \cos[\omega_0 t + MB(t)]$, де M - девіація фази або індекс фазової модуляції. Для такого сигналу

$$\begin{aligned} \overline{\left(\frac{\partial s}{\partial b}\right)^2} &= \frac{1}{2} M^2 U_0^2; P_c = \overline{S^2(B,t)} = \frac{1}{2} U_0^2; \\ G_\varepsilon(f) &= \frac{2N_0}{(M^2 U_0^2)}, g_{\delta i} = \frac{M^2 \alpha}{\Pi^2}, g' = \frac{M^2}{\Pi^2}. \end{aligned} \quad (8.39)$$

При великих індексах модуляції $M \approx \alpha/2$ маємо

$$g_{\Phi M} \approx \frac{\alpha^3}{4\Pi^2}; g'_{\Phi M} \approx \frac{\alpha^2}{4\Pi^2}. \quad (8.40)$$

Як видно, при ФМ виграш залежить від індексу модуляції і пік-фактора повідомлення. Оскільки M може бути більше одиниці, то і виграш у цій системі можна отримати значно більше одиниці. Платою за цей виграш є розширення смуги частот, займаної сигналом. Отримані співвідношення справедливі для малого рівня шуму на вході приймача, оскільки сигнал $S(t, B)$ при ФМ нелінійно залежить від $B(t)$.

Частотна модуляція належить до інтегральних систем модуляції, потенційна завадостійкість якої визначається на підставі співвідношення (8.29). Сигнал при ЧМ можна подати як $S(t, B) = U_0 \cos[\omega_0 t + \Delta\omega \Psi(t)]$, де $\Delta\omega$ - девіація частоти $\Psi(t) = \int_0^t B(\tau) d\tau$. Для такого сигналу

$$\frac{\partial s}{\partial \psi} = -\Delta\omega U_0 \sin[\omega_0 t + \Delta\omega \psi(t)]; \overline{\left(\frac{\partial s}{\partial \psi}\right)^2} = 0,5 \Delta\omega^2 U_0^2.$$

Тоді на підставі виразу (8.29) одержуємо СЩП шуму на виході приймача ЧМ

$$G_{0,\varepsilon}(f) = 2 \left(\frac{2\pi f}{\Delta\omega} \right)^2 \frac{N_0}{U_0^2}, \quad (8.41)$$

а на підставі виразу (8.32) одержуємо вирази для виграшу й узагальненого виграшу при ЧМ

$$g_{\text{ЧМ}} = \frac{3M^2\alpha}{\Pi^2} \approx \frac{3}{4} \frac{\alpha^3}{\Pi^2}; \quad g'_{\text{ЧМ}} = \frac{3M^2}{\Pi^2} \approx \frac{3}{4} \frac{\alpha^2}{\Pi^2}. \quad (8.42)$$

Тут враховано, що при великому індексі модуляції $M = \Delta f / F_c$, а смуга сигналу $F \approx 2MF_c$. При ЧМ так само, як і при ФМ, виграш може бути значно більше одиниці, і досягається це за рахунок розширення смуги частот сигналу (збільшення індексу модуляції).

Частотна і фазова модуляції є прикладами систем, в яких правильність передачі повідомлень при даному рівні завад може бути підвищена не тільки за рахунок збільшення потужності сигналу, як це має місце при лінійних видах модуляції, але і за рахунок розширення смуги частот, займаної сигналом. Всі ці висновки і отримані вище співвідношення для ЧМ, так само як і для ФМ, справедливі лише при малому рівні завад. Ці системи, як буде показано нижче, мають яскраво виражений пороговий ефект.

8.4.1. Системи з піднесучими частотами

У техніці зв'язку нерідко застосовують подвійну модуляцію. При цьому повідомленням, що передаються (первинним сигналом) модулюється допоміжне коливання (піднесуча) з частотою $f_{\text{пн}}$, а потім отриманим модульованим коливанням модулюється інше коливання (несуча) з частотою $f_n \gg f_{\text{пн}}$. У багатоканальних системах подвійна модуляція дозволяє здійснити частотне розділення каналів. В одноканальних системах з піднесучою подвійна модуляція дозволяє понизити вимоги до стабільності несучої частоти, замінивши його більш легко здійснимою вимогою до стабільності піднесучої частоти. Крім того, такі системи, як ЧМ-АМ або ФМ-АМ застосовують у тих випадках, коли не можна використовувати безпосередньо ЧМ або ФМ через селективні завмирання у каналі.

Різноманітність систем модуляції дозволяє створити велике число систем з піднесучими. Практично застосовують системи ОМ-АМ, ФМ-АМ, ЧМ-АМ, ЧМ-ОМ, ОМ-ЧМ, ЧМ-ЧМ та ін. При малому рівні завад вираш систем з піднесучими визначається таким самим методом, що і при звичайній одноразовій модуляції за допомогою формул (8.32). Розглянемо для прикладу систему ФМ-АМ. У цій системі сигнал має вигляд

$$S(t, B) = U_0(1 + m_n \cos[\omega_{nn}t + M_{nn}B(t)])\cos(\omega_n t + \varphi). \quad (8.43)$$

Оскільки система ФМ-АМ належать до прямих систем модуляції, то вираш можна визначити за формулами (8.32) з урахуванням виразу (8.28).

Для сигналу виразу (8.43) маємо

$$\overline{\left(\frac{\partial s}{\partial b}\right)^2} = \frac{1}{2}U_0^2 m_n^2 M_{nn}^2; P_c = \overline{S^2(t)} = \frac{1}{2}U_0^2 \left(1 + \frac{m_n^2}{\Pi^2}\right).$$

Тоді одержуємо узагальнений вираш у системі ФМ-АМ

$$g' = \frac{M_{nn}}{\Pi^2} \frac{m_n^2}{2 + m_n^2} = g'_{\text{ФМ}} g'_{\text{АМ}}. \quad (8.44)$$

З (8.44) випливає, що узагальнений вираш дорівнює добутку вирашів при ФМ і АМ. Легко переконатися, що при будь-якій системі з піднесучою, в якій модуляція несучої є прямою, узагальнений вираш g' дорівнює добутку вирашів першого g'_{nn} і другого g'_n степенів модуляції, тобто $g' = g'_{nn} \times g'_n$.

8.4.2. Системи з імпульсною модуляцією

Згідно з теоремою Котельникова безперервне коливання $B(t)$ може бути передано по лінії зв'язку з необхідною точністю шляхом передачі його окремих миттєвих значень

$$\dots, B(-2\Delta), B(-\Delta), B(0), B(\Delta), B(2\Delta), \dots, \quad (8.45)$$

узятих для моментів часу, що знаходяться на відстані один від одного на величину $\Delta \leq 1/2F_c$, де F_c - найвища частота, що міститься в коливанні $B(t)$.

У системах зв'язку, заснованих на цьому принципі (імпульсних системах), для передачі коливання $B(t)$ використовується періодична послідовність імпульсів

$$f_n(t) = \sum_{k=-\infty}^{\infty} v(t - k\Delta). \quad (8.46)$$

При цьому один з параметрів цієї послідовності змінюється відповідно до зміни миттєвих значень виразу (8.45) передаваного коливання. Таким параметром може бути, наприклад, амплітуда імпульсу, момент його приходу (фаза), його ширина і частота проходження імпульсів. Відповідно види імпульсної модуляції називаються АІМ, ШІМ, ЧІМ.

Модульована послідовність імпульсів на виході першого ступеня модуляції буде записано як

$$f(B, t) = \sum_{k=-\infty}^{\infty} v[B(k\Delta), t - k\Delta]. \quad (8.47)$$

Тут $v(t)$ - форма імпульсу, причому $v(t) = 0$ при $t < 0$.

Для передачі коливання $f(B, t)$ по радіоканалу необхідно застосувати ще один ступінь модуляції. При цьому може бути використана будь-яка з систем модуляції, розглянутих вище. Проте найчастіше в другому ступені модуляції модульована послідовність імпульсів $f(B, t)$ перемножується з гармонійною напругою несучої, внаслідок чого виходить сигнал

$$S(t, B) = f(B, t) a \cos \omega_n t = a \sum_{k=-\infty}^{\infty} v[B(k\Delta), t - k\Delta] \cos \omega_n t. \quad (8.48)$$

Цю операцію перемножування часто називають *амплітудною модуляцією* (хоча правильніше використовувати термін БАМ) і позначають відповідні системи скороченнями типу АІМ-АМ, ФІМ-АМ, ШІМ-АМ і т. д.

Зі всіх систем імпульсної модуляції найбільш завадостійкими є ФІМ і ЧІМ. Оскільки перша з них дещо простіше, то для передачі сигналів по каналах зв'язку використовується майже виключно ФІМ-АМ. Модуляцію АІМ і ШІМ в каналах зв'язку практично не використовують і застосовують лише в процесі обробки сигналів. Тому наведемо тільки відомості про завадостійкість системи ФІМ-АМ [42].

Узагальнений виграш у системі ФІМ-АМ

$$g' = \frac{1}{\Pi^2 k_\phi^2} \left(\frac{\Delta\tau}{\tau_0} \right)^2, \quad (8.49)$$

де k_ϕ - коефіцієнт, залежний від форми імпульсу;

$\Delta\tau$ - максимальна девіація положення імпульсу;

τ_0 - тривалість імпульсу.

Як бачимо, завадостійкість системи ФІМ залежить від форми імпульсу. При оптимальному виборі параметрів системи ФІМ-АМ: $F\tau_0 \approx 1$, $\Delta\tau = 1/4F_c$ одержуємо

$$g'_{opt} = \frac{\alpha^2}{16\Pi^2 k_\phi^2}. \quad (8.50)$$

При $k_\phi^2 = \frac{1}{12}$, що відповідає трикутній формі імпульсів, маємо

$$g'_{opt} = \frac{3}{4} \frac{\alpha^2}{\Pi^2}. \quad (8.51)$$

Це співпадає з виразом (8.42) для системи з частотною модуляцією. Отже, потенційна завадостійкість системи ФІМ-АМ при оптимальному виборі параметрів і системи ЧМ при тому самому значенні α однакова.

8.5. Поріг завадостійкості широкопasmового демодулятора

Всі широкопasmові системи модуляції забезпечують високу завадостійкість за умови, що відношення сигналу до завади $\rho_{вх}$ на вході приймача більше за деяке граничне (порогове) значення $\rho_{пор}$. При $\rho_{вх} < \rho_{пор}$ широкопasmові системи втрачають свої переваги (різко знижується завадостійкість) і зв'язок стає практично неможливим. Значення порогу визначає граничну дальність зв'язку при заданій потужності передавача. Тому важливо визначити це значення і встановити закон зміни відношення сигналу до завади на виході приймача за порогом ($\rho_{вх} < \rho_{пор}$). Для ідеальної системи ($\eta = 1$) згідно з виразом (8.35) при $\rho_{вх} \gg 1$ маємо

$$\rho_{вих} = \rho_{вх}^{\alpha}, \quad (8.52)$$

де $\alpha = F/F_c$.

Для реальної системи згідно з виразом (8.34) одержуємо

$$\rho_{вих} = g\rho_{вх}, \quad (8.53)$$

де g – виграш системи.

Графічно цим рівнянням у логарифмічному масштабі відповідає сімейство прямих. Для ідеальної системи на рис. 8.5 вони зображені безперервними лініями. По осі абсцис відкладені не $\rho_{вх}$, а $\rho'_{вх} = \alpha\rho_{вх} = P_c/N_0F_c$, тобто відношення потужності сигналу на вході приймача до потужності тієї частини шуму, яка лежить у смузї частот, що дорівнює ширині спектра повідомлень F_c (а не в смузї спектра сигналу F). Рівність $F = F_c$ має місце безпосередньо при передачі, а також при однопasmовій модуляції. Для $\alpha = 1$ пряма проходить через початок координат під кутом $\pi/4$; для $\alpha = 1$ ці прямі йдуть крутіше і перетинають вісь абсцис правіше за початок координат.

Для реальних систем відповідні залежності (8.53) є прямими, паралельними прямій, що проходять через початок координат під кутом $\pi/4$, але зсунені на величину узагальненого виграшу $g' = g/\alpha$ (на рис. 8.5 одна з таких кривих для ЧМ зображена штриховою лінією). Якби криві для реальних та ідеальних систем перетиналися, то це б означало, що реальна система стала краще за ідеальну.

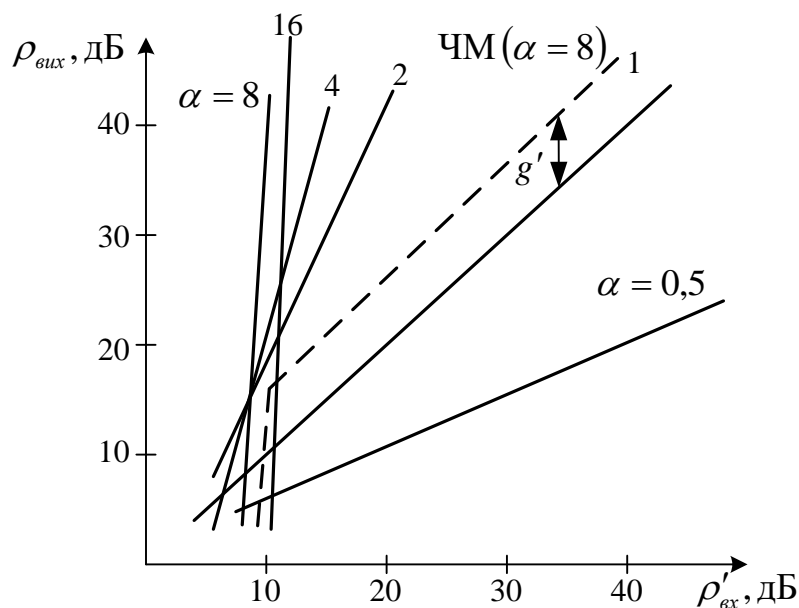


Рис. 8.5. Залежність $\rho_{вих}$ від $\rho_{вх}$ для ідеальної системи зв'язку

Очевидно, що це неможливо, і криві для реальних систем розташовуються завжди не вище відповідних кривих для ідеальних систем. Звідси можна зробити висновок, що системи, які дають виграш при великих $\rho_{вх}$, не можуть зберігати його постійним зі зменшенням $\rho_{вх}$. Починаючи з деякого порогового значення цього відношення $\rho_{вх}$ виграш системи різко зменшується.

Більш того, виграш може стати негативним (програшем), якщо $\rho'_{вх} < \alpha$. Чим більше α , тим сильніше позначається пороговий ефект. Системи з $\alpha \leq 1$ взагалі не схильні до порогових явищ, але вони і при великих $\rho_{вх}$ не дають виграшу. Проте вони дозволяють передавати повідомлення по каналу із

смугою пропускання, що дорівнює або менше ширини спектра повідомлення. Для цього повідомлення повинне бути перетворене на сигнал, спектр якого менше спектра повідомлення.

Щоб отримати високу правильність передачі в такій системі, необхідно мати чимале відношення потужності сигналу до потужності завади в каналі, оскільки при $\alpha < 1$ $\rho_{вих} < \rho_{вх}$. На практиці системи з $\alpha < 1$ застосування поки не знаходять. Для визначення граничних (порогових) значень $\rho_{вх\ пор}$ і $\rho_{вих\ пор}$ необхідно спільно розв'язати два рівняння (8.52) і (8.53), тобто знайти точку перетину кривих для ідеальної і реальної систем. У результаті одержуємо

$$\rho_{вх\ пор} = g^{\frac{1}{\alpha-1}} ; \rho_{вих\ пор} = g^{\frac{\alpha}{\alpha-1}}. \quad (8.54)$$

Як бачимо, виграш g і відповідно порогові значення $\rho_{пор}$ залежать від параметра α , визначуваних виглядом і параметрами модуляції, а також пік-фактором повідомлень, що передаються.

Пояснимо тепер явище порогу геометрично. Для цього скористаємося поняттями простору повідомлень і простору сигналів, введеними раніше. Кожному сигналу $s(t)$, визначуваному кінцевим числом координат, відповідає точка в n -мірному просторі, де $B = 2TF$ - база сигналу, а ансамблю можливих сигналів - деяка область у цьому просторі сигналів. При фіксованій потужності областю можливих сигналів є гіперсфера з радіусом $r_c = \sqrt{BP_c} = \sqrt{2TFP_c}$, а область прийнятих коливань $Z(t) = s(t) + N(t)$ - сфера з радіусом $r_z = \sqrt{2TF(P_c + P_{ш})}$. Навколо кожної точки переданого сигналу утворюється область невизначеності, обумовлена завадою. Якщо завадою є гаусівський шум, то ця область має сферичну форму з радіусом

$$r_{ш} = \sqrt{2TFP_{ш}}. \quad (8.55)$$

Як зазначалося, модуляція є відображенням простору повідомлень на простір сигналів, а демодуляція - зворотним відображенням простору прийнятих сигналів у простір прийнятих повідомлень (оцінок). У загальному випадку розмірність (база) простору повідомлень $2F_cT$ відрізняється від розмірності простору сигналів $2FT$.

Розглянемо деякий відрізок прямої в області простору можливих повідомлень (ця область визначається з урахуванням нормування повідомлення). При модуляції цей відрізок перетвориться на деяку лінію в просторі сигналів, кожна точка якої відповідає певній реалізації повідомлення. Форма і довжина цієї лінії залежать від виду модуляції. Завада, що поклалася на сигнал, викличе зсув точки s в деяку точку $\hat{s} = s + \Delta s$, що відповідає іншому повідомленню $\hat{b}(t) = b + \Delta b$, аналогічно до того, як було показано на рис. 8.4. Очевидно, чим більше $\Delta s / \Delta b$, тим менше помилка в прийнятому повідомленні при тій самій заваді.

Для збільшення завадостійкості (відношення $\Delta s / \Delta b$) необхідно збільшувати довжину лінії сигналів, що відповідає даному відрізку в просторі повідомлень. Це можна зробити, розширюючи використовувану область сигналів (радіус r_c). Але оскільки потужність сигналу обмежена, для цього потрібно збільшувати базу сигналу, розширюючи його спектр. При цьому лінія сигналу може набувати складної звивистої форми.

На рис. 8.6 показаний приклад такої лінії сигналів, причому для наочності використаний двовимірний простір сигналів (при одновимірному просторі повідомлень). Зазначимо, що при лінійній модуляції лінії сигналу і повідомлень подібні і тому виграш в них неможливий. У разі широкосмугових систем (наприклад, ЧМ) лінія сигналів має складну звивисту форму. Довжина лінії збільшується зі збільшенням смуги частот сигналу.

Проте при обмеженій потужності сигналу це подовження веде до зближення різних витків лінії сигналу. Поки завада «невелика» (наприклад, n на рис. 8.6), точка прийнятого коливання $z = s + n$ з великою імовірністю потрапляє в область точки сигналу s , що передається і відтворюється повідомлення \hat{b} ,

що відповідає найближчій точці s , що знаходиться на даному витку лінії сигналу.

Зсув Δs уздовж лінії сигналу малий і визначає нормальну помилку Δb і відповідно $\overline{\varepsilon^2(t)}$ з виразу (8.2). Такі помилки погіршують лише якість оцінки повідомлення. Сильна завада (n' або n'' на рис. 8.6) може перевести точку коливання, що приймається, на інший виток лінії сигналу (точки s' і s''). Такі «перескоки» пов'язані з появою так званих *аномальних* помилок, які не тільки погіршують якість оцінки повідомлення, але роблять її абсолютно помилковою.

Поява аномальних помилок різко погіршує завадостійкість зв'язку і є причиною виникнення порогу завадостійкості в широкосмугових системах модуляції (ЧМ, ФМ, ФІМ і т. п.). Очевидно, в області порогу завадостійкість передачі безперервних повідомлень можна характеризувати імовірністю аномальних помилок $p_{ан}$ (помилки неоднозначності оцінки).

Можна сказати, що імовірність аномальної помилки $p_{ан}$ визначає правильність оцінки повідомлення, а середньоквадратична помилка $\overline{\varepsilon^2(t)}$ визначає точність (якість) цієї оцінки.

Обчислені на основі точнішої теорії залежності $\rho_{вих}$ від $\rho_{вх}$ для ОМ, АМ і ЧМ подані на рис. 8.7, з якого видно відсутність порогу при ОМ і залежність порогу при ЧМ від індексу модуляції. Існує оптимальне значення індексу модуляції, при якому для забезпечення заданої правильності прийому $\rho_{вих}$ витрачається мінімальна потужність сигналу $P_{сmin}$, що дорівнює пороговій потужності $P_{спор}$. Це значення можна знайти з умови $\rho_{вих} = \rho_{вихпор} = \rho_{вхпор} g$. Так, для ЧМ згідно з виразом (8.42) $g = 3\alpha^3 / 4\Pi^2 = 6M^3 / \Pi^2$, тоді $\rho_{вих} = 6M^3 / \Pi^2 = \rho_{вхпор}$. Звідси одержуємо

$$M_{opt} \leq \sqrt[3]{\frac{\Pi^2}{6} \frac{\rho_{вих}}{\rho_{вхпор}}}. \quad (8.56)$$

Таким чином, оптимальний індекс модуляції тим більше, чим вище необхідна правильність передачі ($\rho_{вих}$).

Набуті теоретичні значення порогу для різних видів модуляції є граничними, які в реальних прийमाх можуть бути досягнуті, але не можуть бути перевершені.

За експериментальними даними, поріг у системі ЧМ при звичайному способі прийому настає приблизно при рівності пікових значень сигналу і завади. Це означає, що при великому рівні завад реальна завадостійкість приймача ЧМ значно менше за потенційну.

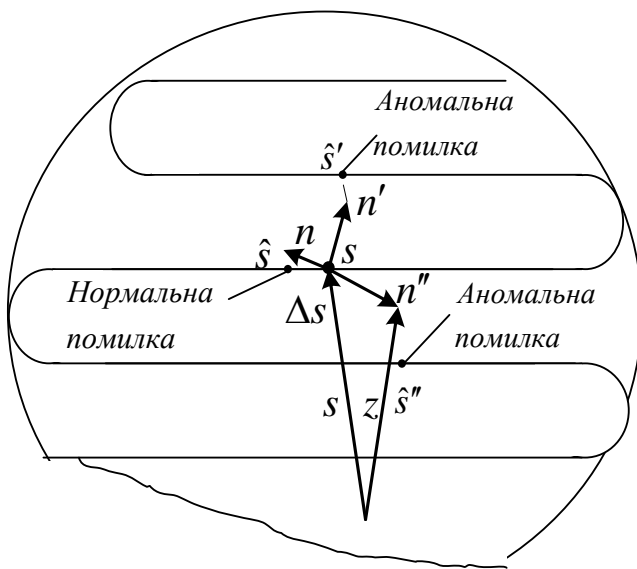


Рис. 8.6. Геометричне трактування появи аномальних помилок

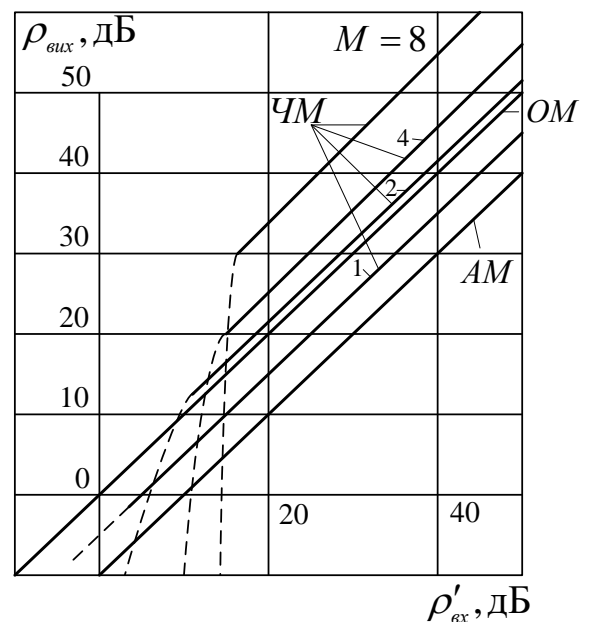


Рис. 8.7. Криві завадостійкості для ЧМ ($M=1,2,4,8$), АМ і ОМ

Отже, є можливість удосконаленням схеми приймача понизити поріг завадостійкості і тим самим збільшити дальність зв'язку при тій самій потужності передавача. Це завдання особливо актуальне для супутникових і космічних систем зв'язку.

Для зниження порогу при ЧМ застосовують різні схеми стежачих демодуляторів, зокрема схему зі зворотним зв'язком за частотою, синхронно-фазовий демодулятор і демодулятор із стежачим фільтром. Мінімальний (граничний) поріг завадостійкості досягається у схемі оптимального демодулятора.

8.6. Оптимальна лінійна фільтрація безперервних сигналів

Лінійну фільтрацію широко використовують у системах передачі інформації для обробки сигналів, не дивлячись на те, що у багатьох випадках необхідна нелінійна обробка. Пояснюється це перш за все простотою реалізації лінійних фільтрів, які порівняно легко синтезуються, і існуванням розвиненої теорії їх побудови, чого не можна сказати про нелінійні фільтри.

Лінійні фільтри є невід'ємною частиною будь-якого приймального пристрою. За їх допомогою здійснюється як додетекторна, так і післядетекторна обробка сигналів.

За допомогою лінійних фільтрів сигнали часто розділяються в багатоканальних системах передачі. Вимоги до цих фільтрів можуть бути вельми різними залежно від їх призначення. Тут розглянемо теорію оптимальної лінійної фільтрації.

Нехай сигнал на вході лінійного фільтра з імпульсною характеристикою $g(t)$ є сумою переданого сигналу $S(t)$ і завади $N(t)$:

$$Z(t) = S(t) + N(t). \quad (8.57)$$

Потрібно знайти таку функцію $g(t)$, яка мінімізує середньоквадратичну помилку:

$$\overline{E^2(t)} = \overline{[\hat{S}(t) - S(t)]^2}, \quad (8.58)$$

де $S(t)$ - оцінка сигналу на виході фільтра.

Тут вважаємо, що час запізнювання сигналу у фільтрі $t_0 = 0$, а середнє значення береться за ансамблями сигналів S і завад N . Вважатимемо, що $S(t)$ і $N(t)$ - стаціонарні взаємно некорельовані процеси з відомими СПМ $G_S(f)$ і $G_N(f)$.

У такій постановці задача була розв'язана незалежно один від одного А.М. Колмогоровим і Н. Вінером у 1942 р.[17], і тому оптимальний (у зазначеному розумінні) лінійний фільтр називають фільтром Колмогорова-Вінера (ФКВ). Вимога до

фізичної реалізованості фільтра, як відомо, зводиться до того, що імпульсна характеристика фільтра повинна задовольняти умову $g(t) = 0$ для всіх $t < 0$. Це обмеження враховується в записі

$$\hat{S}(t) = \int_{\gamma} Z(t - \tau) g(\tau) d\tau, \quad (8.59)$$

де областю інтеграції γ для фільтра, що фізично реалізується, є інтервал $(0, \infty)$, а для фільтрів, що не реалізуються, - $(-\infty, \infty)$.

Іноді завдання фільтрації вирішується в більш загальній постановці: знайти оптимальну оцінку сигналу $S(t + \tau)$ при знаходженні коливання $Z(t) = S(t) + N(t)$ на поточному інтервалі $(0, T)$. Тоді при $\tau = 0$ матимемо завдання поточної фільтрації, при $\tau > 0$ - завдання екстраполяції (фільтрацію з попередженням або передбаченням), а при $\tau < 0$ - завдання інтерполяції (фільтрацію із запізнюванням).

Можна довести, що необхідною і достатньою умовою оптимальної лінійної поточної фільтрації є умова

$$\overline{E(t)Z(t - \tau)} = 0 \text{ для всіх } \tau. \quad (8.60)$$

Це означає, що фільтр потрібно вибрати так, щоб помилка $E(t) = \hat{S}(t) - S(t)$ була некорельована з вхідним сигналом $Z(t)$ у всі моменти часу в області γ . Якби мала місце кореляція між помилкою і приймаємим сигналом, то при подальшій обробці можна було б отримати кращу оцінку.

Подано умову (8.60) як $\overline{[S(t) - \hat{S}(t)] - Z(t - \tau)} = 0$ для всіх τ з γ . Звідси з урахуванням виразу (8.59) маємо $\overline{S(t)Z(t - \tau)} = \overline{\hat{S}(t)Z(t - \tau)} = \int_{\gamma} \overline{Z(t - x)Z(t - \tau)} g(x) dx$ або

$$B_{SZ}(\tau) = \int_{\gamma} B_Z(\tau - x) g(x) dx. \quad (8.61)$$

У тому випадку, коли сигнал $S(t)$ і завада $N(t)$ некорельовані, вираз (8.61) має вигляд

$$B_S(\tau) = \int_{\gamma} [B_S(\tau - x) + B_N(\tau - x)] g(x) dx. \quad (8.62)$$

Це основне інтегральне рівняння теорії лінійної фільтрації називається рівнянням Вінера-Хопфа [44]. Його розв'язком є функція $g(t)$, що мінімізує середній квадрат помилки $\overline{E^2(t)}$.

Слід відрізнити оптимальні лінійні фільтри, визначувані виразами (8.61) або (8.62), від узгоджених фільтрів.

Якщо основне призначення фільтрів, що розглядаються тут, полягає в найкращому відтворенні невідомої форми сигналу, то задача узгоджених фільтрів полягає у формуванні максимально можливого піку сигналу відомої форми у момент відліку на фоні шуму.

Рівняння (8.62) легко розв'язується для фільтрів, що не реалізуються, тобто коли $\gamma \in (-\infty, \infty)$. Для цього випадку, застосувавши перетворення Фур'є до обох частин виразу (8.62), одержуємо його вираз у частотній області

$$G_S(f) = (G_S(f) + G_N(f))K(f). \quad (8.63)$$

Тоді вираз для коефіцієнта ФКВ має вигляд

$$K(f) = \frac{G_S(f)}{G_S(f) + G_N(f)}. \quad (8.64)$$

Доведемо, що дисперсія помилки $E(t) = \hat{S}(t) - S(t + \tau)$ при оптимальній лінійній фільтрації, що не реалізується, в загальній постановці з виразу

$$\overline{E_{min}^2(t)} = \int_{-\infty}^{\infty} \frac{G_S(f)}{G_S(f) + G_N(f)} df. \quad (8.65)$$

Спектральна щільність потужності (СЩП) для випадкового процесу та її енергетичний спектр мають вирази

$$E(t) = \hat{S}(t) - S(t + \tau) = \int_{\gamma} [S(t) + N(\tau)] g(\tau) dt - S(t + \tau);$$

$$G_{\varepsilon}(f) = K^2(f)G_N(f) + \left| \dot{K}(f) - e^{j\omega\tau} \right|^2 G_S(f),$$

де $\dot{K}(f) = K(f)e^{j\varphi(f)}$.

Записавши

$$G_{\varepsilon}(f) = K^2(f)G_N(f) + G_S(f) \left[K^2(f) + 1 - 2K(f)\cos(\varphi - \omega\tau) \right]$$

можна побачити, що $G_{\varepsilon min}(f)$ забезпечується лише при ФЧХ оптимального фільтра $\varphi = \omega t$ і, отже, при передавальній функції

$$\text{фільтра } \dot{K}(f) = \frac{G_S(f)}{G_S(f) + G_N(f)} e^{j\omega\tau}.$$

Вважаючи, що $G_{\varepsilon}(f)_{min} = \frac{G_S(f)G_N(f)}{G_S(f) + G_N(f)}$, мінімальна

дисперсія помилки буде дорівнювати

$$\overline{E_{min}^2} = \int_{-\infty}^{\infty} G_{\varepsilon min}(f) df = \int_{-\infty}^{\infty} \frac{G_S(f)G_N(f)}{G_S(f) + G_N(f)} df. \quad (8.66)$$

Слід зазначити, що помилка $\overline{E_{min}^2} = 0$ тільки у тому випадку, коли $G_S(f)G_N(f) = 0$, тобто коли спектри сигналу і завади не перекриваються. У всіх інших випадках оптимальний фільтр пропускає різні частоти з тією меншою вагою, чим більше відношення $G_N(f)/G_S(f)$ при даній частоті.

Ми тут не обговорюватимемо питання реалізації фільтра, характеристики якого наближаються до характеристики ФКВ, що не реалізується, оскільки відомий варіант фільтра, що завжди реалізується, оптимальний за середньоквадратичним критерієм, визначуваний методом змінних стану.

Результати оптимальної фільтрації можна істотно поліпшити, якщо застосувати так зване передспотворення сигналу з подальшою його корекцією на прийомі. Суть методу

передспотворення полягає в тому, що на боці, який передає, сигнал $s(t)$ пропускається через фільтр з коефіцієнтом передачі $K_1(f)$.

Отриманий таким чином видозмінений сигнал передається по каналу. На приймальному боці включений інший фільтр $K_2(f)$. Характеристики фільтрів $K_1(f)$ і $K_2(f)$ вибираються так, щоб забезпечити мінімум середньоквадратичної помилки.

Розрахунки показують, що передспотворення дає тим більший вигаш, чим менше відносна ширина смуги перекриття спектрів сигналу і завади. Передспотворення дозволяє перерозподілити потужність корисного сигналу в смузі частот каналу так, щоб забезпечити кращі умови узгодження джерела сигналу з каналом (у загальному випадку корисно прагнути до того, щоб сума спектральної щільності потужності сигналу і потужності завади була постійною в межах смуги частот каналу).

Це означає, що передспотворення можна розглядати як деяке «лінійне кодування» безперервного сигналу, що дозволяє зменшити помилку і поліпшити використання пропускної спроможності каналу.

Лінійне передспотворення широко використовується в сучасних системах зв'язку. Характерними в цьому відношенні є системи, в яких використовується частотна модуляція. Згідно з виразом (8.41) щільність потужності шуму на виході ЧМ демодулятора збільшується пропорційно квадрату частоти, так що верхні частотні складові повідомлення схильні до шумів сильніше, ніж нижчі.

Метод передспотворення і подальша корекція дозволяють понизити шум на верхніх частотах і тим самим створити приблизно однакові умови для передачі як нижніх, так і верхніх частот повідомлення.

Слід зазначити, що в результаті передспотворень формується новий сигнал з необхідними властивостями. Так, у радіомовленні і багатоканальному радіорелейному і супутниковому зв'язку з частотною модуляцією використовується передспотворення, близьке до диференціювання.

У цьому випадку на вхід частотного модулятора поступає не первинний сигнал $b(t)$, як це робиться при ЧМ без передспотворення, а його похідна db/dt . Тому пропорційно $b(t)$ змінюється не миттєва частота, а миттєва початкова фаза несучого коливання, тобто формується не ЧМ, а ФМ сигнал.

Оскільки спектр шуму на виході демодулятора ФМ сигналу рівномірний (вираз (8.39)), то тим самим у багатоканальних системах забезпечується однакова завадостійкість у всіх частотних каналах, а в разі радіомовлення - якісніше відтворення мовних і музичних передач.

8.7. Оптимальна лінійна фільтрація безперервних повідомлень

З попереднього пункту стало відомо, що фільтр Колмогорова-Вінера, який не реалізується, є оптимальним для виділення (оцінки) стаціонарного сигналу. Модульовані ж сигнали мають кінцеву тривалість і не є стаціонарними.

На основі методу змінних стану можливий інший підхід, при якому визначають не самі характеристики оптимального фільтра, а диференціальні рівняння, що моделюють цей фільтр. У теорії нелінійної фільтрації такий підхід був застосований Р.Л. Стратоновичем (1959 р.), а в теорії лінійної фільтрації - Калманом і Б'юсі (1960 р.) [44].

Метод змінних стану (диференціальних рівнянь) дозволяє розв'язати більш загальну задачу - виділення з найменшою похибкою $k(t) = \overline{E^2(t)} = \overline{[\hat{B}(t) - B(t)]^2}$ повідомлення $b(t)$ із спостережуваного на кінцевому інтервалі часу сигналу $z(t) = s(t, b) + N(t)$. Умова стаціонарності сигналу і завади в рамках цього методу не є обов'язковою.

Щодо фільтрованого повідомлення $B(t)$ додатково припускають, що воно породжується лінійним стохастичним диференціальним рівнянням. В окремому випадку використовується рівняння [30]

$$\frac{db}{dt} = -\alpha b(t) + \alpha N_1(t), \quad (8.67)$$

де $N_1(t)$ - БГШ (процес, що породжує) з нульовим середнім значенням $\overline{N_1(t)} = 0$ і односторонньою СЦП N_1 .

Фізично це означає, що повідомлення розглядають як результат проходження стаціонарного "білого" шуму $N_1(t)$ через лінійний ланцюг, у даному випадку через інтегруючий ланцюжок RC. Для такого ланцюга постійна $\alpha = 1/(RC)$. Сформоване таким чином повідомлення $V(t)$ є гаусівським марковським процесом з функцією кореляції

$$B_b(\tau) = P_b e^{-\alpha|\tau|}, \quad P_b = \frac{N_1\alpha}{4}. \quad (8.68)$$

Спектральна щільність потужності такого повідомлення

$$G_b(f) = \frac{2\alpha P_b}{\alpha^2 + 4\pi^2 f^2}. \quad (8.69)$$

У більш загальному випадку повідомлення описується системою лінійних диференціальних рівнянь. Великий практичний інтерес являє той факт, що методи диференціальних рівнянь дозволяють синтезувати оптимальний фільтр рекурентним способом, забезпечуючи зручність його реалізації при використанні сучасних ЕОМ або мікропроцесорів.

Розглянемо випадок лінійної фільтрації, коли спостережуваний процес на вході фільтра заданий рівнянням

$$Z(t) = b(t)f(t) + N(t), \quad 0 \leq t \leq T, \quad (8.70)$$

де $f(t)$ - відома функція (несуче коливання);

$b(t)f(t) = s(t, b)$ - корисний сигнал;

$N(t)$ - "білий" гаусівський шум (не обов'язково стаціонарний) з нульовим середнім значенням $\overline{N(t)}=0$ і односторонньою спектральною щільністю N_0 .

Рівняння (8.70) є рівнянням спостереження, а рівняння (8.67) - рівнянням стану. Для отримання рівнянь фільтрації можна виходити з умови оптимальної лінійної фільтрації (вираз (8.60)). На основі цієї умови виводиться рекурентне співвідношення, яке дозволяє отримати таке рівняння:

$$\frac{d\hat{b}}{dt} = -\alpha\hat{b}(t) + \frac{2k}{N_0} f(t)[z(t) - f(t)\hat{b}(t)], \quad (8.71)$$

$$\frac{dk}{dt} = \frac{N_1}{2} - 2\alpha k - \frac{2}{N_0} f^2 k^2. \quad (8.72)$$

Рівняння (8.71) і (8.72) прийнято називати *рівняннями фільтра Калмана* (ФК). Рівняння (8.71) визначає алгоритм формування оцінки, а отже, і структурну схему фільтра, а рівняння (8.72) - помилку фільтрації $k = E^2(t)$.

Структурна схема, що моделює рівняння (8.71), наведена на рис. 8.8. Побудову схеми зручно починати з інтегратора. Для цього позначимо праву частину виразу (8.71) через

$$y(t) = y_1(t) + y_2(t),$$

де $y_1(t) = \frac{2k}{N_0} f(t)[z(t) - f(t)\hat{b}(t)], y_2(t) = -\alpha\hat{b}(t)$.

Тоді рівняння (8.71) запишеться як $d\hat{b}/dt = y(t)$. Звідси випливає, що якщо на вхід інтегратора подати напругу $y(t)$, то на його виході отримаємо оцінку повідомлення $\hat{b}(t)$. Для того щоб сформувати напругу $y_1(t)$, необхідно мати генератор Γ несучого коливання $f(t)$, два перемножувачі, суматор і підсилювач з коефіцієнтом посилення $2k/N_0$.

За допомогою цих пристроїв здійснюються всі операції, що входять у вираз для $y_1(t)$. Напряга $y_2(t)$ забезпечується за допомогою підсилювача з коефіцієнтом посилення α , на вхід якого поступає напруга оцінки $\hat{b}(t)$. Сумарна напруга $y(t) = y_1(t) + y_2(t)$ з виходу суматора поступає на вхід інтегратора, вихід якого представляє оцінку $\hat{b}(t)$. Для змодульованого сигналу, коли $s(t, b) = b(t)$, у виразах (8.71) і (8.72) потрібно враховувати $f(t) = 1$, і схема фільтра Калмана набуде більш простого вигляду (рис. 8.8).

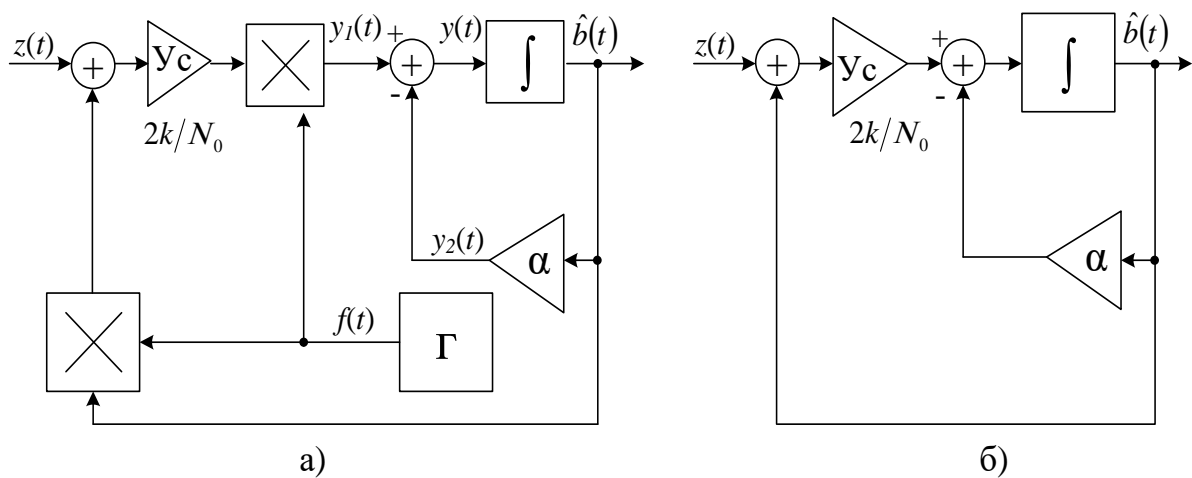


Рис. 8.8. Реалізація фільтра Калмана для гаусівського повідомлення при лінійній модуляції (а) і за відсутності модуляції (б) сигналу

Приклад. Розглянемо фільтрацію гаусівського марковського повідомлення в каналі з амплітудною модуляцією, коли для передачі повідомлення $b(t)$, заданого виразом (8.67), використовується сигнал

$$s(t, b) = b(t) \sin \omega_0 t. \quad (8.73)$$

Це сигнал АМ з пригніченою несучою. Рівняння спостереження в цьому випадку згідно з виразом(8.1)

$$z(t) = b(t) \sin \omega_0 t + N(t). \quad (8.74)$$

Рівняння оцінки згідно (8.71)

$$\frac{d\hat{b}}{dt} = -\alpha\hat{b} + \frac{2k}{N_0} [z(t)\sin\omega_0 t - \hat{b}\sin^2\omega_0 t] \quad (8.75)$$

або

$$\frac{d\hat{b}}{dt} = -\alpha\hat{b} + \frac{k}{N_0} [2z(t)\sin\omega_0 t - \hat{b}(1 - \cos 2\omega_0 t)]. \quad (8.76)$$

Практично, період коливання несучої $T = 2\pi/\omega_0 \ll \tau_k = 1/\alpha$, тобто $\alpha \ll \omega_0$, де τ_k - інтервал кореляції повідомлення $b(t)$. За цих умов доданком, пропорційним $\cos 2\omega_0 t$, можна нехтувати. Тоді маємо

$$\frac{d\hat{b}}{dt} = -\alpha\hat{b} + \frac{k}{N_0} [2z(t)\sin\omega_0 t - \hat{b}] = -\left(\alpha + \frac{k}{N_0}\right)\hat{b} + 2\frac{k}{N_0} z(t)\sin\omega_0 t. \quad (8.77)$$

Рівняння (8.77) можна моделювати лінійним фільтром розімкненого типу (рис. 8.9) з постійною часу $RC = 1/(\alpha + k/N_0)$. Постійна k , що дорівнює помилці фільтрації, визначається з рівняння (8.72).

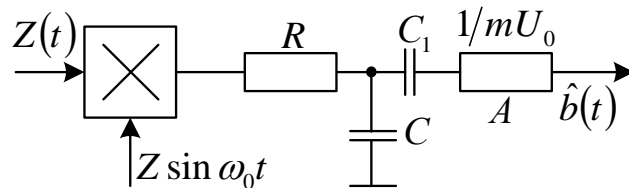


Рис. 8.9. Структурна схема оптимального демодулятора АМ сигналів

Як бачимо, оптимальний фільтр (демодулятор) для сигналів (вираз (8.73)) є схемою когерентного (синхронного) детектора з інтегруючим RC-фільтром. У разі звичайної АМ з несучою, коли $s(t, b) = U_0[1 + mb(t)]\sin\omega_0 t$, синхронний детектор виділяє обвідну $[U_0 + mU_0 \hat{b}(t)]$, тому для отримання на виході оцінки

повідомлення у схему включені розділовий конденсатор, що усуває постійну складову U_0 , і атенюатор A з коефіцієнтом згасання $1/(mU_0)$. Для немодульованого сигналу, коли $f(t)=1$ і $s(t)=b(t)$, рівняння оцінки (8.76) має вигляд

$$\frac{d\hat{b}}{dt} = -\alpha\hat{b} + \frac{2k}{N_0} [z(t) - \hat{b}] = -\left(\alpha + \frac{2k}{N_0}\right)\hat{b} + \frac{2k}{N_0} z(t). \quad (8.78)$$

Це рівняння моделюється інтегруючим фільтром з постійною часу $RC = 1/(\alpha + 2k/N_0)$ і підсилювачем з коефіцієнтом посилення $2k/N_0$. Можна показати, що в цьому випадку, коли сигнал $s(t)$ і шум $N(t)$ стаціонарні, ФК абсолютно еквівалентний ФКВ. Проте для розв'язання багатьох практичних задач обробки сигналів зокрема даних телевимірювань, фільтри Калмана за обчислювальною структурою виявилися зручнішими, ніж варіанти фільтрів Колмогорова-Вінера.

8.8. Загальні відомості про нелінійну фільтрацію

У багатьох випадках сигнал $s(t, b)$ (наприклад при ЧМ, ФМ, ФІМ) є нелінійною функцією повідомлення, що передається, і рівняння спостереження стає нелінійним. Може бути нелінійним і рівняння стану. Основоположні результати з теорії нелінійної фільтрації, з якої зокрема впливають і результати лінійної теорії, отримані Р.Л. Стратоновичем. Подальший розвиток ця теорія отримала в роботах В.І. Тихонова, М.К. Кульмана, Ю.Г. Сосуліна і багатьох інших авторів [44]. У наш час теорія нелінійної фільтрації найбільш розроблена для двох випадків: повідомлення $b(t)$ моделюється гаусівським випадковим процесом або є марковським процесом. Тут зупинимося лише на теорії фільтрації одномірних марковських гаусівських процесів, що породжуються лінійним диференціальним рівнянням (8.67). Рівняння спостереження задане як

$$Z(t) = s(t, b) + N(t), \quad 0 \leq t \leq T, \quad (8.79)$$

де $s(t, b)$ - нелінійна функція $b(t)$.

Як і в попередньому випадку, шум $N(t)$ вважається "білим" гаусівським з нульовим середнім значенням і односторонньою СЩП N_0 .

Не порушуючи спільності в рівнянні (8.67), можна припустити, що $N_1 = 4/\alpha$. Тоді $b(t)$ буде нормованим безрозмірним процесом з одиничною дисперсією (потужністю), а зміни глибини модуляції того або іншого параметра позначаються на коефіцієнті (індексі) модуляції.

Оскільки процес, що описується стохастичним рівнянням (8.67), є марковським (дифузійним), зміни в часі його щільності імовірності $w(b, t)$ визначаються рівнянням Колмогорова-Фокера-Планка [30], яке в даному випадку має вигляд

$$\frac{\partial}{\partial t} w(b, t) = -\frac{\partial}{\partial b} [\alpha b w(b, t)] + \frac{1}{4} N_1 \frac{\partial^2}{\partial b^2} w(b, t). \quad (8.80)$$

Для більшої спільності можна розглядати передачу сигналу в каналі з флюктуючою фазою $\varphi(t)$, яку часто зображують процесом з незалежними приростами, описуваним диференціальним рівнянням

$$\partial \varphi / \partial t = N_2(t), \quad (8.81)$$

де $N_2(t)$ - "білий" гаусівський шум з нульовим середнім значенням і односторонньою спектральною щільністю потужності N_2 .

Всі "білі" шуми, враховані в даній задачі, взаємно незалежні. Маючи в своєму розпорядженні ці апіорні дані, потрібно знайти пристрій, який би з найменшою похибкою відтворював випадкове коливання $b(t)$, що змінювалося в часі.

Якщо як критерій оптимальності використовується критерій середньоквадратичної помилки

$$\overline{E^2(t)} = \min \left[\overline{\hat{B}(t) - B(t)}^2 \right], \quad (8.82)$$

то, як відомо (див. п. 8.2), оптимальною оцінкою є МО апостеріорного розподілу

$$\hat{b}(t) = M[w(b|z)] = \int_B bw(b|z)db . \quad (8.83)$$

При цьому похибку оцінки можна характеризувати апостеріорною дисперсією

$$\sigma_b^2 = \int_B (\hat{b} - b)^2 w(b|z)db . \quad (8.84)$$

Таким чином, для обчислення оптимальної оцінки і її похибки необхідно знати щільність апостеріорного розподілення $w(b|z)$, яка згідно з формулою Байєса (8.6) визначається двома множниками $w(b)$ і $w(z|b)$ (див. вираз (8.20)). Щільність імовірності $w(b)$ фільтрованого процесу $b(t)$, що задовольняє вираз (8.67), визначається з виразу (8.80). Умовна щільність імовірності $w(z|b)$ (функція правдоподібності) легко знаходиться з рівняння спостереження (8.79). Оскільки сигнал $s(t, b)$ є відомою функцією аргументів t і b , а шум має гаусівський розподіл, то і $w(z|b)$ також буде гаусівським. При цьому оцінка $\hat{b}(t)$ є оптимальною не тільки за критерієм мінімуму середньоквадратичної помилки, але і за критерієм максимуму апостеріорної імовірності.

Апостеріорна імовірність містить всю інформацію про передаване повідомлення $b(t)$, яку можна витягнути із спостережуваного сигналу $z(t)$ на інтервалі $(0, T)$ і апріорних відомостей про $b(t)$. У роботах Р.Л. Стратоновича показано, що апостеріорна щільність імовірності $w(z|b)$ реалізації $b(t)$ в кінцевий момент часу спостереження визначається таким нелінійним диференціальним рівнянням [46]:

$$\frac{\partial}{\partial t} w(b|z) = \frac{\partial}{\partial t} [A_1(t,b)w(b|z)] + \frac{1}{2} \frac{\partial^2}{\partial b^2} [A_2(t,b)w(b|z)] + [F(t,b) - \overline{F(t,b)}]w(b|z), \quad (8.85)$$

де A_1 і A_2 - відповідно коефіцієнти зносу і дифузії;

$F(t,b)$ - похідна за часом від логарифма функції правдоподібності

$$F(t,b) = \frac{d}{dt} \ln w(z|b). \quad (8.86)$$

У загальному випадку з точністю до постійних

$$F(t,b) = -\frac{1}{N_0} [z(t) - s(t,b)]^2, \quad (8.87)$$

а стосовно неенергетичних параметрів сигналу (наприклад, для частоти і фази) можна записати

$$F(t,b) = -\frac{2}{N_0} z(t)s(t,b). \quad (8.88)$$

Рівняння (8.85) визначає повну процедуру фільтрації повідомлення на фоні "білого" шуму. Оптимальний приймальний пристрій повинен моделювати рівняння (8.85) і визначати оцінку $\hat{b}(t)$, що відповідає максимуму апостеріорної імовірності. У загальному випадку аналітичне розв'язання цього рівняння виявляється важким завданням. Схеми оптимального приймача при цьому вельми складні. Для отримання простіших схем доцільно використовувати наближені рішення. При чималих відносинах сигнал - шум і великому часі спостереження є підстави вважати, що апостеріорна щільність імовірності приблизно гаусівська, тобто

$$w(b|z) \approx \frac{2}{\sqrt{2\pi\sigma_b^2(t)}} \exp\left[-\frac{[b(t) - \hat{b}(t)]^2}{2\sigma_b^2(t)}\right]. \quad (8.89)$$

Таку апроксимацію $w(b|z)$ в теорії нелінійної фільтрації прийнято називати *гаусівським наближенням*, а алгоритм обробки коливання $z(t)$, що складається при цьому, - квазіоптимальним. Розв'язання задачі в цьому випадку істотно спрощується, оскільки апостеріорна щільність імовірності виразу (8.89) визначається лише двома параметрами: середнім значенням $\hat{b}(t)$, що визначає оптимальну оцінку передаваного повідомлення $b(t)$, і дисперсією $\sigma_b^2(t)$, що характеризує помилку фільтрації.

Якщо вважати коефіцієнт зносу $A_1(b, t)$ лінійною функцією b , а коефіцієнт дифузії $A_2(b, t) = \text{const}$, то після підстановки виразу (8.89) у вираз (8.85) і нескладних обчислень отримаємо для визначення $\hat{b}(t)$ і $\sigma_b^2(t) = k(t)$ систему з двох звичайних нелінійних диференціальних рівнянь:

$$\left. \begin{aligned} \frac{d}{dt} \hat{b}(t) &= A_1(t, \hat{b}) + k(t) \frac{\partial F(t, \hat{b})}{\partial \hat{b}}; \\ \frac{d}{dt} k(t) &= 2k(t) \frac{\partial A_1(t, \hat{b})}{\partial \hat{b}} + k^2(t) \frac{\partial^2 F(t, \hat{b})}{\partial \hat{b}^2} + A_2. \end{aligned} \right\} \quad (8.90)$$

Таким чином, задача оптимальної фільтрації в гаусівському наближенні зводиться до сумісного розв'язання або моделювання рівнянь (8.90). Система, що моделює ці рівняння (нелінійний фільтр), відтворюватиме передане повідомлення $b(t)$ з мінімальною середньоквадратичною помилкою (принаймні, у разі слабких завад, коли виправдано гаусівське наближення).

Для марковського гаусівського процесу, що описується стохастичним диференціальним рівнянням (8.67),

$$A_1(t, \hat{b}) = -ab(t), \quad A_2(t, \hat{b}) = \frac{1}{2} N_1$$

рівняння (8.90) запишуться так:

$$\left. \begin{aligned} \frac{d}{dt} \hat{b}(t) &= -\alpha \hat{b}(t) + k(t) \frac{\partial F(t, \hat{b})}{\partial \hat{b}}; \\ \frac{d}{dt} k(t) &= -2\alpha k(t) + \frac{1}{2} N_1 + k^2(t) \frac{\partial^2 F(t, \hat{b})}{\partial \hat{b}^2}. \end{aligned} \right\} \quad (8.91)$$

Якщо сигнал $s(t, b) = b(t)f(t)$ лінійно залежить від повідомлення, що передається $b(t)$, то згідно з виразом (8.87)

$F(t, b) = -\frac{1}{N_0} [z(t) - f(t)b(t)]^2$ і рівняння (8.91) переходять у рівняння (8.71), (8.72), що описують *лінійний оптимальний* ФК.

Розглянута теорія нелінійної фільтрації може бути узагальнена і на випадок, коли повідомлення, що передається, описується негаусівським марковським процесом. При цьому замість рівняння (8.67) фігурує нелінійне стохастичне диференціальне рівняння.

Іншим узагальненням є передача сигналу, що модулюється повідомленням $b(t)$, яке можна апроксимувати компонентою багатомірного марковського процесу.

Приклад. Розглянемо застосування теорії нелінійної фільтрації для синтезу оптимальних приймачів у системах з фазовою і частотною модуляцією.

Сигнал при ФМ

$$s(t, b) = U_0 \cos[\omega_0 t + \theta(t)], \quad \theta(t) = Mb(t) + \varphi(t), \quad (8.92)$$

де U_0 і ω_0 - апріорі відомі значення амплітуди і частоти;

M - індекс фазової модуляції;

$\varphi(t)$ - процес, що описує флуктуацію початкової фази;

СП $B(t)$ з реалізацією $b(t)$ покладемо нормованим одномірним марковським гаусівським процесом.

Для знаходження рівняння оцінки визначимо функцію $F(t, \hat{b})$. Згідно з виразом (8.88) маємо

$$F(t, \hat{b}) = \frac{2}{N_0} z(t) s(t, \hat{b}) = \frac{2U_0}{N_0} z(t) \cos[\omega_0 t + \hat{\theta}(t)]. \quad (8.93)$$

Запишемо рівняння оцінки, вважаючи спочатку, що флуктуації початкової фази відсутні: $\varphi(t) = \varphi_0$. Підставляючи вираз (8.93) в рівняння (8.91), одержуємо

$$\frac{d\hat{b}}{dt} = -\alpha \hat{b}(t) - \frac{2kMU_0}{N_0} z(t) \sin[\omega_0 t + \hat{\theta}(t)]. \quad (8.94)$$

Побудуємо структурну схему фільтра, що виділяє із спостережуваного коливання $z(t)$ оптимальну оцінку $\hat{b}(t)$. Для цього позначимо $\frac{2kMU_0}{N_0} z(t) \sin[\omega_0 t + \hat{\theta}(t)] = y(t)$ і перепишемо вираз (8.94) як

$$y(t) = \hat{b}(t) + \frac{1}{\alpha} \frac{d\hat{b}}{dt}. \quad (8.95)$$

Легко переконатися, що якщо подати напругу $y(t)$ на інтегруючий RC-ланцюг, де $RC = 1/\alpha$, то напруга на конденсаторі буде дорівнюватиме $\hat{b}(t)$. Для того щоб сформувати напругу $y(t)$, досить мати автогенератор з частотою f_0 , що моделюється за фазою сигналом $\hat{b}(t)$ з індексом модуляції M , і перемножувач, що здійснює множення напруги автогенератора на вхідне коливання $z(t)$. Амплітуда автогенератора $U_z = 2kMU_0/(\alpha N_0)$. Коефіцієнт k визначають розв'язанням нелінійного диференційного рівняння (8.91). У даному випадку для сталого режиму одержуємо

$$k = \frac{\alpha N_0}{M^2 U_0^2} \left(\sqrt{\frac{2M^2 U_0^2 P_b}{\alpha N_0} - 1} \right). \quad (8.96)$$

Таким чином, визначаємо один з можливих варіантів схеми для отримання оптимальної оцінки $b(t)$, зображеної на рис. 8.10. Тут ПГ - підстроюваний генератор, фаза якого модулюється за допомогою елемента, що управляє (УЕ). Це ніщо інше, як схема автопідстроювання фази автогенератора за вхідним коливанням $z(t)$.

Завдяки цьому вона має певною мірою властивості саморегулювання, зокрема малочутлива до точності установлення початкової фази φ_0 автогенератора і його амплітуди. Проте якщо початкова фаза схильна до значних (за величиною і швидкістю) флуктуацій, схема рис. 8.10 стає далеко не оптимальною.

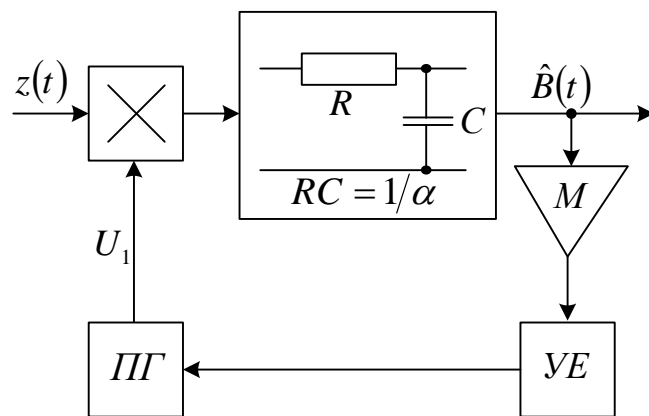


Рис. 8.10. Структурна схема оптимального демодулятора ФМ сигналу з постійною фазою

Якщо фаза $\varphi(t)$ флюктує, то доводиться розв'язувати рівняння сумісної оцінки $\hat{b}(t)$ і $\hat{\varphi}(t)$. У цьому випадку структурна схема оптимального демодулятора ФМ сигналу дещо ускладнюється в порівнянні зі схемою на рис. 8.10, оскільки тепер треба модулювати керований генератор сигналом $\hat{\theta}(t) = M\hat{b}(t) + \hat{\varphi}(t)$.

При ЧМ в $\theta(t) = \Delta\omega \int_0^t b(t)dt + \varphi(t)$. В іншому характер всіх

співвідношень зберігається. Схема оптимального нелінійного фільтра за структурою така сама, як і для ФМ. Відмінність полягає в тому, що в підстроюваному генераторі модулюється не фаза, а частота. Дещо відрізняються також коефіцієнти посилення. Отриманий фільтр є схемою фазового автопідстроювання частоти (ФАПЧ). Зазначимо, що вона аналогічна до схеми на рис. 8.10.

Аналогічно визначають схеми оптимальної нелінійної фільтрації для інших видів модуляції і для складніших каналів.

8.9. Загальні відомості про цифрову передачу безперервних повідомлень

Для передачі безперервних повідомлень можна скористатися дискретним каналом. При цьому необхідно перетворити безперервне повідомлення на *цифровий* сигнал, тобто на послідовність символів, зберігши істотну частину інформації, що міститься в повідомленні, визначувану його епсилон-ентропією.

Типовими прикладами цифрових систем передачі безперервних повідомлень є системи з імпульсно-ковою модуляцією (ІКМ) і дельта-модуляцією (ДМ).

Як зазначалося в п. 2.4, для перетворення безперервного повідомлення на цифрову форму використовуються операції дискретизації і квантування. Отримана таким чином послідовність квантованих відліків кодується і передається по дискретному каналу як будь-яке дискретне повідомлення. На приймальному боці безперервне повідомлення після декодування відновлюється (з тією або іншою точністю).

Основна технічна перевага цифрових систем передачі перед безперервними системами полягає в їх високій завадостійкості. Ця перевага найсильніше виявляється в системах передачі з *багатократною ретрансляцією* (переприйманням) сигналів.

Типові системи подібного типу - кабельні і радіорелейні лінії великої протяжності. У них сигнали передаються по

ланцюгу ретрансляторів, розташованих на таких відстанях один від одного, які забезпечують надійний зв'язок. У таких системах завади і спотворення, що виникають в окремих ланках, як правило, накопичуються.

Припустимо для простоти, що сигнал у кожному ретрансляторі тільки посилюється. Тоді, якщо адитивні завади в кожній ланці статистично незалежні, їх потужність на вході останньої ланки дорівнює сумі потужностей завад всіх ланок.

Якщо система складається з k однакових ланок, для забезпечення заданої правильності зв'язку необхідно забезпечити на вході кожного ретранслятора відношення сигнал-завада в k раз більше, ніж при передачі без ретрансляцій. У реальних системах число ретрансляцій k може досягати декількох десятків, а іноді і сотень; накопичення завад уздовж тракту передачі стає основним чинником, що обмежує протяжність лінії зв'язку.

При цифрових системах передачі для ослаблення ефекту накопичення завад при передачі з ретрансляціями разом з посиленням застосовують регенерацію імпульсів, тобто демодуляцію з відновленням переданих кодових символів і повторну модуляцію на переприймальному пункті.

При використанні регенерації адитивна завада з входу ретранслятора не поступає на його вихід. Проте вона викликає помилки при демодуляції. Помилково прийняті в одному регенераторі символи у такому вигляді передаються і на наступні регенератори, так що помилки все ж таки накопичуються.

При використанні k регенераторів і допустимої імовірності помилки $p \ll 1$ необхідно в демодуляторі кожного регенератора забезпечити імовірність помилки p_1 не більш ніж p/k . Проте на відміну від попереднього випадку для цього зовсім не потрібно збільшувати відношення сигнал-завада на вході регенератора в k раз. Так, при передачі двійкових символів у каналі з адитивним гаусівським шумом і некогерентному прийомі двійкових ортогональних сигналів помилки визначаються у гл.6.

Якщо потрібно забезпечити $p < 10^{-5}$, то в системі без переприймання для цього необхідно мати $h^2 = -21n(2p) = 21,64$. При використанні $k = 1000$ регенераторів слід забезпечити на

кожному з них $p_1 < 10^{-8}$ або $h_2 = 21n(2p)^{-1} \approx 35,45$. Для цього досить підняти потужність сигналу в 1,64 разу, а не в 1000 разів.

При цифровій системі передачі безперервних повідомлень можна, крім того, підвищити правильність застосуванням завадостійкого кодування. Висока завадостійкість цифрових систем передачі дозволяє здійснити практично необмежений за дальністю зв'язок при використанні каналів порівняно невисокої якості.

Іншою істотною перевагою цифрових систем передачі інформації є широке використання в апаратурі перетворення сигналів сучасної елементної бази цифрової обчислювальної техніки і мікропроцесорів. Сукупність пристроїв, що виконують перетворення вхідних цифрових сигналів на вихідні цифрові сигнали на основі заданого алгоритму цифрового перетворення, називають *цифровими фільтрами*.

По суті цифровий фільтр є спеціалізованою ЕОМ або мікропроцесором. Більш того, на цифровій основі можуть бути об'єднані в єдиній системі сигнали передачі даних з сигналами передачі мови і телебачення.

Можливість зведення всіх видів інформації, що передається, до цифрової форми дозволяє здійснити інтеграцію систем передачі і систем комутації. Простота взаємодії цифрового каналу з ЕОМ дозволяє істотно розширити область використання обчислювальної техніки при побудові апаратури зв'язку й автоматизації управління мережами зв'язку.

Зазначені й інші переваги цифрових систем передачі безперервних повідомлень дозволяють припустити, що найближчими роками ці системи знайдуть найбільше застосування, ніж безперервні (аналогові) системи зв'язку. Перейдемо до розгляду структурної схеми цифрового каналу передачі безперервних повідомлень (рис. 8.11).

На відміну від безперервного каналу передачі, у складі цифрового каналу передбачені пристрої для перетворення безперервного повідомлення на цифрову форму – аналого-цифровий перетворювач (АЦП) на боці, що передає, і пристрої перетворення цифрового сигналу на безперервний – цифро-аналоговий перетворювач (ЦАП) на приймальному боці.

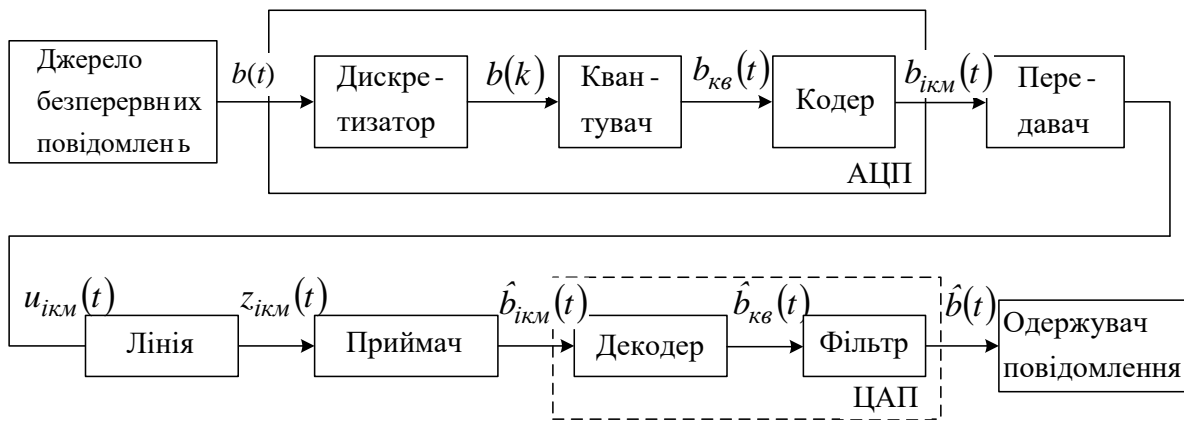


Рис. 8.11. Структурна схема системи цифрової передачі безперервних повідомлень

Перетворення аналог-цифра складається з трьох операцій (рис. 8.12): спочатку безперервне повідомлення піддається дискретизації за часом через інтервали Δ (рис. 8.12, а); отримані відліки миттєвих значень $b(k\Delta)$ квантуються (рис. 8.12, б); нарешті, отримана послідовність квантованих значень $b_{KB}(k\Delta)$ передаваного повідомлення подаються за допомогою кодування як послідовність m -кових кодових комбінацій (рис. 8.12, в).

Таке перетворення називається імпульсно-ковою модуляцією (ІКМ). Найчастіше кодування тут зводиться до запису номера рівня у двійковій системі числення. Надалі розглядатимемо цифрові системи, в яких безперервне повідомлення перетворене на послідовність кодових комбінацій, складених з двійкових символів.

Отриманий з виходу АЦП сигнал ІКМ поступає або безпосередньо в лінію зв'язку, або на вхід передавача (модулятора), де послідовність двійкових імпульсів перетвориться на радіоімпульси.

На приймальному боці лінії зв'язку послідовність імпульсів після демодуляції і регенерації в приймачі поступає на цифро-аналоговий перетворювач ЦАП, призначення якого полягає в зворотному перетворенні (відновленні) безперервного повідомлення за прийнятою послідовністю кодових комбінацій.

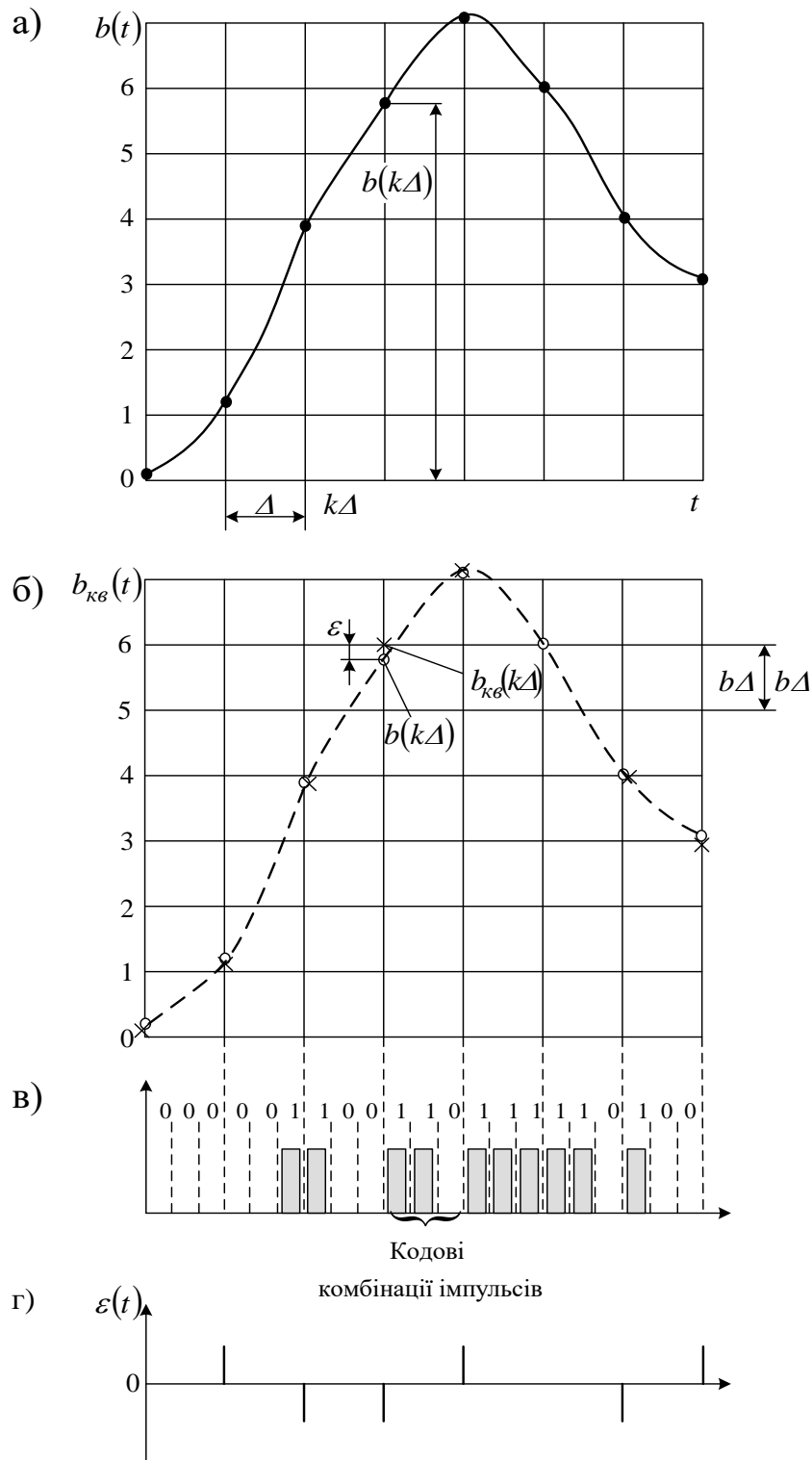


Рис. 8.12. Перетворення безперервного повідомлення на послідовність двійкових імпульсів

До складу ЦАП входять декодуєчий пристрій, призначений для перетворення кодових комбінацій на квантовану

послідовність відліків, і згладжуючий фільтр, поновлюючий безперервне повідомлення за квантованими значеннями.

Перетворення безперервних повідомлень на цифрову форму в системах ІКМ супроводжується *округленням* миттєвих значень до найближчих дозволених рівнів квантування. Похибка $\varepsilon(t)$ перетворення сигналу (рис. 8.12, *в*), що виникає при цьому, є неусувною, але контрольованою (оскільки не перевищує половини кроку квантування) (рис. 8.12, *б*).

Вибравши малий крок квантування, можна забезпечити еквівалентність за заданим ε -критерієм початкового і квантованого повідомлень. Похибку (помилку) квантування, що є різницею між початковим повідомленням і повідомленням, відновленим за квантованими відліками, називають *шумом квантування*.

8.10. Завадостійкість імпульсно-кодової модуляції

Однією з причин, що призводить до відмінності прийнятого повідомлення від переданого в системі з ІКМ, є шум квантування, іншою - завади в каналі, які накладаються на сигнали, що передаються, кодових комбінацій і можуть викликати помилки. Помилки в символах (за відсутності надмірності) призводять до помилкового декодування всієї кодової комбінації.

У результаті помилкового декодування символу дійсно передане дискретне значення повідомлення замінюється іншим можливим (не обов'язково найближчим); похибка залежить від того, які з символів кодової комбінації прийняті з помилкою. Назвемо цю складову шуму *шумом помилкових імпульсів*. Таким чином, при оцінці завадостійкості необхідно враховувати сумарний шум як за рахунок квантування, так і за рахунок помилкових імпульсів при декодуванні.

Шум квантування не пов'язаний з завадами в каналі і цілком визначається вибором числа рівнів квантування. Його можна зробити як завгодно малим, збільшуючи число рівнів. При цьому доведеться збільшувати число кодових символів, що доводяться на кожен відлік, а отже, скорочувати тривалість символу і

розширювати спектр сигналу в каналі. Таким чином, так само, як і при завадостійких аналогових видах модуляції, зниження цього шуму досягається за рахунок розширення спектра сигналу.

Шум помилкових імпульсів є аномальним. Він повністю визначається завадами в каналі і видом модуляції несучої. При розширенні спектра сигналу потужність аномального шуму, як правило, зростає.

8.10.1. Потужність шуму квантування

Вважатимемо, що шум квантування є стаціонарним випадковим процесом $E(t)$ з незалежними значеннями окремих відліків $\varepsilon(k) = b_{KB}(k\Delta) - b(k\Delta)$. Якщо як квантоване (округлене) значення $b(k\Delta)$ приймається найближчий дискретний рівень, то шум квантування $E(k)$ при рівномірному квантуванні з кроком Δb знаходиться в межах $\left[-\frac{\Delta b}{2}, \frac{\Delta b}{2}\right]$. При великому числі рівнів квантування L можна прийняти, що шум квантування має рівномірний розподіл щільності імовірності

$$w(\varepsilon) = \begin{cases} \frac{1}{\Delta b}, & |\varepsilon| < \frac{\Delta b}{2}, \\ 0, & |\varepsilon| > \frac{\Delta b}{2}. \end{cases} \quad (8.97)$$

Тоді $\overline{E(t)} = 0$, а дисперсія шуму квантування (середня потужність) визначається з виразу

$$\overline{E^2(t)} = P_\varepsilon = \int_{-\frac{\Delta b}{2}}^{\frac{\Delta b}{2}} \varepsilon^2 w(\varepsilon) d\varepsilon = \frac{(\Delta b)^2}{12}. \quad (8.98)$$

Відношення середніх потужностей повідомлення і шуму квантування запишеться як

$$P_b/P_\varepsilon = \overline{B^2(t)}/\overline{E^2(t)} = 12 \frac{\overline{B^2(t)}}{(\Delta b)^2}.$$

Виразимо Δb через число рівнів квантування L , вважаючи, що повідомлення нормоване, тобто $-1 < B(t) < +1$. Тоді маємо

$$\Delta b = \frac{B_{max} - B_{min}}{L-1} = \frac{2}{L-1}. \quad (8.99)$$

З іншого боку, згідно з виразом (8.30) $\overline{B^2(t)} = P_b = \frac{1}{\Pi^2}$, де Π - пік-чинник повідомлення. Тоді маємо

$$\frac{P_b}{P_\varepsilon} = \frac{12}{\Pi^2(\Delta b)^2} = \frac{3(L-1)^2}{\Pi^2} = \frac{3(2^n - 1)^2}{\Pi^2}, \quad (8.100)$$

де n - число символів на відлік (розрядність коду) при безнадмірностному двійковому коді.

З виразу (8.100) випливає, що правильність квантованого повідомлення залежить від числа рівнів квантування. Вибираючи його досить великим, можна зменшити відносне значення шуму квантування до будь-якого допустимого значення. У табл. 8.1 наведена залежність відношення P_b/P_ε від числа рівнів квантування при рівномірному розподілі повідомлення $w(b) = 0,5(-1 < b < 1)$, коли $\Pi = \sqrt{3}$.

З табл. 8.1 випливає, що додавання кожного двійкового символу в кодовій комбінації (збільшення розрядності коду) покращує відношення P_b/P_ε приблизно на 6 дБ. З іншого боку, збільшення розрядності вимагає підвищення швидкодії багаторозрядних кодуєчих пристроїв, а також відповідного розширення смуги частот каналу передачі.

Найчастіше розподіл повідомлення не є рівномірним, і в отриманні тут цифри необхідно внести поправку. Для цього дані таблиці слід зменшити за модулем на $20 \lg(\Pi/\sqrt{3})$ дБ. При $\Pi=3$

(телефонні повідомлення) це складає 4,8 дБ, а при $L=10$ (симфонічна музика) - 15,2 дБ. Важливою особливістю шуму квантування, що відрізняє його від адитивних шумів, є те, що він виникає одночасно з появою повідомлення $B(t)$. З технічної точки зору шум, квантування можна трактувати як різновид нелінійних спотворень, що виникають у процесі квантування. Шум квантування не змінюється при ретрансляції сигналів, тобто не накопичується. Його безпосереднє вимірювання складне, і тому для кількісної оцінки використовують розрахункові прийоми, один з яких розглянуто вище.

Таблиця 8.1

Залежність рівнів кодування від n і P_b/P_c

Число рівнів квантування L	8	16	32	64	128	256	512	1024	2048
Число символів у кодовій комбінації n	3	4	5	6	7	8	9	10	11
Відносна потужність шуму квантування, дБ, $-20\lg(L-1)$	-16,9	-23,5	-29,8	-36,0	-42,1	-48,1	-54,2	-60,2	-66,2

Дію шуму квантування на повідомлення, що приймаються, можна помітно зменшити, застосовуючи нерівномірне квантування, при якому великі рівні повідомлення квантуються з великим кроком, а низькі рівні - з меншим (рис. 8.13).

Очевидно, що шум квантування при цьому корельований з повідомленням і має тим меншу миттєву потужність, чим менше рівень повідомлення.

Це дозволяє при тому самому числі рівнів краще розрізнити слабкі відрізки повідомлення. Якщо розподіл імовірності повідомлення такий, що великі рівні зустрічаються значно рідше, ніж малі (що має місце, наприклад, при передачі мови), то нерівномірне квантування зводиться до того, що часто виникаючі

значення $B(t)$ передаються з меншою помилкою квантування, а рідко виникаючі значення $B(t)$ - з більшою помилкою квантування. У результаті усереднювання за всіма значеннями дисперсія помилки квантування буде зменшена.

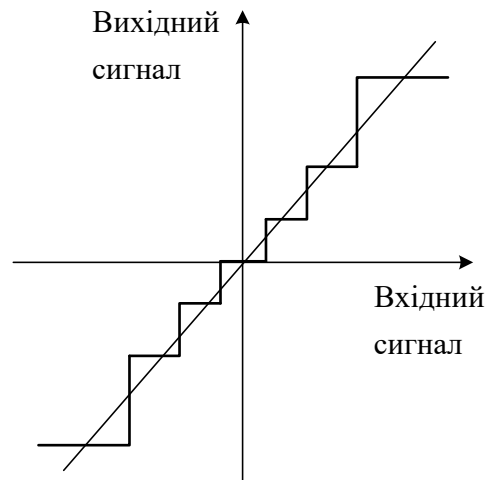


Рис. 8.13. Характеристика квантувача з нерівномірним кроком

Одним з поширених способів нерівномірного квантування є квантування з *компандуванням* сигналу. Компандерна система є комплексом з двох нелінійних перетворювачів - компресора та експандера - з взаємно зворотними характеристиками (рис. 8.14).

Динамічний діапазон вхідного сигналу "стискається" за допомогою *компресора* на боці, що передає, потім сигнал рівномірно квантується. Комбінація цих операцій еквівалентна нерівномірному квантуванню, причому закон зміни кроку квантування визначається нелінійною характеристикою компресора.

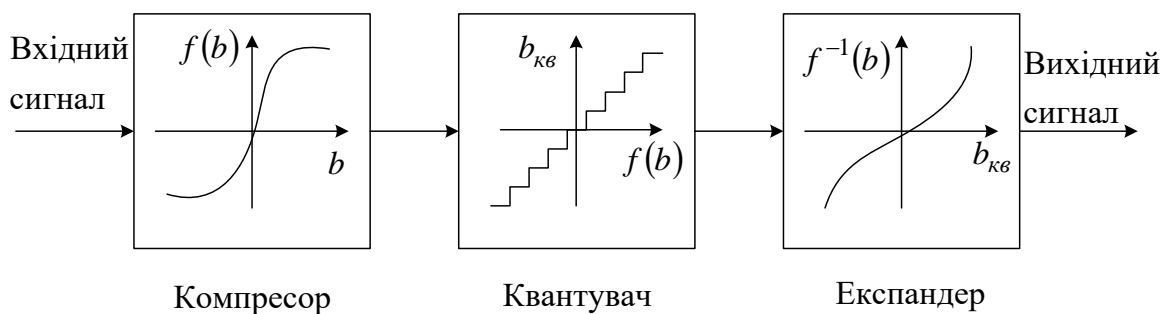


Рис. 8.14. Схема, що пояснює принцип компандування

Після кодування і передачі по лінії зв'язку на приймальному боці проводиться зворотне перетворення, відновлюються відліки з рівномірним кроком квантування. Потім вони піддаються *експандуванню*, при якому відновлюється початковий динамічний діапазон. В наш час найчастіше застосовується компандування в цифровій формі, тобто безпосереднє нерівномірне квантування.

8.10.2. Шум помилкових імпульсів при декодуванні

Позначимо імовірність помилкового прийому одного символу кодової комбінації через p . Ця імовірність залежить від виду модуляції. Припускаючи помилки при прийомі символів незалежними (канал без пам'яті), запишемо вираз для $P(t)$ - імовірність того, що кратність поєднання помилок складе t - як

$$P(t) = C_n^t p^t (1-p)^{n-t}. \quad (8.101)$$

Імовірність того, що кодова комбінація приймається хоч би з однією помилкою при $np \ll 1$,

$$\left[1 - (1-p)^n \right] \approx np. \quad (8.102)$$

При декодуванні кожен символ кодової комбінації залежно від займаного ним місця дає певний "внесок" у декодоване повідомлення. Якщо використовується двійковий код, то помилка в молодшому розряді кодової комбінації викликає похибку у вихідному повідомленні, що дорівнює кроку квантування Δb , помилка в другому символі призводить до появи у вихідному повідомленні похибки $2\Delta b$ і т. д. При цьому середня потужність шуму, обумовленого дією помилкових імпульсів, знаходяться з виразу

$$P_{ni} = \left[1 - (1-p)^n \right] \frac{(\Delta b)^2}{n} \sum_{i=1}^n 2^{2(i-1)} \approx p(\Delta b)^2 \sum_{i=1}^n 2^{2(i-1)}. \quad (8.103)$$

Сума визначає зростаючу геометричну прогресію з показником $q = 2^2 = 4$ і першим членом $a = 1$. Оскільки значення цієї суми дорівнює $\frac{a(q^n - 1)}{q - 1} = \frac{4^n - 1}{3}$, то замість виразу (8.103)

можна написати

$$P_{ni} = p(\Delta b)^2 (4^n - 1)/3. \quad (8.104)$$

При фіксованому значенні $n = \log L$ шум помилкових імпульсів залежить тільки від імовірності помилок p , яка у свою чергу визначається відношенням потужностей сигналу до завади в каналі і видом модуляції.

Як зазначалося, на відміну від шуму квантування шум помилкових імпульсів накопичується при ретрансляції. Проте в правильно спроектованих системах з ІКМ потужність сигналу перевищує порогову, при якій аномальним шумом помилкових імпульсів в порівнянні з шумом квантування можна нехтувати. За цієї умови правильність прийому практично визначається шумом квантування і може бути як завгодно великою, якщо число рівнів достатньо велике.

У системах з ІКМ так само, як і в інших завадостійких (або широкосмугових) системах передачі безперервних повідомлень, має місце поріг завадостійкості, тобто правильність прийому різко погіршується, якщо потужність сигналу впаде нижче порогової. Зі сказаного вище випливає, що порогова потужність збільшується зі зростанням числа ретрансляторів, втім дуже поволі.

При цьому порогова потужність збільшується і із зростанням числа рівнів квантування. Основна причина цього полягає в тому, що чим більше число рівнів, тим більше кодових символів має припадати на один відлік і, отже, тим менше тривалість передачі одного символу.

Оскільки імовірність помилки визначається енергією елемента сигналу, то при скороченні його тривалості доводиться збільшувати його потужність. Втім збільшення порогової потужності також невелике порівняно з відповідним зменшенням

шуму квантування. Так, з переходом від 128 рівнів квантування до 256 шум квантування зменшується на 6 дБ (див. табл. 8.1).

При цьому замість семи символів у кодовій комбінації доводиться передавати вісім, так що тривалість імпульсу зменшиться у $8/7 \approx 1,14$ рази. Для того щоб зберегти колишню імовірність помилки, потрібно збільшити потужність сигналу у 1,14 рази, тобто лише на 0,6 дБ.

Слід зазначити, що слабкий шум помилкових імпульсів, який має місце при роботі над порогом завадостійкості, сприймається (у телефонних системах) як більш-менш рідкісні окремі клацання.

Якщо потужність сигналу впаде і виявиться нижчим за поріг, ці клацання стануть частими і зіллються в суцільний шумовий фон. Аналогічна картина має місце і для аномального шуму в аналогових системах (наприклад ЧМ).

Висока завадостійкість ІКМ систем досягається за рахунок розширення спектра ІКМ сигналу в порівнянні зі спектром початкового повідомлення. Знайдемо спочатку мінімальну ширину спектра ІКМ сигналу при основі коду $m = 2$. Якщо ширина спектра початкового повідомлення дорівнює F_c , то мінімальна частота дискретизації відповідно до теореми Котельникова дорівнює $2F_c$. Кожен відлік після квантування може приймати $L = 2B_{max}/\Delta b + 1$ можливих дискретних значень і замінюється при кодуванні комбінацією з $n = \log L$ двійкових імпульсів. Отже, тривалість кожного імпульсу не може бути більше, ніж $\tau_u = 1/[2F_c \log L]$, а необхідна смуга частот визначається як $F \approx 1/(2\tau_u) = F_c \log L$. При двосмуговій АМ сигнал ІКМ-АМ займатиме смугу частот

$$F = 2F_c \log L. \quad (8.105)$$

Оскільки при ІКМ правильність передачі визначається числом рівнів квантування, то збільшення правильності супроводжується розширенням спектра ІКМ сигналу за логарифмічним законом. Так, збільшення L в 2 рази призводить до збільшення ширини спектру сигналу в

$\log_2 2L / \log_2 L = (1 + 1/\log_2 L)$ разу. Ширина спектра ІКМ сигналу залежить від основи коду m . При $m = 2$ ширина спектра ІКМ сигналу найбільша; при збільшенні m ширина спектра зменшується.

З викладеного випливає, що в системі передачі з ІКМ, як і в завадостійких аналогових системах модуляції, проводиться «обмін» потужності сигналу на смугу частот. Проте в системі з ІКМ цей обмін здійснюється значно ефективніше, ніж у системах з аналоговою модуляцією.

Дійсно, у завадостійких системах модуляції, таких як ЧМ, ФМ, ВІМ, відношення потужності повідомлення до потужності шуму на виході зростає пропорційно квадрату ширини спектра сигналу (якщо потужність сигналу вище за порогову). У системі з ІКМ має місце значно швидше експоненціальне зростання цього відношення.

Дійсно, ширина спектра пропорційна числу n символів в кодовій комбінації, тоді як потужність шуму квантування зменшується відповідно до виразу (8.100) майже пропорційно 2^{2n} . Іншими словами, пропорційно ширині спектра зростає виграш системи, виражений у децибелах.

Як було показано в п. 8.4, такий самий характер залежності правильності від ширини спектру повинен бути в ідеальній системі модуляції, так що в цьому відношенні ІКМ поводить себе як ідеальна система. Докладніший аналіз дозволяє дійти висновку, що при однаковій ширині спектра виграш в ІКМ приблизно на 8 дБ менше, ніж у теоретично ідеальній системі.

У наш час не існує систем модуляції, ближчих до ідеальної, якщо спектр повідомлення, що передається, рівномірний. Тому система з ІКМ широко використовується в тих випадках, коли високу правильність необхідно забезпечити з мінімальною витратою потужності передавача, наприклад у супутникових системах.

8.11. Кодування з передбаченням

У ряді важливих випадків, наприклад при передачі мови, телевізійних зображень, даних телеметрії, між відліками повідомлень, що передаються, є статистичні, зокрема

кореляційні, зв'язки. Наявність таких взаємозв'язків дозволяє підвищити ефективність систем передачі інформації.

Розглянемо один з поширених способів ефективної передачі безперервних повідомлень - спосіб передачі з передбаченням (рис. 8.15). Послідовність корельованих відліків $b(k\Delta) = b(k)$ початкового сигналу подають на один з входів віднімаючого пристрою, а на його інший вхід поступає сигнал передбачення $b_{np}(k)$, сформований з попередніх відліків. Отриманий таким чином сигнал помилки передбачення $\varepsilon(k)$ поступає в тракт передачі.

Оскільки в сигналі помилки саме і містяться нові відомості, що є різницею між дійсним і передбаченим значенням, то такий спосіб передачі називається *передачею з передбаченням*. Для спрощення вважаємо, що завади в каналі помилок не створюють. На приймальному кінці є такий же пристрій передбачення, як і на тому, що передає.

Оскільки він оперує тими самими попередніми відліками, передбачене ним значення нового відліку $\hat{b}_{i\partial}(k)$ буде таким, як і на передавачі. Додавши до нього прийняте значення помилки передбачення $\hat{\varepsilon}(k)$, можна відновити дійсний відлік згідно з формулою

$$\hat{b}(k) = \hat{b}_{i\partial}(k) + \hat{\varepsilon}(k) = b_{i\partial}(k) + \varepsilon(k) = b(k). \quad (8.106)$$

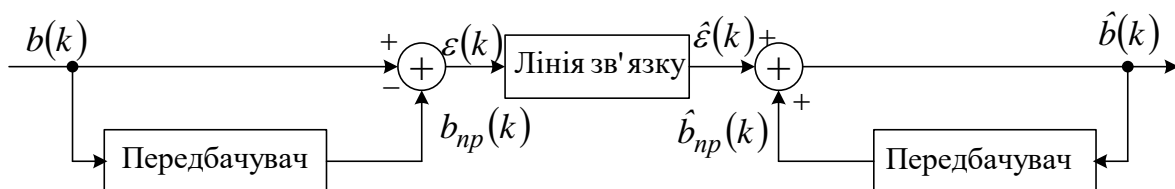


Рис. 8.15. Структурна схема системи зв'язку з передбаченням

Очевидно, що чим сильніше кореляційні зв'язки, тим точніше можна сформулювати сигнал передбачення і, отже, тим менша енергія буде потрібна для передачі сигналу помилки в

порівнянні з початковим сигналом. При сильних кореляційних зв'язках, а також при марковських повідомленнях значення $b_{np}(k)$ може бути передбачене за одним попереднім відліком:

$$b_{np}(k) = f[b(k-1)]. \quad (8.107)$$

Зазвичай можна отримати достатньо малу помилку передбачення

$$b_{np}(k) = b(k-1), \quad (8.108)$$

звідки

$$\varepsilon(k) = b(k) - b(k-1). \quad (8.109)$$

У цьому випадку для середньої енергії сигналу помилки за ансамблем визначається із виразу

$$\begin{aligned} E_\varepsilon &= \overline{[B(k) - B(k-1)]^2} = \overline{B^2(k)} + \overline{B^2(k-1)} - 2\overline{B(k)B(k-1)} = \\ &= 2\overline{B^2(k)} - 2\overline{B(k)B(k-1)} = 2E_b(1 - R), \end{aligned} \quad (8.110)$$

де R - коефіцієнт кореляції між відліками, $R = \overline{B(k)B(k-1)}/E_b$.

З виразу (8.10) випливає, що при $R > 0,5$ енергія сигналу помилки E_ε менше за енергію початкового сигналу E_b . При $R = 0,9$ маємо $E_\varepsilon = 0,2E_b$.

При цифрових системах передачі відліки сигналу помилки підлягають звичайним операціям квантування і кодування. Результатом такого перетворення на боці, що передає, є імпульсно-кодове подання сигналу помилки або сигналу диференційної імпульсно-кової модуляції (ДІКМ).

У наш час відоме велике число варіантів технічного здійснення кодування з передбаченням. Основна їх особливість зводиться до відмінності операцій формування сигналу помилки: в одних системах сигнал помилки формується в аналоговій формі, а потім кодується, в інших - спочатку кодується

початковий аналоговий сигнал, а потім формується сигнал помилки.

У системах з ДКМ зазвичай застосовують нерівномірне квантування сигналу помилки, оскільки найбільш імовірні малі помилки. Оскільки потужність шуму квантування складає певну частку потужності квантованого процесу, а потужність помилки передбачення, як правило, істотно менше від потужності повідомлення, шум квантування при ДКМ менший, ніж при звичайній ІКМ при тому самому числі рівнів.

Як показує аналіз, система ДКМ забезпечує однакову з ІКМ якість передачі при меншому числі символів у кодовій комбінації (на один-два і більше символів залежно від ступеня кореляції відліків і алгоритму передбачення).

Шум помилкових імпульсів також має місце при ДКМ, причому він погіршує правильність прийому більшою мірою, ніж при звичайній ІКМ. Помилковий прийом кодової комбінації веде до помилкового прийому не тільки одного відліку повідомлення, але і ряду подальших відліків, оскільки передбачені значення на приймальному боці відрізнятимуться від передбачених на передавачі.

Тому допустима імовірність помилок при ДКМ менша, ніж при ІКМ, і, отже, порогова потужність сигналу вища. Існують різні способи зменшення впливу «розмноження» помилок.

Кореляція між відліками зростає з скороченням інтервалу між ними. Тому при великій частоті дискретизації число рівнів квантування сигналу помилки можна зменшити до двох і перейти до однорозрядних систем. Такий спосіб кодування називають *дельта-модуляцією* (ДМ).

Квантований сигнал помилки при ДМ визначається з формули

$$\varepsilon_{KB}(k) = \gamma(k)\Delta b, \text{ где } \gamma(k) = \begin{cases} +1, & \varepsilon(k) \geq 0, \\ -1, & \varepsilon(k) < 0. \end{cases} \quad (8.111)$$

По суті сигнал на виході дельта-модулятора (рис. 8.16) містить лише відомості про полярність (знак) сигналу помилки. На приймальному боці інтегратор додає або віднімає величину

кроку дискретизації, забезпечуючи таким чином зменшення похибки між початковими значеннями відліків і відновленими значеннями.

Принцип формування сигналу ДМ показаний на рис. 8.17. Відліки передаваного повідомлення порівнюються з передбаченим квантованим відліком, отриманим у результаті підсумовування всіх попередніх квантованих сигналів помилки:

$$b_{ПРКВ}(k-1) = \sum_{i=0}^{n-1} \varepsilon_{KB}(i) = \Delta b \sum_{i=0}^{n-1} \gamma(i).$$

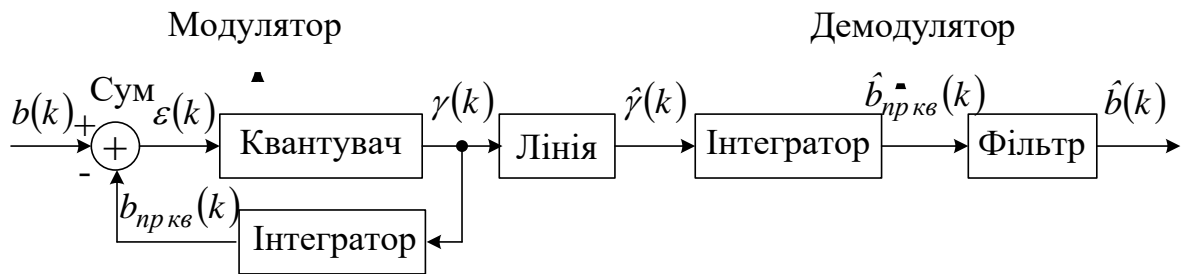


Рис.8.16. Структурна схема системи з дельта-модуляцією

Якщо $b(k) > b_{ПРКВ}(k-1)$ (рис. 8.17, а), квантувач формує значення $\gamma(k)=1$, інакше $\gamma(k)=-1$. Отримані символи γ передаються по тракту зв'язку (рис. 8.17, б). Одночасно імпульси $b\Delta\gamma(i)$ поступають на суматор (інтегратор) для формування квантованого відліку $b_{ПРКВ}(i)$, який, як зазначалося вище, порівнюється з черговим відліком повідомлення.

На виході суматора (Сум) квантований сигнал $b_{ПРКВ}(t)$ має вид ступінчастої функції (рис. 8.17, а). Кожен імпульс +1 збільшує, а кожен імпульс -1 зменшує ступінчасту функцію на один крок квантування, тобто при ДМ сусідні значення ступінчастої функції розрізняються обов'язково на один крок квантування.

Операцію декодування сигналу ДМ на приймальному боці виконує такий же суматор (інтегратор), як і на боці, що передає, на виході якого (за відсутності помилкових імпульсів) виходить

ступінчаста напруга $b_{ПРКВ}(t)$. Після його згладжування за допомогою ФНЧ одержимо функцію $\hat{b}(t)$, достатньо близьку до $b(t)$.

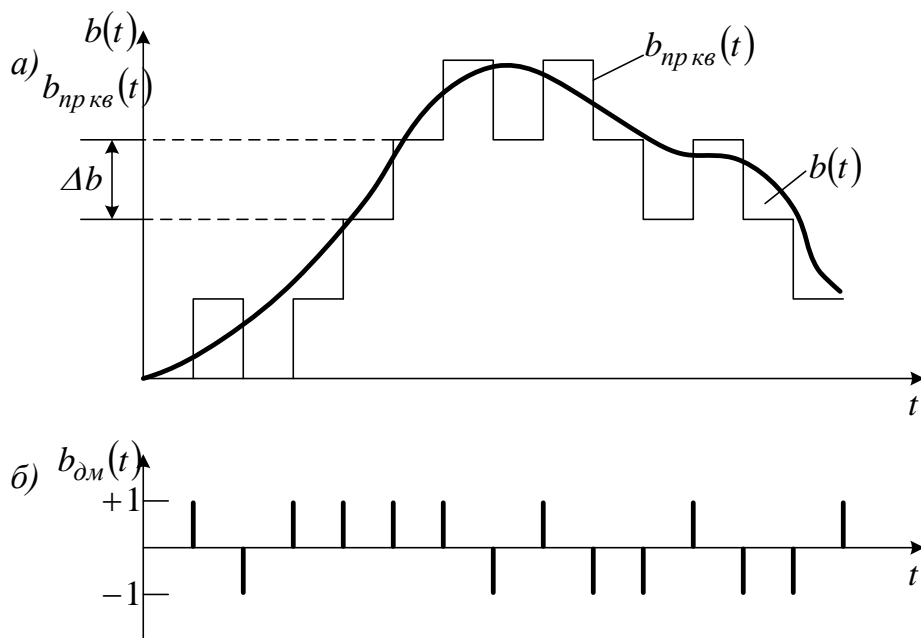


Рис. 8.17. Графіки, що пояснюють принцип формування сигналів дельта-модуляції

Різницею $\hat{b}(t) - b(t)$ є шум квантування. Його значення тим менше, чим вище частота дискретизації і чим менше крок Δb . Проте крок квантування Δb не можна вибирати дуже малим. Інакше виникають додаткові спотворення, звані *перевантаженням за нахилом*, викликані тим, що ступінчаста функція не встигає стежити за швидкими змінами повідомлення $b(t)$ (рис. 8.18).

Якщо прийняти, що максимальне значення крутизни зміни повідомлення складає $|b'(t)|_{\max}$, то для неспотвореної передачі необхідно виконати умову $|b'(t)|_{\max} \Delta \leq \Delta b$.

Якщо практично максимальне значення повідомлення за модулем не перевищує $|b(t)|_{\max}$, а число різних рівнів квантування, що визначає шум квантування, дорівнює L , то можна записати

$$\Delta \leq \frac{2}{L-1} \frac{|b(t)|_{max}}{|b'(t)|_{max}}. \quad (8.112)$$

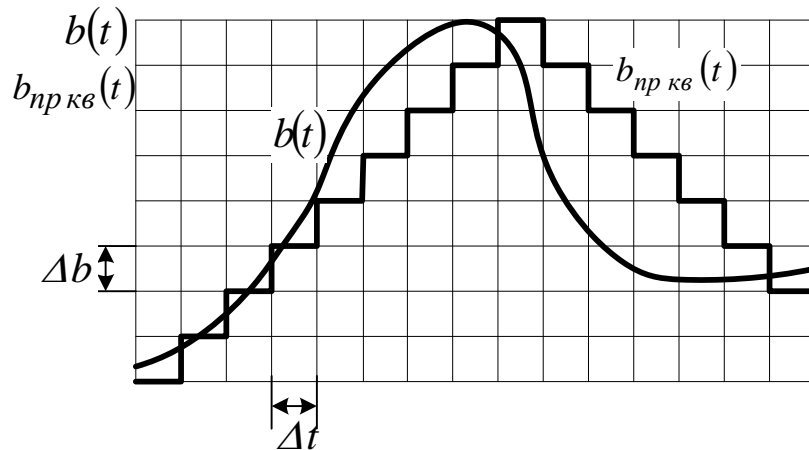


Рис. 8.18. Графік, що пояснює виникнення спотворень через перевантаження за нахилом при дельта-модуляції

У порівнянні з ІКМ і ДІКМ сигнали ДМ мають значно вищу частоту проходження відліків. Проте на кожен відлік при ДМ передається один імпульс, а при ІКМ - декілька, залежно від числа рівнів.

Як показує аналіз, при однаковій правильності передачі частота проходження імпульсів при ІКМ і ДМ також приблизно однакова. Тому обидві ці системи займають приблизно однакову смугу частот. Істотною перевагою систем передачі з ДМ є порівняльна простота кодуєчих і декодуєчих пристроїв.

Крім того, шум помилкових імпульсів при ДМ менший, ніж при ІКМ (при тій же імовірності помилки в каналі), оскільки при ДМ кожна помилка змінює рівень сигналу тільки на $\pm 2\Delta b$. До недоліків ДМ можна віднести явище розмноження помилок, що створюються помилковими імпульсами (таке, як при ДІКМ), а також можливість перевантаження за нахилом.

Зазначимо, що останніми роками успішно розробляються численні різновиди ДІКМ і ДМ, зокрема адаптивні системи, в яких для зменшення шумів квантування використовують змінний крок квантування залежно від поточних статистичних характеристик передаваного повідомлення.

Як приклад на рис. 8.19, *а* наведена структурна схема кодера адаптивної ДМ. Характерною особливістю є наявність у ланцюзі зворотного зв'язку вирішального пристрою, управляючого величиною кроку квантування Δb .

Якщо знак приростів $\gamma(k)$ залишається незмінним протягом трьох-чотирьох інтервалів дискретизації, то це означає наявність перевантаження (рис. 8.19, *б*). Вирішальний пристрій подвоює амплітуду імпульсів, що поступають на вхід інтегратора 1.

Якщо і в цьому випадку знак приросту (сигналу помилки) не зміниться, то розмір кроку Δb знову подвоюється і т. д. При зміні знаку приросту розмір кроку квантування зменшується. Неважко зрозуміти, що до складу пристрою управління розміром кроку квантування повинні входити аналізатор щільності одиниць і імпульсний підсилювач з керованим коефіцієнтом посилення. На виході інтегратора 2 при зміні коефіцієнта посилення в залежності від щільності одиниць формуватиметься ступінчаста напруга з кроком квантування, що адаптивно змінюється (рис. 8.19, *б*).

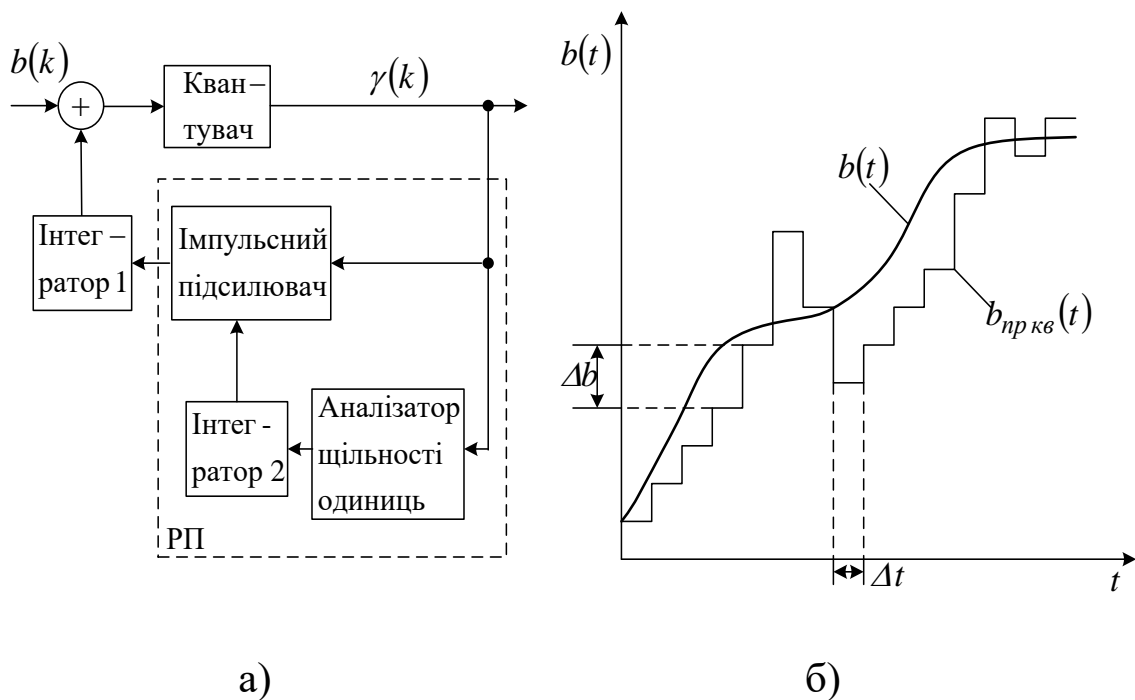


Рис. 8.19. Структурна схема кодера адаптивної ДМ (а), зміна кроку квантування при адаптивній ДМ (б)

Висновки до глави 8

1. Оптимальну оцінку безперервного параметра часто знаходять за критерієм мінімуму середньоквадратичної помилки або максимуму апостеріорного розподілу. Часто ці критерії призводять до оцінки за алгоритмом максимальної правдоподібності.

2. У системах передачі безперервних повідомлень для обробки сигналів широко використовується фільтрація (у загальному випадку нелінійна) з метою отримання якнайкращої (за заданим критерієм) оцінки повідомлень, що передаються.

3. Якість оцінки повідомлень зазвичай визначається середньоквадратичною помилкою або відношенням потужності сигналу до завади на виході приймача $\rho_{\text{вих}}$. При порівняльному аналізі зручними показниками є виграш g і узагальнений виграш g' системи.

4. Квазіоптимальним приймачем (демодулятором) безперервних повідомлень є стежачий кореляційний пристрій (нелінійний фільтр), який може бути реалізований за допомогою фільтра зі змінними параметрами.

5. У системах з лінійними видами модуляції (АМ, ОМ, БМ) оптимальним демодулятором є синхронний детектор з оптимальним лінійним фільтром.

6. Завадостійкість систем передачі безперервних повідомлень істотно залежить від способу модуляції. Так, при ЧМ узагальнений виграш g' пропорційний квадрату індексу модуляції M і може бути набагато більше одиниці. При ОМ і БМ цей виграш дорівнює одиниці, а при АМ – менше одиниці.

7. Ширококутові системи модуляції забезпечують високу завадостійкість лише при відношенні сигналу до шуму на вході приймача $\rho_{\text{вх}}$, більшого за деяке граничне (порогове) значення $\rho_{\text{пр}}$. При $\rho_{\text{вх}} < \rho_{\text{пр}}$ перешкодостійкість цих систем різко знижується.

8. Для стаціонарного сигналу і шуму фільтр Колмогорова-Вінера є еквівалентним фільтру Калмана. Проте реалізація фільтра Калмана за обчислювальною структурою зручніша та простіша.

9. Цифрові методи передачі безперервних повідомлень забезпечують більш високі характеристики якості в порівнянні з аналоговими методами.

10. Серед цифрових систем передачі безперервних повідомлень (мова, музика, телебачення і т. п.) найбільшого поширення набули ІКМ, ДІКМ, ДМ, адаптивні системи ДІКМ і ДМ.

11. Якість цифрових систем передачі безперервних повідомлень характеризується як середньоквадратичною помилкою квантування, так і середньоквадратичною помилкою імпульсів.

Контрольні питання та задачі

8.1. Знайти оптимальну оцінку амплітуди сигналу γ , якщо сигнал, що приймається, поданий як $\gamma S(t, \Theta)$, де Θ – фаза сигналу випадкова і може вважатися рівномірно розподіленою на інтервалі $(-\pi, \pi)$.

8.2. Які умови задовольняє добра оцінка безперервного параметра?

8.3. Як пояснити, що оптимальний прийом безперервних повідомлень зводиться до сумісної оптимальної оцінки багатьох параметрів.

8.4. Чим відрізняється шум на виході оптимального приймача для прямих і непрямих систем модуляції?

8.5. Від чого залежить відношення сигнал-завада на вході і виході оптимального приймача для ідеальної системи передачі безперервних повідомлень і в разі передачі цих повідомлень за допомогою ІКМ?

8.6. У чому причина появи порогового ефекту в широкосмугових системах зв'язку?

8.7. Визначте порогове значення амплітуди ЧМ сигналу виходячи з умови, що лише з вірогідністю 0,001 вона перевищується амплітудою стаціонарного гаусівського шуму з дисперсією $\sigma^2 = 0,1$ Вт.

8.8. На вхід когерентного приймача поступає адитивна суміш $Z(t) = S(t) + N(t)$, де $N(t)$ – "білий" гаусівський шум із

спектральною щільністю N_0 , $S(t) = \gamma B(t) \sin \omega_0 t$ – сигнал АМ з пригніченою несучою, а $B(t)$ – випадковий модулюючий процес із СЩП $G_b(\omega) = 2\alpha / (\alpha^2 + \omega^2)$, $-\infty < \omega < \infty$. Опорне коливання $f(t) = \sqrt{2} \sin \omega_0 t$. Визначте передавальну функцію оптимального (за середньоквадратичним критерієм) ФНЧ $\dot{K}(f)$ і обчисліть величину середньоквадратичної помилки $\overline{E_{min}^2}$.

8.9. У чому суть кодування безперервних повідомлень з передбаченням і в чому особливості системи з ДМ?

8.10. У чому причина виникнення спотворень при ДМ через «перевантаження по нахилу»?

ГЛАВА 9

ПРИНЦИПИ БАГАТОКАНАЛЬНОГО ЗВ'ЯЗКУ

9.1. Багатоканальна система передачі повідомлень

Практика побудови сучасних телекомунікаційних систем і мереж показує, що найдорожчими ланками трактів передачі є лінії зв'язку (кабельні, волоконно-оптичні, стільникового мобільного радіозв'язку, радіорелейного й супутникового зв'язку й ін.). Оскільки економічно недоцільно використовувати дорогу лінію зв'язку для передачі інформації єдиної пари абонентів (від джерела до одержувача повідомлень і назад при дуплексному зв'язку), то виникає завдання побудови багатоканальних систем передачі, які забезпечуватимуть передачу великої кількості повідомлень різних джерел інформації зі спільної лінії зв'язку.

Зрозуміло, багатоканальна передача можлива в тих випадках, коли пропускна спроможність лінії C' не менша від сумарної продуктивності джерел інформації: $C' \geq \sum_{k=1}^N H'_k$, де

H'_k – продуктивність k -го джерела, а N – кількість каналів (незалежних джерел інформації). Багатоканальні системи так само, як й одноканальні, можуть бути аналоговими й цифровими. Для уніфікації аналогових багатоканальних систем за основний або стандартний канал приймають канал тональної частоти (канал ТЧ), який забезпечує передачу повідомлень з ефективно переданою смугою частот 300...3400 Гц, що відповідає основному спектру телефонного сигналу.

У цифрових системах передачі найбільшого поширення набули основні цифрові канали зі швидкістю 64 кбіт/с. Багатоканальні аналогові системи формуються об'єднанням каналів ТЧ у групи, як правило, кратні 12 каналам. Цифрові системи передачі (ЦСП), використовувані на мережах зв'язку, формуються відповідно до прийнятих ієрархічних структур.

Європейська плезіохронна (PDH) ієрархія ґрунтується на первинній ЦСП типу ІКМ-30 зі швидкістю передачі групового сигналу 2048 кбіт/с; в основу північноамериканської системи

покладена первинна система ІКМ-24 зі швидкістю групового сигналу 1544 кбіт/с [43].

Наприкінці 80-х рр. ХХ ст. була розроблена синхронна цифрова ієрархія, єдина для всіх країн світу (SDH). Відповідно до неї, використовуючи віртуальні контейнери, формується синхронний транспортний модуль STM-1 зі швидкістю передачі 155,52 Мбіт/с, на базі якого утворюються синхронні модулі вищих порядків STM-4, STM-16, STM-64, STM-256 [43].

На схемі рис. 9.1 зображено загальний принцип побудови багатоканальної системи передачі (БСП) повідомлень.

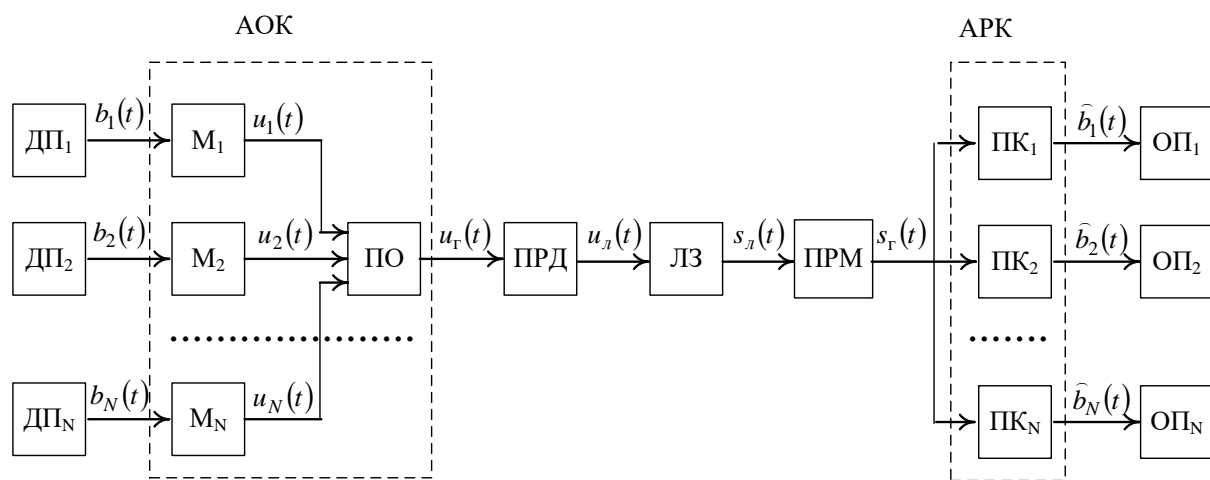


Рис. 9.1. Структурна схема багатоканальної системи передачі повідомлень: ДП – джерело повідомлень; М – модулятор; ПО – пристрій об’єднання; ПРД – передавач групового сигналу; ЛЗ – лінія зв’язку; ПРМ – приймач групового сигналу; ПК – приймач каналний; ОП – одержувач повідомлень

Тут первинні сигнали кожного джерела (Дж) $b_1(t)$, $b_2(t)$, ..., $b_i(t)$, ..., $b_N(t)$ за допомогою індивідуальних перетворювачів M_1 , M_2 , ..., M_i , ..., M_N перетворюються на відповідні каналні сигнали $u_1(t)$, $u_2(t)$, ..., $u_i(t)$, ..., $u_N(t)$. Сукупність каналних сигналів на виході пристрою об’єднання (ПО) утворює груповий сигнал $u_{\Gamma}(t) = \sum \{u_i(t)\}$. У випадку роздільної системи ущільнення це об’єднання зводиться до звичайного знаходження суми:

$$u_{\Gamma}(t) = \sum_{i=1}^N \{u_i(t)\}. \quad \text{З урахуванням частотного діапазону}$$

направляючої системи (лінії зв'язку) сигнал $u_{\Gamma}(t)$ за допомогою групового передавача (ПРД) перетворюється на лінійний сигнал $u_{\Delta}(t)$, котрий і надходить у лінію зв'язку (ЛЗ). Спочатку будемо вважати, що завада в каналі відсутня, а канал не вносить спотворення в сигнал. Тоді прийнятий лінійний сигнал буде дорівнювати $s_{\Delta}(t) = \gamma u_{\Delta}(t)$, де γ – коефіцієнт передачі каналу, його для простоти можна вважати рівним 1.

Вважаємо, що припізнення в каналі враховано системою синхронізації. На приймальному кінці ЛЗ лінійний сигнал s_{Δ} за допомогою групового приймача Π може бути знову перетворений на груповий сигнал $s_{\Gamma}(t) = \gamma u_{\Gamma}(t)$. Канальними або індивідуальними приймачами $\Pi_1, \Pi_2, \dots, \Pi_k, \dots, \Pi_N$ із групового сигналу виділяються відповідні канальні сигнали $s_i(t) = \gamma u_i(t)$, $i = \overline{1, N}$, які перетворюються на індивідуальні сигнали $\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_i(t), \dots, \hat{b}_N(t)$.

Канальні перетворювачі M_i разом із пристроєм об'єднання утворюють *апаратуру об'єднання (уцільнення) каналів* АОК.

Індивідуальні приймачі Π_i системи, виконуючи звичайну операцію перетворення канальних сигналів $s_i(t)$ на відповідні первинні сигнали $\hat{b}_i(t)$, повинні забезпечувати виділення сигналів $s_i(t)$ із групового сигналу $s_{\Gamma}(t)$ із припустимими спотвореннями. Апаратуру індивідуальних приймачів, що забезпечують цю операцію, називають *апаратурою розподілу каналів* (АРК).

Груповий передавач, лінія зв'язку та груповий приймач Π складають *груповий тракт передачі*, що разом з апаратурою об'єднання й розподілу каналів складає систему *багатоканального зв'язку*.

9.2. Основні положення лінійної теорії розподілу сигналів

Розглянемо основні властивості сигналів, придатних для незалежної передачі інформації в системах багатоканального зв'язку. Для того щоб розподілюючі пристрої могли розрізняти сигнали окремих каналів, повинні існувати певні ознаки, властиві тільки сигналу даного каналу. Такими ознаками в загальному випадку можуть бути параметри переносника, наприклад амплітуда, частота або фаза у випадку модуляції синусоїдального переносника, положення у часі, тривалість або форма сигналу при модуляції імпульсних переносників. Відповідно будуть розрізнятися й способи розподілу сигналів: частотний, часовий, фазовий, розподіл за формою сигналів та ін.

Нехай необхідно організувати одночасну роботу N індивідуальних каналів по спільному груповому тракту. Будемо вважати, що груповий тракт придатний для передачі сигналів кожного i -го каналу $u_i(t)$.

Припустимо, що сигнал i -го каналу запишеться як

$$u_i(t) = C_i \Psi_i(t), \quad (9.1)$$

де $\Psi_i(t)$ – функція переносника;

C_i – деякий коефіцієнт, що відображає передане повідомлення.

Для суми всіх каналних сигналів (групового сигналу) маємо

$$u_{\Gamma}(t) = \sum_{i=1}^N u_i(t) = \sum_{i=1}^N C_i \Psi_i(t). \quad (9.2)$$

Після перетворення групового сигналу на лінійний $u_{\Pi}(t)$ останній надходить у тракт передачі. На прийомному кінці $s_{\Pi}(t)$ знову перетворюється на груповий $s_{\Gamma}(t)$.

Для розподілу N каналних сигналів на прийомному боці необхідно буде мати відповідну кількість N розподілюючих

пристроїв, причому кожен k -й розподілюючий пристрій повинен виконувати операцію виділення k -го сигналу.

Дію прийомного пристрою k -го каналу будемо позначати оператором розподілу Π_k . В ідеальному випадку k -й прийомний пристрій повинен реагувати («відгукуватися») тільки на сигнал $s_k(t)$ і не реагувати на сигнали всіх інших каналів. Додатковою вимогою до оператора будемо вважати його лінійність, тобто він повинен задовольняти принцип незалежності дії (суперпозиції):

$$\Pi_k(s_i + s_k) = \Pi_k(s_i) + \Pi_k(s_k). \quad (9.3)$$

Операцію розподілу сигналів можна сформулювати у математичному вигляді. Позначимо через $\hat{s}_k(t)$ відгук, тобто результат впливу оператора Π_k прийомного пристрою k -го каналу на груповий сигнал $s_r(t)$: $\pi_k\{s_r(t)\} = \hat{s}_k(t)$.

На вході кожного k -го приймача діє сума сигналів усіх N каналів. Щоб прийомний пристрій Π_k був «чутливим» тільки до сигналів $s_k(t)$, необхідне виконання умови:

$$\hat{s}_k(t) = \Pi_k\left\{\sum_{i=1}^N s_i(t)\right\} = \sum_{i=1}^N \Pi_k\{s_i(t)\} = \begin{cases} s_k(t), & i = k, \\ 0, & i \neq k. \end{cases} \quad (9.4)$$

Інакше кажучи, для всіх i і k має виконуватися умова

$$\Pi_k\{s_i(t)\} = \begin{cases} \hat{s}_k(t), & i = k, \\ 0 & i \neq k. \end{cases} \quad (9.5)$$

Підставляючи вираз (9.1) в вираз (9.5), одержуємо

$$\Pi_k(t)\{C_i\Psi_i(t)\} = \begin{cases} C_k\Psi_k(t), & i = k, \\ 0 & i \neq k, \end{cases} \text{ й, отже, } \hat{s}_k(t) = C_k\Psi_k(t).$$

Отримані результати можуть бути узагальнені також для випадку, коли відгук розподілюючого пристрою на сигнал $s_k(t)$ буде мати іншу форму. Важливо, щоб величина відгуку була однозначно пов'язана з переданим сигналом.

В окремому випадку відгуком на сигнал $s_k(t)$ може бути просто деяке число γ_k , однозначно пов'язане з коефіцієнтом C_k :

$$\hat{s}_k(t) = \Pi_k \{s_\Gamma(t)\} = \Pi_k \left\{ \sum_{i=1}^N C_i \Psi_i(t) \right\} = \sum_{i=1}^N \Pi_k \{C_i \Psi_i(t)\} = \gamma_k \quad (9.6)$$

$$\text{або } \Pi_k \{C_i \Psi_i(t)\} = \begin{cases} \gamma_k, & i = k, \\ 0 & i \neq k. \end{cases}$$

Фізичне значення отриманих виразів (9.5) і (9.6) зводиться до того, що приймач Π_k має виборчі властивості стосовно сигналів $s_k(t)$. Оскільки дія приймачів Π_k у виразах (9.5) і (9.6) описується лінійним оператором Π_k , то відповідні пристрої розподілу реалізуються лінійними колами, а теорію розподілу називають лінійною.

У реальних умовах при розподілі сигналів виникають перехідні завади, які будуть обговорені нижче при аналізі конкретних прикладів розподілу сигналів.

9.3. Умова лінійного розподілу сигналів

Вплив лінійного оператора Π_k на функцію $s_\Gamma(t)$ можна подати як скалярний добуток

$$\Pi_k s_\Gamma(t) = \int_{-\infty}^{\infty} s_\Gamma(\tau) \eta_k(t, \tau) d\tau, \quad (9.7)$$

де $\eta_k(t, \tau)$ – деяка вагова функція, що відповідає операторові Π_k .

Уперше визначення виборчих властивостей приймача виразів (9.5) і (9.6) було запропоновано в 1935 р. Д.В. Агєєвим, він довів, що необхідною й достатньою умовою розподілу сигналів виразу (9.2) за допомогою лінійних пристроїв є умова лінійної незалежності, яка визначається тотожністю [48]

$$C_1 \Psi_1(t) + C_2 \Psi_2(t) + \dots + C_k \Psi_k(t) + \dots + C_N \Psi_N(t) \equiv 0 \quad (9.8)$$

і може виконуватися в тому єдиному випадку, коли всі коефіцієнти одночасно дорівнюють нулю.

Дійсно, відповідно до визначення виборчих властивостей виразів (9.5) і (9.6) сигнали, які розподіляються, і прийомні пристрої повинні задовольняти умову лінійного розподілу, тобто

$$\Pi_k \{C_i \Psi_i(t)\} = \gamma_{ik}, \quad i, k = 1, 2, \dots, N, \quad (9.9)$$

де γ_{ik} – відгук розподільного пристрою на сигнал $s_i(t)$, причому при $\gamma_{ik} = 0, i \neq k$ і $\gamma_{ik} \neq 0$.

Якщо застосувати оператор Π_k до обох частин тотожності (9.8), то, беручи до уваги вираз (9.9) і принцип суперпозиції, одержуємо

$$\Pi_k \left\{ \sum_{i=1}^N C_i \Psi_i(t) \right\} = \sum_{i=1}^N C_i \Pi_k \{ \Psi_i(t) \} = C_k \gamma_{kk} = 0. \quad (9.10)$$

Але значення γ_{kk} не дорівнюють тотожно нулю, отже, дорівнюють нулю всі C_k для $k = 1, 2, \dots, N$. Інакше кажучи, умова лінійного розподілу (9.5) буде виконуватися лише тоді, коли каналні сигнали лінійно незалежні.

Окремим випадком лінійно незалежних сигналів є ортогональні сигнали.

У загальному випадку необхідною й достатньою умовою лінійної незалежності ансамблю сигналів є відмінність від нуля визначника Грама; визначника, що поданий як матриця [42]:

$$\Gamma(\Psi_1, \Psi_2, \dots, \Psi_N) = \begin{vmatrix} (\Psi_1, \Psi_1) & (\Psi_1, \Psi_2) & \dots & (\Psi_1, \Psi_N) \\ (\Psi_2, \Psi_1) & (\Psi_2, \Psi_2) & \dots & (\Psi_2, \Psi_N) \\ \dots & \dots & \dots & \dots \\ (\Psi_N, \Psi_1) & (\Psi_N, \Psi_2) & \dots & (\Psi_N, \Psi_N) \end{vmatrix}, \quad (9.11)$$

де (Ψ_i, Ψ_k) – скалярний добуток сигналів $\psi_i(t)$ та $\psi_k(t)$.

Визначник виразу (9.11) дорівнює нулю, якщо функції $(\psi_1(t), \psi_2(t), \dots, \psi_N(t))$ лінійно залежні і більші від нуля для лінійно незалежних функцій. Він дорівнює добутку квадратів норм функцій $\psi_i(t)$, якщо вони попарно ортогональні. Ортонормована система завжди лінійно незалежна, для неї визначник Грама дорівнює одиниці.

Нагадаємо, що в n -мірному евклідовому просторі, зокрема в просторі сигналів, представлених усіченим рядом Котельникова, коли $n = B = 2FT$ (B – база сигналу), можна вибрати не більше B лінійно незалежних сигналів. Вищевикладена теорія лінійного розподілу побудована без урахування адитивних завад у каналі. У цих умовах теоретично байдуже, чи вибрати систему переносників $\{\Psi_i(t)\}$ ортогональних або просто лінійно незалежних сигналів, у будь-якому випадку їх максимальна кількість не перевищує бази сигналу n . Однак якщо врахувати вплив завад, то перевага ортогональних сигналів стає незаперечною, тому що тільки в цьому випадку енергія сигналу використовується повністю. Окремим випадком ортогональних сигналів є сигнали зі спектрами, що не перекриваються, а також сигнали, що не перекриваються в часі.

9.4. Основні способи розподілу сигналів

9.4.1. Частотний розподіл каналних сигналів

Розглянемо основні етапи перетворення сигналів, спектри яких займають смуги частот, що не перекриваються (рис. 9.2). Спочатку відповідно до переданих повідомлень первинні індивідуальні сигнали $b_1(t), b_2(t), \dots, b_N(t)$ зі спектрами $S_1(f), S_2(f), \dots, S_N(f)$ модулюють переносники - піднесучі частоти f_k кожного каналу. Ця операція виконується за допомогою модуляторів M_1, M_2, \dots, M_N каналних перетворювачів. Отримані на виході частотних фільтрів $\Phi_1, \Phi_2, \dots, \Phi_N$ спектри $S_l(f)$ каналних сигналів займають відповідно смуги частот $\Delta f_1, \Delta f_2, \dots, \Delta f_N$, які для спрощення будемо вважати такими, що дорівнюють ширині спектру повідомлень F_c (використовується

односмугова модуляція). Будемо вважати, що спектри повідомлень фінітні.

Тоді можна вибрати піднесучі частоти f_k так, що Δf_1 смуги $\Delta f_2, \dots, \Delta f_N$ попарно не перекриваються. За цієї умови сигнали s_k $k = 1, 2, \dots, N$ взаємно ортогональні. Спектри $\dot{S}_1(f), \dot{S}_2(f), \dots, \dot{S}_N(f)$ підсумовуються в 1-му пристрої об'єднання сигналів, і їхня сукупність $\dot{S}_r(f)$ надходить на 2-й груповий модулятор M .

Сумарна смуга частот групового сигналу $\Delta f = NF$. У груповому модуляторі спектр $\dot{S}_r(f)$ за допомогою коливання несучої частоти f_0 переноситься в область частот, відведена для передачі даної групи каналів, у такий спосіб груповий сигнал $s_r(t)$ перетворюється на лінійний сигнал $s_{\text{л}}(t)$. При цьому може використовуватися будь-який вид модуляції, що забезпечує необхідну завадостійкість передачі (див. на рис. 9.1).

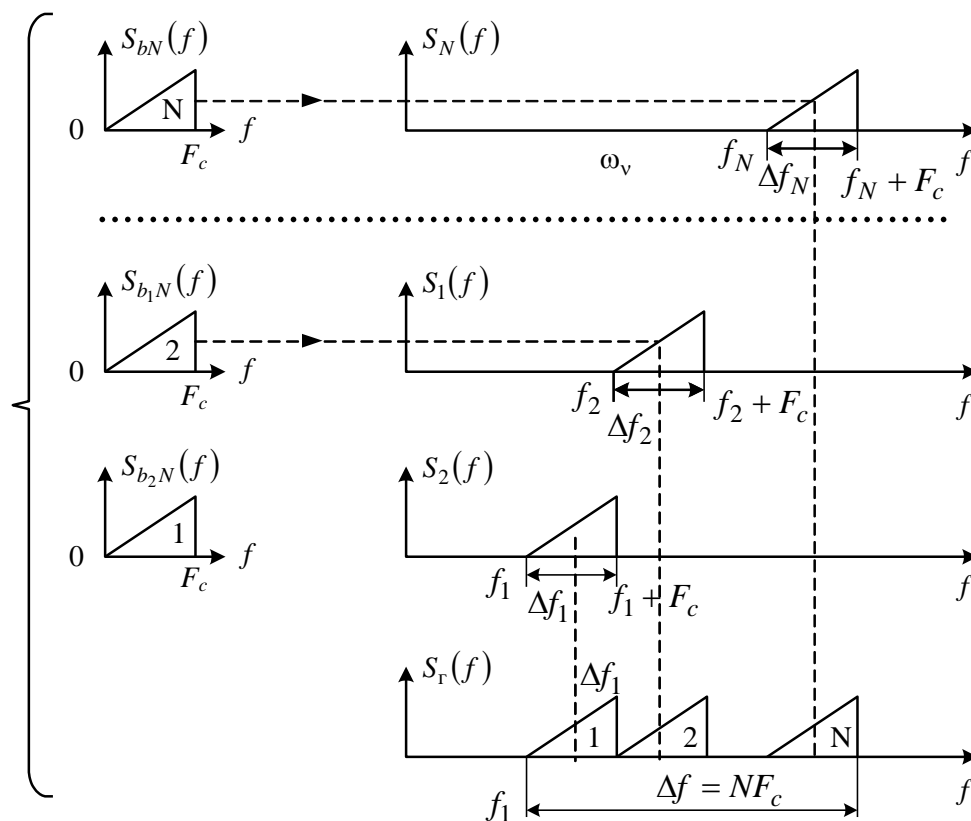


Рис. 9.2. Утворення спектра групового сигналу при багатоканальній передачі повідомлень з частотним розподілом каналів

На приймальному боці тракту лінійний сигнал надходить на груповий демодулятор (приймач Π), котрий перетворює спектр лінійного сигналу на спектр групового сигналу $\dot{S}_\Gamma(f)$. Спектр групового сигналу потім за допомогою каналних приймачів Π_k і частотних фільтрів Φ_k , що входять до них, знову розділяється на окремі смуги Δf_k й потім за допомогою демодуляторів D_k перетворюється на спектри повідомлень $\dot{S}_{bk}(f)$, призначених одержувачам.

Для того щоб без взаємних завад розподілити сигнали при частотному розподілі, кожний з фільтрів Φ_k повинен пропускати без ослаблення лише ті частоти $f \in \Delta f_k$, які належать сигналу даного каналу; частоти сигналів всіх інших каналів $f \in \Delta f_k$ фільтр повинен послаблювати в певну кількість разів.

Математично частотний розподіл ідеальними смуговими фільтрами (СФ) за аналогією з виразом (9.7) можна подати як

$$s_k(t) = \int_{-\infty}^{\infty} s_\Gamma(\tau) g_k(t - \tau) d\tau, \quad (9.12)$$

де $g_k(t)$ – імпульсна характеристика ідеального смугового фільтра, що пропускає без спотворень сигнал із шириною спектра Δf_k .

Вираз (9.12) збігається з виразом (9.7) при ваговій функції $\eta_k(t - \tau) = g_k(t - \tau)$. Згорті вираз (9.12) у частотній області відповідає перемножуванню спектра групового сигналу $\dot{S}_\Gamma(f)$ на Π – подібну передатну функцію i -го фільтра (рис. 9.3).

Отже, для ідеального виділення сигналів при частотному розподілі необхідне виконання двох умов: повне зосередження енергії сигналів $s_k(t)$ у межах відведених смуг Δf_k та ідеальність характеристик СФ. Однак жодна із цих умов принципово не здійсненна.

Результатом цього є взаємні завади при розподілі сигналів. Для зниження взаємних перехідних завад між каналами до

припустимого рівня доводиться вводити захисні частотні інтервали $\Delta f_{\text{зах}}$ (рис. 9.4).

Так, у сучасних системах багатоканального телефонного зв'язку кожному каналу виділяється смуга частот 4 кГц, хоча частотний спектр переданих мовних сигналів обмежується смугою від 300 до 3400 Гц, тобто ширина спектра становить 3,1 кГц.

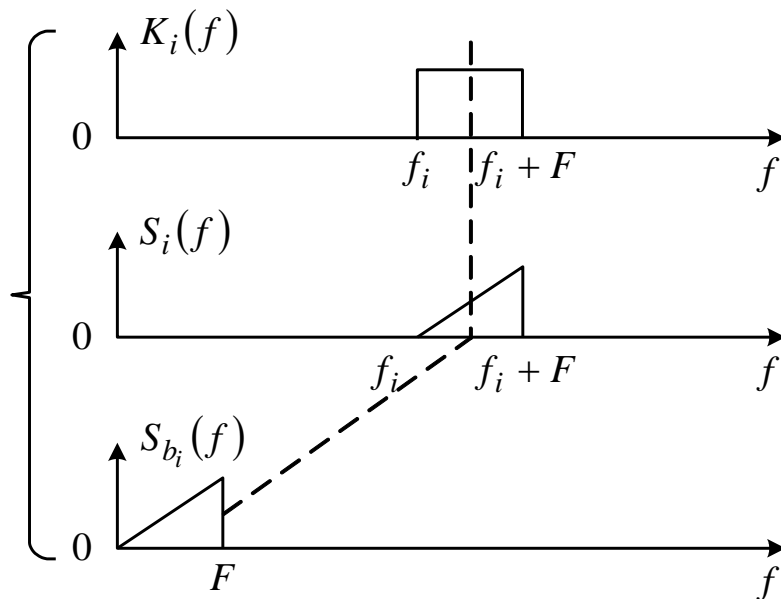


Рис. 9.3. Виділення сигналів окремих каналів за допомогою частотної фільтрації

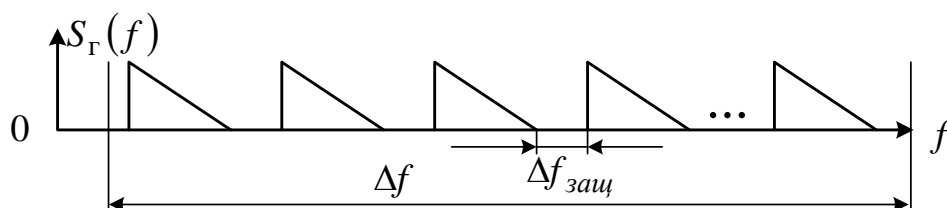


Рис. 9.4. Введення захисних частотних інтервалів роздільних смугових фільтрів між окремими каналами

Між смугами частот сусідніх каналів передбачені інтервали шириною по 0,9 кГц, призначені для зниження рівня взаємних завад при розподілі сигналів. Це означає, що в багатоканальних системах із частотним розподілом каналів (ЧРК) ефективно

використовується лише близько 80 % смуги пропускання тракту передачі. Крім того, необхідно забезпечити дуже високий ступінь лінійності всього тракту групового сигналу, що є істотним недоліком способу ЧРК.

9.4.2. Часовий розподіл каналів

При часовому розподілі каналів (РКЧ) груповий тракт за допомогою комутатора $K_{пер}$ надається по черзі для передачі сигналів кожного каналу багатоканальної системи (рис. 9.5).

Спочатку передається сигнал 1-го каналу, потім наступного, і так далі до останнього каналу з номером N , після чого знову підключається 1-й канал, і процес періодично повторюється із частотою дискретизації $f_{\Delta} = \frac{1}{\Delta}$, де Δ - інтервал дискретизації згідно з теоремою В.О. Котельникова.

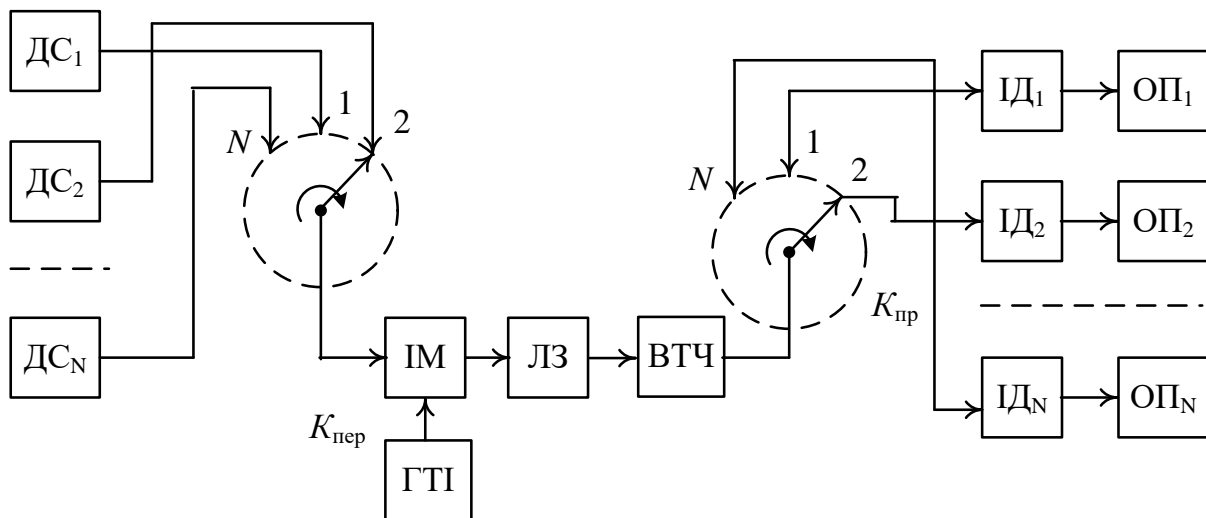


Рис. 9.5. Структурна схема багатоканальної системи передачі повідомлень із часовим розподілом каналів: ДС – джерело сигналу; $K_{пер}$ – комутатор передачі; ІМ – імпульсний модулятор; ГТІ – генератор тактових імпульсів; ЛЗ – лінія зв’язку; ВТЧ – виділювач тактової частоти; $K_{пр}$ – комутатор прийому; ІД – імпульсний детектор; ОП – отримувач повідомлень

На прийомному кінці встановлюється аналогічний комутатор K_{np} , що підключає груповий канал по черзі до приймачів різних каналів. Приймач i -го каналу підключається тільки на час передачі i -го сигналу й відключається на весь інший час, поки передаються сигнали інших каналів. Для нормальної роботи системи необхідно забезпечити синхронне й синфазне перемикання каналів на передавальному і приймальному боках. Часто для синхронізації використовують один з інформаційних каналів БСП.

Як каналні сигнали в системах РКЧ використовуються послідовності, модульованих імпульсів, що не перекриваються в часі, (наприклад, за амплітудою на рис. 9.6) $s_1(t), s_2(t), \dots, s_i(t), \dots, s_N(t)$.

Сукупність каналних сигналів утворює груповий сигнал $s_r(t)$. На рис. 9.6, а для спрощення зображені тільки сигнали $i = 1$ та $k = 2$ каналів. Груповий сигнал $s_r(t)$ надходить на комутатор K_{np} , дію якого можна ототожнювати з «часовим фільтром» (ключем), передатний коефіцієнт якого (рис. 9.6, б) дорівнює одиниці тільки на інтервалах дії імпульсів обраного каналу:

$$K_i(t) = \begin{cases} 1 & t \in \Delta t_i; \\ 0 & t \notin \Delta t_i. \end{cases} \quad (9.13)$$

У результаті часової фільтрації на виході i -приймача виділяться лише імпульси i -го каналу (рис 9.6,в). Отримані після демодуляції повідомлення $b_i(t)$ надходять до i -го одержувача.

При часовому розподілі взаємні завади обумовлені двома причинами. Перша полягає в тому, що лінійні спотворення, які виникають за рахунок обмеження смуги частот і неідеальності АЧХ і ФЧХ будь-якої фізично здійснюваної системи зв'язку, порушують форму імпульсів сигналів. Дійсно, якщо при передачі модульованих імпульсів кінцевої тривалості обмежити спектр, то імпульси «розпливуться» і замість імпульсів кінцевої тривалості одержуємо процеси, нескінченно протяжні в часі. При РКЧ це призведе до того, що імпульси одного каналу будуть накладатися

на імпульси інших каналів (рис. 9.7). Інакше кажучи, між каналами виникають взаємні перехідні завади. Можливі такі завади і через порушення синхронізації.

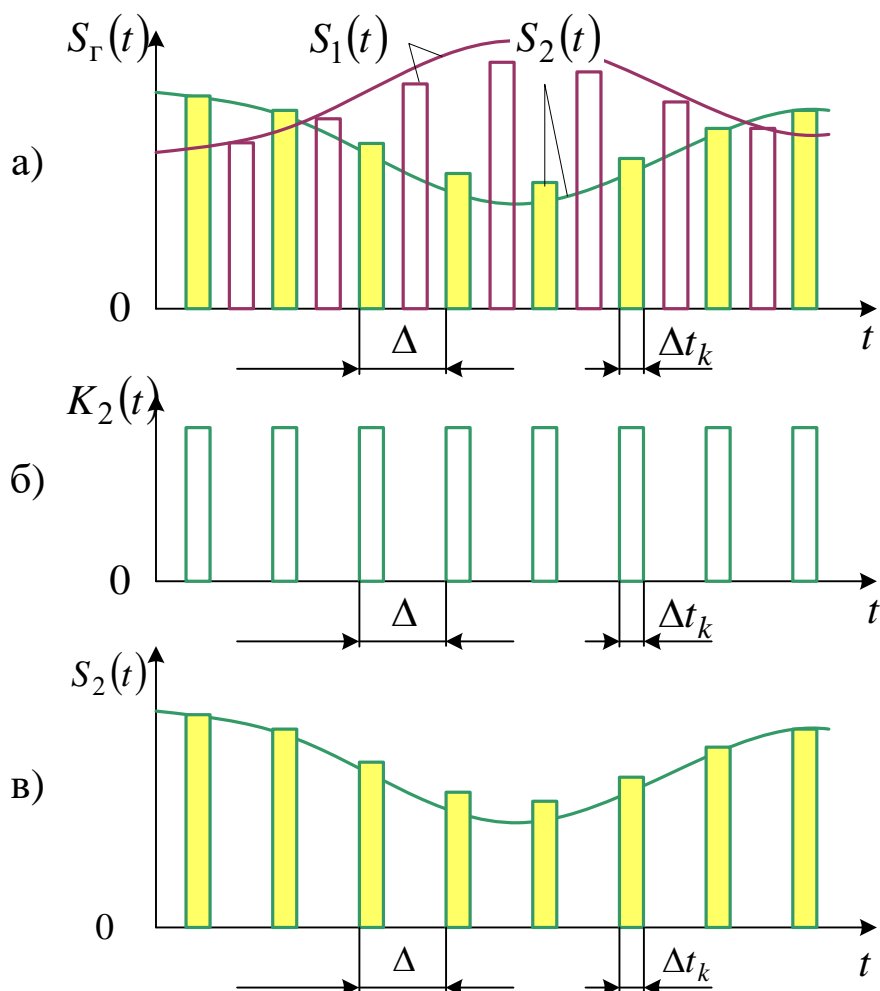


Рис. 9.6. Принцип виділення каналного сигналу $S_2(t)$ із групового $S_r(t)$

Для зниження рівня взаємних завад доводиться вводити «захисні» часові інтервали. Це вимагає зменшення тривалості імпульсу кожного каналу i , як наслідок, розширення спектра сигналів. Так, у багатоканальних системах телефонії смуга ефективно використовуваних частот $F_e = 3100$ Гц відповідно до теореми відліків про мінімальне значення частоти дискретизації

$$f_d = \frac{1}{\Delta} = 2F_e = 6300 \text{ Гц.}$$

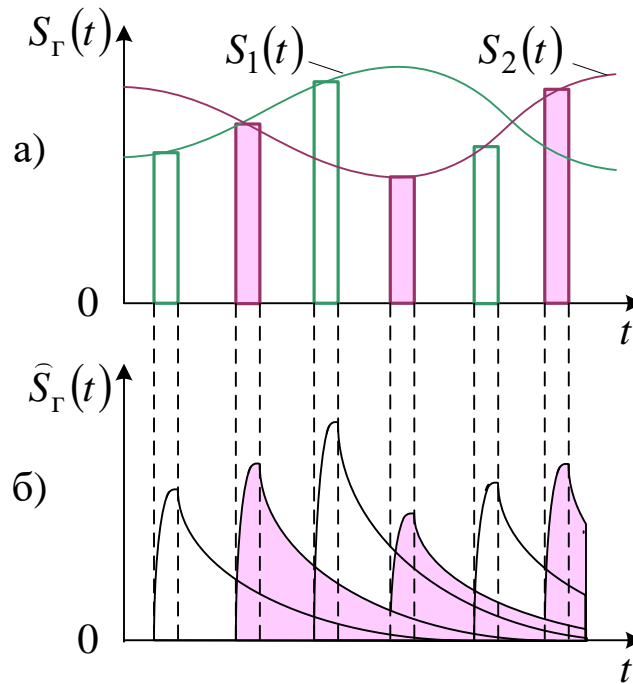


Рис. 9.7. Виникнення взаємних завад між каналами при РКЧ

Однак у реальних системах частоту проходження імпульсів вибирають трохи більшою (з деяким запасом): $f_{\partial} = 8$ кГц. Для передачі таких імпульсів в одноканальному режимі буде потрібна смуга частот не менше $F_{\partial} = 4$ кГц. При РКЧ сигнал кожного каналу займає однаковий інтервал часу, який визначається в ідеальних умовах відповідно до теореми відліків із співвідношення (без урахування каналу синхронізації)

$$\Delta_N = \frac{\Delta}{N} = \frac{1}{2NF_{\partial}} = \frac{1}{2F_{заг}}, \quad (9.14)$$

де $F_{заг} = NF_{\partial}$.

Хоча теоретично системи із РКЧ і ЧРК дозволяють одержати однакову ефективність використання частотного спектра, проте у реальних умовах системи РКЧ трохи поступаються системам ЧРК за цим показником.

Однак системи із РКЧ мають значну перевагу перед системами із ЧРК, пов'язану з тим, що при передачі сигналів різних каналів у різний час у них відсутні перехідні завади

нелінійного походження. Крім того, апаратура із РКЧ значно простіша, ніж з ЧРК, де для кожного індивідуального каналу потрібні свій модулятор на передачі й розділовий смуговий фільтр на прийомі. Немаловажною перевагою систем із РКЧ є менший пік-фактор. Системи із РКЧ застосовуються при передачі безперервних повідомлень за допомогою аналогових видів імпульсної модуляції (ФІМ, ШІМ), особливо в цифрових системах з ІКМ, ДІКМ, ДМ, АДІКМ [43].

Слід зазначити також, що сумарна потужність $P_{заг}$ прийнятого сигналу $s(t)$, необхідна для забезпечення заданої правильності за наявності флукуаційних завад (як при ЧРК, так і при РКЧ), в ідеальному випадку в N раз більша, ніж потужність P при одноканальній передачі з тим же виглядом модуляції: $P_{заг} = NP$. Це легко зрозуміти, оскільки при додаванні незалежних сигналів потужності складаються. У дійсності ж через перехідні завади правильність прийому в багатоканальних системах при виконанні цієї умови трохи нижча, ніж в одноканальних. Збільшуючи потужність сигналу в багатоканальній системі із частотним розподілом, не можна знижувати рівень перехідних завад, оскільки при цьому зростає й потужність останніх, а в ряді випадків потужність завад нелінійного походження зростає навіть швидше, ніж потужність сигналу.

9.4.3. Розподіл каналних сигналів за фазою

Розглянемо множину сигналів

$$\{s_i(t) = A_i \sin(\omega_0 t + i\Delta\varphi)\}, i = \overline{1, N}, \quad (9.15)$$

де A_i – коефіцієнти, що відображають передані повідомлення;

$\Delta\varphi$ – різниця фаз між сигналами сусідніх каналів, $\Delta\varphi = \frac{2\pi}{N}$.

Сигнали виразу (9.15) розрізняються за фазою й імовірно можуть бути використані для побудови багатоканальної системи. Однак виявляється, що серед N переносників вигляду

$$\{\Psi_i(t) = \sin(\omega_0 t + i\Delta\varphi)\}, i = \overline{1, N} \quad (9.16)$$

лише два будь-яких $\Psi_i(t)$ і $\Psi_k(t)$ можуть бути лінійно незалежними. Щоб переконатися в цьому, досить скористатися визначником Грама (вираз (9.11)). Нехай $N = 2$, тобто

$$\Psi_1(t) = \sin(\omega_0 t + \Delta\varphi), \Psi_2(t) = \sin(\omega_0 t + \Delta\varphi) = \Psi_1(t). \quad (9.17)$$

Тоді скалярні добутки $(\Psi_1, \Psi_1) = (\Psi_2, \Psi_2) = 0,5$; $(\Psi_1, \Psi_2) = (\Psi_2, \Psi_1) = 0,5 \cos \varphi$ і, отже, для визначника Грама маємо [19]

$$\begin{aligned} \Gamma(\Psi_1, \Psi_2) &= \begin{vmatrix} (\Psi_1, \Psi_1) & (\Psi_1, \Psi_2) \\ (\Psi_2, \Psi_1) & (\Psi_2, \Psi_2) \end{vmatrix} = \\ &= \frac{1}{4} (1 - \cos^2 \Delta\varphi) = \frac{1}{4} \sin^2 \Delta\varphi \neq 0. \end{aligned} \quad (9.18)$$

З виразу (9.18) випливає, що при кожному $\Delta\varphi \neq 0$ або π сигнали $s_1(t)$ та $s_2(t)$ можуть бути розділені. Цього не відбудеться при $N = 3$:

$$\begin{aligned} \Psi_1(t) &= \sin(\omega_0 t + \Delta\varphi), \Psi_2(t) = \sin(\omega_0 t + 2\Delta\varphi), \\ \Psi_3(t) &= \sin(\omega_0 t + 3\Delta\varphi). \end{aligned} \quad (9.19)$$

Для скалярних добутків маємо

$$\begin{aligned} (\Psi_1, \Psi_1) &= (\Psi_2, \Psi_2) = (\Psi_3, \Psi_3) = \frac{1}{2}; \\ (\Psi_1, \Psi_2) &= (\Psi_2, \Psi_1) = (\Psi_2, \Psi_3) = \frac{1}{2} \cos \Delta\varphi; \\ (\Psi_1, \Psi_3) &= (\Psi_3, \Psi_1) = \frac{1}{2} \cos 2\Delta\varphi. \end{aligned}$$

Підставивши значення скалярних добутків у визначник Грама для $N = 3$, одержимо

$$\Gamma(\Psi_1, \Psi_2, \Psi_3) = \frac{1}{8} \left(1 - (\sin^2 \Delta\varphi + \cos^2 \Delta\varphi) \right) (1 - \cos 2\Delta\varphi) = 0.$$

З множини сигналів, що розрізняються за фазою (вираз (9.15)) лінійно незалежні тільки будь-які два сигнали. Тому на одній несучій частоті ω_0 при довільних значеннях амплітуд A_i і A_k та фаз φ_i і φ_k можна забезпечити лише двоканальну передачу.

Повертаючись до визначника Грама для двох переносників Ψ_i і Ψ_k , знаходимо, що переносники із виразу (9.17) ортогональні при різниці фаз $\Delta\varphi = \frac{\pi}{2}$. Ортогональні переносники використовуються, зокрема, у квадратурній системі передач двох незалежних каналів.

9.5. Розподіл сигналів за формою

Загальною ознакою розподілу сигналів може служити форма сигналів. Сигнали, що розрізняються за формою, можуть передаватися одночасно, мати частотні спектри, які перекриваються, і проте такі сигнали можна розділити, якщо виконується умова їхньої лінійної незалежності або умова ортогональності.

Нехай переносниками обрані імпульси, послідовність яких утворює, наприклад, степеневий ряд

$$\left\{ \Psi_1(t) = 1, \Psi_2(t) = t, \Psi_3(t) = t^2, \dots \right\}, \quad (0 \leq t \leq T). \quad (9.20)$$

Для групового сигналу, припустивши, що передана інформація міститься в коефіцієнтах C_1, C_2, \dots , можна записати

$$s_T(t) = \left[C_1 + C_2 t + \dots + C_N t^{N-1} \right], \quad (0 \leq t \leq T). \quad (9.21)$$

Члени ряду (9.20) лінійно незалежні, і, отже, жоден з каналних сигналів $C_i t^{i-1}$ не може бути утворений лінійними

комбінаціями всіх інших сигналів. Це легко зрозуміти, звернувши увагу на те, що багаточлен вигляду з виразу (9.21) може бути тотожно дорівнювати нулю тільки в тому випадку, коли всі його коефіцієнти C_i дорівнюють нулю.

Для розподілу таких сигналів можна застосувати загальний метод відповідно до умови лінійної незалежності переносників з виразу (9.20). Так, при двоканальній передачі для інтервалу $0 \leq t \leq T$ можна записати

$$s_{\Gamma}(t) = s_1(t) + s_2(t) = C_1 + C_2 t = C_1 \Psi_1(t) + C_2 \Psi_2(t).$$

Якщо вагові функції у виразі (9.7) вибрати так:

$$\left. \begin{aligned} \eta_1(t) &= a_{11} \Psi_1(t) + a_{12} \Psi_2(t) = \frac{4}{T} - 6 \frac{t}{T}; \\ \eta_2(t) &= a_{21} \Psi_1(t) + a_{22} \Psi_2(t) = -\frac{6}{T} + 12 \frac{t}{T}, \end{aligned} \right\} \quad (9.22)$$

то в результаті операцій проектування вектора s_{Γ} на напрямки координатних осей η_1 і η_2 одержимо (при $T = 1$)

$$\left. \begin{aligned} \pi_1(s_{\Gamma}) &= \int_0^T s(t) \eta_1(t) dt = 4C_1(\Psi_1, \Psi_1) - 6C_1(\Psi_1, \Psi_2) + \\ &+ 4C_2(\Psi_2, \Psi_1) - 6C_2(\Psi_2, \Psi_2) = C_1; \\ \pi_2(s_{\Gamma}) &= \int_0^T s(t) \eta_2(t) dt = 12C_1(\Psi_1, \Psi_2) - 6C_1(\Psi_1, \Psi_1) + \\ &+ 12C_2(\Psi_2, \Psi_2) - 6C_2(\Psi_2, \Psi_1) = C_2, \end{aligned} \right\} \quad (9.23)$$

де $(\Psi_1, \Psi_1) = 1$, $(\Psi_2, \Psi_2) = \frac{1}{3}$, $(\Psi_1, \Psi_2) = (\Psi_2, \Psi_1) = \frac{1}{2}$.

Операції системи (9.23) виконуються розподільним пристроєм, зображеним на рис. 9.8. На відміну від пристрою розподілу ортогональних сигналів тут додається пристрій формування вагових функцій η_1 і η_2 , котрі з функцій $\Psi_1(t)$ і $\Psi_2(t)$

утворюють лінійні комбінації як з виразу (9.22). У загальному випадку використовується ітеративний метод ортогоналізації Грама-Шміда, що полягає в побудові ортонормованої системи $\{\eta_i(t)\}$ за заданою лінійно незалежною системою $\{\Psi_i(t)\}$ використовуючи допоміжні ортогональні вектори відповідно до співвідношення

$$e_i(t) = \Psi_i(t) - \sum_{k=1}^{i-1} (\Psi_i \eta_k) \eta_k(t),$$

де $\eta_k(t) = \frac{e_i(t)}{\|e_i\|}, i = 1, 2, \dots$

Перестановка вихідних векторів $\{\Psi_i\}$ призводить до різних ортонормованих систем $\{\eta_i(t)\}$, а через те, що процедура ітеративна, її можна застосовувати для побудови як кінцевовимірних ортогональних базисів, так і рахункових ортонормованих систем у нескінченно вимірних функціональних просторах $L_2(T)$.

При формуванні сигналів, що розрізняються за формою, як переносники застосовуються різні ортогональні функції, отримані на основі ортогоналізації степеневого ряду (вираз (9.20)). Головним недоліком отриманих у такий спосіб класичних поліномів є їхній континуальний характер, що пов'язано з необхідністю використання аналогової техніки для формування таких сигналів.

Зовсім по-іншому формуються ортогональні сигнали на основі дискретних ортогональних послідовностей як функції Уолша, Радемахера, що розглянуті в п. 2.3.

Через можливість застосування до функцій Уолша логічних операцій вони є досить перспективними при розробленні багатоканальних цифрових систем передачі з розподілом за формою, а також в апаратурі формування й перетворення сигналів на базі мікропроцесорної техніки.

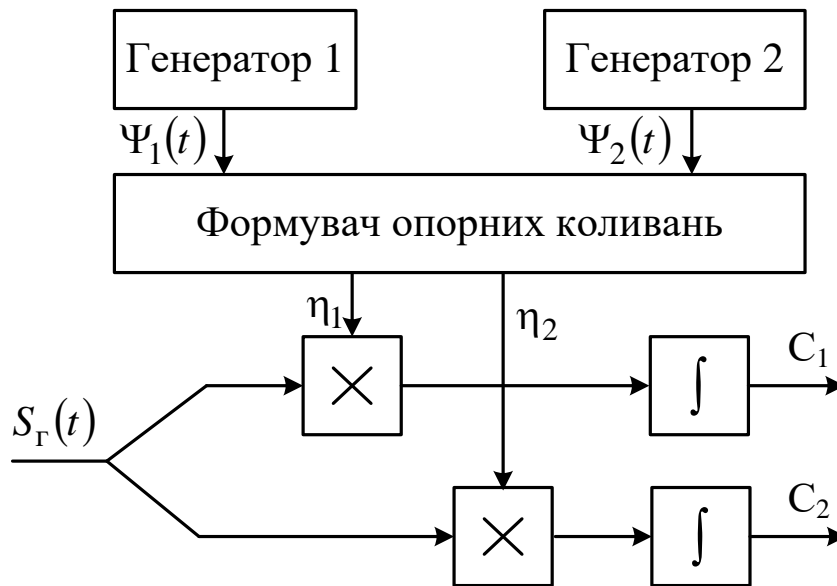


Рис. 9.8. Структурна схема лінійного розподілу сигналів за формою

У сучасній техніці багатоканального зв'язку при передачі дискретних повідомлень широке застосування знаходять системи з ортогональними індивідуальними сигналами, що перекриваються в часі та за частотою. Вони розподіляються в місці прийому за допомогою узгоджених пасивних фільтрів або еквівалентних їм активних кореляційних схем, забезпечуючи оптимальний прийом на фоні шуму.

9.6. Системи передачі із шумоподібними сигналами

Розглянуті вище системи багатоканальної передачі з ортогональними та лінійно незалежними сигналами вимагають для нормальної роботи системи синхронізації, котра необхідна для точного збігу спектра сигналу зі смугою пропускання при ЧРК; точного збігу часових інтервалів передачі сигналів окремих каналів при РКЧ; точного визначення моментів початку й кінця тактового інтервалу в системах з розподілом активними фільтрами сигналів за формою; точного установлення моменту відліку в системі з узгодженими фільтрами.

У ряді випадків здійснити точну синхронізацію важко. З подібними ситуаціями доводиться зустрічатися, наприклад, при організації оперативного зв'язку між рухомими об'єктами

(автомобілями, поїздами, літаками та ін.). Таке завдання виникає також при організації оперативного зв'язку, коли штучні супутники Землі використовуються як ретранслятори. У всіх цих випадках можуть бути використані системи асинхронного багатоканального зв'язку, коли сигнали всіх абонентів передаються в спільній смузі частот, а канали не синхронізовані між собою в часі.

Оскільки в таких системах за каналами не закріплені ні частотні смуги, ні часові інтервали й час роботи кожного каналу довільний, то такі системи називають системами з вільним доступом до лінії зв'язку або системами з незакріпленими каналами. У системах з вільним доступом кожному каналу (абонентіві) присвоюється певна форма сигналу, яка і є відмітною ознакою, «адресою» даного абонента. На відміну від звичайного розподілу за формою, де умова ортогональності сигналів виконуються лише тоді, коли тактові інтервали всіх каналів жорстко синхронізовані, щоб мати змогу повного лінійного розподілу сигналів у системах з вільним доступом, ортогональність або лінійна незалежність повинні зберігатися при будь-яких часових зсувах сигналів. Це означає, що для будь-якої пари сигналів $s_i(t)$ і $s_k(t)$ повинна виконуватися умова

$$T \int_t^{t+T} s_i(t) s_k(t - \tau) dt \equiv 0 \quad (9.24)$$

при $0 \leq \tau \leq T$, де T – тривалість елементарного сигналу, а інтегрування відбувається на будь-якому інтервалі від t до $t + T$.

Строго кажучи, умова (9.24) виконується тільки у випадку, коли сигнали $s_k(t)$ являють собою "білий" шум, тобто мають необмежену ширину спектра й нескінченну дисперсію; для реальних сигналів вона нездійсненна. Разом з тим можна сформулювати такі сигнали, для яких умова (9.24) виконується приблизно в тому розумінні, що

$$T \int_t^{t+T} s_i(t) s_k(t - \tau) dt \ll \|s_i\|^2 \quad \text{при } 0 \leq \tau \leq T, \quad (9.24,a)$$

тобто скалярні добутки сигналів при будь-якому зсуві за часом набагато менші від енергії елементарного сигналу. Такі сигнали можна назвати майже ортогональними. За своїми властивостями майже ортогональні сигнали наближаються до "білого" шуму, тому їх часто називають широкосмуговими шумоподібними (ШШС). Їхні кореляційні функції й спектри щільності потужності близькі до аналогічних характеристик "квазібілого" шуму. Шумоподібні сигнали належать до класу складних сигналів, база яких $B = 2FT \gg 1$, і є подальшим розвитком сигналів, що розрізняються за формою.

Теорії шумоподібних сигналів і питанням їхнього використання в різних системах зв'язку присвячені роботи Л.Є. Варакіна [11].

Найпоширенішим прикладом технічної реалізації майже ортогональних шумоподібних сигналів можуть бути певним чином сформовані псевдовипадкові послідовності дискретних, зокрема двійкових, радіоімпульсів. База сигналів при цьому визначається кількістю імпульсів у послідовності. Кожному каналу присвоюється одна з множин майже ортогональних двійкових послідовностей, яка і є «адресою» каналу. Це призводить до назви «асинхронні адресні системи зв'язку» (ААСЗ).

Важливою перевагою ААСЗ є те, що немає потреби мати центральну комутаційну станцію; у всіх абонентів є прямий доступ один до одного без частотної перебудови прийомних і передавальних пристроїв (рис. 9.9). Тут досить набрати «адресу» абонента, що викликається, тобто змінити «форму» імпульсної адресної послідовності.

У системах із закріпленими каналами з ЧРК і РКЧ додавання хоча б одного нового абонента виявляється можливим лише при виключенні одного з тих, що є в системі. Значно простіше це завдання вирішується в системах ААСЗ. Тут внаслідок вільного доступу до лінії зв'язку можуть вести передачу будь-які N_a активних абонентів із загального числа N абонентів системи зв'язку.

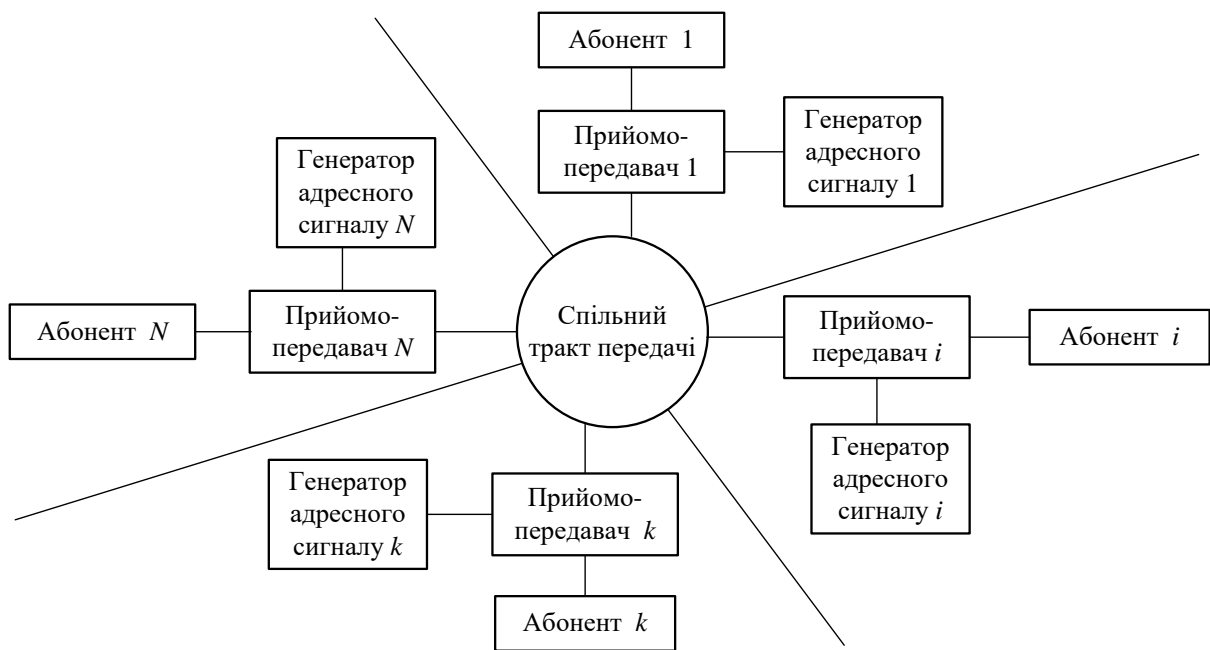


Рис. 9.9. Структурна схема багатоканальної асинхронно-адресної системи зв'язку

При визначенні числа N_a потрібно враховувати, що внаслідок неповної ортогональності каналів в ААСЗ неминучі перехідні завади («шуми неортогональності»), рівень яких зростає зі збільшенням N_a . Тому кількість одночасно працюючих абонентів повинна бути обмежена. Припустиме значення N_a зростає зі збільшенням бази сигналу, тому що чим більша база, тим точніше виконується умова (9.24, а).

Залежно від часу активності абонентів (тобто від частки часу, яку займає k -й канал для передачі повідомлень) можна організувати, наприклад, 1000-канальну систему зв'язку, у якій одночасно проводиться передача будь-яких 50 абонентів з тисячі. У таких системах легко реалізуються резерви пропускної спроможності, які виникають за рахунок малоактивних абонентів. Вивчивши статистику повідомлень, переданих по кожному каналу, можна встановити припустиме число каналів у системі N , при якому забезпечується нормальна робота N_a активних каналів.

Приклади шумоподібних сигналів. На практиці широко використовуються методи синтезу сигналів із заданими автокореляційними (АКФ) та взаємно кореляційними функціями

(ВКФ). Якщо розглядати послідовності з n імпульсів прямокутної форми, які можуть приймати значення ± 1 , то простим перебором можна знайти такі послідовності, для яких $\frac{B(0)}{E_1} = n$, $\frac{\max|B(\tau \neq 0)|}{E} = \frac{1}{n}$, $E = nE_1$, де E – енергія всього сигналу, E_1 – енергія одного елемента. Серед них назовемо насамперед послідовності коду Баркера (табл. 9.1) [19].

Послідовності Баркера мають близьку до ідеальної форми автокореляційну функцію: абсолютне значення бічних пелюстків не перевищує $\frac{1}{n}$ основного. На рис. 9.10, *a* наведена послідовність (яка називається також кодом) Баркера для $n = 11$ і її автокореляційна функція (рис. 9.10, *б*).

Таблиця 9.1

Структури послідовностей коду Баркера

Кількість імпульсів	Номер імпульсу в послідовності													Максимум нормованого модуля АКФ	
	1	2	3	4	5	6	7	8	9	10	11	12	13	основний	побічні
2	+1	-1												1	1/2
3	+1	+1	-1											1	1/3
4	+1	+1	-1	+1										1	1/4
5	+1	+1	+1	-1	+1									1	1/5
7	+1	+1	+1	-1	-1	+1	-1							1	1/7
11	+1	+1	+1	-1	-1	-1	+1	-1	-1	+1	-1			1	1/11
13	+1	+1	+1	+1	+1	-1	-1	+1	+1	-1	+1	-1	+1	1	1/13

Прийом послідовності $s_1(t)$ («адреса» 1 каналу) на рис. 9.10, *a* виконується узгодженим трансверсальним фільтром зображеним на рис. 9.11. Імпульси послідовності Баркера з кількістю знаків $n = 11$ надходять спочатку на фільтр УФП, узгоджений з одиничним прямокутним імпульсом, а потім у лінію затримки (ЛЗ), яка має відводи через проміжки Δ , далі на фазоінверсні (–) і фазозберігаючі (+) каскади з однаковими коефіцієнтами передачі, складаючий й вирішуючий пристрій (ВП).

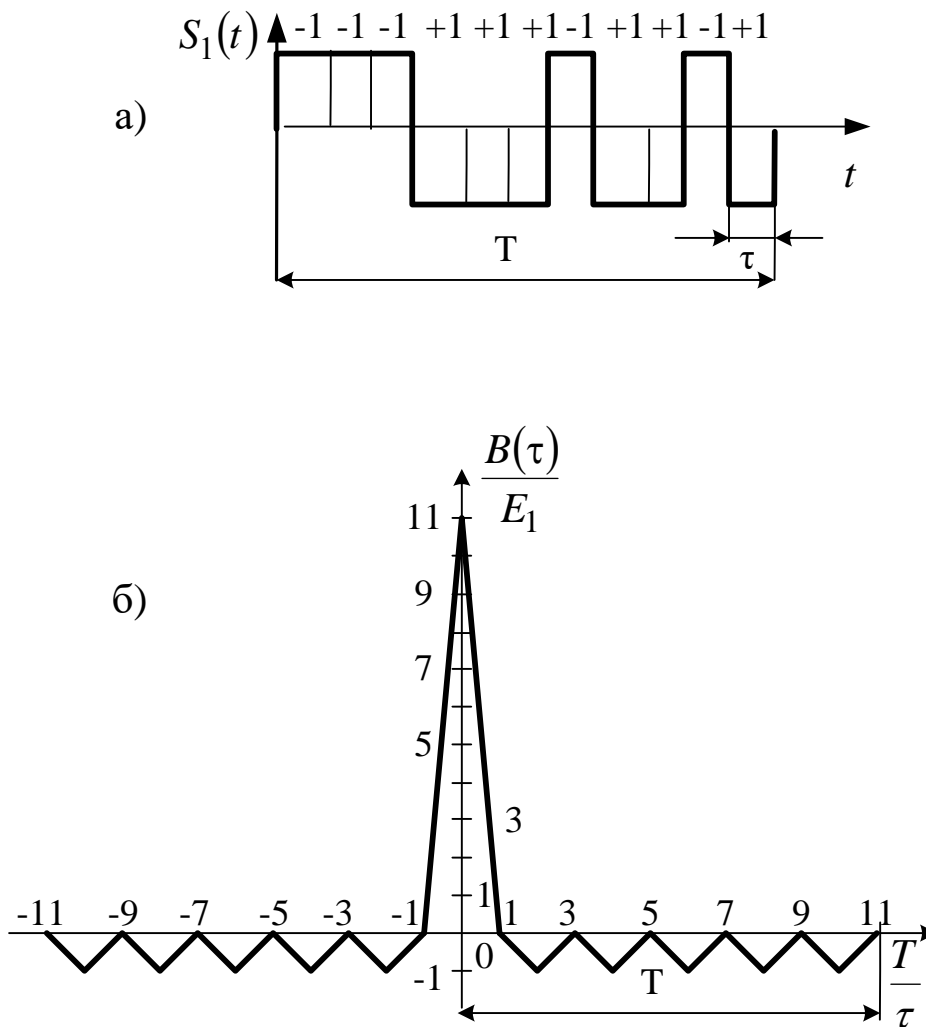


Рис. 9.10. Реалізація послідовності Баркера (а) та її автокореляційна функція (б)

Фазоінверсні та фазозберігаючі каскади ввімкнені в порядку, що відповідає зворотному порядку чергування біполярних імпульсів n -послідовності (рис. 9.10, а): кількість цих каскадів дорівнює кількості елементів послідовності. Перший каскад ввімкнений перед лінією затримки, останній – на її кінці.

При прийомі n -послідовність проходить по ЛЗ, і в момент, коли всі імпульси послідовності співпадут за знаком з вагами, ввімкненими між відводами ЛЗ і складаючим пристроєм, всі імпульси складуться синфазно, на виході ВП з'явиться найбільший імпульс – узгоджений фільтр зафіксує адресу 1-го каналу.

При всіх інших зсувах підсумовування відбувається не у фазі (з різними знаками), і на виході ВП з'являються рівні, що не перевищують за модулем $\frac{1}{n}$ максимального значення. Оскільки функції взаємної кореляції між послідовностями також мають найбільші значення, що не перевищують $\frac{1}{n}$, то послідовність адреси чужого каналу не може викликати помилкового спрацьовування вирішуючого пристрою 1-го каналу.

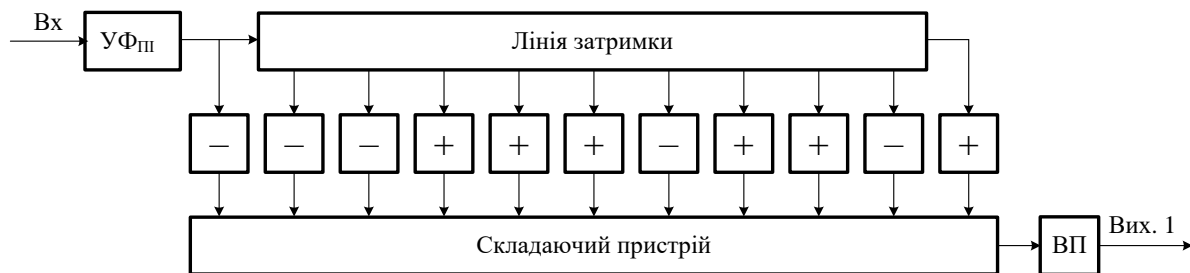


Рис. 9.11. Узгоджений фільтр для послідовності Баркера

Дослідження показують, що послідовностей Баркера із «залишками» величини $\frac{1}{n}$ для $n > 13$ не існує. Тому для більших n доводиться задовольнятися послідовностями, які мають «залишки» більші, ніж $\frac{1}{n}$.

Трохи гірші автокореляційні функції порівняно з послідовностями Баркера, але все ж таки досить придатні для використання в якості адресних сигналів мають лінійні рекурентні M -послідовності (ЛРП), або, як їх ще називають, лінійні послідовності зсувного регістра, формуючого послідовність максимальної довжини. Для ЛРП відношення головного максимуму до максимального бічного пелюстка автокореляційної функції зростає приблизно як \sqrt{n} , де n – кількість імпульсів у послідовності.

Лінійним рекурентним послідовностям притаманна властивість хаотичності, що полягає в наступному. Якщо з періоду ЛРП, що містить $n = 2^\mu - 1$ членів, вибрати можливі відрізки μ членів у кожному, то, по-перше, серед цих відрізків не

буде тих, що співпадають, і, по-друге, серед них знайдуться будь-які комбінації з $+1$ та -1 , що складаються з μ членів (окрім забороненої комбінації, яка складається тільки з $+1$). Ці властивості подібні до властивостей випадкових біполярних послідовностей; тому ЛРП часто називають псевдовипадковими або шумоподібними послідовностями.

До того ж автокореляційна функція ЛРП має форму, подібну до автокореляційної функції "квазібілого" шуму з обмеженим спектром. ЛРП формуються генераторами двійкових імпульсів з використанням зсувного регістра. Прийом ЛРП може здійснюватися як узгодженими фільтрами, так і кореляторами. При передачі псевдовипадкових послідовностей по високочастотному каналу зазвичай застосовується фазова або відносна фазова модуляція.

З інших способів формування ШШС для асинхронно-адресних систем зв'язку згадаємо про спосіб формування сигналів за допомогою частотно-часової матриці (ЧЧМ). Уже зазначалося, що в системах з ортогональними сигналами енергія кожного сигналу повністю відокремлюється від енергії інших сигналів. Це положення стає найбільш виразним, якщо звернутися до частотно-часових діаграм системи зв'язку при частотному розподілі (рис. 9.12, *а*) і при часовому розподілі сигналів (рис. 9.12, *б*). Тут кожному каналу відводиться певна область частотно-часового простору; положення площадки можна розглядати як «адресу» абонента.

Однак, частотно-часову область F і T можна розділити на площадки іншим способом: адресу кожного каналу можна сформуванати з набору «елементарних площадок» частотно-часової площини (рис. 9.13). Кожен двійковий інформаційний символ (рис. 9.13, *а*) передається за час T і відображається певною послідовністю імпульсів, які мають різні частоти (рис. 9.13, *б*) у спільній смузі F . Двійкову інформацію в послідовність можна ввести, змінюючи один з параметрів елементарного радіоімпульсу. Ці адресні набори імпульсів складаються на основі їхнього подання у вигляді ЧЧМ (рис. 9.13, *в*); до них висуваються звичайні вимоги автокореляційних функцій (з малими бічними пелюстками) і малих значень взаємної кореляції.

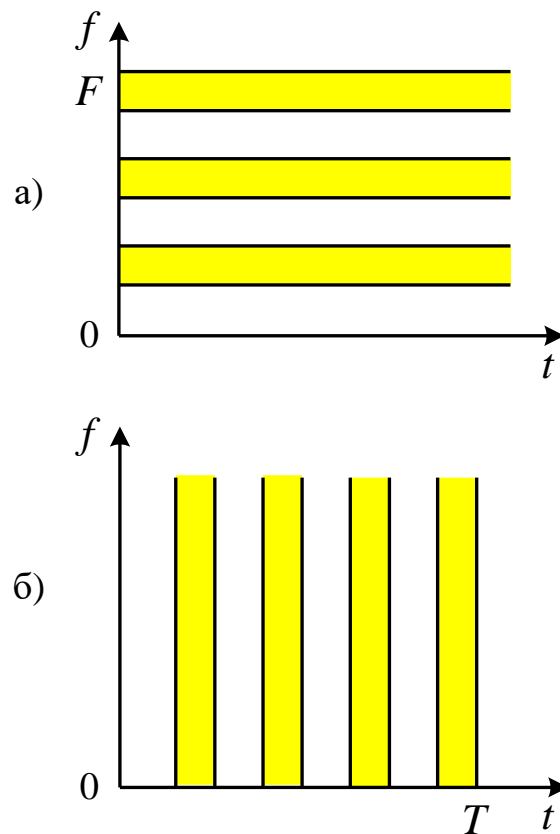


Рис. 9.12. Принцип частотного (а) і часового (б) розподілу каналів

Зміна часового положення імпульсів і розходження в їх частотах дозволяють порівняно простими технічними засобами одержати кілька тисяч частотно-часових коливань (адрес). Не всі комбінації ЧЧМ використовують як адресні сигнали; серед них зустрічаються й такі, які не мають необхідних кореляційних властивостей. Максимальна кількість адрес, для яких рівень бічних пелюстків кореляційних функцій не перевищує $\frac{1}{\sqrt{FT}}$, наближається до FT . Сигнали ЧЧМ також є різновидом сигналів, що різняться за формою. Їх також можна розподілити як узгодженими фільтрами, так і кореляторами.

У технічній літературі є опис великої кількості різних систем зв'язку з вільним доступом. Найхарактернішими з них є системи на 1000...1500 каналів з 50...100 активними абонентами.

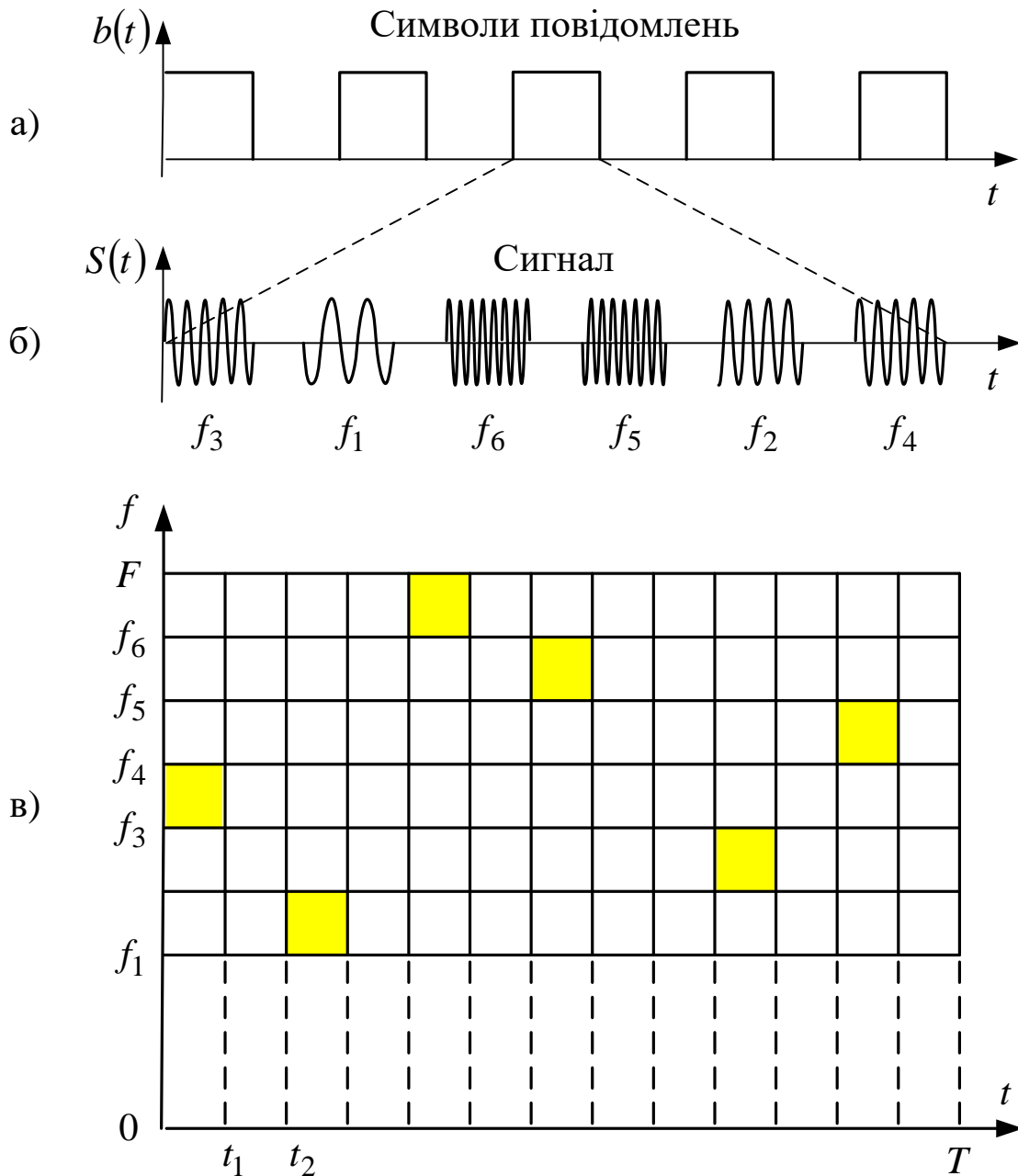


Рис. 9.13. Приклад утворення багатоканального широко-
 смугового сигналу за допомогою частотно-часової матриці:
 а) послідовність двійкових інформаційних сигналів;
 б) подання одного двійкового імпульсу за допомогою
 послідовного набору радіоімпульсів різних частот;
 в) подання сигналу у вигляді ЧЧМ

9.7. Комбінаційний розподіл сигналів

9.7.1. Принцип комбінаційного ущільнення

При багатоканальній передачі дискретних повідомлень наряду з ЧРК, РКЧ і розподілом сигналів за формою використовується комбінаційний спосіб формування групового сигналу. Сутність цього способу полягає в наступному.

Нехай необхідно організувати передачу N незалежних дискретних повідомлень по загальному груповому тракту. Якщо елемент i -го повідомлення може набувати одне з m_i можливих значень ($i = 1, 2, \dots, N$), то загальна кількість значень, які може приймати елемент N -канального джерела, в якому поєднуються вихідні N джерел, буде дорівнює $M = \prod_{i=1}^N m_i$. При однакових значеннях $m_i = m$ маємо

$$M = m^N. \quad (9.25)$$

Таким чином, при комбінаційному ущільненні кожне поєднання каналних повідомлень відображається елементом групового повідомлення з основою коду (вираз (9.25)), тобто використовуючи основу коду, можна одночасно передавати інформацію від N індивідуальних джерел, які працюють із основою коду m . Якщо, зокрема, $m = 2$ (двійкові коди), а кількість каналів $N = 2$, то групове повідомлення b_T може набувати чотири можливих значення, які відповідають різним комбінаціям нулів та одиниць в обох каналах, при $N = 3$ кількість різних комбінацій буде дорівнювати $M = 8$ і т. д. Завдання тепер зводиться до передачі деяких чисел b_T , які визначають номер комбінації. Ці числа можуть передаватися за допомогою сигналів дискретної модуляції будь-якого вигляду. Розподіл сигналів, що ґрунтується на розходженні в комбінаціях сигналів різних каналів, називається комбінаційним. Структурна схема багатоканальної системи з комбінаційним (кодовим) об'єднанням і розподілом каналів (РК) зображена на рис. 9.14.

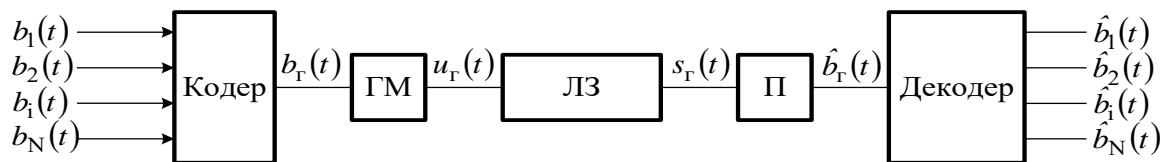


Рис. 9.14. Структурна схема багатоканальної системи з комбінаційним ущільненням і розподілом сигналів

Зі схеми видно, що первинні повідомлення $b_1(t)$, $b_2(t)$, ..., $b_N(t)$ від N джерел надходять на вхід кодера, який виконує функції пристрою об'єднання каналів. Отримане групове повідомлення $b_{\Gamma}(t)$ перетворюється за допомогою групового модулятора (ГМ) на груповий сигнал $u_{\Gamma}(t)$ і надходить у груповий тракт. На прийомному боці після демодуляції в приймачі (П) і декодування в декодері формуються каналні повідомлення, що відповідають N первинним повідомленням.

9.7.2. Принцип комбінаційного розподілу сигналів

Типовим прикладом комбінаційного ущільнення є система двократної частотної модуляції (ДЧМ), у якій для передачі чотирьох комбінації сигналів двох джерел (каналів) використовують чотири різні частоти: f_1 , f_2 , f_3 , f_4 . У прийнятих тут аббревіатурах ДЧМ і ДФМ перші букви означають «двократні». При двократній фазовій модуляції (ДФМ) кожній комбінації повідомлень 1-го і 2-го джерела відповідає певне значення фази групового сигналу φ_1 , φ_2 , φ_3 або φ_4 (табл. 9.2).

Як принцип комбінаційного розподілу розглянемо приклад розподілу каналів при ДЧМ (рис. 9.15). Тут прийнятий сигнал подається на фільтри Φ_1 , Φ_2 , Φ_3 , Φ_4 , підключені до діодів D_1 , ..., D_8 , що виконують функції детекторів, які попарно працюють на загальні навантаження.

При передачі частоти f_1 напруга з виходу Φ_1 підводиться через діоди D_1 і D_2 до вхідних затискачів 0_1 і 0_2 апаратів 1-го та 2-го каналів за табл. 9.2. При передачі частоти f_3 напруга з фільтра Φ_3 підключається через діоди D_5 і D_6 відповідно до затискачів 0_1 і 0_2 . Всі інші з'єднання на схемі рис. 9.15 виконані

також відповідно до табл. 9.2. При оптимальному прийомі для розподілу сигналів на частотах f_1, f_2, f_3, f_4 використовують замість смугових фільтрів узгоджені.

Таблиця 9.2

Відповідність частот і фаз номерам кодових комбінацій

Канал 1 (джерело 1)	0	1	0	1
Канал 2 (джерело 2)	0	0	1	1
Номер комбінації	1	2	3	4
ДЧМ	f_1	f_2	f_3	f_4
ДФМ	φ_1	φ_2	φ_3	φ_4

Порівняння системи ДЧМ зі звичайною двоканальною ЧМ системою з ЧРК показує, що обидві системи займають практично однакову смугу частот, однак потужність сигналу, необхідна для забезпечення заданої ймовірності помилки при ДЧМ, майже вдвічі менша, ніж при ЧРК. Істотно меншою виявляється й пікова потужність при ДЧМ. Тому в системах з обмеженою енергетикою комбінаційне ущільнення за методом ДЧМ знаходить широке застосування.

Комбінаційні системи ДФМ на практиці реалізуються у вигляді подвійної відносної фазової модуляції ПВФМ за тими ж принципами, за якими замість абсолютних систем ФМ використовуються відносні - ВФМ.

Аналогічно можна формувати сигнали комбінаційного ущільнення для великої кількості каналів - багаторазову частотну модуляцію (БЧМ), багаторазову відносну фазову модуляцію (БВФМ) та ін.

У випадку БЧМ при виборі частот, що забезпечують ортогональність системи переданих сигналів, зайнята смуга частот при зростанні N збільшується експоненціально. Імовірність помилки в кожному каналі зі збільшенням N також зростає, але дуже повільно. Тому такі системи застосовують у тих випадках, коли використовуваний канал має більші частотні ресурси, але його енергетичні можливості обмежені.

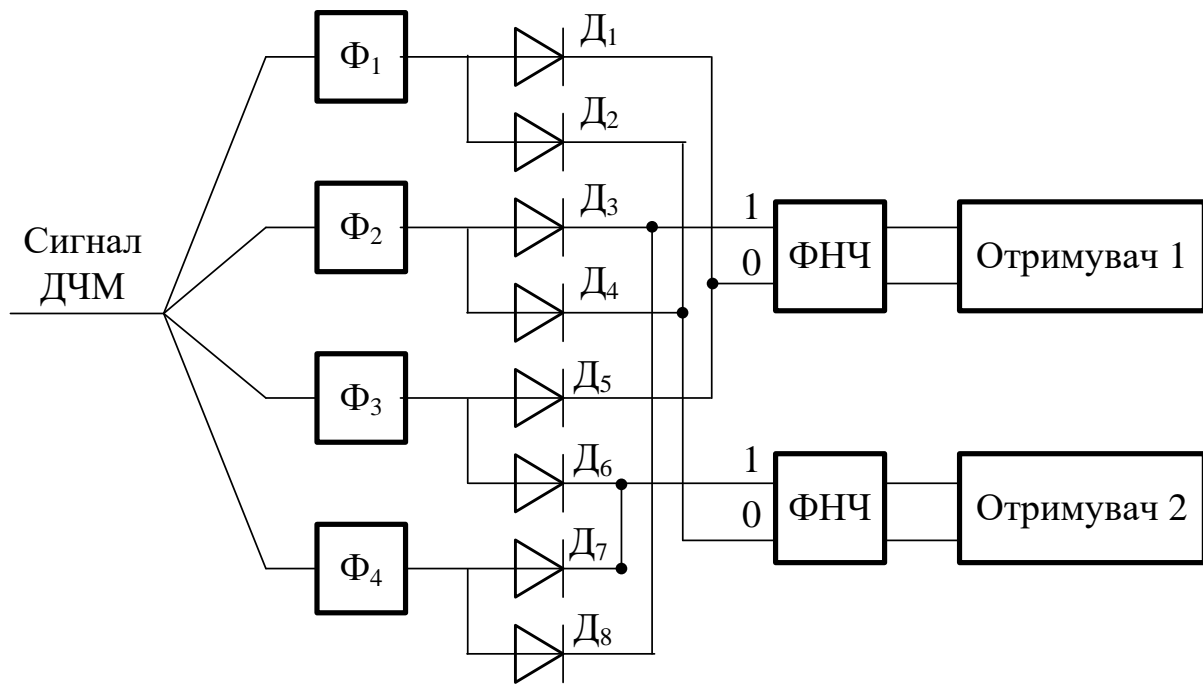


Рис. 9.15. Структурна схема прийому сигналів ДЧМ

У випадку БВФМ, навпаки, зайнята смуга частот з ростом N майже не збільшується, але ймовірність помилки збільшується дуже швидко і для збереження необхідної правильності необхідно збільшувати потужність сигналу. Такі системи придатні в тих ситуаціях, коли існують тверді обмеження смуги пропускання каналу, а потужність сигналу практично не лімітована.

Розглянуті приклади систем комбінаційного ущільнення на основі сигналів БЧМ і БФМ по суті є окремими випадками систем передачі з багатопозиційними сигналами. При БЧМ одержуємо багаточастотні сигнали, а при БФМ - багатофазні.

Також можна модулювати одночасно кілька параметрів переносника, наприклад амплітуду й фазу, частоту й фазу та ін.

Останнім часом великий інтерес являють сигнали амплітудно-фазової модуляції (АФМ), які можна реалізувати схемою квадратурної модуляції. У системах АФМ впродовж інтервалу передачі одного елементарного сигналу його фаза й амплітуда приймають значення, вибрані з ряду можливих дискретних значень амплітуд і фаз. Кожна комбінація значень

амплітуди й фази відображає один з багатопозиційних сигналів групового сигналу з алфавітом коду $M = 2^N$.

Сигнали АФМ можна формувати, наприклад, шляхом багаторівневої амплітудної й фазової модуляції двох квадратурних (зсунутих за фазою на $\pi/2$) коливань несучої частоти. Якщо для модуляції синфазної (їй присвоєний номер 1) і квадратурної (їй присвоєний номер 2) складових використати значення рівнів $C_{1,2} = \pm 1$, то отриманий сигнал АФМ-4 (рис. 9.16) буде відповідати ДФМ. Якщо ж для модуляції як у синфазному, так і у квадратурному каналі використовуються чотирирівневі сигнали $C_{1,2} = \pm 1, \pm 3$, то при цьому виходить 16-позиційна АФМ, яку можна подати як [43]

$$S_{AFM-16}(t) = \{A_i \cos(\omega_0 t + \varphi_i)\} \quad i = \overline{1,16}. \quad (9.26)$$

Розташування сигнальних точок 16-позиційної КАФМ на амплітудно-фазовій площині зображене на рис. 9.17, де точками показані положення кінців вектора A_i при різних значеннях C_1 й C_2 . У принципі для кожного числа M можна будувати різні ансамблі сигналів. Так, крім наведеної на рис. 9.16 і 9.17 квадратурної мережі, розробляються ансамблі сигналів на основі трикутної мережі, різні варіанти кругових розташувань сигнальних точок та ін.

В останні роки успішно розвивається теорія сигнально-кодових конструкцій (СКК), спрямована на підвищення швидкості передачі й завадостійкості при істотних обмеженнях на енергетику й зайняту смугу частот [41].

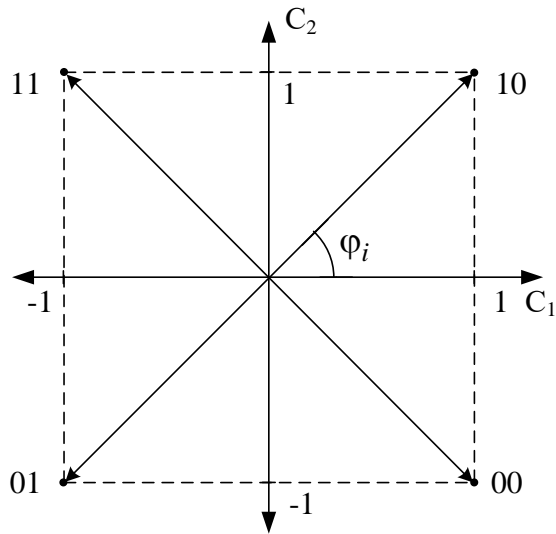


Рис. 9.16. Розташування сигнальних точок сигналів АФМ-4 або ДФМ

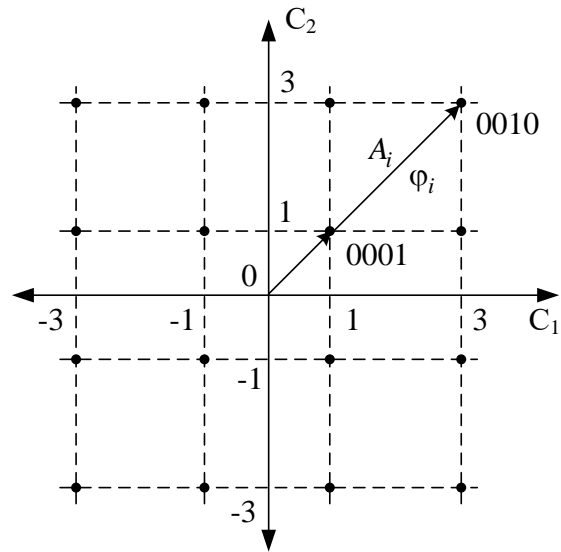


Рис. 9.17 Розташування сигнальних точок при 16-позиційній АФМ

9.8. Кодовий розподіл сигналів

Ефективність систем передачі дискретних повідомлень можна істотно підвищити, застосовуючи багатопозиційні сигнали і коригувальні коди. Вибір сигналів і кодів у цих випадках є визначальним для побудови високоефективних кодеків і модемів. Це досягається збільшенням ширини спектра сигналу, тобто за рахунок частотної надмірності. У всіх цих системах різко виражений поріг завадостійкості.

На рис. 9.18 наведені криві ефективності $\beta = f(\gamma)$ для аналогових і цифрових систем передачі безперервних повідомлень. Ці криві розраховувалися за формулами

$$\beta = \frac{g \log_2 \rho_{вих}}{\alpha \rho_{вих}}; \quad \gamma = \frac{\log_2 \rho_{вих}}{\alpha}. \quad (9.27)$$

Для аналогових систем цифри на кривих показують значення $\alpha = F/F_c$, а цифри в дужках значення $\rho_{вих}$. У пороговій області криві для ЧМ зображені пунктиром. Роботу в цій області можна реалізувати при слідкуючому прийомі.

Для цифрових систем передачі безперервних повідомлень розрахунки проводилися з урахуванням мінімально необхідних смуг каналу F . Практично це відповідає каналу, у якому міжсимвольні завади повністю скоректовані. Цифри на кривих показують основу коду. Криві розраховані для ймовірності помилки $\rho_b = 10^{-5}$, що відповідає відношенню сигналу до шуму $\rho_{вих} \approx 36$ дБ при рівності потужностей шуму квантування й шуму помилкових імпульсів $\overline{E_{кв}^2} = \overline{E_{пi}^2}$ [44].

З рис. 9.18 видно, що ефективність реальних систем істотно нижча від межі Шеннона. Характер обміну між β і γ (вигляд кривих ефективності) залежить від виду модуляції (сигналу) і коду.

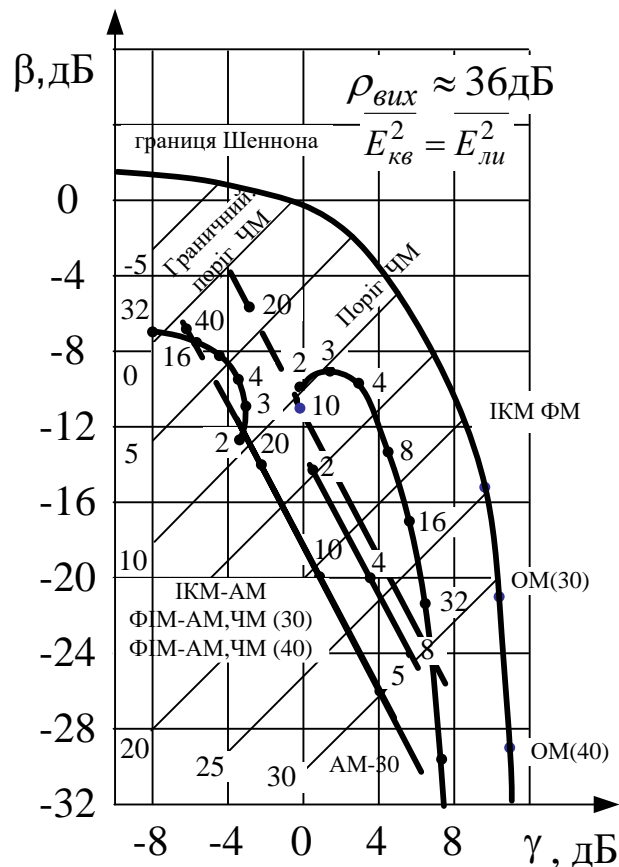


Рис. 9.18. Криві енергетичної та частотної ефективності аналогових і цифрових систем зв'язку

У системах космічного зв'язку визначальною є вимога, відповідно до якої необхідно забезпечити найкраще

використання потужності сигналу при заданій правильності передачі інформації.

Цю умову найповніше задовольняють цифрові системи із ФМ і ВФМ. Ефективність цих систем можна істотно підвищити за рахунок коригувальних кодів. У системах провідного зв'язку найважливішим показником є частотна ефективність. Тут визначальним є вимога, відповідно до якої необхідно домогтися найкращого використання смуги частот каналу при заданій правильності передачі. Цій умові найповніше відповідає односмугова модуляція (ОМ).

Аналогові системи з ОМ, АМ й вузькосмугова ЧМ забезпечують високу частотну ефективність γ при порівняно низькій енергетичній ефективності β . Застосування цих систем доцільно в каналах з гарною енергетикою (при великих значеннях $\rho_{вх}$) або в тих випадках, коли необхідне значення $\rho_{вх}$ мале. Цифрові системи забезпечують високу β -ефективність при досить добрій γ -ефективності. У каналах з обмеженою енергетикою (при малих значеннях $\rho_{вх}$) переваги цифрових систем особливо помітні.

При високій якості передачі, коли необхідні значення $\rho_{вх}$ великі, широкосмугові системи з ЧМ і цифрові системи забезпечують приблизно однакову ефективність.

У багатоканальних системах ефективність зв'язку знижується через недосконалість системи розподілу сигналів. Для таких систем можна користуватися розрахунковою формулою [44]

$$\eta = \eta_m \eta_p = \sum_{i=1}^n R_i / C, \quad (9.28)$$

де η_m — усереднена ефективність методів модуляції по всіх каналах, $\eta_m = \frac{1}{\eta} \sum_{i=1}^n R_i / \frac{1}{n} \sum_{i=1}^n C_i$.

9.9. Використання багатопозиційних сигналів

Ансамбль сигналів $\{s_i(t)\}_{i \in \overline{0, m-1}}$ на відрізку $[0, T]$ можна подати як

$$s_i(t) = \sum_{j=1}^n a_{ij} \varphi_j(t), \quad (9.29)$$

де $\{\varphi_i(t)\}$ - система базисних ортонормованих функцій за умови

$$\int_{-T/2}^{T/2} \varphi_i(t) \varphi_j(t) dt = \begin{cases} 0 & i \neq j, \\ 1 & i = j, \end{cases}$$

n – кількість вимірювань (відліків) на інтервалі T (для фінітних сигналів, енергія яких майже повністю зосереджена в смузї F).

Коефіцієнти розкладання у виразі (9.29) визначаються як проекції вектора s_i , на координатні осі φ_i :

$$a_{ij} = \int_0^T s_i(t) \varphi_j(t) dt.$$

Геометрично кожному сигналу ансамблю $s_i(t)$ відповідає точка (або вектор) в n -мірному просторі з координатами $(a_{i1}, a_{i2}, \dots, a_{in})$, $i = \overline{0, m-1}$. Енергія сигналу $s_i(t)$ дорівнює квадрату норми:

$$E_i = \|s_i\|^2 = \int_0^T s_i^2(t) dt = \sum_{j=1}^n a_{ij}^2, \quad (9.30)$$

а відстань між сигналами становить

$$d_{ij} = \sqrt{\int_0^T (s_i(t) - s_j(t))^2 dt} = \sqrt{E_i + E_j - 2R_{ij} \sqrt{E_i E_j}}, \quad (9.31)$$

де R_{ij} – коефіцієнт взаємної кореляції розглянутих сигналів,

$$R_{ij} = \frac{1}{\sqrt{E_i E_j}} \int_0^T s_i(t) s_j(t) dt.$$

Надалі будемо розглядати ансамблі сигналів з однаковою енергією, тобто $E_i = E$. Прикладом сигналів, у яких сигнальні точки (вектори) розташовуються на прямій, є двійкові протилежні сигнали. Їм відповідає дві симетрично розташовані відносно початку координат точки на прямій з координатами $(0, \sqrt{E})$ й $(0, -\sqrt{E})$. Відстань між сигналами дорівнює $d_{12} = 2\sqrt{E}$, а коефіцієнт кореляції $R_{12} = -1$.

Двійкові ортогональні сигнали є прикладом сигналів, у яких сигнальні точки розташовуються в площині. Їм відповідають два ортогональних вектори на одній площині з координатами $(\sqrt{E}, 0)$ і $(0, \sqrt{E})$. Відстань між сигналами дорівнює $d_{12} = 2\sqrt{E}$, а коефіцієнт кореляції $R_{12} = 0$.

Найпоширенішими багатопозиційними сигналами є ортогональні, біортогональні та симплексні. Якщо сигнальні точки вибрати на лініях, що збігаються з ортами φ_j на відстанях \sqrt{E} від початку координат, то одержимо систему ортогональних сигналів. Кількість сигналів у такому ансамблі $m = n$.

Так, якщо прийняти

$$\varphi_f(t) = \begin{cases} \sqrt{\frac{2}{T}} \sin \omega_j t, & t \in \left[-\frac{T}{2}, T\right], \\ 0, & t \notin \left[-\frac{T}{2}, T\right], \end{cases}$$

то згідно з виразом (9.28) одержимо

$$s_k(t) = \sqrt{\frac{2E}{T}} \sin \omega_k t, \quad 0 < t < T, \quad k = \overline{1, m-1},$$

тобто сигнали є відрізками гармонійних коливань різних частот $\omega_1, \omega_2, \dots, \omega_n$, які задовольняють умову ортогональності. Це відомі сигнали багаточастотної модуляції (БЧМ).

Ортогональні сигнали утворюють еквідистантну систему: відстані між будь-якими двома сигнальними точками однакові й дорівнюють $d_{ij} = \sqrt{2E}$. Перспективним варіантом ЧМ сигналів є частотно-маніпульовані сигнали з безперервною фазою (ЧМБФ).

Біортогональні сигнали утворюються за таким правилом: до кожного ортогонального сигналу додається протилежний. Тут кількість сигналів $m = 2n$. Найпростішим з біортогональних є ансамбль з $m = 4$.

Сигнали мають однакові енергії, а сигнальні точки розташовуються на однаковій відстані від початку координат. На амплітудно-фазовій площині вони утворюють квадратну мережу (на рис. 9.16). При виборі в якості базисних функцій вигляду

$$\varphi_1(t) = \sqrt{2E/T} \cos \omega_0 t ; \quad \varphi_2(t) = \sqrt{2E/T} \sin \omega_0 t$$

сигнали цього ансамблю мають вигляд

$$\begin{aligned} s_1(t) &= \sqrt{2E/T} \cos(\omega_0 t + \pi/4); & s_2(t) &= \sqrt{2E/T} \cos\left(\omega_0 t + \frac{3}{4}\pi\right); \\ s_3(t) &= \sqrt{2E/T} \cos(\omega_0 t - 3\pi/4); & s_4(t) &= \sqrt{2E/T} \cos\left(\omega_0 t - \frac{\pi}{4}\right) \end{aligned}$$

і відрізняються тільки початковими фазами. Це сигнали з фазовою модуляцією та кількістю позицій $m = 4$ (ФМ-4). Відстань між найближчими сигнальними точками цього ансамблю дорівнює $\sqrt{2E}$, а між протилежними сигналами $-\sqrt{2E}$.

Відомі сигнали з амплітудно-фазовою модуляцією (АФМ-4) утворюють кругову мережу: три сигнали рівномірно розподілені по колу, а четвертий розташований у центрі кола (рис. 9.16). У тому ж базисі вони можуть бути подані як

$$s_1(t) = \sqrt{2E/T} \cos \omega_0 t; \quad s_2(t) = \sqrt{2E/T} \cos(\omega_0 t + 2\pi/3);$$

$$s_3(t) = \sqrt{2E/T} \cos(\omega_0 t + 4\pi/3), \quad s_4(t) = 0.$$

Симплексні сигнали знаходяться один від одного на однаковій відстані. У n -мірному просторі вони утворюють правильний симплекс із кількістю вершин $m = n + 1$. У двовимірному просторі сигнальні точки лежать у вершинах рівностороннього трикутника (рис. 9.19). Відстань між сигнальними точками симплексного ансамблю становить $d = \sqrt{2(n+1)E/n}$. При $n = 1$ симплексні сигнали співпадають з протилежними. Для ансамблів з більшим алфавітом m ($n \gg 1$) симплексні сигнали за своїми властивостями і зокрема за завадостійкістю близькі до ортогонального $d = \sqrt{2E}$.

Багатопозиційні сигнали з фазовою модуляцією (ФМ) утворюють кругову мережу з рівномірним розподілом точок по колу. На рис. 9.19 показане також розташування сигнальних точок у 8 - позиційній системі з амплітудно-фазовою модуляцією (АФМ-8).

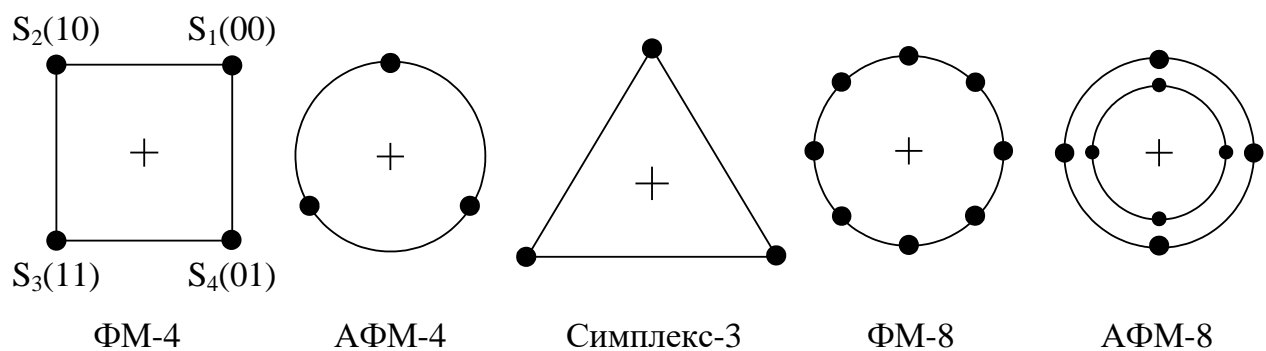


Рис. 9.19. Приклади ансамблів сигналів, відображуваних на амплітудно-фазовій площині

Побудову ансамблів багатопозиційних сигналів можна здійснити й на основі двійкових послідовностей. Для цього як правило, використовують елементарну матрицю Адамара (А), повторенням якої тричі в додатній й один раз у від'ємній формі

можна збільшити розміри матриці кожного разу удвічі й одержати матрицю Б, яка являє собою ансамбль ортогональних сигналів з $m = 4$ [41]

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}; B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}; V = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix}.$$

Кожен рядок цієї матриці (послідовність двійкових символів) утворює один сигнал. Незавжно перевірити, що ці рядки (сигнали) взаємно ортогональні. Доповнюючи матрицю Б інверсіями рядків, одержуємо матрицю В, що являє собою ансамбль $m = 8$ біортогональних сигналів. Аналогічно будуються ансамблі з більшою кількістю сигналів m . Симплексні сигнали також можуть бути отримані на основі двійкових послідовностей.

В асинхронно-адресних системах широко використовуються ансамблі «майже ортогональних» сигналів, які також формуються на основі двійкових послідовностей. Це відомі рекурентні псевдовипадкові M -послідовності.

Раніше були отримані вирази для ймовірності помилки при $m = 2$. Для недвійкових систем $m > 2$ одержати такі прості вирази не вдається. Для деяких ансамблів сигналів у результаті чисельного інтегрування отримані графіки залежності $p = f(h_e^2)$, які наведені на рис. 9.18. Для наближених обчислень можна скористатися верхньою оцінкою: $p_m < (m-1)p_2$, де p_2 – імовірність помилки у двійковій системі, що використовує сигнали $s_0(t)$ і $s_1(t)$.

Варто нагадати, що p_m – це ймовірність помилкового прийому сигналу, а не символу коду. Кожен багатопозиційний

сигнал містить інформацію про блок, в якому присутня певна кількість (ν) двійкових символів. Так, при $\nu = 2$ (випадок ФМ-4) буде чотири блоки 01, 10, 00, 11. Багатопозиційні сигнали, що відповідають цим блокам, вибираються так, щоб мінімізувати ймовірність помилки на двійковий символ (на біт) p_b .

Очевидно, помилки частіше будуть відбуватися за рахунок переходів в області сусідніх сигналів. Тому блоки, що відповідають сусіднім сигналам, повинні відрізнятися найменшою кількістю двійкових символів. Цю умову задовольняє маніпуляційний код Грея, приклад якого наведений на рис. 9.16. Тут протилежним сигналам присвоєні протилежні блоки 11 та 00, а сусіднім (найближчим) 10 та 01. У цьому випадку перехід з будь-якої сигнальної точки в сусідню область призводить до помилки в одному двійковому символі.

При рівноімовірних ортогональних багатопозиційних сигналах ймовірність помилки на біт обчислюється за такою формулою [20]:

$$p_b = \frac{2^{\nu-1}}{2^{\nu-1}} p_m, \quad (9.32)$$

де p_m – ймовірність помилкового прийому багатопозиційного сигналу ортогонального ансамблю;

ν – кількість інформаційних символів.

При тому самому способі прийому різні ансамблі сигналів будуть забезпечувати різну завадостійкість. Пояснюється це особливостями розташування границь областей правильного прийому, які оточують кожен сигнал.

Мінімум середньої ймовірності помилки досягається при розміщенні границь на рівних відстанях від сусідніх сигнальних точок.

Пошук найкращого ансамблю зводиться до знаходження такого розташування сигнальних точок, при якому області сигналів мають найбільшу величину, є найближчими одна до одної за розмірами і наближаються за формою до сфер. Це відома в багатовимірній геометрії задача найщільнішого укладання однакових куль у заданому об'ємі.

Таке розташування забезпечує однакову ймовірність помилки будь-якого сигналу (області сигналів однакові) і мінімальну середню енергію сигналів (області найщільніше впаковані).

На рис. 9.20 наведені графіки β і γ -ефективності для деяких ансамблів багатопозиційних сигналів. Значення енергетичної ефективності β визначалися за кривими завадостійкості для заданого значення ймовірності помилки на біт $p_b = 10^{-5}$ відповідно до виразу (9.27), а частотна ефективність γ визначалася відповідно до виразу

$$\gamma = \frac{\log_2 m}{TF} = \frac{2 \log_2 m}{n} \quad (9.33)$$

Можна виділити два класи багатопозиційних сигналів. До першого віднесемо так звані щільні сигнали, коли при зростанні об'єму ансамблю m при фіксованій мірності n відстань між сигнальними точками зменшується, а питома ефективність γ згідно з виразом (9.33) зростає при відповідному зниженні енергетичної ефективності β .

Залежності β і γ таких сигналів (багатопозиційні ФМ і АФМ) наведені на рис. 9.20.

До другого класу віднесемо біортогональні (БС), симплексні (СС) і ортогональні (ОС) сигнали – це приклади «рознесених» сигналів, коли при збільшенні відстані між сигнальними точками збільшується й енергетична ефективність за рахунок зниження частотної ефективності γ .

Центральне місце на рис. 9.20 займає система із сигналами ФМ-4, які належать до класу біортогональних при $m = 4$. Із простих систем – це найбільш ефективна система ($\gamma = 2$, $\beta = -9,6$ дБ, $\eta \approx 0,47$).

У цифрових мережах система ФМ-4 є найпоширенішою і прийнята як стандарт. Тому при порівняльній оцінці ефективності систем доцільно прийняти за еталон систему з ФМ-4.

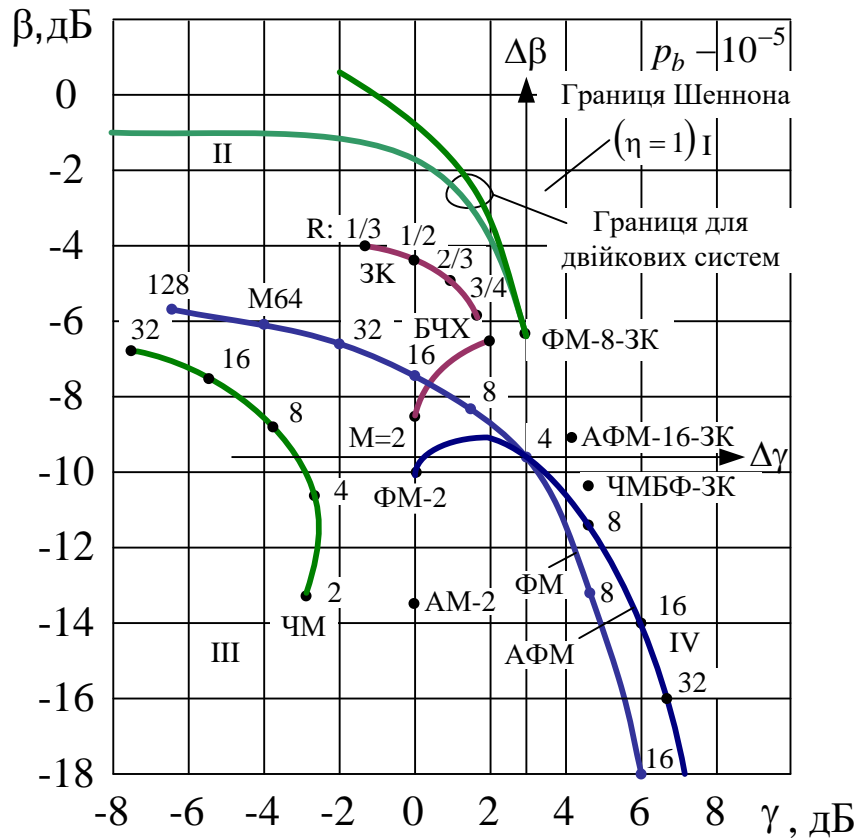


Рис. 9.20. Залежності β і γ систем з багатопозиційними сигналами і коректувальними кодами

Якщо початок координат перенести в точку, що відповідає ФМ-4, то в новій системі координат по вертикальній осі буде йти відлік енергетичного виграшу $\Delta\beta$ для розглянутих систем у порівнянні із ФМ-4, а по горизонтальній осі – виграш $\Delta\gamma$ у питомій швидкості.

У цій системі координат всі можливі системи зв'язку можна умовно розділити на чотири групи, що відповідають чотирьом квадрантам на площині: 1) малоефективні системи (III квадрант), які мають відносно ФМ-4 програш за β і γ (наприклад, АМ-2, ЧМ-2); 2) системи з високою енергетичною ефективністю (II квадрант), які забезпечують виграш за β й програш за γ (системи з коригувальними кодами); 3) системи з високої частотною ефективністю (IV квадрант), які забезпечують виграш за γ й програш за β (системи з багатопозиційними ФМ і АФМ сигналами) і 4) високоефективні системи (I квадрант), які

дозволяють одержати одночасно виграш за обома показниками β і γ (складні сигнально-кодові конструкції).

Наведені на рис. 9.20 залежності дозволяють кількісно оцінити обмінний виграш (програш) різних систем. Так, застосування біртогональних сигналів з $m=16$ дозволяє одержати енергетичний виграш $\Delta\beta = 2,4$ дБ в обмін на зниження питомої швидкості β і γ в 2 рази (3 дБ). Обмін енергетичної ефективності на частотну можна здійснити за допомогою багатопозиційних сигналів із ФМ. Однак ефективнішими є АФМ-сигнали.

9.10. Кориговальні коди

Наряду з багатопозиційними сигналами для підвищення ефективності СПІ широко використовуються завадостійкі коди. Застосування коригувальних кодів дозволяє підвищити правильність передачі повідомлень або при заданій правильності підвищити енергетичну ефективність системи. Останнє є особливо важливим для систем з малою енергетикою (систем супутникового та космічного зв'язку).

На практиці використовуються як блокові, так і безперервні гратчасті коди. Серед блокових кодів найцікавішими є циклічні коди (зокрема коди БЧХ), а серед безперервних – згорткові коди. Розрахунок енергетичної ефективності коду (ЕЕК) визначається за формулами, наведеними у гл. 7.

На рис. 9.20 наведені β - і γ -залежності для циклічного коду (БЧХ) і для згорткового коду (ЗК) з декодуванням за алгоритмом Вітербі. Як бачимо, застосування циклічного коду дозволяє одержати енергетичний виграш $\Delta\beta = 2...3$ дБ, а згорткового коду - $\Delta\beta = 5...6$ дБ в обмін на зниження частотної ефективності в 2 рази (3 дБ).

Застосування каскадних кодів, як показують розрахунки, дозволяє одержати ще більший енергетичний виграш та значно наблизитися до граничної кривої для двійкових систем. Енергетичний виграш $\Delta\beta$ від застосування завадостійкого кодування тим більший, чим вища необхідна правильність передачі.

Для безперервного постійного каналу з "білим" гаусівським шумом при необхідній імовірності помилки $p_b = 10^{-5}$ граничний енергетичний вигравш кодування $\Delta\beta$ в порівнянні із ФМ без кодування й при оптимальному когерентному прийомі становить приблизно 10 дБ. При сучасній елементній базі витрати на реалізацію пристроїв кодування і декодування значно скоротилися.

У той же час вартість енергетичних витрат каналу практично не змінилася. Таким чином, «ціна» вигравшу $\Delta\beta$ за рахунок кодування може бути значно меншою від ціни того самого вигравшу, отриманого за рахунок збільшення енергетики каналу (потужності сигналу або розмірів антен).

Зазначимо, що вибір способів кодування і модуляції залежить від характеристик каналу. Поліпшення цих характеристик (наприклад, вимірюванням завад і спотворень сигналу з їхньою наступною компенсацією) знижує втрати в каналі й створює кращі умови для застосування коригувальних кодів.

Найбільш розробленими є коди для m СК з незалежними помилками. Для будь-якого іншого каналу ці коди малоефективні. Зокрема при передачі дискретних повідомлень по каналу з повільними завмираннями виникає групування помилок. Дискретний канал у цьому випадку буде істотно відрізнятися від m СК – він буде каналом з пам'яттю. У цих умовах можна вибрати код, який виправляє пачки помилок.

Для каналів із змінними параметрами цей код повинен бути адаптивним. У цих випадках перед кодуванням (або тільки перед декодуванням) проводиться вимірювання параметрів каналу й відповідно до їх результатів змінюється алгоритм кодування (або декодування). Вимірювання параметрів каналу може здійснюватися, наприклад, за допомогою спеціального іспитового (зондувального) сигналу [41] або на підставі результатів обробки попередніх відрізків сигналу. У деяких умовах можливі методи кодування, оптимальні або близькі до оптимальних для широкого класу каналів.

Інший шлях полягає в тому, що вихідні канали з пам'яттю приводяться до деякого «стандартного» дискретного каналу, для

якого оптимальні коди відомі. Найпростішим прикладом цього є метод декореляції помилок за допомогою рознесення (перемеження) символів, що входять в один кодовий блок. Цей метод дозволяє перетворити різноманітні дискретні канали приблизно на m СК. Тому в багатьох випадках для ефективного використання кодування може виявитися доцільним за допомогою модема й каналних пристроїв перетворювати вихідні реальні канали на канал, близький до m СК.

Окрім рознесення (перемеження) символів, можна здійснювати й інші перетворення, що мають за мету додати потоку символів властивості, схожі на випадкову послідовність рівноймовірних і незалежних двійкових символів. Така операція називається *скремблюванням*. Його здійснюють пропускаючи потік символів через регістр зсуву із зворотними зв'язками, у якому ці символи «переміщуються» і перетворюються.

Зрозуміло, що характеристика скремблера повинна бути оборотною. Тоді прийняті після демодулятора символи піддаються зворотній операції - дескремблюванню для відновлення вихідної послідовності.

Скремблювання часто застосовують для підвищення надійності системи синхронізації. Але воно також поліпшує роботу декодера, особливо в системах із зворотним зв'язком.

У каналах із завмираннями часто використовують сигнали складної структури, такі, наприклад, як багаточастотні, частотно-часові, шумоподібні й ін. Так, при використанні частотно-часових сигналів за допомогою багаточастотних модемів утворюється двовимірний канал. Передані й прийняті символи в такому каналі розміщуються у вигляді матриці, у якій рядки відповідають різним інтервалам часу, а стовпці - різним частотам.

Варіюючи параметри модема, можна, не змінюючи загальної швидкості передачі інформації, перерозподіляти пам'ять у дискретному каналі, наприклад збільшувати кореляцію між символами за часом і зменшувати кореляцію між символами, близькими за частотою. У таких каналах зручно застосовувати каскадні коди при різних алгоритмах декодування.

Застосування модемів, що створюють двовимірний дискретний канал, є корисним і для безперервних каналів, у яких

мають місце зосереджені й імпульсні завади великої інтенсивності. У цих випадках помилки мають тенденцію вражати символи, що передаються на одній піднесучій (при зосереджених завадах) або в одному інтервалі часу (при імпульсних завадах). Відомі коди, які дозволяють успішно виправляти такі гратчасті конфігурації помилок.

Прикладом узгодженого підходу до кодування й модуляції з урахуванням властивостей безперервного каналу є системи зі зворотним зв'язком. Невід'ємним елементом цих систем є кодування з виявленням помилок.

Практично майже всі діючі системи передачі даних по каналах зі змінними параметрами використовують вирішуючий зворотний зв'язок. У цьому випадку за результатами «вимірювання» стану каналу за допомогою коду автоматично здійснюється зміна режиму роботи системи шляхом перепитування.

Кодування й декодування з виявленням помилок у каналі зі змінними параметрами здійснюється значно ефективніше, ніж безпосереднє виправлення помилок. Однією із причин цього є той факт, що оптимальний алгоритм декодування з виявленням помилок зовсім не залежить від характеристик каналу, імовірність невиявленої помилки мало залежить від цих характеристик і визначається головним чином структурою коду.

Найважливішою проблемою на шляху створення високоефективних систем передачі інформації є проблема узгодження модемів і кодеків з урахуванням статистичних властивостей безперервного каналу. Кодування й модуляцію необхідно розглядати як єдиний процес формування найкращого сигналу, а демодуляцію й декодування - як процес найкращої обробки сигналів.

9.11. Використання сигнально-кодових конструкцій

Багатопозиційні сигнали зі щільним упакуванням (наприклад ФМ, АФМ) забезпечують високу частотну ефективність γ за рахунок зниження енергетичної ефективності β . З іншого боку, коригувальні коди дозволяють підвищити

енергетичну ефективність при певному зниженні питомої швидкості. Кожний із цих способів дає вигоду за одним показником в обмін на погіршення іншого. Разом з тим у багатьох випадках важливим є одночасне підвищення як енергетичної, так і частотної ефективності.

Вирішення цього завдання можливе при використанні ансамблів багатопозиційних сигналів разом із завадостійким кодуванням. При цьому, очевидно, необхідно сформулювати такі сигнальні послідовності, точки яких у багатовимірному просторі щільно упаковані (щоб забезпечити високу частотну ефективність) і достатньо рознесені (щоб забезпечити досить високу енергетичну ефективність).

Такі сигнальні послідовності, побудовані на базі завадостійких кодів і багатопозиційних сигналів, називаються сигнально-кодovими конструкціями (СКК). В якості завадостійких кодів, як правило, використовуються в СКК згорткові та каскадні коди, а в якості багатопозиційних сигналів - сигнали ФМ, АФМ і ЧМБФ.

Пристрій, що реалізує СКК, складається з кодека, модема й узгоджувальних пристроїв. Для узгодження кодека двійкового завадостійкого коду й модема m -позиційних сигналів часто використовується маніпуляційний код Грея, при якому більшій відстані, за Хемінгом, між блоками кодovих символів відповідає більша відстань, за Евклідом, між сигналами.

Код Грея, ввімкнений між завадостійким кодеком і модемом, перетворює m -позиційний канал без пам'яті на двійковий канал з пам'яттю довжиною $\nu = \log_2 m$ символів. Однак код Грея не є оптимальним. Двійкове подання каналних символів вимагає, як правило, нерівного захисту за допомогою коригувального коду. Це обумовлено тим, що використовувані в каналах ансамблі багатопозиційних сигналів у більшості випадків виявляються в місці прийому нееквідистантними.

Нееквідистантними є також і відповідні їм набори двійкових символів маніпуляційного коду. У наш час відомі й інші способи узгодження джерел повідомлень і каналів. Зокрема способи, що ґрунтуються на ієрархічному розбитті ансамблю сигналів на набір вкладених підансамблів з монотонно-зростаючими

відстанями між ними й підбором кодів для кожного рівня ієрархії так, щоб вирівняти результуючі відстані.

Більш ефективним у цьому напрямку є спосіб побудови СКК на основі узагальненого каскадного кодування. У цьому випадку зовнішні завадостійкі коди узгоджуються із внутрішніми кодами, у якості яких використовуються вкладені підансамблі сигналів [18].

Можна будувати СКК і на основі багатомірних сигналів, що дозволяє збільшити кількість позицій сигналу без істотного зменшення відстані між ними. Однак варто пам'ятати, що побудова більш досконалих СКК пов'язана з неминучим ускладненням їхньої реалізації.

Показники ефективності СКК визначаються такими співвідношеннями:

$$\beta = \beta_m + \Delta\beta_{\text{кк}}; \quad \gamma = \gamma_m \gamma_k = \frac{2R \log m}{n},$$

де β_m і γ_m – показники ефективності системи модуляції (модема);

$\Delta\beta_{\text{кк}}$ – енергетичний вигравш кодування (кодека);

γ_k – частотна ефективність кодека.

Результати розрахунків показують (рис. 9.20), що застосування СКК дозволяє одержати одночасно вигравш як в енергетичній, так і в частотній ефективності й у будь-якому разі вигравш за одним показником, не погіршуючи іншого. Так, система ФМ-8-ЗК при використанні перфорованого згорткового коду зі швидкістю $R = 2/3$ забезпечує енергетичний вигравш $\Delta\beta = 2,8$ дБ без зниження питомої швидкості γ , а система АФМ-16-ЗК при $R = 1/2$ й кодовому обмеженні $\nu = 3$ – вигравш у питомій швидкості $\Delta\gamma = 2$ дБ без зниження енергетичної ефективності β . Інформаційна ефективність цих систем $\eta = 0,6 \dots 0,7$.

Значний інтерес являють сигнально-кодові конструкції ЧМБФ-ЗК, побудовані на основі ЧМБФ-сигналів і згорткових кодів. Фазові зміни ЧМБФ сигналів мають вигляд регулярних

граток аналогічно гратчастій діаграмі ЗК. Це дозволяє поєднати в системі ЧМБФ-ЗК процедури демодуляції й декодування шляхом обробки сигналів на прийомі за єдиними сигнально-кодовими гратками з використанням алгоритму Вітербі (АВ) або алгоритму Кловського-Ніколаєва (АКН) [24].

9.12. Оптимізація систем зв'язку

9.12.1. Загальна методологія оптимізації системи

Якість роботи систем передачі інформації характеризується сукупністю великої кількості показників, основними з яких є завадостійкість, швидкість, пропускна спроможність, дальність дії, електромагнітна сумісність, маса й габарити апаратури, вартість, екологічна сумісність.

Сукупність показників якості системи можна записати вектор

$$Q = \{q_1, q_2, \dots, q_n\}. \quad (9.34)$$

Найкращою (оптимальною) вважається така система S , якій відповідає найбільше (найменше) значення деякої функції [37]

$$\varphi \{q_1, q_2, \dots, q_n\} \quad (9.35)$$

від окремих показників якості q_1, q_2, \dots, q_n . Величина Q називається ефективністю або узагальненим показником якості системи, а функція $\varphi \{q_1, q_2, \dots, q_n\}$ – цільовою функцією системи (критерієм якості).

Одним із центральних моментів методики оптимального проектування або порівняння систем є формування оцінок ефективності – цільових функцій системи.

Такі оцінки вкрай необхідні при системних дослідженнях, пов'язаних з такими завданнями, як вибір кращої системи із числа існуючих, оцінка рівня розробки системи у порівнянні з сучасними світовими зразками, визначення оптимального варіанта нової (проектованої) системи й т. ін.

У найпростіших випадках ефективність систем оцінюють за окремими найбільш суттєвим параметрам, наприклад за швидкістю, шириною смуги частот каналу, відношенням сигналу до шуму й т. д.

У загальному випадку необхідним є системний підхід, при якому ефективність оцінюється в цілому за сукупністю параметрів. При цьому насамперед необхідно врахувати усі найістотніші параметри систем. Прагнення врахувати всі параметри, у тому числі дрібні та другорядні, призводить до ускладнення цільової функції (критерію якості) і робить важкодоступними для огляду результати оцінювання. Разом з тим надлишкове обмеження кількості врахованих параметрів може призвести до того, що критерій виявиться занадто грубим.

Будь-яка оцінка ефективності систем проводиться з метою прийняття певного рішення. Так, при проектуванні потрібно визначити сукупність параметрів системи, при яких досягається найбільша ефективність.

Кількісна оцінка ефективності повинна задовольняти певні вимоги. Вона має досить повно характеризувати систему в цілому й мати ясний фізичний зміст, мати необхідну гнучкість й універсальність. Оцінювання ефективності системи повинно бути конструктивним - придатним як для аналізу, так і для синтезу систем. Нарешті, оцінювання ефективності повинно бути досить простим для обчислення й зручним для практичного використання. Розповсюдженим є оцінювання ефективності у вигляді лінійної функції

$$\Psi = \sum_{i=1}^n \lambda_i \alpha_i, \quad (9.36)$$

де n – кількість врахованих, (показників) параметрів;

λ_i - вагові коефіцієнти;

α_i - відносні значення врахованих параметрів.

При такому визначенні параметрів, що входять у суму (9.36), величину Ψ можна визначити в межах від 0 до 1. Найкращою буде та система, для якої величина Ψ більша.

Вибір вагових коефіцієнтів λ_i певною мірою є довільним. Те саме стосується й кількості врахованих параметрів n . Однак їх частка може бути доведена до мінімуму розробкою раціональної методики знаходження цих коефіцієнтів (наприклад методики експертних оцінок). Абсолютні значення ваг не мають значення; істотними є тільки відносні ваги.

Сучасні складні системи зв'язку не завжди можуть бути повністю охарактеризовані одним єдиним показником. Оцінка за кількома показниками може бути більш повною й разом з тим більш предметною, що дозволяє охарактеризувати різні властивості системи.

Цілком очевидно, що велика кількість показників є неприйнятною. Потрібно мати кілька показників, які характеризують основні найсуттєвіші властивості системи: інформаційні, технічні, економічні й т. ін. У багатьох випадках досить обмежитися двома показниками, наприклад завадостійкістю й швидкістю передачі, частотною й енергетичною ефективністю, технічним ефектом і витратами.

Остаточне рішення, як правило, ґрунтується не тільки на кількісних даних розрахунку, але й на досвіді, інтуїції й інших евристичних категоріях, а також на додаткових міркуваннях, які не могли бути враховані при побудові математичної моделі.

У загальному випадку завдання оптимізації СПІ зводиться до знаходження максимуму цільової функції $\Psi(S) = \max \varphi(q_1, q_2, \dots, q_n)$ при варіації системи S (її структури або значень її параметрів) з урахуванням вихідних даних та обмежень на структуру й параметри системи.

Якщо задано цільову функцію $\Psi(S)$ й визначена сукупність допустимих систем (або їхніх варіантів) S_1, S_2, \dots, S_N , то оптимізація зводиться до задачі дискретного вибору з кінцевої кількості N заданих систем, тобто до вибору системи, якій відповідає найбільше (найменше) із значень

$$\Psi_1 = \varphi(S_1), \Psi_2 = \varphi(S_2), \dots, \Psi_N = \varphi(S_N) \dots \quad (9.37)$$

Більш складним завданням є завдання оптимізації (синтезу) структури системи. Якщо структура системи може бути досить

повно описана відомими функціями з кінцевою кількістю параметрів, то завдання зводиться до оптимізації цих параметрів. В окремому випадку, коли цільова функція та всі функції, що визначають обмеження, лінійно залежать від параметрів x_1, x_2, \dots, x_m , завдання зводиться до лінійного програмування. У деяких випадках вдається вирішити завдання аналітично на підставі методів функціонального аналізу.

У загальному вигляді вирішення завдання оптимізації СПІ може виявитися складним і мало придатним для прийняття рішення. Тому, як правило, застосовують поетапну процедуру оптимізації. Спочатку, наприклад, здійснюється оптимізація за інформаційними параметрами, а потім – за техніко-економічними показниками. На першому етапі визначається структурна схема системи, це дозволяє оцінити її основні потенційні характеристики, вибрати способи модуляції й кодування, метод обробки сигналу в приймачі. Потім визначаються алгоритми функціонування й параметри окремих блоків системи (модема, кодека каналу, кодека джерела й т. ін.). Завершальним етапом є конструювання системи.

Техніко-економічний аналіз базується як мінімум на двох показниках: ефекті й витратах [44]. При цьому основними принципами визначення ефективності СПІ може бути принцип максимуму ефекту Ψ або принцип мінімуму витрат W .

За витрати звичайно приймаються приведені річні витрати на одиницю продукції (у нашому випадку - вартість передачі одного біта в секунду).

9.12.2. Оптимізація системи передачі інформації

Корисним ефектом (продуктом) у СПІ є кількість інформації, що доставляється споживачеві в одиницю часу (швидкість передачі) при заданій правильності передачі, тобто середня швидкість передачі по каналу в мережі зв'язку при ймовірності помилки $p \leq p_{\text{доп}}$. Цю швидкість прийнято називати пропускною спроможністю системи й позначати C'_c , на відміну від пропускної спроможності каналу C' , за Шенноном. Якщо C' – поняття теоретичне, що характеризує граничні можливості

каналу, то C'_c – характеристика технічна, вона залежить від реальних характеристик і апаратури даної системи.

За визначенням, пропускна спроможність дискретного каналу зв'язку дорівнює

$$C'_c = \frac{I}{T_{\Pi}} = \eta C' = \eta_{\text{кд}} \eta_{\text{к}} \eta_{\text{м}} C', \quad (9.38)$$

де I – кількість бітів інформації, яка передана по каналу в мережі зв'язку за час T_{Π} , $T_{\Pi} = T_1 + T_3$,

T_1 – час передачі інформації (повідомлення);

T_3 , – час затримки, включаючи час очікування;

$\eta_{\text{кд}}$ – ефективність кодека джерела, $\eta_{\text{кд}} = 1 - \chi_{\text{дп}}$;

$\chi_{\text{дп}}$ – надлишковість джерела повідомлення;

$\eta_{\text{к}}$ – ефективність каналу, обчислена з урахуванням коригувального коду, виду модуляції та втрат у каналі, $\eta_{\text{к}} = \eta_{\text{кк}} \eta_{\text{м}}$;

$\eta_{\text{кк}}$ – ефективність кодека каналу;

$\eta_{\text{м}}$ – ефективність модуляції.

З урахуванням виразу (9.38) маємо

$$C'_c = (1 - \chi_{\text{дп}}) \frac{\gamma \eta_{\text{м}} C'}{\log\left(\frac{\gamma}{\beta + 1}\right)}, \quad (9.39)$$

де

$$\beta = \beta_{\text{м}} + \Delta\beta_{\text{кк}} - \Delta\beta_{\text{пк}};$$

$$\gamma = \gamma_{\text{м}} + \Delta\gamma_{\text{кк}} - \Delta\gamma_{\text{пк}}.$$

При оптимізації СПІ вираз (9.39) для C'_c можна прийняти як цільову функцію $\Psi_1 = C'_c = \varphi(\chi_{\text{дп}}, \beta, \gamma, \eta_{\text{м}}, p, C')$. Тоді завдання буде полягати в тому, щоб знайти таку систему зв'язку, яка доставляє максимум цієї функції при заданих умовах і обмеженнях.

Математично – це завдання нелінійного, а в ряді випадків лінійного програмування. У деяких окремих випадках завдання вирішується аналітично, як завдання пошуку екстремуму функціонала. У тих випадках, коли потрібно забезпечити задану досить високу величину C'_c , вибір системи здійснюється за допомогою аналізу (порівняння) можливих варіантів, які задовольняють задані вимоги.

Необхідна величина C'_c в цих випадках досягається компромісним вибором показників, які входять у вираз (9.39), з урахуванням техніко-економічних вимог.

Завдання оптимізації СПІ виникає як при розробці нових, так і при вдосконаленні існуючих систем. У багатьох випадках воно ставиться як завдання підвищення ефективності СПІ. Рішення такого завдання не є однозначним. Високе (або необхідне) значення C'_c згідно з виразом (9.39) може бути досягнуто різними шляхами.

Розглянемо це на прикладі системи передачі дискретних повідомлень (СПДП). Будемо вважати, що мережа зв'язку, у якій повинна працювати розглянута СПДП, відома (задана її ефективність η_m). Відоме зазвичай і джерело повідомлень (задана його надлишковість $\chi_{дп}$). Заданою є й необхідна правильність (помилка) передачі $p \leq p_{доп}$.

Пропускна спроможність каналу C' є інформаційним ресурсом системи. Вона звичайно задається або вибирається на основі існуючих стандартів. Тут при виборі можливі варіанти. Відповідно до формули Шеннона величина $C' = F \log(r + 1)$ повністю визначається енергетичним ресурсом $r = \frac{P_c}{P_{ш}}$ і частотним ресурсом F . Вибір смуги частот каналу F значно обмежений і регламентується міжнародними угодами.

Що ж до енергетичного ресурсу r , то він залежить від потужності передавача $P_{прд}$ і шумової температури $T_{ш}$ приймача,

а в радіосистемах - і від коефіцієнта підсилення антен G_A :
 $r = AP_{\text{прд}} \frac{G_A}{T_{\text{ш}}}$, де A – постійний коефіцієнт.

Звідси впливає можливість варіювання величинами $P_{\text{прд}}$, G_A і $T_{\text{ш}}$ для одержання необхідного значення C' . Так, застосування вузьконаправлених антен дозволяє істотно поліпшити енергетику каналу при заданих параметрах передавача і приймача.

При обраному значенні C' й заданих значеннях $\eta_{\text{м}}$ і $\chi_{\text{дп}}$ підвищення ефективності СПІ зводиться до підвищення ефективності каналу $\eta_{\text{к}}$, яка залежить від енергетичної ефективності β і питомої швидкості γ . Тоді для заданої ймовірності помилки $p_{\text{доп}}$ і розрахованої величини енергетики каналу r за обмінними $\beta\gamma$ -характеристиками (рис. 9.20) можна вибрати вид модуляції й спосіб кодування.

При цьому необхідно враховувати, що високі показники ефективності β і γ можуть бути досягнуті або за рахунок вибору високоефективних систем модуляції й кодування, або за рахунок зниження втрат у каналі. Практично це означає або ускладнення сигнально-кодових систем (кодеків і модемів), або ускладнення каналних пристроїв, включаючи систему обробки сигналів (компенсаторів завад, оптимальних приймачів і т. ін.). Виникає питання: якому з цих шляхів надати перевагу? Код для каналу чи канал для коду? Очевидно, у реальних умовах необхідно шукати компромісні рішення.

Після того, як обрана система, за показниками β і γ згідно з формулою (9.39) обчислюється C'_c . Визначивши в такий спосіб прийнятні варіанти систем за технічним ефектом $\Psi = C'_c$, розрахуємо витрати W для цих варіантів.

Отримані варіанти розрахунку C'_c і W дозволяють здійснити техніко-економічний аналіз за принципом мінімуму витрат W або максимуму ефекту C'_c .

Варто мати на увазі, що при аналізі систем за двома показниками (у нашому випадку за C'_c та W) рішення, що обертає на максимум (мінімум) один показник, не перетворює на

максимум (мінімум) інший показник. Тому кращою вважається така система, яка забезпечує максимум (мінімум) одного показника при заданому значенні другого показника.

Для більш докладного аналізу зручно результати розрахунків подавати графічно. Нехай нам необхідно зробити вибір з кількох варіантів систем A_1, A_2, \dots, A_r , керуючись двома показниками C'_c і W . Кожному варіанту системи відповідає точка на площині з координатами C'_{ci} та W_i (рис. 9.21). З рис. 9.21 видно, що умову максимуму C'_c при заданому значенні W і відповідно умову мінімуму W при заданому значенні C'_c задовольняють системи, які лежать на лівій і верхній границях області можливих (прийнятних) варіантів (штрихова лінія на рис. 9.21). Решта варіантів, що лежать нижче й правіше від штрихової кривої, цю умову не задовольняють і можуть бути відразу відкинуті.

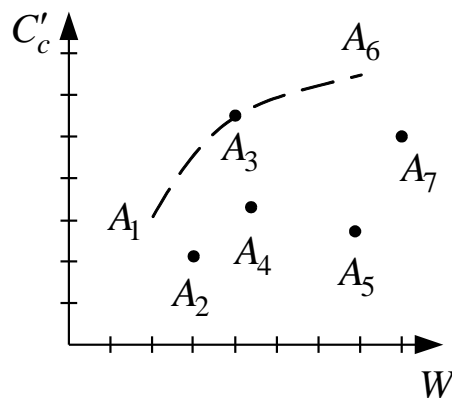


Рис. 9.21. Графічне зображення залежності ефекту системи C'_c від витрат W

Залишається проаналізувати тільки три варіанти: A_1, A_3, A_6 . Варіант A_6 забезпечує найбільшу ефективність C'_c , але вимагає більших витрат W . Варіанти A_3, A_6 мало відрізняються за ефективністю. Затрати при реалізації варіанта A_3 значно менші, і тому, очевидно, варто віддати перевагу варіанту A_3 .

Для прийняття остаточного рішення необхідно враховувати не тільки результати розрахунку, але і цілий ряд інших вимог, пов'язаних з надійністю, екологічністю, технологічністю й т. ін.

Висновки до глави 9

1. Головною проблемою багатоканальної передачі повідомлень є підвищення ефективності використання високовартісних трактів передачі (ліній зв'язку). З цією метою тракт передачі надається для одночасної і незалежної передачі сигналів від великої кількості джерел повідомлень до такої самої кількості одержувачів. Груповий сигнал, що передається по тракту передачі, формується з канальних сигналів, які задовольняють, як правило, умову лінійної незалежності або ортогональності.

2. До основних способів розподілу канальних сигналів належать частотний і часовий розподіл, за фазою і формою, кодовий розподіл. Гранична кількість каналів у багатоканальній системі при одночасній незалежній передачі визначається базою групового сигналу.

3. Пропускна спроможність систем багатоканального зв'язку обмежується не тільки потужністю шуму в каналі, але також потужністю взаємних завад між каналами. Тому збільшити пропускну спроможність багатоканальної системи за рахунок збільшення потужності канальних сигналів не можна. Для зниження рівня канальних завад доводиться вводити захисні проміжки, що призводить до зниження ефективності використання багатоканальних трактів.

4. Найперспективнішими є цифрові багатоканальні системи з часовим розподілом у поєднанні зі статистичним кодуванням і збільшенням кількості користувачів.

5. Характерною рисою системного аналізу є перехід від аналізу окремих частин (пристроїв) системи до аналізу альтернативних варіантів побудови системи як єдиного цілого.

6. У загальному випадку ефективність будь-якої технічної системи визначається кількістю і якістю виданої продукції. У системах зв'язку такою продукцією є передана інформація, кількість якої визначається середньою швидкістю передачі бітів за секунду, а якість – величиною помилки.

7. Найважливішим показником ефективності систем зв'язку є інформаційна ефективність $\eta = R_1 C'$, яка визначає ступінь використання системою пропускну спроможності каналу, а також

показники $\beta = \frac{R_i}{\frac{P_c}{N_0}}$ і $\gamma = R_i/F$, що характеризують відповідно

використання каналу за потужністю (енергетична ефективність) і частотою (частотна ефективність).

8. Залежності між показниками β і γ носять обмінний характер: збільшення одного показника пов'язане зі зменшенням другого і навпаки. Існує гранична залежність між β і γ при $\eta = 1$ (межа Шеннона). Ця залежність відображає найкращий обмін між показниками β і γ у безперервному каналі. У реальних системах ($\eta < 1$) обмін між β і γ залежить від способів модуляції та кодування.

9. Обмінні $\beta\gamma$ -діаграми дозволяють порівняти системи між собою й оцінити ступінь їхнього наближення до ідеальної шеннонівської системи, дозволяють зробити вибір способу модуляції й кодування при заданих умовах, визначити енергетичний і частотний виграш у порівнянні з «еталонною» системою (наприклад ФМ-4).

10. Аналогові системи ОМ, АМ та вузькосмугова ЧМ забезпечують високу частотну ефективність γ при порівняно низькій енергетичній ефективності β . Цифрові системи забезпечують високу енергетичну ефективність при порівняно гарній частотній ефективності. При високій якості передачі цифрові системи й широкосмугова ЧМ забезпечують приблизно однакову ефективність. У багатоканальних системах найефективнішим є метод часового розподілу сигналів, метод частотного розподілу й метод розподілу сигналів за формою.

11. У системах передачі дискретних повідомлень енергетичну ефективність можна істотно підвищити застосовуючи коригувальні коди, а за рахунок застосування багатопозиційних сигналів підвищити частотну ефективність. Застосування каскадних сигнально-кодових конструкцій на основі коригувальних кодів і багатопозиційних сигналів дозволяє підвищити одночасно як енергетичну, так і частотну ефективність. Ефективними, зокрема, є конструкції на основі згорткових кодів і багатопозиційних сигналів із ФМ, АФМ, ЧМБФ.

12. Для скорочення надлишковості джерела безперервних повідомлень широко використовується диференціальне кодування (кодування із передбаченням), що дозволяє істотно підвищити ефективність СП. Так, АДКМ у поєднанні з інтерполяцією мови дозволяє знизити швидкість цифрового потоку з 64 кбіт/с при ІКМ до 16 і навіть 9,6 кбіт/с, а в поєднанні з вокодерами – до 2,4 кбіт/с. Найбільша ефективність досягається при спільному кодуванні джерела й каналу.

13. У високоефективних СП ($\eta > 0,6$) кодек джерела, кодек каналу й модем повинні бути добре узгоджені між собою з урахуванням характеристик безперервного каналу. Кодування й модуляцію варто розглядати як єдиний процес побудови найкращого сигналу, а демодуляцію й декодування – як найкращий спосіб обробки сигналу.

14. Завдання оптимізації СП зводиться до знаходження такого варіанта системи, при якому споживачеві в одиницю часу доставляється максимальна кількість бітів інформації при заданій правильності передачі. Економічним показником при цьому є наведені річні витрати або вартість передачі одного біта в секунду. Зіставлення ефекту й витрат дозволяє вибрати найкращий варіант системи при заданих умовах та обмеженнях.

Контрольні питання та задачі

9.1. Сформулюйте основні переваги багатоканальної передачі і поясніть її структурну схему.

9.2. Які вимоги повинні задовольняти каналні сигнали при формуванні групового сигналу системи багатоканальної передачі?

9.3. Запишіть умови лінійної незалежності сигналів і поясніть її фізичну суть.

9.4. У чому полягає відмінність між ортогональними і лінійно незалежними сигналами?

9.5. Визначте ширину спектра групового сигналу 12-ти каналної системи односмугової ЧРК за умови, що кожне з каналних повідомлень займає смугу частот 300...3400 Гц, а захисний проміжок складає 30 % від рознесення між піднесучими частотами.

9.6. Доведіть, що енергетична ефективність β при $p = const$

для ФМ-2 і ФМ-4 однакова, а частотна ефективність γ відрізняється в два рази (3 дБ).

9.7. Як визначається кількість і якість продукції в системах зв'язку? Як визначаються показники інформаційної, енергетичної і частотної ефективності системи зв'язку?

9.8. Покажіть, що при переході від симплексних сигналів до ортогональних програш за енергетичною і частотною ефективностями порівняно невеликий. Як цей програш залежить від числа сигналів?

9.9. Визначте енергетичний виграш ідеальної системи в порівнянні з ФМ-2 при $\gamma = 1$. Як залежить цей виграш від допустимої імовірності помилки p ? Вказівка: скористайтеся

наближеним виразом для інтеграла $Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{-\frac{x^2}{2}} dx \approx \frac{1}{2} e^{-\frac{\alpha^2}{2}}$,

де для ФМ-2 $\alpha^2 = \frac{2E_b}{N_0}$.

9.10. У гаусівському каналі з ФМ-2 використовується код БЧХ, що коректує (127, 84). Визначте за кривими енергетичний виграш кодування при імовірності помилки $p_b = 10^{-5}$ і обчисліть величину програшу за частотною ефективністю.

9.11. На основі загального визначення ефективності систем розділення сигналів отримайте співвідношення для η_p з урахуванням захисних смуг при частотному розподілі сигналів і захисних часових інтервалів при часовому розподілі сигналів.

9.12. Користуючись графіками рис. 9.18, обчисліть інформаційну ефективність систем ІКМ-ФМ і ІКМ-ЧМ при основі коду $m = 4$ і порівняйте частотну ефективність цих систем.

9.13. Інформаційна ефективність η аналогової ЧМ досягає найбільшого значення в районі порогу $\rho \approx 10$ дБ (рис. 9.18). Обчисліть значення η_{max} , а також відповідні β та γ при $\rho_{вих} = 40$ дБ.

9.14. Доведіть, що при використанні широкосмугових сигналів з базою $2TF \gg 1$ вплив зосереджених за спектром завад зменшується пропорційно базі сигналу.

9.15. Одним із способів підвищення ефективності зв'язку за наявності замирань є рознесений прийом. Визначте енергетичний виграш при переході від одиночного до здвоєного і строєного прийому в каналі з повільними релеєвськими замираннями.

Загальні висновки

Завершивши вивчення даного навчального посібника, можна підвести деякі підсумки, щоб оцінити засвоєний матеріал і намітити шляхи подальшої творчої роботи.

Світ стрімко входить у нову економічну епоху, головною характеристикою якої, не дивлячись на світову кризу, є процеси вироблення (генерації) інформації, її обробки, зберігання, передачі і використання. Конвергенція інформаційних і телекомунікаційних технологій стала основою створення Інтернету – глобальної інфокомунікаційної системи, що перетворила всі сторони економічних, соціальних і політичних відносин розвинених країн. Сьогодні рівень доступності населенню ресурсів мережі Інтернет та інформаційних послуг на її основі розглядається як найважливіша характеристика розвитку і вважається підставою економічного і соціального прогресу країн, умовою ліквідації безробіття, бідності, хвороб.

Із змісту даного посібника зрозуміло, що дисципліна «Теорія електричного зв'язку» використовує поняття і методи з різних наукових областей, перш за все математики, фізики, інформатики, теорії електричних ланцюгів, технічної електродинаміки, мікроелектроніки і схемотехніки, обчислювальної техніки. Всі ці поняття і методи утворюють у курсі ТЕЗ певну єдність і повинні розглядатися як щось ціле в рамках системного підходу, що відповідає вимогам сучасної науки і техніки.

Досі теорія електричного зв'язку розвивалася так, що методи аналізу завжди випереджали методи синтезу. Проте, останніми роками ця тенденція істотно змінюється, головним чином під впливом бурхливого впровадження комп'ютерних технологій у практику наукових досліджень. У наш час розроблення нових систем (пристроїв) формування й обробки сигналів з необхідними параметрами якості найчастіше базується на способі, що містить три етапи: модель – алгоритм – програма.

Повсюдний перехід до цифрових методів обробки сигналів у телекомунікаційних системах при використанні мікропроцесорів призводить до інтеграції радіотехнічних пристроїв засобів обчислювальної техніки.

Працями вітчизняних і зарубіжних вчених та інженерів створені і подалі розробляються фундаментальні методи аналізу, синтезу й оптимізації інформаційних систем. Розвивається математичне моделювання процесів передачі й обробки повідомлень і сигналів. Розробляються й упроваджуються нові методи модуляції, методи аналізу нестационарних безперервних каналів, неоднорідних асиметричних дискретних каналів, неоднорідних асиметричних дискретних каналів з пам'яттю.

Розвивається теорія багатопроменевих каналів, заводостійкі методи й алгоритми передачі інформації, в яких враховується ненадійність апаратури і відхилення характеристик пристроїв від ідеальних. Розробляються математичні моделі складних заводових ситуацій.

Інтенсивно розвиваються коректувальне кодування і оптична обробка сигналів у реальному часі. Подальший розвиток отримують дослідження систем із зворотним зв'язком, адаптивних систем, інформаційно-вимірювальних систем, систем збору й обробки інформації радіотелеметрії, систем штучного інтелекту, робастних і експертних систем.

Основні проблеми і перспективи розвитку фундаментальних розділів дисципліни «Теорія електричного зв'язку» такі:

- подальше розроблення і широке впровадження цифрових методів передачі повідомлень і часового розділення каналів, використання в теорії і техніці зв'язку нових способів і засобів цифрової обчислювальної техніки, мікропроцесорних систем, систем і мереж стільникового зв'язку;
- зростання питомої ваги алгоритмічних і програмних методів управління процесорами передачі інформації;
- розроблення і застосування нових методів формування й обробки широкопasmових і шумоподібних сигналів для систем передачі інформації в складній заводовій обстановці;
- розроблення і застосування нових методів оптичної обробки сигналів;
- широке використання детекторів якості сигналів коректуючого кодування і логічної фільтрації;
- застосування адаптації для оперативної корекції каналів, усунення надмірності повідомлень, статистичного розділення каналів, компенсації завод;

- розроблення методів оцінки ефективності передачі інформації з позицій системного підходу, нових інформаційних технологій;

- застосування цифрового і статистичного моделювання процесів передачі повідомлень для аналізу і синтезу інформаційних систем;

- широке впровадження агрегатно–модульного моделювання і функціонально–модульного принципу побудови інформаційних систем.

Всі перераховані проблеми при їх осмисленні можуть бути певною мірою предметом досліджень найбільш підготовлених студентів в процесі їх участі в науково–дослідницькій роботі на кафедрі, при написанні рефератів, конкурсних і магістерських робіт, а також при їх підготовці для навчання в аспірантурі.

Глибокий і всебічний розвиток теорії електричного зв'язку тісно пов'язані з практичними потребами інформаційної техніки, творче освоєння її методів вченими та фахівцями дозволить підвищувати якість передачі інформації й ефективність експлуатованих і проєктованих телекомунікаційних систем і мереж у народному господарстві, у тому числі і на транспорті.

Бібліографічний список

1. Андреев В.С. Теория нелинейных электрических цепей: Учеб. пособие для вузов. – М.: Радио и связь, 1982. – 280 с.
2. Банкет В.Л., Дорофеев В.М. Цифровые методы в спутниковой связи. – М.: Радио и связь, 1988. – 240 с.
3. Баранник В.В., Ковтун И.В., Королева Н.А. Межкадровое полиадическое кодирование изображений с выявлением серий одинаковых элементов // Открытые информационные и компьютерные интегрированные технологии. – 2006. – Вып. 35. – С. 86 - 91.
4. Баскаков С.И. Радиотехнические цепи и сигналы: Учеб. для вузов. – М.: Высшая школа, 1988. – 448 с.
5. Батаєв О.П. Методика оцінки ефективності рангового виявлення радіосигналів в умовах дії хаотичних імпульсних завад// Зб. наук. праць. –Харків: УкрДАЗТ, 2008. –Вип. 98. – С. 11-19.
6. Батаєв О.П., Воліченко О.С. Математична модель дискретного двійкового каналу без пам'яті багатоканальної системи передачі при впливанні на нього імпульсної завади // Інформаційно–керуючі системи на залізничному транспорті. – Харків: УкрДАЗТ, 2007. – Вип. 4. – С. 113-116.
7. Батаєв О.П., Дешин Р.И., Поляков В.П. Методика оценки вероятности ошибки в каналах с быстрыми замираниями сигналов на базе функций Уолша // Радиотехника. - 2003. - Вып. 133. – С. 182-190.
8. Батаєв О.П., Колесников Б.Г., Корогодін А.М. и др. Системы и средства связи: Учеб. для вузов / Под ред. Б.Г. Колесникова. - МО СССР, 1990. - 588 с.
9. Беллами Дж. Цифровая телефония: Пер. с англ. / Под ред. А.Н. Берлина, Ю.Н. Чернышова. - М.: Эко-Тренд, 2004. - 640 с.
10. Бокер П. Передача данных. Техника связи в системах телеобработки данных: Пер. с нем. / Под ред. Д.Д. Кловского. - М.: Связь, 1980. - Т.1 - 2.
11. Варакин Л.Е. Системы с шумоподобными сигналами. - М.: Радио и связь, 1985. – 384 с.

12. Витерби А.Д., Омура Д.К. Принципы цифровой связи и кодирование: Пер. с англ. / Под ред. К.Ш. Зигангирова. - М.: Радио и связь, 1982. – 526 с.

13. Возенкрафт Дж., Джекобс И. Теоретические основы техники связи: Пер. с англ. / Под ред. Р.Л. Добрушина - М.: Мир, 1969. – 640 с.

14. Волков В.М., Головин Э.С., Кудряшов В.А. Электрическая связь и радио на железнодорожном транспорте: Учеб. для студентов вузов ж.-д. трансп. - М.: Транспорт, 1991. - 311 с.

15. Галлагер Р. Теория информации и надежная связь: Пер. с англ. / Под ред. М.С. Пинскера и Б.С. Цыбакова. - М.: Сов. радио, 1974. – 720 с.

16. Гуаэр Дж. Оптические системы связи: Пер. с англ. / Под ред. А.И. Ларкина. - М.: Радио и связь, 1989. – 501 с.

17. Гоноровский И.С. Радиотехнические цепи и сигналы: Учеб. для вузов. - М.: Радио и связь, 1986. – 512 с.

18. Денисенко А.Н. Сигналы. Теоретическая радиотехника. Справочное пособие. - М.: Горячая линия - Телеком, 2005. – 704 с.

19. Зюко А.Г., Кловский Д.Д., Назаров М.В., Финк Л.М. Теория передачи сигналов: Учеб. для вузов. - М.: Радио и связь, 1986. – 302 с.

20. Зюко А.Г. Помехоустойчивость и эффективность систем связи. - М.: Связь, 1972. – 359 с.

21. Информационные технологии в радиотехнических системах: Учеб. пособие / В.А. Васин, И.Б. Власов, Ю.М. Егоров и др.; Под ред. И.Б. Федорова. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2003. – 672 с.

22. Каганов В.И., Битюков В.К. Основы радиотехники и связи: Учеб. пособие для вузов. – М.: Горячая линия – Телеком, 2007. – 542 с.

23. Кларк Дж. мл., Кейн Дж. Кодирование с исправлением ошибок в системах цифровой связи: Пер. с англ. / Под ред. Б.С. Цыбакова. - М.: Радио и связь, 1987. – 392 с.

24. Кловский Д.Д. Передача дискретных сообщений по радиоканалам: - 2-е изд. - М.: Радио и связь, 1982. – 304 с.

25. Кловский Д.Д., Шилкин В.А. Теория электрической связи: Сб. задач и упражнений. - М.: Радио и связь, 1990. – 280 с.
26. Ковтун І.В., Корольова Н.А. Обґрунтування необхідності використання міжкадрового поліадичного кодування відеоданих у системах транкінгового зв'язку // Зб. наук. праць. – Харків: УкрДАЗТ, 2006. – Вип. 78 – С. 102 - 112.
27. Колесник В.Д., Полтырев Г.Ш. Курс теории информации. - М.: Наука, 1982. – 416 с.
28. Коржик В.И., Финк Л.М., Щелкунов К.Н. Расчет помехоустойчивости передачи дискретных сообщений: Справочник. - М.: Радио и связь, 1981. – 231 с.
29. Котельников В.А. Теория потенциальной помехоустойчивости. – М.; Л.: Госэнергоиздат, 1956. – 152 с.
30. Левин Б.Р. Теоретические основы статистической радиотехники. - М.: Радио и связь, 1989. – 653 с.
31. Назаров М.В., Прохоров Ю.Н. Методы цифровой обработки и передачи речевых сигналов. - М.: Радио и связь, 1985. – 176 с.
32. Нефедов В.И. Основы радиоэлектроники и связи: Учеб. для вузов. – 2-е изд., перераб. и доп. – М.: Высш. шк., 2002. – 510 с.
33. Основы теории информации и кодирования / И.В. Кузьмин, В.А. Кедрус. – 2-е изд., перераб. и доп. – К.: Вища шк. Головне вид-во, 1986. – 238 с.
34. Панфілов І.П. та ін. Теорія електричного зв'язку: Підруч. для студентів вищ. навч. закл. I та II рівнів акредитації за напрямом "Телекомунікації" / І.П. Панфілов, В.Ю. Дирда, А.В. Капацін. – К.: Техніка, 1998. – 328 с.
35. Першин В.Т. Основы современной радиоэлектроники: Уч. пособие. – Ростов н/Д.: Феникс, 2009. – 541 с.
36. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки: Пер. с англ. / Под ред. Р.Л. Добрушина и С.И. Самойленко. - М.: Мир, 1976. – 596 с.
37. Перегудов Ф.И., Тарасенко Ф.П. Введение в системный анализ: Учеб. пособие для вузов. – М.: Высшая школа, 1989. – 368 с.

38. Помехоустойчивость и эффективность систем передачи информации / Под ред. А.Г. Зюко. - М.: Радио и связь, 1985. – 279 с.
39. Прокис Джон. Цифровая связь: Пер. с англ. / Под ред. Д.Д. Кловского. - М.: Радио и связь, 2000. – 800 с.
40. Поляков П.Ф. Прием сигналов в многолучевых каналах. - М.: Радио и связь, 1986. – 248 с.
41. Скляр Бернард. Цифровая связь. Теоретические основы и практическое применение: Пер. с англ. – 2-е изд. - М.: Издательский дом «Вильямс», 2003. – 1104 с.
42. Стеклов В.К., Беркман Л.Н. Теорія електричного зв'язку: Підруч. для вузів. – К.: Техніка, 2006. – 550 с.
43. Телекоммуникационные системы и сети : Учеб. пособие / Под ред. В.П. Шувалова – М.: Горячая линия-Телеком, 2005. - 3т.
44. Теория электрической связи: Учеб. для вузов / А.Г. Зюко, Д.Д. Кловский, В.И. Коржик, М.В. Назаров; Под ред. Д.Д. Кловского. - М.: Радио и связь, 1999. – 432 с.
45. Тимченко О.В. Методи різницевого кодування форми сигналів у системах передачі мовної інформації. – Львів: Вид-во Української академії друкарства, 2006. – 320 с.
46. Тихонов В.И. Статистическая радиотехника. - М.: Радио и связь, 1982. – 624 с.
47. Тихонов В.И., Харисов В.Н. Статистический анализ и синтез радиотехнических устройств и систем: Учеб. пособие для вузов. - М.: Радио и связь, 1991. – 608 с.
48. Финк Л.М. Теория передачи дискретных сообщений. - М.: Сов. радио, 1970. – 727 с.
49. Харкевич А.А. Избранные труды. Т.3. Теория информации. Оpozнание образов. - М.: Наука, 1972. – 524 с.
50. Шеннон К. Работы по теории информации и кибернетике: Пер. с англ. / Под ред. Н.А. Железнова. - М.: Иностранная литература, 1963. – 829 с.
51. Шварц М. Сети связи. Протоколы. Моделирование и анализ: Пер. с англ. / Под ред. В.И. Неймана. - М.: Наука, 1992. - Т. 1-2.

Прості розривні функції

№ пп	Назва функції*	Аналітичний вираз функції	Графічна форма функції	Зв'язок між функціями
1	Функція знаку $\text{sign}(t)$ (сигнум- функція)	$\text{sign}(t) = \begin{cases} 1; t > 0, \\ 0; t = 0, \\ -1; t < 0. \end{cases}$		-
2	Одинична функція (функція Хевісайда) $\sigma(t)$	$\sigma(t) = \begin{cases} 1; t > 0, \\ 1/2; t = 0, \\ 0; t < 0. \end{cases}$		$\sigma(t) = \frac{1}{2} + \frac{1}{2} \text{sign}(t);$ $\sigma(t) = \int_{-\infty}^t \delta(t) dt.$
3	Дельта- функція (функція Дірака), $\delta(t)$	$\sigma(t) = \begin{cases} \infty; t = 0, \\ \int_{-\infty}^{\infty} \delta(t) dt = 1. \\ \int_{-\infty}^{\infty} S(t) \delta(t_0) dt = S(t_0). \end{cases}$		$\delta(t) = \frac{d\sigma(t)}{dt}$
4	Прямо- кутний відеоімпульс з одиничною висотою, $\text{rect}(t/\tau_i)$	$\text{rect}(t/\tau_i) = \begin{cases} 1; t \leq \tau_i/2, \\ 0; t > \tau_i/2. \end{cases}$		$\text{rect}(t/\tau_i) = \sigma\left(t + \frac{\tau_i}{2}\right) - \sigma\left(t - \frac{\tau_i}{2}\right).$

Примітка. *Функції можуть мати довільний аргумент, наприклад частоту ω .

Додаток 3

Таблиця Д.3.1

Значення інтегралів імовірності

Табульовані значення функції Крампа $F(v) = \frac{2}{\sqrt{2\pi}} \int_0^v \exp(-t^2/2) dt$

Інтеграл Лапласа $\Phi(v) = 0,5[1 + F(v)]$

v	$F(v)$	v	$F(v)$	v	$F(v)$
0,00	0,0000	1,25	0,7887	2,50	0,9876
0,05	0,0399	1,30	0,8064	2,55	0,9892
0,10	0,0797	1,35	0,8230	2,60	0,9907
0,15	0,1192	1,40	0,8385	2,65	0,9920
0,20	0,1585	1,45	0,8529	2,70	0,9931
0,25	0,1974	1,50	0,8664	2,75	0,9940
0,30	0,2358	1,55	0,8789	2,80	0,9949
0,35	0,2737	1,60	0,8904	2,85	0,9956
0,40	0,3108	1,65	0,9011	2,90	0,9963
0,45	0,3473	1,70	0,9109	2,95	0,9968
0,50	0,3829	1,75	0,9199	3,00	0,99730
0,55	0,4177	1,80	0,9281	3,10	0,99806
0,60	0,4515	1,85	0,9357	3,20	0,99863
0,65	0,4843	1,90	0,9426	3,30	0,99903
0,70	0,5161	1,95	0,9488	3,40	0,99933
0,75	0,5467	2,00	0,9545	3,50	0,99953
0,80	0,5763	2,05	0,9596	3,60	0,99968
0,85	0,6047	2,10	0,9643	3,70	0,99978
0,90	0,6319	2,15	0,9684	3,80	0,99986
0,95	0,6579	2,20	0,9722	3,90	0,99990
1,00	0,6827	2,25	0,9756	4,00	0,99994
1,05	0,7063	2,30	0,9786	4,417	$1 \cdot 10^{-5}$
1,10	0,7287	2,35	0,9812	4,892	$1 \cdot 10^{-6}$
1,15	0,7499	2,40	0,9836	5,327	$1 \cdot 10^{-7}$
1,20	0,7699	2,45	0,9857	6,00	$1 \cdot 10^{-8}$

Додаток 6

Таблиця Д.6.1

Значення логарифмів

Таблиця значень логарифмів цілих чисел

x	$\log_2 x$	x	$\log_2 x$
1	0	1024	10
2	1	2048	11
3	1,58496	4096	12
4	2	8192	13
8	3	16384	14
16	4	32768	15
32	5	65536	16
64	6	131072	17
128	7	262144	18
256	8	524288	19
512	9	1048576	20

Таблиця Д.6.2

Значення величин $-p \log_2 p$

p	$-p \log_2 p$	p	$-p \log_2 p$	p	$-p \log_2 p$
0,01	6,643	0,12	3,059	0,23	2,120
0,02	5,644	0,13	2,943	0,24	2,059
0,03	5,059	0,14	2,836	0,25	2,000
0,04	4,644	0,15	2,737	0,26	1,943
0,05	4,322	0,16	2,644	0,27	1,889
0,06	4,059	0,17	2,556	0,28	1,836
0,07	3,936	0,18	2,474	0,29	1,786
0,08	3,644	0,19	2,396	0,30	1,737
0,09	3,474	0,20	2,322	0,31	1,690
0,10	3,322	0,21	2,252	0,32	1,644
0,11	3,184	0,22	2,184	0,33	1,599

Додаток 7

Таблиця Д.7.1

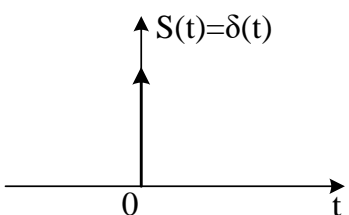
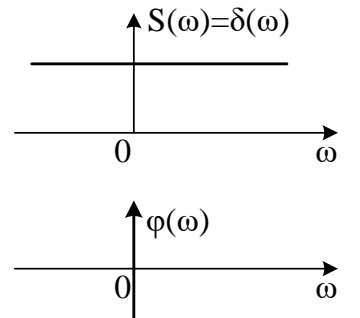
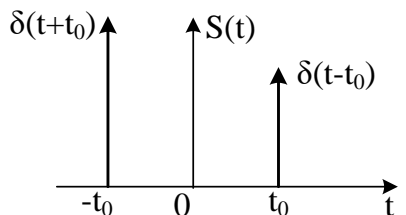
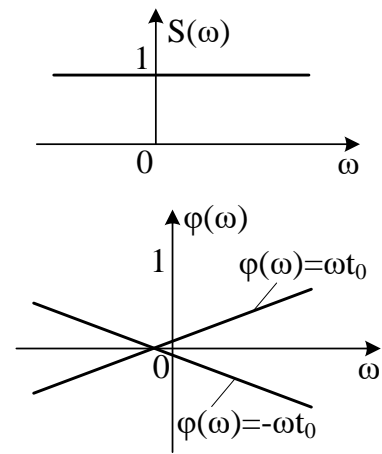
Поліноми, що не приводяться

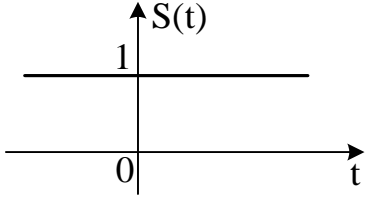
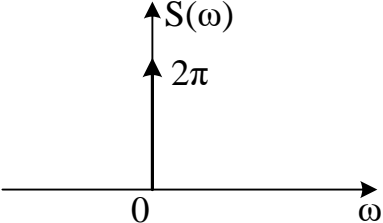
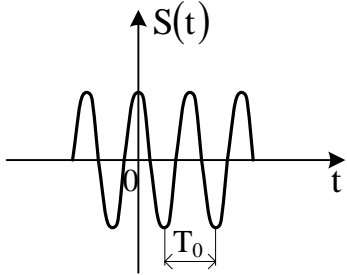
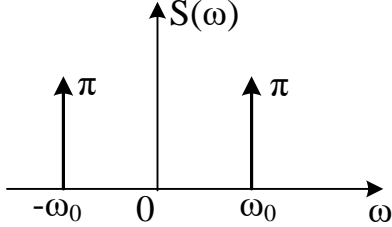
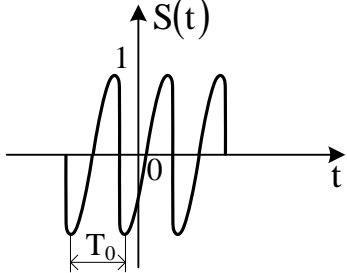
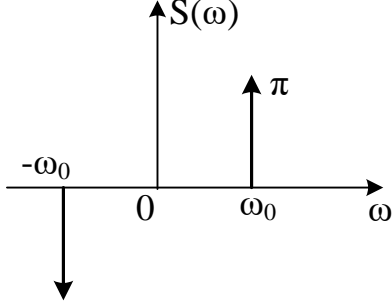
Степінь створюючого полінома r	Вид полінома, що не приводиться $P(x)$	Значність коду $n = l = 2^r - 1$
1	$x+1$	1
2	x^2+x+1	3
3	x^3+x^2+1 x^3+x+1	7
4	x^4+x+1 x^4+x^3+1 $x^4+x^3+x^2+x+1$	15
5	x^5+x^2+1 x^5+x^3+1 $x^5+x^4+x^2+1$ $x^5+x^3+x^2+x+1$ $x^5+x^4+x^2+x+1$ $x^5+x^4+x^3+x+1$ $x^5+x^4+x^3+x^2+1$	31
6	x^6+x+1 x^6+x^3+1 x^6+x^5+1 $x^6+x^4+x^2+x+1$ $x^6+x^4+x^3+x+1$ $x^6+x^5+x^2+x+1$ $x^6+x^5+x^3+x^2+1$ $x^6+x^5+x^4+x+1$ $x^6+x^5+x^4+x^2+1$	63
7	x^7+x+1 x^7+x^6+1	127
8	$x^8+x^4+x^3+x^2+1$ $x^8+x^6+x^5+x^4+1$	255
9	x^9+x^4+1 x^9+x^5+1	511
10	$x^{10}+x^3+1$ $x^{10}+x^7+1$	1023

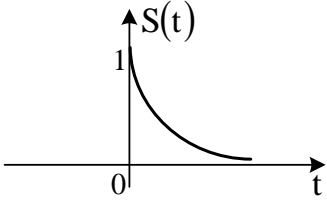
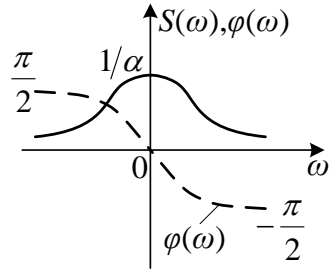
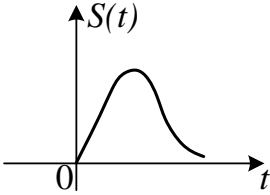
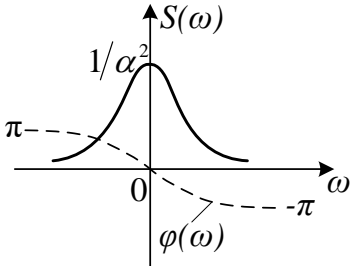
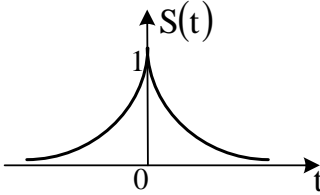
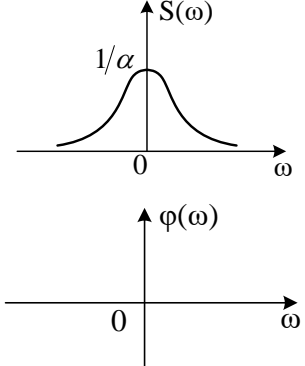
Додаток 2

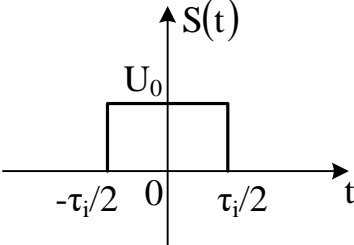
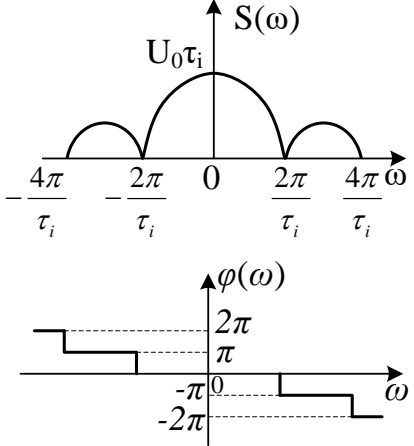
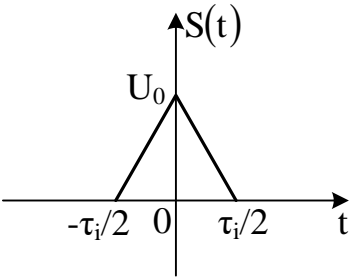
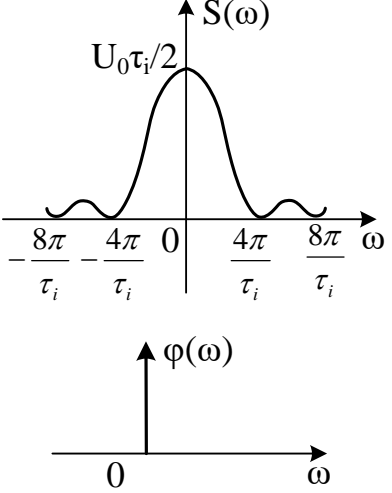
Таблиця Д.2.1

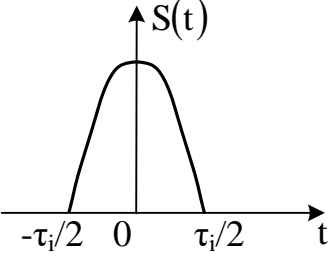
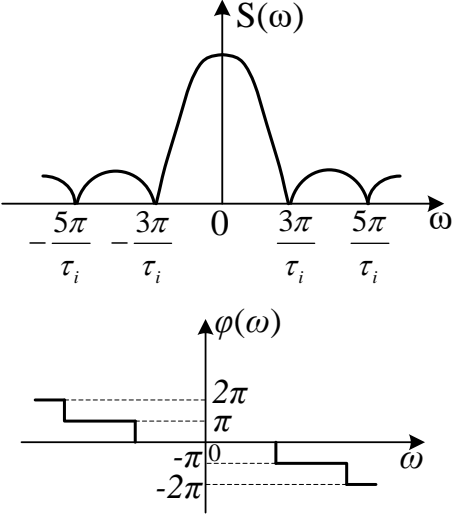
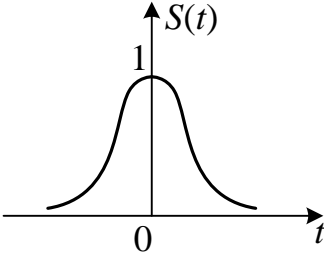
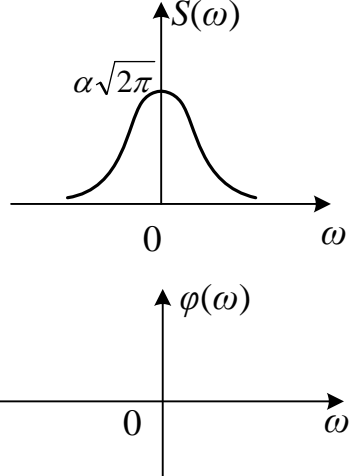
Спектри та спектральна щільність сигналів різної форми

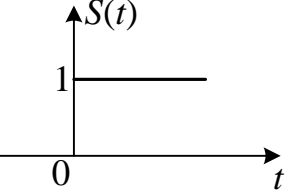
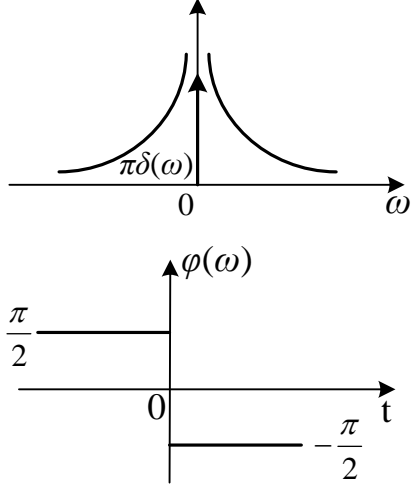
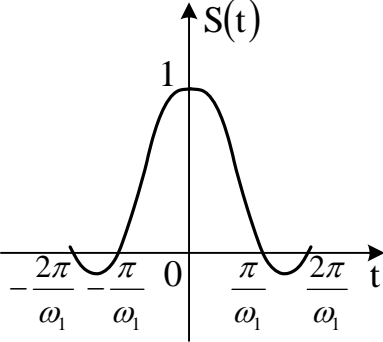
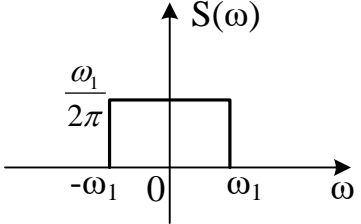
№ пп	Сигнал $S(t)$		$S(j\omega)$	
	Аналітичний вираз	Графічне зображення	Аналітичний вираз	Графічне зображення АЧХ, ФЧХ
1	2	3	4	5
1	$S(t) = \delta(t),$ $t = 0$		$S(j\omega) = 1$	
2	$S(t) = \delta(t - t_0),$ $t = t_0$		$S(j\omega) = 1e^{-j\omega t_0}$	

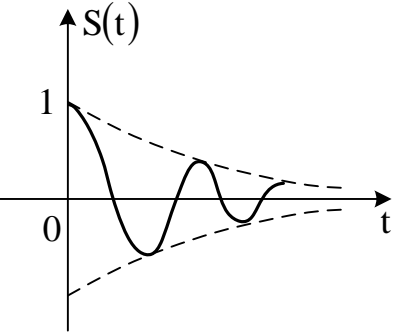
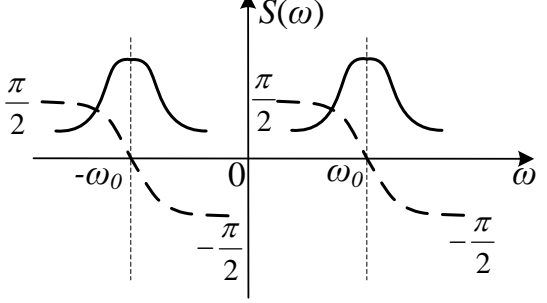
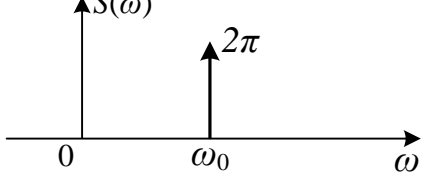
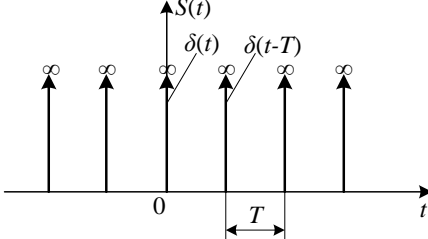
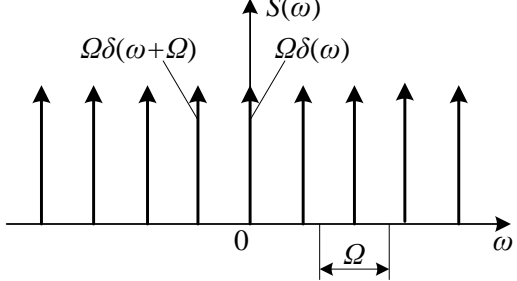
1	2	3	4	5
3	$S(t) = 1 = \text{const},$ $-\infty < t < \infty$		$S(j\omega) = 2\pi\delta(\omega)$	
4	$S(t) = \cos \omega_0 t,$ $T_0 = \frac{2\pi}{\omega_0} = \frac{1}{f_0},$ $-\infty < t < \infty$		$S(j\omega) = \pi\delta(\omega - \omega_0) + \pi\delta(\omega + \omega_0)$	
5	$S(t) = \sin \omega_0 t,$ $T_0 = \frac{2\pi}{\omega_0} = \frac{1}{f_0},$ $-\infty < t < \infty$		$S(j\omega) = \frac{\pi}{j}\delta(\omega - \omega_0) - \frac{\pi}{j}\delta(\omega + \omega_0)$	

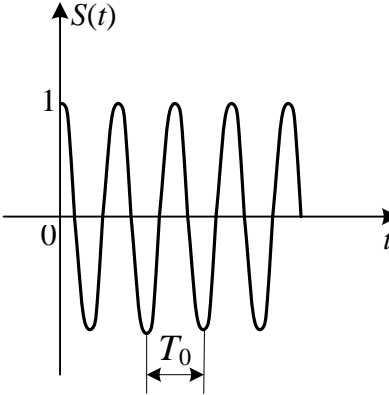
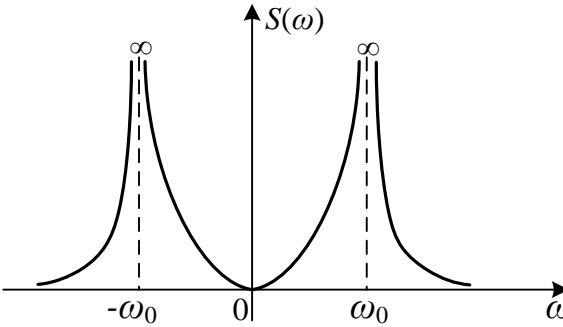
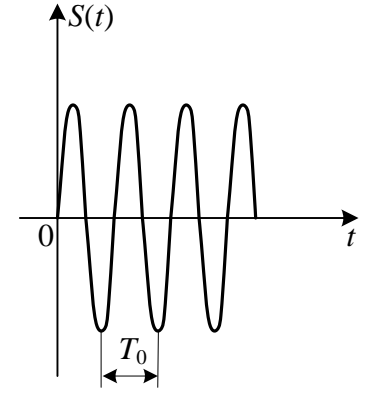
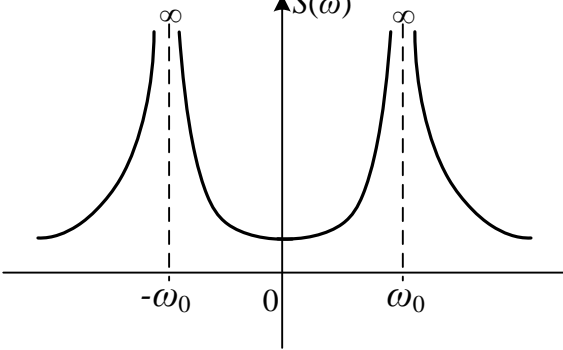
1	2	3	4	5
6	$S(t) = 1 \cdot e^{-\alpha t} \times \sigma(t),$ $t \geq 0$		$S(j\omega) = \frac{1}{\alpha + j\omega}$	
7	$S(t) = t e^{-\alpha t} \sigma(t),$ $t \geq 0$		$S(j\omega) = \frac{1}{(\alpha + j\omega)^2}$	
8	$S(t) = 1 \cdot e^{-\alpha t },$ $-\infty < t < \infty$		$S(j\omega) = \frac{2\alpha}{\alpha^2 + \omega^2}$	

1	2	3	4	5
9	$S(t) = \begin{cases} U_0, & t \leq \frac{\tau_n}{2}, \\ 0, & t > \frac{\tau_n}{2}. \end{cases}$		$S(j\omega) = U_0 \tau_n \times \frac{\sin \omega \tau_n / 2}{\omega \tau_n / 2}$	
10	$S(t) = U_0 \begin{cases} \left(1 - \frac{ t }{\tau_n/2}\right), & t \leq \tau_n/2, \\ 0, & t > \tau_n/2. \end{cases}$		$S(j\omega) = \frac{U_0 \tau_n}{2} \times \left(\frac{\sin \omega \tau_n / 4}{\omega \tau_n / 4} \right)^2$	

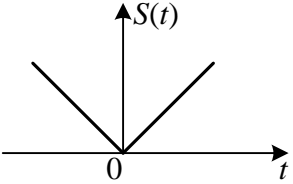
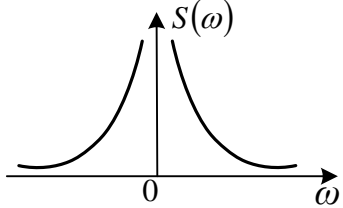
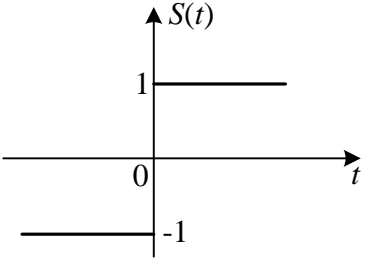
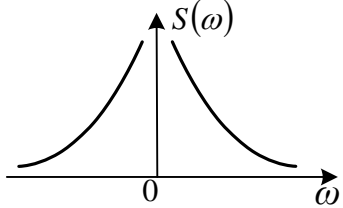
1	2	3	4	5
11	$S(t) = U_0 \begin{cases} \cos \omega_0 t, & t \leq \tau_n/2, \\ 0, & t > \tau_n/2. \end{cases}$		$S(j\omega) = \frac{2\tau_n/2}{\pi} \times \frac{\cos \omega \tau_n/2}{\left(\frac{\pi}{4}\right)^2 - \omega^2}$	
12	$S(t) = e^{-t^2/2\alpha^2}, \quad -\infty < t < \infty$		$S(j\omega) = \alpha\sqrt{2\pi} e^{-\frac{\alpha^2\omega^2}{2}}$	

1	2	3	4	5
13	$S(t) = \sigma(t),$ $t \geq 0$		$S(j\omega) = \pi\delta(\omega) + \frac{1}{j\omega}$	
14	$S(t) = \frac{\sin \omega_1 t}{\omega_1 t} =$ $= \text{sinc}(\omega_1 t),$ $-\infty < t < \infty$		$S(j\omega) = U_0 \begin{cases} \omega_1/2\pi, & \omega \leq \omega_1, \\ 0, & \omega > \omega_1. \end{cases}$	

1	2	3	4	5
15	$S(t) = e^{-\alpha t} \cos \omega_0 t \sigma(t),$ $t \geq 0$		$S(j\omega) = \frac{1}{2} \times$ $\times \left[\frac{1}{\alpha + j(\omega - \omega_0)} + \frac{1}{\alpha + j(\omega + \omega_0)} \right]$	
16	$S(t) = e^{j\omega_0 t},$ $-\infty < t < \infty$		$S(j\omega) = 2\pi \delta(\omega - \omega_0)$	
17	$S(t) = \sum_{h=-\infty}^{\infty} \delta(t - hT),$ $T = 2\pi/\Omega,$ $-\infty < t < \infty$		$S(j\omega) =$ $= \Omega \sum_{n=-\infty}^{\infty} \delta(\omega - n\Omega)$	

1	2	3	4	5
18	$S(t) = \cos \omega_0 t \sigma(t),$ $T_0 = 2\pi/\omega_0,$ $t \geq 0$		$S(j\omega) = \frac{\pi}{2} [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] + \frac{j\omega}{\omega_0^2 - \omega^2}$	
19	$S(t) = \sin \omega_0 t \sigma(t),$ $T_0 = 2\pi/\omega_0,$ $t \geq 0$		$S(j\omega) = \frac{\pi}{2j} [\delta(\omega - \omega_0) - \delta(\omega + \omega_0)] + \frac{\omega_0}{\omega_0^2 - \omega^2}$	

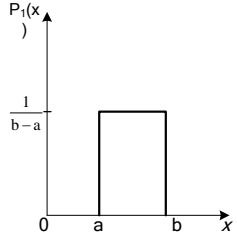
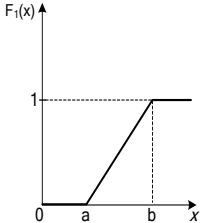
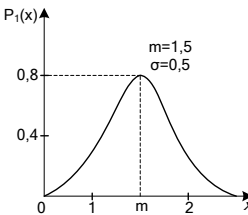
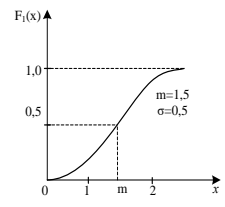
Продовження табл. Д.2.1

1	2	3	4	5
20	$S(t) = t ,$ $-\infty < t < \infty$		$S(j\omega) = -\frac{2}{\omega^2}$	
21	$S(t) = \text{sign}(t) =$ $= \begin{cases} 1, & t > 0, \\ -1, & t < 0 \end{cases}$		$S(j\omega) = \frac{2}{j\omega}$	

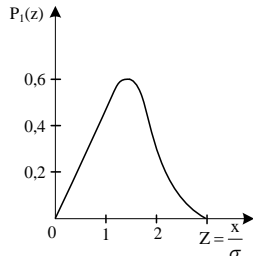
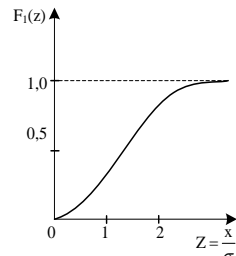
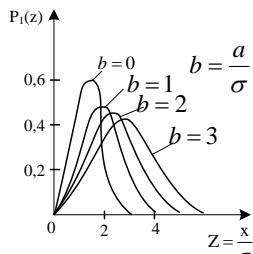
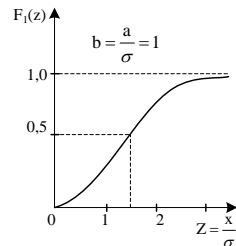
Додаток 4

Таблиця Д.4.1

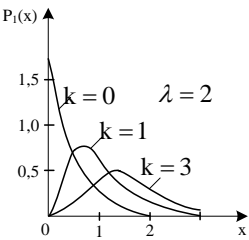
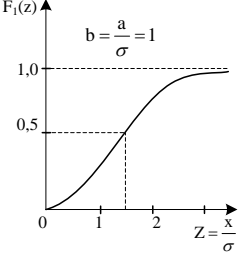
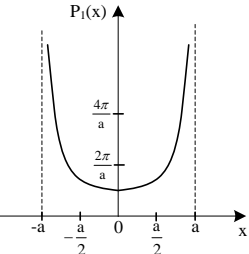
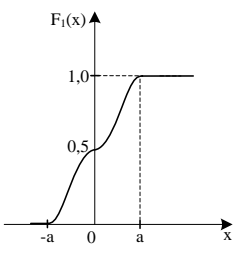
Функції розподілу безперервних випадкових величин

№ п/п	Функції розподілу	Область значень випадкової величини	Аналітичний вираз щільності імовірності $p_1(x)$	Визначальні параметри	Графік щільності імовірності $p_1(x)$	Аналітичний вираз функції розподілу $F_1(x)$	Графік функції розподілу $F_1(x)$	Характеристична функція $\Theta_1(jv)$
1	2	3	4	5	6	7	8	9
1	Рівномірний прямокутний	$a < x < b$	$\frac{1}{b-a}$	a, b		$\begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a < x < b, \\ 1, & x > b. \end{cases}$		$\frac{e^{jbv} - e^{jav}}{jv(b-a)}$
2	Гаусівський (нормальний)	$-\infty < x < \infty$	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{x-m}{2\sigma^2}\right]$	m, σ		$\frac{1}{\sigma\sqrt{2\pi}} \times \int_{-\infty}^x \exp\left[-\frac{(t-m)^2}{2\sigma^2}\right] dt = \Phi\left(\frac{x-m}{\sigma}\right)$		$\exp(jmv - \sigma^2 v^2 / 2)$

Продовження табл. Д.4.1

1	2	3	4	5	6	7	8	9
3	Релея	$0 < x < \infty$	$\frac{x}{\sigma^2} \exp(-x^2/2\sigma^2)$	σ		$\begin{cases} 0, x < 0, \\ 1 - \exp(-x^2/2\sigma^2), \\ x > 0 \end{cases}$		$1 + ja\nu \sqrt{\frac{\pi}{2}} W\left(\frac{a\nu}{\sqrt{2}}\right)$ $W(z) = \exp(-z^2) \times$ $\times \left(1 + \frac{2j}{\sqrt{\pi}} \times \int_0^z \exp(-t^2) dt \right) -$ <p>табульований інтеграл імовірності від комплексного аргументу</p>
4	Узагальнений закон Релея (Райса)	$0 < x < \infty$	$\frac{x}{\sigma^2} \exp\left(-\frac{x^2 + a^2}{2\sigma^2}\right) \times$ $\times I_0\left(\frac{ax}{\sigma^2}\right)$	a, σ		$\begin{cases} 0, x < 0, \\ e^{\frac{a^2}{2\sigma^2}} \sum_{k=0}^{\infty} \frac{x}{(k!)^2} \left(\frac{a^2}{2\sigma^2}\right)^k \times \\ \times \Gamma\left(k+1; \frac{x^2}{2\sigma^2}\right), x > 0. \end{cases}$		

Продовження табл. Д.4.1

1	2	3	4	5	6	7	8	9
5	Ерланга к-го порядку	$0 < x < \infty$	$\frac{\lambda^{k+1}}{\Gamma(k+1)} x^k e^{-\lambda x}$	λ, k (к – ціле чис- ло)		$\begin{cases} 0, x < 0, \\ \frac{\Gamma(k+1; \lambda x)}{\Gamma(k+1)}, x > 0 \end{cases}$		$\left(1 - \frac{j\nu}{\lambda}\right)^{-(k+1)}$
6	Аркси- нуса	$-a < x < a$	$\frac{1}{\pi\sqrt{a^2 - x^2}}$	a		$\begin{cases} 0, x < -a, \\ \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x}{a}, \\ -a < x < a, \\ 1, x > a. \end{cases}$		$\frac{1}{\pi} \int_{-a}^a \frac{e^{j\nu x}}{\sqrt{a^2 - x^2}} dx -$ - $I_0(a\nu)$, де $I_0(z)$ - функція Бесселя нульо- вого порядку першого роду

Додаток 5

Таблиця Д.5.1

Функції розподілу дискретних випадкових величин

№ п/п	Функції розподілу	Область змінення значень випадкової величини	Аналітичний вираз закону розподілу	Визначальні параметри	Графік закону розподілу	Характеристична функція $\Theta_1(jv)$
1	2	3	4	5	6	7
1	Біноміальний (Бернуллі)	$k=0,1,2,\dots,n$	$p_n(k) = C_n^k p^k (1-p)^{n-k}$	n, p		$[1 + p(e^{jv} - 1)]^n$
2	Пуассона	$k=0,1,2,\dots$	$p_n(k) = \frac{\lambda^k}{k!} e^{-\lambda}$	λ		$e^{\lambda(e^{jv} - 1)}$

Продовження табл. Д.5.1

1	2	3	4	5	6	7
3	Рівномірний	$k=0,1,2,\dots,n$	$p(k) = \frac{1}{n}$	n		$\frac{e^{jv}(1 - e^{jvn})}{n(1 - e^{jv})}$
4	Полна	$k=0,1,2,\dots$ при $\alpha=0$; $k=1,2,\dots$ при $\alpha>0$;	$p(k) = p_0 \left(\frac{\lambda}{1 + \alpha\lambda} \right)^k \times$ $\times \frac{1(1 + \alpha) \dots [1 + (k - 1)\alpha]}{k!},$ $\alpha \geq 0, \lambda > 0,$ $p_0 = p(0) = (1 + \alpha\lambda)^{-1/\alpha}$	α, λ		$[1 + \alpha\lambda(1 - e^{jv})]^{-1/\alpha}$

