

Інформація, котра у вас є, насправді не та, яку вам хотілося б отримати.

Інформація, яку вам хотілося б отримати, насправді не та, котра вам потрібна.

Інформація, яка вам насправді потрібна, вам недоступна.

Інформація, яка в принципі вам доступна, коштує більше, ніж ви можете за неї заплатити.

Закони інформації Спенсера

Theoretical basics of INFORMATION- MEASUREMENT SYSTEMS

SECOND EDITION, REVISED AND SUPPLEMENTED

Textbook

*Edited by Member of the National Academy
of science of Ukraine V. Babak*

Kyiv 2017

Теоретичні основи ІНФОРМАЦІЙНО- ВИМІРЮВАЛЬНИХ СИСТЕМ

ВИДАННЯ ДРУГЕ, ПЕРЕРОБЛЕНЕ ТА ДОПОВНЕНЕ

Підручник

*За редакцією чл.-кор. НАН України,
В. П. Бабака*

Київ 2017

УДК 681.518.3 (075.8)

ББК з960.7я7

T741

Автори:

В.П. Бабак, С.В. Бабак, В.С. Єременко, Ю.В. Куц, Л.М. Щербак

Рецензенти:

Б.С. Стогній, д-р техн. наук, професор, академік НАН України,
В.П. Малайчук, д-р техн. наук, професор

Рекомендовано до друку вченою радою Університету новітніх технологій, м. Київ, і кафедрами інформаційно-вимірювальних систем Національного авіаційного університету та приладів і систем неруйнівного контролю Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Теоретичні основи інформаційно-вимірювальних систем: Підручник /
T741 В.П. Бабак, С.В. Бабак, В.С. Єременко та ін.; за ред. чл.-кор. НАН
України В.П. Бабака / 2-е вид., перероб. і доп. – К.: Ун-т новітніх
технологій; НАУ, 2017. – 496 с.
ISBN 978-966-932-027-8

У підручнику систематизовано викладено основні поняття та принципи побудови інформаційно-вимірювальних систем, які базуються на математичному апараті перетворення вимірювальних сигналів, технологіях вимірювань, передавання та опрацювання сигналів і даних, використанні сучасних інформаційно-комунікаційних каналів передавання інформації, методах досліджень метрологічних характеристик тощо. Наприкінці кожної глави розміщено основні висновки, запитання для самоконтролю, а також ключові слова українською та англійською мовами.

Для студентів технічних спеціальностей вищих навчальних закладів, аспірантів, наукових і інженерно-технічних працівників, спеціалістів у галузі вимірювань.

УДК 681.518.3 (075.8)

ББК з960.7я7

ISBN 978-966-932-027-8

© В. П. Бабак, С. В. Бабак, В. С. Єременко,
Ю. В. Куц, Л. М. Щербак, 2017

© Університет новітніх технологій, м. Київ, 2017

© Національний авіаційний університет, 2017

Зміст

Список скорочень	11
Передмова	13
Розділ 1. Завдання функціонування інформаційно - вимірювальних систем	15
1.1. Основні завдання вимірювань	16
1.2. Засади створення ІВС	21
1.3. Класифікація та апаратно-програмна реалізація ІВС..	29
Основні висновки	36
Запитання для самоконтролю	37
Conclusions	38
.....	
Розділ 2. Кількісні оцінки вимірювальної інформації	39
2.1. Вимірювальна інформація, її функції та властивості ...	40
2.2. Кількісні характеристики інформації	43
2.3. Міри інформації	45
2.4. Ентропія та її властивості	49
2.5. Продуктивність і надмірність джерела інформації	52
2.6. Зв'язок інформації з характеристиками сигналів	68
Основні висновки	71
Запитання для самоконтролю	72
Conclusions	73
Розділ 3. Сигнали та їх математичні моделі	75
3.1. Види сигналів і їх математичні моделі	76
3.2. Випадкові сигнали та завади	94
3.3. Числові характеристики сигналів і завад	102
3.4. Математичні моделі сигналів з обмеженим спектром	106
Основні висновки	117
Запитання для самоконтролю	120
Conclusions	121
Розділ 4. Перетворення сигналів	123
4.1. Дискретні сигнали та їх перетворення	124
4.2. Аналого-цифрове перетворення	141
4.3. Цифро-аналогове перетворення	156
Основні висновки	163
Запитання для самоконтролю	165
Conclusions	166

Розділ 5. Статистичне опрацювання даних вимірювань	169
5.1. Термінологія статистичного опрацювання даних вимірювань	170
5.2. Методи статистичного опрацювання даних вимірювань	177
5.3. Основи регресійного аналізу	187
5.4. Статистичне опрацювання випадкових кутів	191
5.5. Типові методи опрацювання результатів вимірювань..	202
Основні висновки	209
Запитання для самоконтролю	211
Conclusions	212
Розділ 6. Статистичне оцінювання характеристик випадкових процесів	215
6.1. Статистичного оцінювання характеристик стаціонарних процесів	216
6.2. Статистичне оцінювання стаціонарності випадкової послідовності	227
6.3. Статистичне оцінювання часових характеристик стаціонарної послідовності	230
6.4. Статистичне оцінювання спектральних характеристик стаціонарної послідовності	237
Основні висновки	243
Запитання для самоконтролю	245
Conclusions	247
Розділ 7. Перетворювачі ІВС	249
7.1. Класифікація та основні характеристики вимірювальних перетворювачів.....	250
7.2. Різновиди сенсорів	254
7.3. Перетворювачі на операційних підсилювачах та їх похибки	278
Основні висновки	293
Запитання для самоконтролю	294
Conclusions	295
Розділ 8. Аналогові та цифрові інтерфейси	297
8.1. Аналогові інтерфейси та їх метрологічні характеристики	298
8.2. Обчислювальні ресурси комп'ютеризованих систем ...	310
8.3. Цифрові інтерфейси	318

8.3.1. Паралельні інтерфейси	321
8.3.2. Послідовні інтерфейси	324
8.3.3. Інтерфейс USB	329
Основні висновки	334
Запитання для самоконтролю	336
Conclusions	336
.....	
Розділ 9. Передавання інформації в каналах ІВС	339
9.1. Системи передавання даних	340
9.2. Види інформаційних каналів, їх математичні моделі та характеристики	345
9.3. Швидкість передавання інформації в каналах зв'язку..	358
9.4. Синтез елементів систем. Оптимальний приймач	368
9.5. Багатоканальні мережі передавання даних	380
9.6. Завадостійкість систем передавання інформації	389
.....	
Основні висновки	394
Запитання для самоконтролю	395
Conclusions	395
.....	
Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу	397
10.1. Методи несанкціонованого доступу до ресурсів інформаційних систем	398
10.2. Засоби захисту від несанкціонованого доступу	416
10.3. Моделювання систем і процесів захисту інформації	422
10.4. Протидія мережевому несанкціонованому доступу...	429
Основні висновки	445
Запитання для самоконтролю	448
Conclusions	448
.....	
Розділ 11. Метрологічне забезпечення ІВС	451
11.1. Зміст метрологічного забезпечення ІВС	452
11.2. Основні метрологічні характеристики ІВС	458
11.3. Методи визначення метрологічних характеристик ІВС	467
Основні висновки	488
Запитання для самоконтролю	490
Conclusions	491
Список літератури	494

Contents

List of abbreviations	11
Preface	13
Chapter 1. Functional tasks of information-measurement systems	15
1.1. Typical tasks of measurement	16
1.2. IMS.....	21
1.3. Classifikations IMS	29
Main conclusions (Ukrainian)	36
Self-control questions	37
Main conclusions (English)	38
Chapter 2. Quantitative evaluation of measurement information	39
2.1. Measurement information and its properties	40
2.2. Quantitative characteristics of information	43
2.3. Measures of information	45
2.4. Entropy and its properties	49
2.5. Productivity, redundancy and interference immunity of information source	52
2.6. Link between information and signals characteristics	68
Main conclusions (Ukrainian)	71
Self-control questions	72
Main conclusions (English)	73
Chapter 3. Signals and their mathematical models	75
3.1. Types of signals and their mathematical models	76
3.2. Random signals and interferences	94
3.3. Numerical characteristics of signals and interferences	102
3.4. Signals with bounded spectrum	106
Main conclusions (Ukrainian)	117
Self-control questions	120
Main conclusions (English)	121
Chapter 4. Signals transformations	123
4.1. Discrete signals transformations	124
4.2. Analog-to-digital signals conversion	141
4.3. Digital-to-analog signals conversion	156
Main conclusions (Ukrainian)	163
Self-control questions	165

Main conclusions (English)	166
Chapter 5. Statistical of measurement data processing	169
5.1. Notions of statistical of measurement data processing	170
5.2. Methods processing of experimental data measuring	177
5.3. Basis of regression analysis	187
5.4. Statistical processing of random angles	191
5.5 Typical methods of measurement data processing	202
Main conclusions (Ukrainian)	209
Self-control questions	211
Main conclusions (English)	212
Chapter 6. Statistical evaluation of random processes' characteristics	215
6.1. Typical features of statistical evaluation of stationary processes characteristics	216
6.2. Statistical evaluation of random sequence stationary based on measurement data	227
6.3. Statistical evaluation of temporary characteristics of stationary sequence based on measurement data	230
6.4. Statistical evaluation of spectral characteristics of stationary sequences based on measurement data	237
Main conclusions (Ukrainian)	243
Self-control questions	245
Main conclusions (English)	247
Chapter 7. Converters IMS	249
7.1. Converters classification and characteristics	250
7.2. Main types of sensors	254
7.3. Converters uncertainty at operational amplifier and them errors	278
Main conclusions (Ukrainian)	293
Self-control questions	294
Main conclusions (English)	295
Chapter 7. Analog and digital interfaces	297
7.1. Analog interfaces and metrological characteristics	298
7.2. Calculable resources of the computer systems	310
7.3. Digital interfaces	318
7.2.1. Parallel interface	321

7.2.2. Serial interface	324
7.2.3. USB interface	329
Main conclusions (Ukrainian)	334
Self-control questions	336
Main conclusions (English)	336
Chapter 9. Transmission of information data in IMS	
channels	339
9.1. Data transmission systems	340
9.2. Types of information channels, their mathematical models and characteristics	345
9.3. Information transmission speed in systems channels	358
9.4. Synthesis of systems elements. Optimal receiver	368
9.5. Multichannel data transmission networks. Separation of information channels	380
9.6. Information transmission systems' interference immunity ..	389
Main conclusions (Ukrainian)	394
Self-control questions	395
Main conclusions (English)	395
Chapter 10. Protection of measurement data from unauthorized	
access	397
10.1. Methods of unauthorized access to information systems resources	398
10.2. Means of protection from unauthorized access	416
10.3. Modelling of information protection systems and processes	422
10.4. Counteracting to network unauthorized access	429
Main conclusions (Ukrainian)	445
Self-control questions	448
Main conclusions (English)	448
Chapter 11. Metrological provision of systems	451
11.1. Content of systems metrological provision	452
11.2. Basic metrological characteristics of systems	458
11.3. Methods to determine metrological characteristics of systems hardware and software means	467
Main conclusions (Ukrainian)	488
Self-control questions	490
Main conclusions (English)	491
References	494

СПИСОК СКОРОЧЕНЬ

АІ	– аналоговий інтерфейс
АІМ	– амплітудно-імпульсна модуляція
АЛП	– арифметико-логічний пристрій
АОС	– акустооптичний сенсор
АЦП	– аналого-цифровий перетворювач
АЧХ	– амплітудно-частотна характеристика
БАК	– безпілотний авіаційний комплекс
БЛМ	– модель Белла - Лападула
БПЛА	– безпілотний літальний апарат
БС	– блок синхронізації
ВК	– вимірювальний канал
ВП	– вимірювальний перетворювач
ВС	– вимірювальна система
ГІ	– генератор імпульсів
ДНФ	– диз'юнктивна нормальна форма
ДПФ	– дискретне перетворення Фур'є
ДОН	– джерело опорної напруги
ДРР	– дисциплінний розподіл ресурсів
ДУ	– диференціальне рівняння
І	– інтегратор
ІВТ	– інформаційна вимірювальна технологія
ІВС	– інформаційно-вимірювальна система
ІКС	– інформаційно-комунікаційна система
ІНС	– ідеальний низькочастотний сигнал
ІПС	– ідеальний полосовий сигнал
К	– компаратор
КПДП	– контролер прямого доступу до пам'яті
КС	– комп'ютерна система
МЗ	– метрологічне забезпечення
МЛ	– мажоритарна логіка
МНК	– метод найменших квадратів
МОНПТ	– метал-окисел-напівпровідниковий польовий транзистор
МП	– мікропроцесор
МХ	– метрологічні характеристики
НЗЗ	– негативний зворотний зв'язок
НСД	– несанкціонований доступ
НСП	– невиключена систематична похибка
ОЗП	– оперативний запам'ятовуючий пристрій
ОМР	– одиниця молодшого розряду
ОП	– операційний підсилювач
ПАХ	– поверхневі акустичні хвилі
ПВЗ	– пристрій вибірки і запам'ятовування
ПВП	– первинний вимірювальний перетворювач

ПЗП	– постійний запам'ятовуючий пристрій
ПК	– персональний комп'ютер
ПНЧ	– перетворювач напруга-частота
ПЗ	– програмне забезпечення
ПП	– приймач-передавач
ППП	– перетворювач профілю поверхні
ПІВС	– поліметрична інформаційно-вимірювальна система
РЗП	– реєстр загального призначення
РПН	– реєстр послідовного наближення
РЛІ	– реверсивний лічильник імпульсів
РУ	– різницеве рівняння
С	– сенсор
СД	– система діагностики
СІ	– система ідентифікації
ЗВТ	– засоби вимірювальної техніки
СК	– система контролю
СКВ	– середньоквадратичне відхилення
СРО	– система розпізнавання образів
СЗД	– система збирання даних
СТМ	– скануючий тунельний мікроскоп
СК	– система керування
ТЗ	– технічне завдання
ТІВС	– телеметрична інформаційно-вимірювальна система
ТО	– термометр опору
ПВЗ	– пристрій вибірки / зберігання
ФМ	– фазова модуляція
ФНЧ	– фільтр нижніх частот
ФЧХ	– фазочастотна характеристика
ХіМПТ	– хімічний польовий транзистор
ЦАП	– цифро-аналоговий перетворювач
ЧМ	– частотномодульований сигнал
ШІМ	– широтно-імпульсна модуляція
ШПФ	– швидке перетворення Фур'є
ЕОМ	– електронна обчислювальна машина
USB	– універсальна послідовна шина

ПЕРЕДМОВА

З моменту публікації підручника «Теоретичні основи інформаційно-вимірювальної техніки» (1983 р.) відомого вченого в галузі вимірювань П.П. Орнатського пройшло більше 30-ти років. Автори цього видання уявляють всі труднощі викладення матеріалів з інформаційно-вимірювальних систем (ІВС). У виданні збережена методологія викладення матеріалів згаданого вище підручника – теоретичні основи адаптовані до функціонування основних структурних модулів ІВС.

Теоретичні основи ІВС використовують методи і результати досліджень **низки** природничих наук, включаючи сучасну математику і фізику, технічних наук для вирішення різнопланових завдань вимірювань.

Теоретичні основи ІВС умовно можна розділити на два напрями: основи розроблення і створення ІВС як технічної системи та основи функціонування ІВС як апаратно-програмних засобів вимірювання, що відображають характерні специфіку та особливості спільності і конкретики проведення процесів вимірювання.

У цій книзі (2-е видання) викладені, в основному, матеріали другого напрямку досліджень функціонування ІВС.

Процес вимірювання є спільною реалізацією теорії, методології та практики, що використовує засоби вимірювання, інтелектуальний ресурс дослідників при постановці завдань вимірювання, обґрунтування методу вимірювання, виконання вимірювального експерименту, розроблення інформаційного, алгоритмічного та програмного забезпечення для опрацювання даних вимірювань, отримання результату і визначення **точності** вимірювань.

Класичні операції процесу вимірювання доповнені операціями захисту даних вимірювань, тим самим відображені сучасні тенденції забезпечення достовірності вимірювальної інформації.

Процес вимірювання характеризується послідовним застосуванням двох мір: фізичної міри (одиниці або шкали) при експериментальному її порівнянні зі значенням вимірюваної величини, а також нормованої імовірнісної міри для статистичного оцінювання результату вимірювання і його якості (характеристики точності).

ІВС є продуктом використання науково-технічного, інформаційного, інтелектуального, виробничого, економічного та інших ресурсів, і в інтегральному вигляді відображає рівень розвитку і можливості інформаційно-вимірювальної техніки в цілому при вирішенні науково-прикладних завдань метрології.

Сучасні ІВС можуть бути реалізовані як:
реальні апаратно-програмні засоби вимірювань;

віртуальні ІВС з використанням сучасних інформаційно-вимірювальних технологій моделювання, включаючи імітаційне, математичне, фізичне, з широким використанням засобів комп'ютерної техніки для проведення імітаційного вимірювального експерименту.

Віртуальні ІВС в порівнянні з реальними мають ряд переваг, їх реалізація вимагає менших часових, виробничих і фінансових витрат. Однак, в ряді випадків такі ІВС через інформаційну невизначеність об'єкта вимірювань не можуть забезпечити отримання об'єктивних достовірних результатів і тому потрібно провести натурний вимірювальний експеримент з використанням реальних систем.

Потенційні можливості сучасних систем на етапі інформаційного розвитку суспільства значно зросли, що сприяє як розширенню предметних областей їх застосування, так і їх використання для підвищення ефективності відомих і вирішення нових науково-прикладних завдань вимірювань.

У прогнозах розвитку ІВС автори солідарні з іншими колегами - фахівцями з вимірювань. Ніякі удосконалення інформаційних технологій вимірювань, включаючи комп'ютерні та інтелектуальні, не привели, не приводять і не можуть очевидно привести в майбутньому до розширення номенклатури **вимірювань величин** поки відсутні відповідні сенсори, що формують первинну інформацію при безпосередній взаємодії їх з об'єктом дослідження. Подальший розвиток ІВС і використання їх в різних галузях науки і техніки, включаючи квантову метрологію та нанотехнології, в значній мірі буде визначатися розробленням нових принципів функціонування та створення на їх основі нових видів сенсорів.

Автори висловлюють щире подяку рецензентам - академіку НАН України, доктору технічних наук, професору Б.С. Стогнію (Інститут електродинаміки НАН України) та доктору технічних наук, професору, завідувачу кафедри Дніпропетровського національного університету імені О. Гончара В.П. Малайчуку за зауваження та рекомендації в написанні книги.

Київ, весна 2017

Від авторів



ЗАВДАННЯ ФУНКЦІОНУВАННЯ ІНФОРМАЦІЙНО- ВИМІРЮВАЛЬНИХ СИСТЕМ

1

- 1.1. Основні завдання вимірювань
- 1.2. Засади побудови ІВС
- 1.3. Класифікація та апаратно-програмна реалізація ІВС

1.1. Основні завдання вимірювань

Роль вимірювань у пізнанні світу. Перші відомості про вимірювання відносяться до давніх часів становлення життя і свідомої діяльності людини на Землі. Як лічба і обчислювання, вимірювання послужили відправною точкою появи математичних дисциплін, в першу чергу геометрії, а в подальшому стали потужним інструментом розвитку науки і техніки. Історія земної цивілізації – це шлях, який пройшло людство від найпростіших вимірювань оснований на використанні органів почуттів людини, до *метрології* – науки про вимірювання та їх застосування. Metrologia була і лишається основою методології наукового пізнання світу.

Для практики вимірювань в структурі метрології виділяють наступні напрями:

теорію вимірювань включаючи теоретичні основи (теорію) інформаційно-вимірювальних систем (ІВС);

прикладну метрологію, яка включає методи, способи, технічні, інформаційні, програмні та інші засоби забезпечення єдності вимірювань;

законодавчу метрологію, що включає законодавчі акти, стандарти, рекомендації тощо, які регламентуються і контролюються державою для забезпечення єдності вимірювань.

Теорія ІВС є одним із основних напрямів теорії вимірювань, базується на положеннях і методах низки природничих і технічних наук, в тому числі, теорії сигналів і систем, передачі інформації, теорії ймовірностей і випадкових функцій, функціонального аналізу, обчислювальної математики, математичної статистики і квантової механіки. Вони є теоретичним базисом і всіх інших засобів забезпечення вимірювань, які виконують лише частину функцій ІВС.

Об'єкти, предмети і завдання вимірювань. Вимірювання у практичній діяльності суспільства зустрічаються так часто і у таких різноманітних формах, що виникає хибна думка про їх універсальність і простоту реалізації – ілюзія можливості виміряти все і в будь-який час. Але в дійсності виникають значні труднощі починаючи з формалізації самого процесу вимірювань, прогнозу його результату, оцінювання характеристик точності результату (похибки, невизначеність) та ін.

Об'єктами вимірювання є множина об'єктивно існуючих явищ і ефектів, тіл і речовин різної природи у різних предметних областях. З метою їх упорядкування введемо наступне.

Об'єктами вимірювання в широкому сенсі є об'єкти різної природи, стан і перебіг процесів в яких відображаються за певними правилами числами з множини дійсних чисел.

Об'єктами вимірювань у вузькому сенсі є відповідні величини, які згідно існуючих нормативних документів визначаються так.

Величина – властивість явища, тіла або речовини, яка може бути виражена кількісно у виді числа з вказівкою розмірності, як основи для порівняння.

Значення величини може змінюватися у просторі і часі .

Таке доповнення по суті є аксіомою вимірювань і використовується в теорії ІВС. Динаміка зміни значень і характеристик величин дає можливість їх практичного застосування під час формування первинної інформації в дослідженнях різних об'єктів, властивостей змінюваних у просторі і в часі. Характерними прикладами є вимірювання просторово-часових характеристик електромагнітних і вібраційних полів. Для опису явищ і процесів, а також властивостей матеріальних тіл використовуються різні величини, число яких сягає декількох тисяч: електричні, магнітні, механічні, акустичні, оптичні, хімічні, біологічні та ін. На сьогодні існує така кількісна оцінка вимірюваних величин:

число електричних і магнітних величин не перевищує 100;

число неелектричних величин постійно збільшується і на початок ХХІ ст. перевищило 4000.

Значення величини при вимірюваннях прийнято представляти у вигляді числа, вектора, функції, як із зазначенням розмірності одиниці величини, так і безрозмірні. Розмірні значення величин називають *іменованими*, наприклад, *іменоване число*.

Предметами, що досліджуються в теорії вимірювань, є моделі, методи, технічні, інформаційні, програмні та інші засоби забезпечення єдності вимірювань.

До основних завдань вимірювань належать наступні.

1. *Завдання відображення* – здійснення гомоморфного відображення основних властивостей, характеристик об'єкта дослідження у вимірювальну інформацію за умови, що операції і співвідношення з елементами і модулями об'єкта гомоморфні операціям з даними вимірювань.

2. *Завдання однорідності і стаціонарності міри*, яка забезпечує визначення кількісного вмісту вимірювальної величини з заданою точністю за результатами вимірювань у просторі і часі.

3. *Завдання захисту вимірювальної інформації* – полягає у використанні відповідних засобів захисту інформації під час

формування, передавання, опрацювання, реєстрації та поданні замовнику даних вимірювань з метою мінімізації впливу як природних завад, так і навмисних загроз та несанкціонованого доступу до даних.

Вирішення вказаних завдань вимірювань поєднуються на основі застосування фізичних і математичних моделей досліджуваних об'єктів, процесів і явищ. Незважаючи на той факт, що такі моделі є вторинними, тобто є копіями, відображеннями, які не в повній мірі відповідають реальним вимірюванням, їх роль в методології вимірювань дуже вагома і має фундаментальний характер. Кожна з моделей, фізична і математична, доповнюють одна одну і дає можливість підвищення ефективності процесу вимірювань, наприклад, обґрунтувати шляхи підвищення точності результатів вимірювань.

Означення математичної моделі інформації для вимірювань об'єкта досліджень формулюється наступним чином.

Математична модель інформації для вимірювань об'єкта досліджень – це сукупність знань, припущень, гіпотез, початкових і граничних умов, які гомоморфно відображають основні властивості, характеристики в просторі і в часі досліджуваної інформації, побудованої у вигляді цілісної, несуперечливої, логічно витриманою структури, сформульованої з використанням необхідних математичних об'єктів, термінів і символів для вирішення заданого класу завдань вимірювань.

В теорії ІВС використовуються різні види моделі. Такі моделі і їх характеристики, у загальному випадку є просторово-часовими і використовуються як класифікаційні ознаки різних видів ІВС. Серед моделей сигналів для вимірювань виділяють: детерміновані і випадкові; однорідні і багатовимірні; циклічні та шумові і інші. Для створення моделей використовують методи функціонального аналізу, теорії випадкових функцій і математичної статистики, математичного і комп'ютерного моделювання, теорії вимірювань, сигналів і систем.

Окрім моделей детермінованих сигналів широко використовують і моделі випадкових величин і сигналів, а саме випадкова величина $\xi(\omega)$ і випадковий процес $\xi(\omega, t)$ або їх комбінації з детермінованими моделями. У зв'язку зі зростанням складності завдань вимірювань постійно збільшується і розмірність моделей, в тому числі і за рахунок просторових координат x , $(x, y) = \mathbf{p}$, $(x, y, z) = \mathbf{r}$. У більшості випадків вимірювань значень і просторово-часових характеристик досліджуваних сигналів використовуються наступні моделі, загальний вид яких в залежності від аргументів наведено в табл. 1.1.

В табл.1.1 використанні наступні позначення: ω – елементарна випадкова подія з простору подій Ω ; T – часовий інтервал спостереження, R – множина дійсних чисел.

Таблиця 1.1

Аргументи Види функцій	$\omega \in \Omega$ $t \in T$	$\omega \in \Omega$ $x \in R$ $t \in T$	$\omega \in \Omega$ $(x, y) = \mathbf{p} \in R^2$ $t \in T$	$\omega \in \Omega$, $(x, y, z) = \mathbf{r} \in R^3$ $t \in T$
Детерміновані	число $\alpha \in R$, одновимірна функція $u(t) \in R$	двовимірна функція $u(x, t) \in R$	тривимірна функція $u(\mathbf{p}, t) \in R$	чотиривимірна функція $u(\mathbf{r}, t) \in R$
Випадкові	величина $\xi(\omega) \in R$, процес $\xi(\omega, t) \in R$	одновимірне просторове поле $\xi(\omega, x, t) \in R$	двовимірне просторове поле $\xi(\omega; \mathbf{p}; t) \in R$	тривимірне просторове поле $\xi(\omega; \mathbf{r}; t) \in R$

З метою використання вказаних моделей в завданнях вимірювань, а саме в обмежених областях простору G і скінченних інтервалах часу T , застосовується наступна багатовимірна індикаторна функція

$$I(\mathbf{r}, t) = \begin{cases} 1, & \mathbf{r} \in G, t \in T \\ 0, & \mathbf{r} \notin G, t \notin T, \end{cases}$$

Використання математичної моделі виду $\xi(\omega, \mathbf{r}, t) \cdot I(\mathbf{r}, t)$ дає можливість проводити дослідження з визначенням і просторово-часових характеристик сигналів за умови проведення вимірювань у різних місцях простору на скінченних інтервалах часу. Це забезпечує також проведення подальшого порівняльного аналізу результатів вимірювань з метою перевірки адекватності запропонованих моделей, прогнозу динаміки зміни основних характеристик об'єкта досліджень в просторі і часі та інше.

Використання міри. Вирішення завдання забезпечення однорідності і стаціонарності міри в теорії ІВС в кожному конкретному випадку вимірювання має свою специфіку і характерні особливості. В той же час можна виділити загальні властивості використання і види міри. Це наступні міри:

M1 – міра одиниці величини;

M2 – міра одиниці просторових координат місця проведення вимірювань;

M3 – міра одиниці часу проведення вимірювань;

M4 – ймовірна нормована міра результату і точності вимірювань при статистичному опрацюванні даних вимірювань;

M5 – міра захисту інформації конкретного процесу вимірювань.

У випадку вимірювань $n(n>1)$ величин є n мір одиниць досліджуваних величин. Перші три міри є фізичними мірами (у загальному випадку маємо $n+2$ мір), які реалізуються відповідними еталонами мір з заданою точністю. Однорідність і стаціонарність таких мір визначаються характеристиками технічних пристроїв і систем їх формування. Це дає можливість в певній мірі управляти забезпеченням однорідності і стаціонарності фізичних мір, при проведенні процесу вимірювань в різних місцях простору і в різні часи, забезпечуючи при цьому порівняння результатів вимірювання і таким чином виконати вимоги єдності вимірювань

Ймовірна нормована міра є нефізичною мірою, а мірою сукупності дії різних випадкових факторів на значення і характеристики даних і результату вимірювань під час їх проведення. Використання ймовірної міри при статистичному опрацюванні даних вимірювання дає можливість підвищити точність результату вимірювань у порівнянні з точністю даних вимірювань.

Міра захисту інформації при вимірюваннях є комплексною. Міра формується значною кількістю факторів, дія більшості з яких має випадковий характер. Це дає можливість таку міру визначати як ймовірнісну, яка може бути застосована як для окремих операцій, наприклад передача даних вимірювань по каналах зв'язку, реєстрації результату вимірювань так і для всього процесу вимірювання у цілому.

На рис. 1.1 наведено схематичне зображення використання вказаних мір при перетворенні вимірювальної інформації функціональними модулями ІВС.

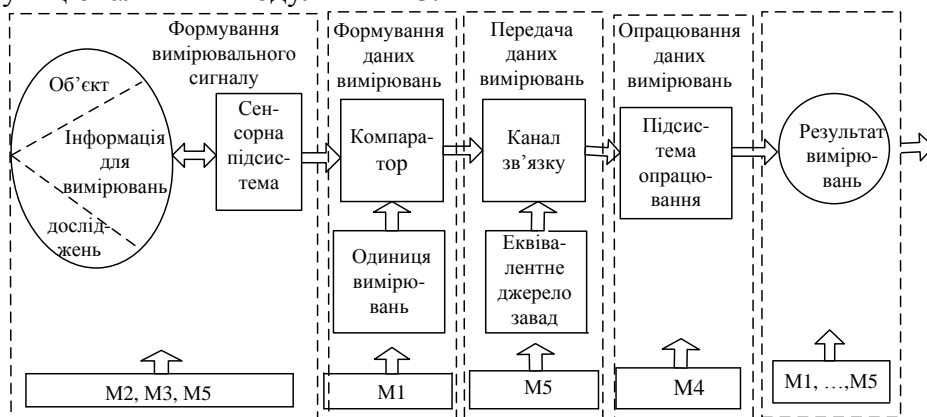


Рис. 1.1. Схематичне зображення використання різних мір в структурі ІВС

Частинний випадок захисту інформації – підвищення завадостійкості вимірювань, застосовується практично для всіх процесів вимірювань. У структурі ІВС, зазвичай, є модуль фільтрації сигналів, що забезпечує підвищення співвідношення сигнал/завада на своєму виході.

Значення точності вказаних п'яти мір трансформуються відповідно у значення точності даних і результатів вимірювань у відповідних модулях ІВС.

Інформаційні вимірювальні технології. Виділення класу інформаційних вимірювальних технологій (ІВТ) на множині інформаційних технологій (ІТ) аналогічно виділенню класу ІВС на загальній множині технічних систем. Саме специфіка завдань вимірювань в ІВС дала можливість виділити класифікаційні ознаки ІВТ. Такими ознаками є:

отримання первинної вимірювальної інформації – вимірювального сигналу в результаті взаємозв'язку і взаємодії сенсорної підсистеми ІВС з об'єктом дослідження;

забезпечення порівняння рівня параметрів та характеристик вимірювального сигналу з розмірами прийнятих одиниць вимірювань і формування даних вимірювань;

перетворення і опрацювання даних вимірювань ІВС із заданою точністю;

оцінювання і представлення результату вимірювань.

В теорії ІВС динаміка розвитку ІВТ на сьогодні є найбільш інтенсивною і цей факт обумовлений наступним:

складністю завдань вимірювань, збільшення інформативності вимірювань;

розширенням кола предметних областей використання результатів вимірювань;

необхідністю удосконалення відомих і створення нових сенсорних підсистем ІВС для перетворення досліджуваних величин;

необхідністю розширення формату і ресурсів вимірювальних експериментів при дослідженнях динаміки у просторі і часі динаміки стану і режимів функціонування технічних об'єктів і систем, фізичних явищ і процесів.

В табл. 1.2 наведена класифікація видів ІВТ, яка базується на інтегральній ознаці розмірності первинної вимірювальної інформації – вимірювального сигналу (одновимірні і багатовимірні). Ця класифікація має аналоги в класі відомих ІТ і визначає потужність і ресурси процесу вимірювань.

Для *статичних* вимірювань моделі сигналів, значення і просторові характеристики досліджуваних величин розглядаються як незмінні в часі. Це значно спрощує реалізацію відповідних ІВТ, але суттєво зменшує інформативність і вагомість отриманих результатів вимірювань.

Таблиця 1.2

Види технологій	Види моделей вимірювання сигналів
Інформаційна технологія вимірювань 1D (ІТВ–1D)	$\alpha, u(x, t), \xi(\omega), \xi(\omega, t); \xi(\omega; x; t)$
Інформаційна технологія вимірювань 2D (ІТВ–2D)	$u(\mathbf{r}, t), \xi(\omega; \mathbf{r}; t)$
Інформаційна технологія вимірювань 3D (ІТВ–3D)	$u(\mathbf{r}, t), \xi(\omega; \mathbf{r}; t)$

Наведена класифікація відповідає частинному випадку вимірювань однієї вимірюваної величини. Загальний випадок одночасного вимірювання більшої кількості однорідних і різнорідних величин є складнішим, але методологія створення і представлення моделей лишається незмінною для кожної вимірюваної величини.

1.2. Засади створення ІВС

Ефективність використання технічних систем у більшості практичних випадків оцінюється на етапі їх експлуатації, але інформативність і ресурси ІВС визначаються і реалізуються під час їх створення. На цьому етапі основні положення і методи теорії ІВС використовуються у повному обсязі.

Етапи життєвого циклу ІВС. З метою більш глибокого розуміння важливості основних положень теорії ІВС для створення систем вимірювань розглянемо ілюстративну схему послідовності етапів життєвого циклу ІВС представлену на рис. 1.2.

Початковий етап. На початку створення ІВС (далі системи) виконується значний обсяг робіт з обґрунтування, розробки та узгодження технічного завдання на створювану систему між замовником і розробником системи. В технічному завданні на систему наведені всі технічні характеристики, в першу чергу, метрологічні, показники надійності, терміни виконання розробки системи, проведення випробувань, сертифікації та передачі системи в експлуатацію, вартість розробки системи. На цьому етапі використовуються останні досягнення теорії і практики вимірювань для оцінювання можливості

розв'язання актуальних і важливих науково-технічних завдань вимірювань конкретної предметної області.



Рис. 1.2. Ілюстрація етапів життєвого циклу ІВС

Етап проектування передбачає виконання значного обсягу теоретичних, імітаційних (моделюючих) та експериментальних (лабораторних) досліджень, на основі результатів яких обґрунтовується оптимальна апаратно-програмна структура системи. Для цього використовується ряд критеріїв (основні з яких спрямовані на досягнення заданої точності вимірювань), розробляються електричні схеми основних апаратних модулів і підсистем, зміст і структури алгоритмічно-програмного забезпечення, технічна та конструкторська документація на виготовлення дослідних і серійних зразків системи, проводяться роботи з підготовки виробництва.

На цьому етапі обґрунтовуються, обчислюються і оцінюються розрахункові характеристики окремих модулів, підсистем і системи в цілому. Такі характеристики є базовими, використовуються на наступних етапах життєвого циклу системи та можуть доповнюватися даними лабораторних і натурних випробувань, експлуатації, коригуватися, уточнюватися, підтверджуватися статистичними гіпотезами з використанням критеріїв згоди. Отже на етапі проектування створюється методологія втілення подальших етапів життєвого циклу системи.

В табл. 1.3 наведені основні результати досліджень, які отримуються на етапі проектування на основі використання напрямів і методів ряду природничих і технічних наук.

Таблиця 1.3

Напрями, методи	Результати використання	Обґрунтування, обчислення, оцінювання результатів досліджень
Функціональний аналіз, обчислювальна математика, теорія ймовірностей, теорія випадкових функцій		Одновимірні та багатовимірні моделі сигналів, характеристики сигналів, алгоритми оцінювання характеристик сигналів, результати комп'ютерних вимірювальних експериментів, результати функціонування різних варіантів систем.
Теорія сигналів і систем		Оптимізація структур систем, результатів аналізу перетворень сигналів та опрацювання даних вимірювань, методик випробувань модулів, підсистем і систем в цілому на відповідність технічному завданню, резервування модулів, підсистем для забезпечення метрологічної надійності систем.
Теорія вимірювань		Оцінки результатів і невизначеності вимірювань, метрологічних характеристик, методик вимірювань.
Теорії інформації, методи захисту інформації		Оцінки кількості вимірюваної інформації, апаратно-програмні засоби захисту інформації
Математична статистика		Алгоритми перевірки адекватності моделей оцінки результату і невизначеність вимірювань, статистичні критерії

Автоматизація проектування є важливим фактором, що впливає на рівень працездатності складних технічних систем, оскільки дозволяє вирішити наступні завдання: істотно знизити кількість проектних помилок, підвищити якість конструкторської і технологічної документації, створити умови для багатоваріантного проектування і вибору оптимального варіанту, скоротити терміни проектування, вивільнивши певний час на доопрацювання технічних рішень, узгодження параметрів модулів та підсистем, проведення випробувань.

Особливе значення для підвищення якості та надійності систем має використання під час розроблення програмного забезпечення систем високих технологій програмування, розвинених системи тестування, проведення автономних і комплексних випробувань програмного забезпечення ще до виготовлення технічних систем.

Етап виготовлення та випробування дослідних зразків модулів, підсистем ІВС. На етапі реалізуються:

методика і відповідні плани проведення випробувань з використанням засобів вимірювань для отримання експериментальних даних, характеристик і параметрів інформаційних сигналів, метрологічних характеристик дослідних модулів, підсистем;

відповідна корекція алгоритмічно-програмного забезпечення для визначення метрологічних характеристик модулів і підсистем, а також характеристик математичних моделей за результатами порівняльного аналізу розрахункових і експериментальних характеристик.

За результатами досліджень проводиться коригування проектної та конструкторської документації на розробку серійного зразка системи.

На *етапі створення серійного зразка, сертифікації та передачі ІВС в експлуатацію* виконуються такі основні роботи:

підготовка виробництва для виготовлення серійних зразків системи з заданими метрологічними характеристиками;

реалізуються методики налагоджування різних модулів і підсистем та перевірки функціонування системи для виконання заданих функцій;

випробування і сертифікація серійного зразка системи з використанням метрологічно атестованих вимірювальних (апаратних і програмних) засобів.

Етап експлуатації ІВС є найбільш тривалим в часі і передбачає такі роботи:

транспортування, налагодження і запуск в експлуатацію системи;

моніторинг функціонування системи в заданих режимах з використанням підсистем вимірювань, контролю, діагностики і управління;

формування бази даних вимірювань функціонування системи з наступним статистичним опрацюванням даних вимірювань характеристик, параметрів інформаційних сигналів;

реалізація методики довгострокового моніторингу та визначення прогнозованих метрологічних характеристик системи;

реалізація методики авторського нагляду та ремонту системи;

коригування математичного і програмного забезпечення варіантів модернізації діючої системи, проведення моделюючих експериментів і формування рекомендацій модернізації системи з метою підвищення ефективності її використання.

Заключний етап життєвого циклу ІВС є найбільш наукомістким в життєвому циклі ІВС і полягає у формуванні даних досліджень всіх етапів, їх систематизації та теоретичному аналізу з питань коректності використання математичного апарату, узагальнення результатів функціонування конкретного виду ІВС. Таке узагальнення результатів

досліджень метрологічних характеристик слугує відповідним науково-технічним, інформаційним, математичним і програмним забезпеченням для створення нових перспективних ІВС.

На даному етапі виконується також роботи з утилізації ІВС як системи, яка вичерпала свій технічний ресурс.

Точність вимірювань. Відомо, що результат вимірювань без оцінки його точності втрачає практичне значення. Історія вимірювань підтвердила, що основною метою їх розвитку є підвищення точності вимірювань. На початку становлення земної цивілізації (оцінюється умовно в декілька сотень тисяч років) для вимірювань було достатньо природних органів почуття людини. Перші засоби вимірювань з'явилися приблизно декілька тисячоліть назад і пройшли шлях від простого вимірювального шнура з вузлами до гігантського п'ятиповерхового «Атласу» – сучасного вимірювального комплексу протонного прискорювача в Женеві, сенсорна підсистема якого з 140 мільйонами мікроскопічних сенсорів фіксує ~800 млн. зіткнень протонів в секунду під час проведення фізичних ядерних досліджень.

Ще в середині XIX століття англійський фізик А. Майкельсон, аналізуючи відкриття, зроблені в ході проведення фізичних вимірювальних експериментів, висловив гіпотезу про те, що нові відкриття в фізиці будуть можливі за умови забезпечення відносних похибок вимірювань не більше за 10^{-6} . В ряді випадків така гіпотеза, підтвердилась. Сучасні засоби вимірювань дають можливість отримати відносну похибку результату вимірювань, що наближається до потенціально можливих значень $10^{-17} \dots 10^{-18}$.

Зупинимось на аналізі сучасного стану оцінювання точності результатів вимірювань. На сьогодні у Міжнародній метрології відбувається еволюційний природний перехід оцінювання якості результатів вимірювань від *концепції похибок* (умовна назва) до *концепції невизначеності* (прийнята з 1993 року, як GUM 1993). Таке співставлення має аналогію з переходом від одної міжнародної системи одиниць до іншої.

Відомо, що так звану *концепцію похибок* називають ще класичною теорією похибок. Вона базується на постулаті існування істинного (точного) значення вимірюваної величини. Абсолютні значення похибки результату вимірювань визначають як різницю між результатом вимірювання величини і її істинним значенням. Використання *концепції похибок* дало змогу отримати важливі результати теорії і практики вимірювань які до початку XX століття сприяли розвитку вимірювань і метрології в цілому.

Невизначеність результату вимірювань. Сам термін «невизначеність вимірювань» означає сумнів стосовно достовірності результату вимірювання. Принциповою відмінністю даного показника якості результату вимірювань від традиційного полягає в наступному: достовірність результатів вимірювань з урахуванням похибок базується на отриманих в процесі проведення вимірювального експерименту даних вимірювань, відомих характеристиках засобів вимірювань та умов проведення експерименту, іншої апріорної інформації. В концепції невизначеності такі поняття, як *похибка* і *істинне значення вимірюваної величини* замінюються поняттями *невизначеність* і *оціночне значення вимірюваної величини*, а поділ (класифікація) видів похибок за різними ознаками замінюється на поділ за способом *оцінювання невизначеностей вимірювань за типом А*, який ґрунтується на використанні методів математичної статистики, і *за типом В* – на використанні і інших методів.

Відповідність між показниками якості результатів вимірювань за концепціями похибок та невизначеності подано на рис. 1.3.

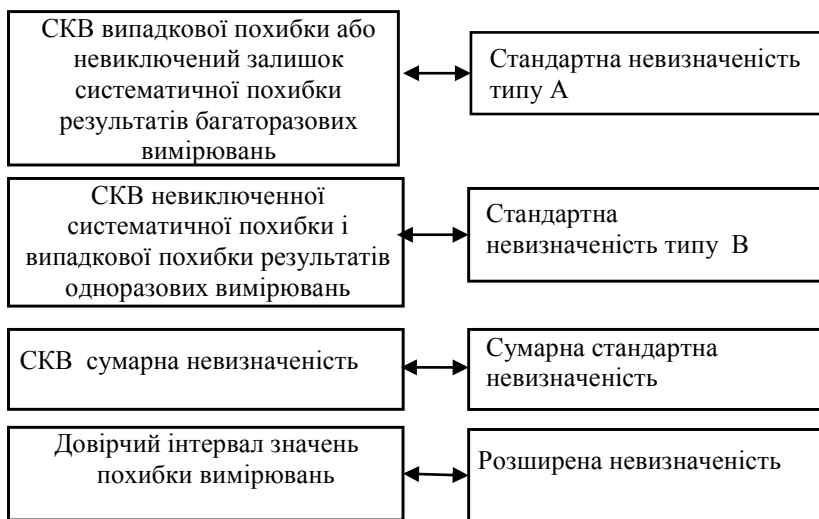


Рис.1.3. Відповідність оцінок похибок і невизначеності результатів вимірювань

Коротко зупинимося на методах оцінювання і способах відображення об'єктів невизначеності результатів вимірювань.

За *методом оцінювання* виділяють невизначеності:

типу А, до яких належать складові оцінювані статистичними методами опрацювання даних багаторазових вимірювань;

типу B , до яких відносяться складові, що оцінюються іншими методами опрацювання даних вимірювань.

За способом відображення **виділяють**:

стандартну невизначеність результату прямих вимірювань як стандартне середньоквадратичне відхилення (СКВ);

сумарну невизначеність як стандартну невизначеність результату непрямих вимірювань, яку отримують зі значень інших величин, пов'язаних з вимірюваною величиною функціональною залежністю;

розширену невизначеність як інтервальну оцінку вимірювань, яка оцінюється добутком стандартної невизначеності на коефіцієнт охоплення, який залежить від виду закону розподілу і рівнів заданих ймовірностей;

відносну невизначеність – відношення стандартної, сумарної або розширеної невизначеності до оцінки вимірюваної величини.

1.3. Класифікація і апаратно-програмна реалізація ІВС

З історії розвитку ІВС. Розвиток ІВС, починаючи з 50-х років ХХ століття, характеризується створенням систем декількох поколінь.

Для *першого покоління* ІВС характерним є формування методології і концепції створення ІВС як системної структури сумісного функціонування засобів формування, опрацювання і передачі даних вимірювань. Це були системи централізованого циклічного отримання вимірювальної інформації з елементами обчислювальної техніки. Відповідний період розвитку таких систем (50-60 роки ХХ ст.) іменують *періодом детермінізму*, оскільки для дослідження процесу вимірювань використовувався здебільшого аналітичний апарат функціонального аналізу і обчислювальної математики.

Друге покоління ІВС пов'язано з використанням адресного збору вимірювальної інформації і її опрацюванням засобами обчислюваної техніки. Елементну базу ІВС складали мікроелектронні компоненти малої та середньої степені інтеграції. Цей період (70 роки ХХ ст.) іменують *періодом стохастичності* оскільки ряд науково-технічних завдань вимірювань було вирішено завдяки використанню теорії випадкових процесів і математичної статистики.

Третє покоління ІВС широко застосовує мікропроцесори і промислові функціональні електронні модулі, сумісні між собою за інформаційними, метрологічними, конструктивними, енергетичними та експлуатаційними характеристикам. Це період (80 роки ХХ ст.) створення просторово-розподілених і програмованих ІВС.

Четверте покоління ІВС (90-і роки ХХ століття) пов'язане зі стрімким розвитком обчислювальної техніки і створенням гнучких програмованих ІВС. В таких системах основну роль відіграє користувач-дослідник, який має можливість змінювати програму функціонування ІВС для вирішення відповідних задач вимірювання.

П'яте покоління ІВС розвивається на початку ХХІ століття і характеризується створенням адаптивних, віртуальних і інтелектуальних систем вимірювань з використанням персональних комп'ютерів і сучасних інформаційних вимірювальних технологій.

Класифікація ІВС. У нормативних документах метрології класифікація ІВС не наведена і це пояснюється наступним:

у класифікації ІВС необхідно чітко виділити класифікаційні ознаки, а це для всіх видів і поколінь ІВС зробити практично неможливо;

класифікацію ІВС можна лише умовно провести на певному етапі розвитку ІВС в зв'язку з динамікою змін елементної бази та ІВТ.

Наведена в таблиці 1.4 класифікація ІВС враховує тенденції сучасного інформаційного етапу розвитку систем вимірювань.

Таблиця 1.4

Класифікація ІВС Ознаки класифікації	Види ІВС
Функціональне призначення	Системи вимірювальні, автоматичного контролю, технічної діагностики, ідентифікації, розпізнавання образів, управління, телеметричні, статистичні, поліметричні.
Режими (алгоритм) функціонування	Постійний, адаптований до динаміки змін величин і умов роботи об'єкта досліджень, віртуальний на основі методів математичного і комп'ютерного моделювання, інтелектуальний з можливостями перебудови структури і виконання функцій вимірювань.
Кількість і динаміка у просторі і часі вимірювальних величин	Одноканальні, багатоканальні, статичні, динамічні; реалізації ІВТ: ІВТ-1D, ІВТ-2D і ІВТ-3D

Значна кількість ознак класифікації ІВС, які використовувались раніше, в тому числі: вид вихідної інформації; вид структурно-

функціональної схеми системи; види діалогового режиму оператор-система та інші інтегровані в класифікації ІВС, що наведена в табл. 1.4.

Апаратно-програмна реалізація ІВС. Аналізуючи функціонування різних видів ІВС, відзначимо, що всі системи в своєму складі мають вимірювальний канал, як базову і типову підсистему ІВС. Саме вимірювальний канал (ВК) є найбільшою структурною одиницею ІВС, метрологічні характеристики (МХ) якої нормуються.

Вимірювальний канал – це конструктивно або функціонально виділена частина ІВС, що виконує закінчену функцію від сприйняття вимірюваної величини до отримання результату її вимірювань, який відображається числом або відповідним йому кодом чи формування відповідного аналогового сигналу, один з параметрів якого є функцією значення вимірюваної величини.

Апаратна частина ВК. Структура ВК в більшості випадків є типовою і являє собою послідовне з'єднання різного роду модулів засобів вимірювань. Узагальнена структура ВК ІВС представлена на рис. 1.4.

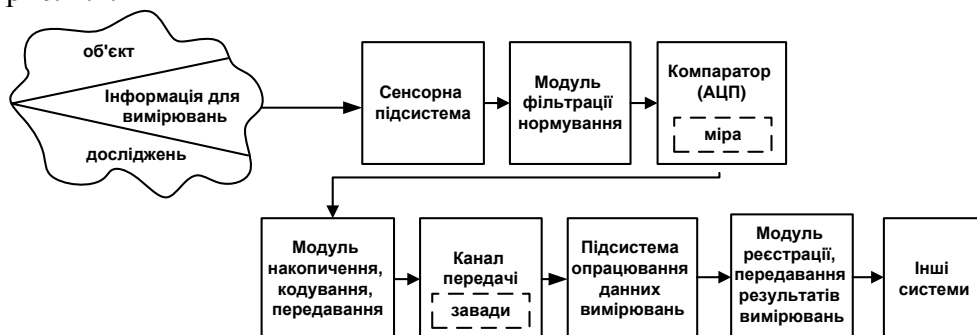


Рис. 1.4. Узагальнена структура ВК ІВС

У сенсорну підсистему ВК входять сенсори (первинні вимірювальні перетворювачі), підсилювачі і фільтри, які у сукупності перетворюють інформацію для вимірювань об'єкта досліджень у вимірювальний сигнал, зазвичай електричний. У теорії ІВС виділений науково-технічний напрям досліджень – сенсорна електроніка. У рамках цього напрямку розробляються нові сенсори фізичних величин, хімічні сенсори, біосенсори, радіаційні, оптичні та оптоелектронні сенсори, акустоелектронні сенсори, наносенсори і ін. За особливостями функціонування сенсори поділяють на: лінійні або нелінійні; інерційні або безінерційні; з інваріантними (незалежні від зміни певних параметрів і характеристик об'єкту досліджень) або неінваріантними в часі і/або в просторі. Тому під час оцінювання функціонування ВК

ІВС, в першу чергу, аналізують сенсор як один з основних модулів ВК.

Другий модуль ВК забезпечує підсилення, фільтрацію (узгоджену, оптимальну) з метою зменшення впливу завад, коригування (лінеаризацію) функції перетворення сенсора, зменшення адитивних і мультиплікативних похибок перетворень, нормування вимірювальних сигналів.

У сучасних ІВС як компаратор здебільшого використовуються аналогово-цифрові перетворювачі (АЦП). Характеристики АЦП, в першу чергу, частота дискретизації і крок (рівень) квантування за інтенсивністю вимірювальних сигналів, узгоджуються із завданням вимірювання і вибираються на основі аналізу метрологічних характеристик існуючих АЦП.

У загальному випадку отримані з виходу АЦП дані вимірювань в модулі ВК накопичуються, кодуються з метою захисту інформації для подальшого передавання по каналу передачі даних вимірювань.

Як канали зв'язку використовуються: оптоволоконні, лінії зв'язку; радіоканали; високовольтні лінії електропередачі з відповідною каналоутворюючою апаратурою; акустичні і гідроакустичні канали та ін.

Підсистема опрацювання даних вимірювань з відповідним інтерфейсом забезпечує приймання, декодування і опрацювання даних вимірювань з метою отримання результату вимірювань.

Програмна частина ІВС. Розглядаючи ІВС як єдиний апаратно-програмний комплекс виходять з того, що «м'яке обладнання», до складу якого входить функціональне інформаційне, математичне і програмне забезпечення (відповідно ІЗ, МЗ і ПЗ) не є самостійними об'єктами і працюють тільки у складі ІВС. Тому працездатність «м'якого обладнання» здебільшого оцінюється не окремо, а за результатами роботи ІВС в цілому.

Програмна частина системи має характерні особливості:

вона не схильна до зношування;

якщо виявлені в процесі налагодження і дослідної експлуатації ПЗ помилки усуваються, а нові не вносяться, то такий об'єкт розглядається як *нестаріюча* система (за термінологією теорії надійності);

працездатність програм значною мірою залежить від використуваної вхідної інформації, оскільки від неї залежить результат реалізації програми; якщо вхідна інформація сама містить дефекти, програма видасть помилковий результат і за відсутності помилок програми;

якщо помилки, що виникають під час роботи, не діагностувати і не усувати, то помилки ПЗ будуть мати систематичний характер;

працездатність ПЗ залежить від області застосування ІВС; у випадку розширення діапазону функціонування або зміні області застосування ІВС працездатність може істотно змінюватися без зміни ПЗ.

Ефективність функціонування апаратної частини ІВС оцінюють шляхом опрацювання статистичних даних за результатами експлуатації певної кількості однотипних зразків ІВС. Можливості використання такого підходу для програмного продукту обмежені, оскільки копії програмного забезпечення ідентичні і разом з їх копіюванням вносяться і однакові дефекти – проектні помилки. Існує й інша можливість використання попереднього досвіду функціонування ПЗ. Характеристика числа допущених проектних помилок є досить стійким показником якості роботи сформованого колективу програмістів і використовуваних ними засобів проектування ПЗ. Реєстрація відомостей про проектні помилки у всіх раніше виконаних проекта та їх відповідне статистичне опрацювання дає можливість отримати заслуговуючі на довіру вихідні дані для оцінювання працездатності нового ПЗ. Якщо такі дані відсутні, то використовують більш загальні відомості про процес проектування ПЗ або дані результатів налагодження ПЗ розроблюваного проекту. Щоби за цими даними оцінити працездатність ПЗ, розробляють і досліджують відповідні моделі узгоджені з певним етапом життєвого циклу програми.

На ранніх стадіях проектування використовують опис алгоритмів за входами і виходами (опис «чорного ящика») або структуру алгоритму як сукупність структурних елементів і опис кожного з них за входами і виходами (опис «білого ящика»). Після розроблення текстів програм, можна використовувати параметри програм: словник мови програмування, кількість операцій, операндів, використовуваних підпрограм, локальних міток і т. ін.

У процесі налагодження та експлуатації ПЗ, коли з'являються статистичні дані про виявлені дефекти, вихідне число дефектів як одну з важливих характеристик якості програмування оцінюють методами математичної статистики.

Моделі працездатності ПЗ і методи її оцінювання поділяються на дві групи:

моделі і методи проектного оцінювання працездатності, що ґрунтуються на вихідних даних, які можна отримати до початку налагодження та експлуатації програм;

моделі і методи статистичного оцінювання працездатності ПЗ, що ґрунтуються на результатах налагодження на етапах дослідної або нормальної експлуатації ІВС.

Як вихідні дані використовуються структурна схема функціонального ПЗ для кожної функціональної самостійної операції, а також опис входів і виходів кожного структурного елемента, міжмодульних і зовнішніх зв'язків комплексу алгоритмів і програм. Типова структура функціонального ПЗ має у своєму складі ПЗ верхнього і нижнього рівнів. У свою чергу типова структура функціонального ПЗ нижнього рівня спрямована на опрацювання даних, введення і виведення даних і т. ін.

Кожен алгоритм може бути поділений на секції (модулі) певного розміру відповідно до рекомендацій технології програмування. На ранніх етапах проектування в умовах значної невизначеності до структурних характеристик додають ще рівень використовуваних мов програмування. На більш пізніх етапах, коли вже розроблені тексти програм, можуть бути використані методики проектного оцінювання програмних модулів, а в подальшому і ПЗ в цілому.

Типова структура програмного забезпечення ІВС представлена на рис. 1.5. За функціональним призначенням всі модулі ПЗ поділяються на чотири групи: взаємодії з модулями системи, виділення первинних інформативних параметрів, статистичного опрацювання даних і керування процесом вимірювання.

Програмні модулі взаємодії з компонентами системи забезпечують програмне узгодження цих модулів з ПЗ системи в цілому. Взаємодія ПЗ із апаратними засобами ІВС реалізується шляхом послідовного виклику функцій, які відповідають різним завданням – ініціалізації окремих підсистем, отриманню службових даних, налаштуванню динамічного діапазону та частоти опитування, запуску процедури збирання даних, очікуванню відповіді, отриманню вибірок інформаційного сигналу та завершенню роботи.



Рис. 1.5. Типова структура програмного забезпечення ІВС

Крім того ці модулі реалізують такі операції:
пошук підключеного пристрою;
отримання ідентифікатора пристрою;
конфігурування та запуск процесу перетворення;
перевірка поточного стану процесу збору інформації та передачі сигналу про його завершення;
зчитування отриманих даних та передачу їх для подальшого опрацювання;
зупиненку процесу збирання даних та завершення сеансу роботи з пристроєм.

Програмні модулі виділення первинних інформативних параметрів забезпечують накопичення первинних даних і їх передавання в інші модулі ПЗ системи для подальшого статистичного оброблення, побудови розподілів інформаційних полів, графічного представлення даних (наприклад, відео та теплових зображень), локалізації місць з перевищенням граничних значень контрольованого параметра, автоматичного розпізнавання класів даних, що відповідають різним станам об'єкта досліджень. Зокрема модуль накопичення первинних даних виконує зберігання вхідних інформаційних масивів даних, зменшення їхньої розмірності за рахунок відбору певних діагностичних ознак, які використовуються для подальшого опрацювання, що дає можливість підвищити ефективність опрацювання даних та роботи системи в цілому, а також нормування вхідних даних для приведення його до заданого діапазону значень.

Модуль опрацювання графічних (відео, теплових полів, розподіл параметрів електромагнітн полів у просторі) даних реалізує наступні алгоритми роботи із зображеннями – підвищення просторової розрізнявальної здатності, підвищення контрастності, фільтрації зображень, виділення геометричних примітивів, прив'язки до еталонів. Крім того, даний модуль дозволяє проводити геометричні вимірювання з використанням зображень в 2D і 3D просторах для локалізації координат і розмірів об'єктів та його окремих зон (наприклад, дефектів).

Модуль розпізнавання класів даних здійснює класифікацію вхідних даних за певними заданими ознаками, розбиття множини векторів-ознак на кластери та розпізнавання різних класів (образів). Такі модулі можуть розроблятися на основі нейромережових

технологій, які дозволяють будувати правила прийняття рішення за мінімально можливого обсягу експериментальних даних.

Суттєве функціональне навантаження в ІВС мають модулі статистичного опрацювання експериментальних даних.

Модуль попереднього цензурування даних вимірювань призначено для зменшення впливу вибірових значень з надлишковими похибками та виявлення прогресуючих чи періодичних трендів. Перша задача вирішується фільтрацією даних, наприклад віконною чи медіанною, в залежності від розмірності вхідного вектора, або застосуванням статистичних критеріїв, які для заданої довірчої ймовірності дозволяють розрахувати граничні вибірові значення.

Ідентифікація законів розподілу експериментальних даних необхідна для підтвердження факту їх однорідності та вибору моделі інформаційних параметрів та характеристик, а також може бути використана як діагностична ознака. Модуль ідентифікації законів розподілу може використовувати різні статистичні інструменти, наприклад направлені критерії перевірки на гауссовість, на рівномірність, апроксимацію законів розподілу лінійних даних кривими Пірсона тощо.

Модуль визначення вибірових статистик оцінює статистичні параметри експериментальних даних з урахуванням визначеного закону розподілу для подальшого використання в розрахунках характеристик вимірюваних параметрів чи полів або прогнозу стану об'єкта вимірювання (контролю). Поля інформаційних параметрів можна будувати як в декартових, так і в полярних координатах.

Програмний модуль визначення характеристик полів дозволяє отримувати зображення полів у 3D просторі, а також досліджувати зміну їх характеристик у часі.

Модуль розрахунку прогнозу здійснює побудову регресійної моделі поведінки вимірюваних параметрів в часі (наприклад, на основі сплайн - апроксимації експериментальних даних).

Модулі керування процесом вимірювання забезпечують загане керування режимами роботи ІВС, в тому числі і в інтерактивному режимі, відображення результатів вимірювань із застосуванням 2D і 3D графіки, налаштування та калібрування окремих складових ІВС та проведення їх діагностики.

Характерні властивості і приклад апаратно-програмної реалізації сучасних ІВС. Аналіз різноманітних і різнопланових завдань вимірювань широкого кола предметних областей застосування різних видів ІВС не дає практичної можливості представити схематично узагальнену структуру ІВС, тобто у кожному конкретному випадку

апаратно-програмна реалізація ІВС має свою специфіку і характерні особливості. Але послідовність операцій перетворення інформації для вимірювання об'єкта досліджень у результат вимірювань на основі використання функціональних модулів формування і опрацювання даних вимірювань ІВС дає можливість визначити апаратний склад системи. У залежності від кількості та динаміки у просторі і часі вимірювальних величин (n величин), загальна структура ІВС включає n вимірювальних каналів, апаратна і програмна частина одного з них була описана вище.

Апаратна частина ІВС. Наведемо перелік апаратних модулів, які можуть бути включені в структури різних видів ІВС:

множина різних сенсорів або сенсорних модулів, які розміщені стаціонарно, або можуть переміщатись у просторі по заданому закону;

множина вимірюваних перетворювачів;

групи АЦП, ЦАП та аналогових пристроїв допускового контролю;

множина цифрових пристроїв, яка включає пристрої формування імпульсів, перетворення кодів, комутатори, спеціалізовані цифрові обчислювані пристрої, пристрої пам'яті, пристрої порівняння кодів, канали цифрового зв'язку, універсальні обчислювальні пристрої – мікропроцесори, мікроконтролери тощо;

група цифрових пристроїв виводу, відображення і реєстрації;

інтерфейсні пристрої, які включають системи шин, інтерфейсні вузли і інтерфейсні пристрої аналогових блоків, які забезпечують приймання сигналів управління і передачі інформації про стан блоків;

пристрої управління, які забезпечують формування, приймання сигналів управління і їх передачу на виконуючі пристрої, для здійснення впливу на об'єкт дослідження.

Наведені вище функціональні модулі з'єднуються між собою безпосередньо або через стандартні інтерфейси за умови сумісності їх інформаційних, метрологічних, конструктивних, енергетичних та експлуатаційних характеристик.

Для прикладу на рис. 1.6 наведено структурну схему сучасної ІВС у складі безпілотного авіаційного комплексу (БАК), який включає безпілотний літальний апарат (БПЛА) та наземний пункт управління. Така ІВС забезпечує дистанційне визначення характеристик довкілля просторово розгалужених промислових об'єктів в автоматизованому режимі, безлюдних технологій вимірювань та контролю, функціонування систем у важкодоступних місцях та за умов

техногенної, у тому числі радіаційної, небезпеки, багатофункціональність (адаптацію) призначення за рахунок зміни модулів сенсорних підсистем, проведення динамічних вимірювань у 3D просторових координатах.

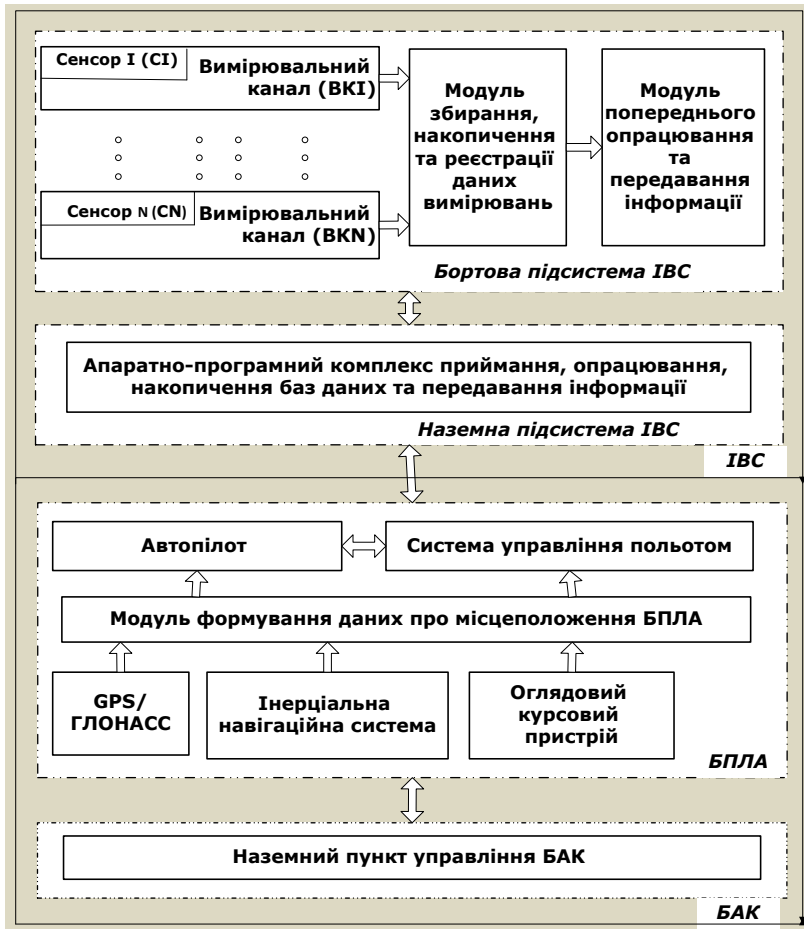


Рис. 1.6. Структурна схема безпілотного авіаційного комплексу з ІВС

Перспективи створення та використання ІВС в умовах постійного

загального розвитку науки, техніки і виробництва та досліджень взаємодії космосу, атмосфери Землі та об'єктів життєдіяльності людини структурно ілюструється рис. 1.7.



Рис. 1.7. Ілюстрація шляхів розвитку ІВС

Необхідність розроблення і удосконалення ІВС обумовлює потребу постійного розвитку теоретичних засад їх створення і функціонування, що стимулює процес включення нових областей знань, необхідних для вирішення актуальних завдань вимірювань.

Основні висновки

Вимірювання відіграють важливу роль у розвитку економіки, науки і техніки. Теоретичні основи ІВС є одним із напрямів теорії вимірювань, а ІВС є основним засобом забезпечення єдності вимірювань на практиці.

Об'єктами вимірювань є значна множина явищ, процесів, систем різної природи в значній кількості предметних областях.

Предметами вимірювання є моделі, методи, способи, технічні, інформаційні, програмні та інші засоби забезпечення єдності вимірювань.

Існують три основні завдання вимірювань: гомоморфне відображення інформації для вимірювань об'єкта досліджень у фізичній та математичній моделі, використання однорідності у просторі і стаціонарності у часі мір при проведенні процесів вимірювання та забезпечення захисту інформації. Вирішення вказаних завдань вимірювань об'єднують на базі застосування фізичних і математичних моделей. Математичні моделі вимірюваних сигналів у загальному виді відображають динаміку у просторі і часі вимірювальних величин (фізичних величин).

Класифікація ІВС базується на використанні трьох основних ознак: функціонального призначення; режимів (алгоритмів) функціонування; кількості і динаміки у просторі і часі вимірювальних величин.

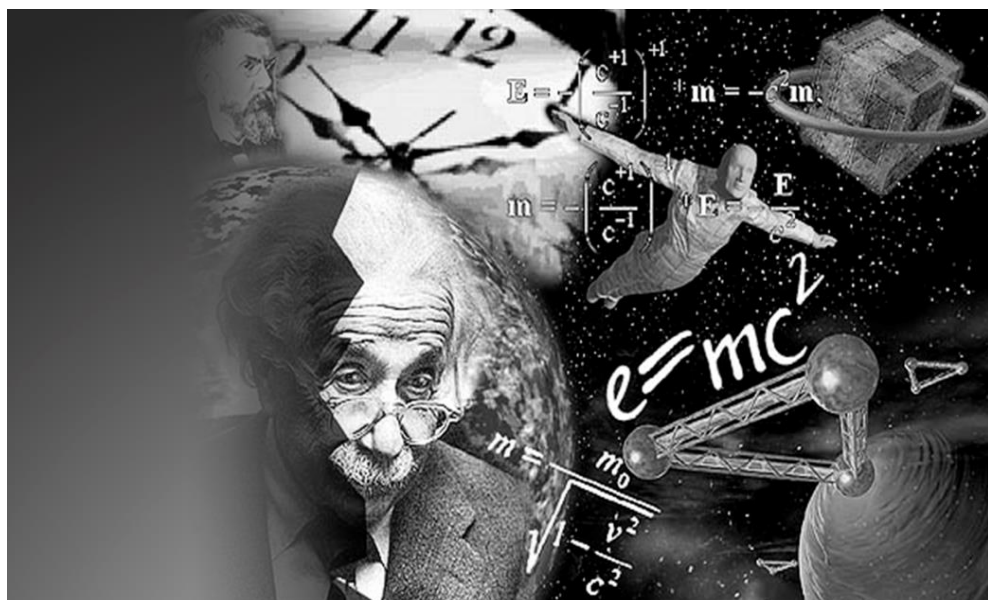
Запитання для самоконтролю

- 1. Дайте визначення метрології і трьох основних напрямів дослідження метрології.*
- 2. Що таке досліджувана величина? Назвіть конкретні види таких величин.*
- 3. Сформулюйте три основні завдання вимірювань.*
- 4. Дайте визначення математичної моделі.*
- 5. Наведіть класифікаційні ознаки сучасних інформаційно-вимірювальних технологій.*
- 6. Наведіть і розкрийте зміст основних етапів життєвого циклу ІВС.*
- 7. Проведіть порівняльний аналіз оцінювання точності вимірювань на основі використання Концепції похибок і Концепції невизначеності.*
- 8. Навести класифікацію ІВС з використанням відповідних ознак класифікації.*
- 9. Провести аналіз апаратної і програмної частини ВК ІВС.*
- 10. Навести основні модулі апаратної частини сучасних ІВС.*
- 11. Розкрити перспективи використання БАК при створенні нових ІВС.*

Conclusions

Ключові слова

вимірювання	measurement
процес вимірювання	measurement process
засіб вимірювання	measurement tool
інформаційно-вимірювальна система	information-measurement system
апаратно-програмний засіб вимірювання	hardware-software measurement tool
невизначеність вимірювань	measurement uncertainty
результат вимірювань	measurement result
точність вимірювання	measurement error



КІЛЬКІСНІ ОЦІНКИ ВИМІРЮВАЛЬНОЇ ІНФОРМАЦІЇ

2

- 2.1. Вимірювальна інформація, її функції та властивості
- 2.2. Кількісні характеристики інформації
- 2.3. Міри інформації
- 2.4. Ентропія та її властивості
- 2.5. Продуктивність і надмірність джерела інформації
- 2.6. Зв'язок інформації з характеристиками сигналів

2.1. Вимірювальна інформація, її функції та властивості

Термін *інформація* походить від латинського *informatio*, що означає *роз'яснення, інформування, викладення*. У широкому значенні інформація - це загальнонаукове поняття, яке охоплює обмін інформацією між людьми, обмін сигналами між живою і неживою природою, людьми і пристроями. У цьому сенсі інформація є відображенням реального світу за допомогою відомостей (даних, повідомлень).

Поняття «вимірювальна інформація» має кілька визначень, залежних від поглядів на неї. Ми будемо дотримуватися тлумачення, згідно з яким поняття «вимірювальна інформація» об'єднує характеристики об'єктів живої і неживої природи, які є потенційним джерелом відомостей різного роду для людей, а також і самі відомості, які отримують люди.

Як робоче пропонуємо менш коротке визначення: *вимірювальна інформація* – це сукупність відомостей про навколишній світ, об'єкти і явища навколишнього середовища, а також процеси, які відбуваються в них, відомості про їх параметри, властивості і стан, які зменшують ступінь невизначеності, неповноти знань про них .

Методи аналізу вимірювальної інформації ілюструє рис. 2.1.



Рис. 2.1. Методи аналізу вимірювальної інформації

Згідно синтаксичному підходу здійснюється структурний аналіз вимірювальної інформації (кількість повідомлень, форма їх подання, статистичні характеристики появи тощо).

Семантичний підхід полягає в аналізі змістовності вимірювальної інформації і впливу її на людину.

Прагматичний підхід пов'язаний з оцінюванням корисності, цінності отриманої вимірювальної інформації.

Важливо відзначити динамічні властивості вимірювальної інформації. Об'єкти вимірювальної інформації навколишнього світу постійно змінюються з різною швидкістю. Ці зміни мають випадковий характер щодо моментів прояву, характеристик, діапазону зміни характеристик і законів їх зміни. Вимірювальна інформація як відображення цих процесів також має динамічний, випадковий характер. Завжди необхідна оновлена вимірювальна інформація, покликана усувати невизначеність ситуації в даний момент.

Узагальненою функцією інформації є забезпечення виживання, розвитку людини і суспільства в складному і мінливому

Розділ 2. Кількісні оцінки вимірювальної інформації

навколишньому світі. Вона складається з ряду часткових функцій, які мають різну ступінь виявлення в залежності від зовнішніх умов і характеру діяльності людини. Найчастіше ми зустрічаємося з проявами таких функцій інформації:

управлінської, присутньої у всіх сферах повсякденного життя і покликаної допомогти людині у виборі варіанту своєї поведінки або цілеспрямованого впливу на об'єкти і процеси реального світу;

комунікаційної, присутньої під час обміну інформацією між людьми і спрямованої на організацію взаємодії між ними;

пізнавальної, зумовлюється потребою в інформації заради загального розвитку, отримання спеціальності і, в загальному, для задоволення потягу до нового безвідносно до його прагматичного значення;

психологічної, позначається при формуванні певного емоційного настрою за допомогою деяких видів інформації, способів її подання людині.

Узагальненою властивістю інформаційних функцій є їх спрямованість на усунення невизначеності нашого уявлення про стан об'єктів і процесів реального світу, які нас цікавлять, або на даний момент, або після деяких впливів на них внутрішніх або зовнішніх сил. Саме на усунення невизначеності зорієнтовані процеси збирання, передавання і опрацювання інформації. Ступінь зміни невизначеності ситуації покладена в основу кількісної міри інформації.

Адекватність інформації - це певний рівень відповідності образу, який створюється за допомогою отриманої інформації, реальному об'єкту, процесу, явищу і ін. На практиці навряд чи можна досягти повної адекватності інформації. Завжди існує деяка ступінь невизначеності. Від ступеня адекватності інформації реальному стану об'єкта залежить правильність прийнятих людиною рішень.

Адекватність інформації може виражатися в трьох формах: *синтаксичній, семантичній, прагматичній*.

Синтаксична адекватність відображає формально-структурні характеристики інформації без її змістовного наповнення. На синтаксичному рівні враховуються тип носія і спосіб подання інформації, швидкість передачі і обробки, розміри кодів подання інформації, надійність і точність перетворення цих кодів і т.п. Формалізовану інформацію, яка розглядається тільки з синтаксичних позицій, зазвичай називають даними, оскільки при цьому абстрагуються від змістовного аспекту. Ця форма адекватності сприяє сприйняттю зовнішніх структурних, тобто синтаксичних, характеристик інформації.

Семантична (змістова) адекватність визначає іншу ступінь відповідності образу досліджуваного об'єкта самому об'єкту.

Семантичний аспект передбачає врахування змістовного наповнення інформації. На цьому рівні аналізуються ті відомості, які відображає інформація, і розглядаються змістовні зв'язки між ними.

Ця форма адекватності дає можливість формувати поняття і уявлення, виявляти значення, зміст інформації і узагальнювати її.

Прагматична (споживча) адекватність відображає співвідношення між інформацією та її споживачем, ступінь відповідності інформації тій меті управління, яка на її основі реалізується. Виявляються прагматичні властивості інформації тільки при наявності єдності інформації (об'єкта), користувача і цілі управління. Прагматичний аспект розгляду пов'язаний з цінністю, корисністю інформації при підготовці споживачем рішення для досягнення поставленої мети. Відповідно до цього аналізуються споживчі властивості інформації.

Ця форма адекватності пов'язана з практичним використанням інформації та її відповідності цільової функції роботи системи.

Вимірювальна інформація характеризує співвідношення між джерелом вимірювальної інформації (об'єктом дослідження), сполученням і його споживачем.

Якщо мова йде про автоматизовану роботу з вимірюваною інформацією за допомогою деяких технічних пристроїв, перш за все цікавляться не змістом, а джерелом повідомлення та кількістю символів, яке містить це повідомлення.

Інформаційні об'єкти - це предмети, процеси, явища матеріальної або нематеріальної природи, які розглядаються з точки зору їх інформаційних властивостей.

Джерело вимірювальної інформації або повідомлення - це фізичний об'єкт (система або явище), що формує передане повідомлення. Саме повідомлення – це значення або зміна деякої фізичної величини, яка відображає стан об'єкта (системи або явища).

При роботі з інформацією завжди існують її джерело і споживач (одержувач). Засоби і процеси, які забезпечують передачу повідомлень від джерела інформації до її споживача, називаються інформаційно-комунікаційними системами. Така система в широкому розумінні – як комплекс організаційно-технічних заходів, інформаційних технологій та інформаційних ресурсів, призначених для забезпечення інформаційних процесів, зокрема створення, поширення, використання, зберігання та знищення інформації.

Щодо *комп'ютерного опрацювання даних* під інформацією розуміють деяку послідовність символічних позначень (букв, цифр, закодованих графічних образів і звуків тощо), яка несе змістовне

Розділ 2. Кількісні оцінки вимірювальної інформації

навантаження. Кожен новий символ в такій послідовності збільшує *інформаційний обсяг повідомлення*, а самі повідомлення кодуються для подальшого опрацювання засобами обчислювальної техніки.

2.2. Кількісні характеристики інформації

У теорії інформації під *кількістю інформації* розуміють міру зменшення невизначеності знання. Знаходження такої міри вимагає оцінки і обліку кількості переданої інформації.

В цьому розумінні *кількість інформації* числова характеристика сигналу, яка не залежить від його форми і змісту і характеризує невизначеність, яка зникає після отримання повідомлення у вигляді даного сигналу. В цьому випадку кількість інформації залежить від вірогідності отримання повідомлення про ту або іншу подію.

Інформаційний об'єм повідомлення (інформаційна місткість повідомлення) – кількість інформації в повідомленні, виміряна в бітах, байтах або похідних величинах.

Для абсолютно достовірної події (подія обов'язково відбудеться, тому її ймовірність дорівнює 1, кількість інформації в повідомленні про неї дорівнює 0. Чим неймовірніша подія, тим більшу інформацію про нього несе повідомлення. Лише при рівноймовірних відповідях, відповідь “так” чи “ні” несе один біт інформації.

Г. Хартлі запропонував формулу для обчислення кількості інформації про об'єкт, який може знаходитися в одному з рівноймовірних N станів:

$$I = \log_2 N. \quad (2.1)$$

Формулу для обчислення кількості інформації для подій з різною ймовірністю запропонував К. Шеннон в 1948 р. У цьому випадку кількість інформації визначається за формулою:

$$I = -\sum_{i=1}^N p_i \log_2 p_i, \quad (2.2)$$

де N - кількість можливих подій $\{x_i, i = \overline{1, N}\}$; $p_i = P(x_i)$ - ймовірність окремої події x_i .

Кількісна міра інформації. Кількісна міра інформації (Г. Хартлі і К. Шеннон) враховує тільки кількість N повідомлень, та ймовірності $\{x_i\}$ їх надходження на вхід каналу передачі. Сукупність повідомлень подають у вигляді наступної системи

$$X = \{(x_1, p_1), \dots, (x_n, p_n)\}, \quad \sum_{i=1}^N p(x_i) = 1, \quad (2.3)$$

де: x_i – окремі повідомлення (або їх типи, класи), $p(x_i)$ – апіорна ймовірність появи повідомлень x_i . В результаті передавання повідомлення x_i буде отримане повідомлення y_j . Апостеріорна (післядослідна) ймовірність присутності x_i в y_j дорівнює $P(x_i/y_j)$. В основу міри кількості інформації встановлені зміни ймовірності появи повідомлень від апіорного значення $P(x_i)$ на вході каналу до апостеріорного значення $P(x_i/y_j)$ на виході каналу, пов'язані із викривленнями інформації у каналі зв'язку. Порівнюючи ймовірність $P(x_i)$ і $P(x_i/y_j)$ можна встановити міру кількості інформації, переданої до споживача. Зручною мірою виявився логарифм відношення апостеріорної ймовірності до апіорної.

Кількість інформації, що міститься в події y_j щодо події x_i визначається за формулою

$$I(x_i; y_j) = \log \left[P(x_i / y_j) / P(x_i) \right]. \quad (2.4)$$

Основою логарифма найчастіше приймають: 2, $e \approx 2,72$, або 10 (може бути 8, 16 та ін). Залежно від основи змінюються одиниці вимірювання кількості інформації (біт – двійкові од.; ніт – натуральні од.; хартлі (діт) – десяткові од.)

Властивості кількості інформації:

1. *Властивість симетрії.* Інформація, що міститься в y_j щодо x_i , дорівнює інформації, що міститься в x_i щодо y_j . Це стає очевидним, якщо чисельник і знаменник в (2.4) помножити на $P(y_j)$ і виконати перетворення:

В результаті вказаної властивості величину $I(x_i, y_j)$ називають *кількістю взаємної інформації* між x_i і y_j .

2. *Властивість адитивності.* Інформація, що міститься в парі символів y_j, z_k щодо x_i , дорівнює сумі інформації, що міститься в y_j щодо x_i і інформації, що міститься в z_k щодо x_i за умови, що значення y_j відоме

$$I(x_i; y_j, z_k) = I(x_i; y_j) + I(x_i; z_k / y_j) \quad (2.5)$$

Вираз (2.4) з урахуванням (2.5) приймає вид

$$I(x_i; y_j) = I(x_i) - I(x_i / y_j),$$

де $I(x_i / y_j) = -\log P(x_i / y_j)$ – умовна власна інформація.

2.3. Міри інформації

Для кількісної оцінки інформації досить часто застосовуються синтаксична, семантична та прагматична міри інформації (рис. 2.2).

Синтаксична міра інформації - оперує із знеособленою інформацією, що не передає смислового відношення до об'єкту. На

Розділ 2. Кількісні оцінки вимірювальної інформації

синтаксичному рівні враховуються тип носія і спосіб представлення інформації, швидкість передавання та опрацювання, розміри кодів представлення інформації.

Для визначення такої кількісної міри інформації вводять два параметри: *об'єм даних* V_d і *кількість інформації* I .

Об'єм даних (V_d) - інформаційний об'єм повідомлення або об'єм пам'яті, необхідний для зберігання повідомлення без яких-небудь змін.

Об'єм даних V_d у повідомленні вимірюється кількістю символів (розрядів). Одиниця вимірювання залежить від системи числення.

Кількість інформації (I) в повідомленні про об'єкт, який може знаходитися в одному з рівноімовірних N станів, визначають, згідно вище зазначеної формули Р. Хартлі (2.1).

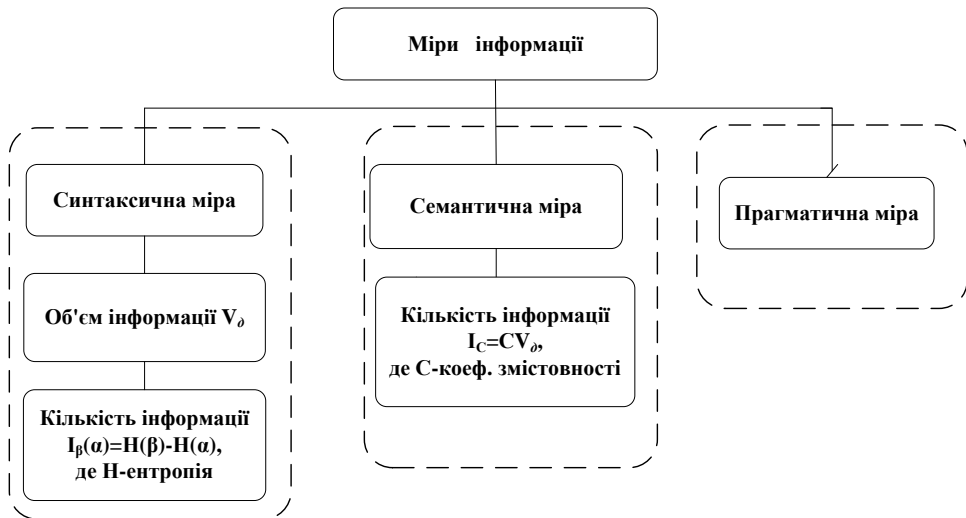


Рис. 2.2 Класифікація мір інформації

Одиниці вимірювання інформації залежать від вживаної системи числення: у двійковій системі одиниця вимірювання – біт (bit – binary digit – двійковий розряд), у десятковій – діт.

Біт – кількість інформації, яка потрібна для розрізнення двох рівноімовірних повідомлень. Біт – дуже мала одиниця вимірювання. На практиці частіше застосовується одиниця – *байт*, що дорівнює *восьми бітам*. Саме вісім біт потрібно для того, щоб закодувати будь-який з 256 символів алфавіту клавіатури комп'ютера ($256 = 2^8$).

Широко використовуються наступні похідні одиниці вимірювання інформації:

$$1 \text{ Кілобайт (Кбайт)} = 1024 \text{ байт} = 2^{10} \text{ байт},$$

$$1 \text{ Мегабайт (Мбайт)} = 1024 \text{ Кбайт} = 2^{20} \text{ байт},$$

1 Гігабайт (Гбайт) = 1024 Мбайт = 2^{30} байт.

1 Терабайт (Тбайт) = 1024 Гбайт = 2^{40} байт,

1 Петабайт (Пбайт) = 1024 Тбайт = 2^{50} байт.

За одиницю інформації можна було б вибрати кількість інформації, необхідної для розрізнення, наприклад, десяти рівноймовірних повідомлень. Це буде не двійкова (біт), а десяткова (діт) одиниця інформації. У комп'ютерній практиці слово "біт" використовується також як одиниця вимірювання об'єму пам'яті. Елемент пам'яті розміром в 1 біт може знаходитися в двох станах ("увімкнено" та "вимкнено") і в нього може бути записана одна цифра (0 або 1).

Кількість інформації (I) на синтаксичному рівні визначається через поняття ентропії системи.

Нехай до отримання інформації споживач має деякі попередні (апріорні) відомості про систему y . Мірою його непоінформованості про систему є функція $H(y)$, яка в той же час служить і мірою невизначеності стану системи.

Після отримання повідомлення x одержувач дістав додаткову інформацію $I(x)$, що зменшила його апіорну непоінформованість так, що невизначеність стану системи після отримання повідомлення x стала дорівнювати $H(x)$. Тоді *кількість інформації* $I(x)$ про систему, отриману в повідомленні x є наступною

$$I(x) = H(y) - H(x). \quad (2.6)$$

тобто *кількість інформації* вимірюється величиною зміни (зменшення) невизначеності стану системи.

Ентропія системи $H(x)$, яка має N можливих станів, визначається за формулою К. Шеннона у виді

$$H(y) = -\sum_{i=1}^N p_i \log p_i,$$

де p_i - ймовірність того, що система знаходиться в i -му стані. Для випадку, коли всі стани системи рівноймовірні $p_i = 1/N$, її ентропія визначається так

$$H(y) = -\sum_{i=1}^N (1/N) \log (1/N) = \log N$$

Приклад. Часто інформація кодується числовими кодами в тій або іншій системі числення, особливо це актуально при представленні інформації в комп'ютері. Природно, що одна і та ж кількість розрядів в різних системах числення може передати різне число станів об'єкту, що відображається, що можна представити у вигляді співвідношення $N = m^n$ де N – число множини станів, що відображаються; m – основа системи числення (різноманітність символів, вживаних в алфавіті); n – число розрядів (символів) в повідомленні. Допустимо, що по каналу зв'язку передається n -розрядне повідомлення, що

Розділ 2. Кількісні оцінки вимірювальної інформації

використовує m різних символів. Оскільки кількість кодових комбінацій буде $N = m^n$, то при рівномірності появи будь-якої з них кількість інформації, абонент приймає наступне повідомлення $I = \log N = n \log m$.

Якщо за основу логарифму прийняти m , то $I=n$. В даному випадку кількість інформації (за умови повного апіорного незнання абонентом змісту повідомлення) буде дорівнювати об'єму даних $I = V_d$, отриманих по каналу зв'язку.

Коефіцієнт (ступінь) інформативності (лаконічність) повідомлення визначається відношенням кількості інформації до об'єму даних у виді

$$Y = I/V_d.$$

Для підвищення інформативності розробляються спеціальні методи оптимального кодування інформації.

Семантична міра інформації - служить для вимірювання змісту інформації. Для цього використовується *тезаурусна міра*, яка враховує здатність одержувача повідомлення його сприйняти.

Тезаурусом називають сукупність відомостей, які має в своєму розпорядженні користувач або система.

Максимальну кількість семантичної інформації I_c споживач набуде при узгодженні її змісту S зі своїм тезаурусом $S_p = S_{popt}$, коли інформація, що поступає, зрозуміла користувачу і несе йому раніше невідомі (відсутні в його тезаурусі) відомості. Кількість семантичної інформації в повідомленні є

відносною і залежить від підготовленості одержувача. Одна і та ж інформація може мати зміст для фахівця і бути абсолютно даремною для дилетанта.

Залежно від співвідношень між смисловим змістом інформації S і тезаурусом користувача S_p змінюється кількість семантичної інформації I_c , що сприймається користувачем і включається їм надалі в свій тезаурус. Характер такої

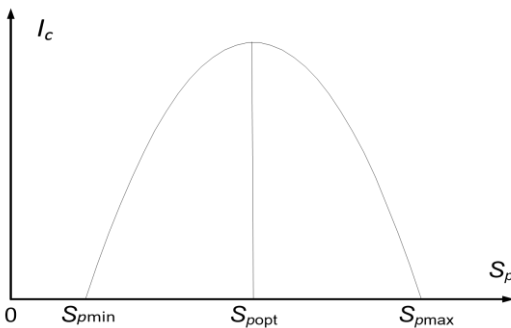


Рис. 2.3. Залежність кількості семантичної інформації, що сприйнята споживачем, від його тезауруса

залежності проілюстрований на рис. 2.3.

Розглянемо два граничні випадки, коли кількість семантичної інформації I_c дорівнює 0:

при $0 \leq S_p \leq S_{pmin}$ користувач не сприймає, не розуміє поступаючу інформацію;

при $S_p > S_{pmin}$ користувач все знає, і поступаюча інформація йому не потрібна.

Відносною мірою кількості семантичної інформації може служити коефіцієнт змістовності C , який визначається як відношення кількості семантичної інформації до її об'єму

$$C = I_c / V_d.$$

Прагматична міра інформації – визначає корисність інформації для досягнення користувачем поставлених цілей. Стосовно економічної інформації її корисність можна визначити зростанням економічних показників організації, обумовлених її використанням. Це може бути скорочення товарних запасів, збільшення швидкості обороту засобів, підвищення якості ухвалення управлінських рішень і т.д.

Цінність інформації доцільно вимірювати в тих же самих одиницях (або близьких до них), в яких вимірюється цільова функція.

2.4. Ентропія та її властивості

Ентропія визначає міру невизначеності множини всієї безлічі повідомлень на вході системи і обчислюється як середня кількість власної інформації їх повідомлень

$$I(X) = -\sum p(x_i) \log p(x_i) = H(X).$$

Властивості ентропії:

1. Ентропія $H(X)$ додатня: $H(X) > 0$.

2. Ентропія $H(X) < \log N$.

3. Величина $\log N = D$ називається інформаційною пропускною спроможністю алфавіту (інформаційною місткістю алфавіту).

4. Якщо $N = 2$, то $P(x_1) = p$, $P(x_2) = 1-p$, $H(X) = -p \log_2 p - (1-p) \log_2 (1-p)$.

5. Максимум

$H(X) = -\log_2 0,5 = \log_2 2 = 1$ – місткість двійкового алфавіту дорівнює 1 біт.

Графік залежності $H(X)$ від значення ймовірності p показано на рис. 2.4.

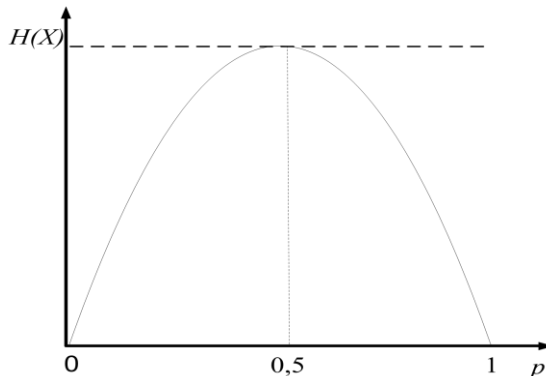


Рис. 2.4. Графік залежності $H(X)$ від значення ймовірності p

Характеристики джерела інформації - кількість інформації і ентропія, які відносяться до одного джерела, що виробляє потік незалежних або простих повідомлень є *характеристиками джерела без пам'яті*.

Розділ 2. Кількісні оцінки вимірювальної інформації

Однак у реальних умовах незалежність елементарних повідомлень, що виробляються джерелом, – явище досить рідкісне. Частіше буває саме навпаки – детермінований або статистичний зв'язок між елементами повідомлення одного або декількох джерел.

При передаванні та зберіганні даних часто також мають справу з декількома джерелами, що формують статистично зв'язані один з одним повідомлення. Повідомлення, що виробляються такими джерелами, називаються *складними повідомленнями*, а самі джерела – *джерелами з пам'яттю*.

Очевидно, що при визначенні ентропії й кількості інформації в повідомленнях, елементи яких статистично зв'язані, не можна обмежуватися тільки безумовними ймовірностями – необхідно обов'язково враховувати також умовні ймовірності появи окремих повідомлень.

Визначимо *ентропію складного повідомлення*, що виробляється двома залежними джерелами (подібним же чином визначається ентропія складного повідомлення, що виробляється одним джерелом з пам'яттю):

Нехай повідомлення першого джерела приймають значення $x_1, x_2, x_3, \dots, x_k$ з ймовірностями, відповідно, $P(x_1), P(x_2), \dots, P(x_k)$, повідомлення другого – значення y_1, y_2, \dots, y_m з ймовірностями $P(y_1), P(y_2), \dots, P(y_m)$. Спільну ентропію двох джерел X і Y можна визначити за формулою

$$H(X, Y) = -\sum_{i=1}^k \sum_{j=1}^m P(x_i, y_j) \log P(x_i, y_j), \quad (2.7)$$

де $P(x_i, y_j)$ – ймовірність спільної появи повідомлень x_i і y_j . Оскільки спільна ймовірність $P(x_i, y_j)$ за формулою Байєса визначається як

$$P(x_i, y_j) = P(x_i)P(y_j/x_i) = P(y_j)P(x_i/y_j), \quad (2.8)$$

то вираз (2.7) з урахуванням (2.8) можна записати в наступному вигляді

$$\begin{aligned} H(X, Y) &= l - \sum_{i=1}^k \sum_{j=1}^m P(x_i)P(y_j/x_i) \log \{P(x_i)P(y_j/x_i)\} = \\ &= -\sum_{i=1}^k P(x_i) \log P(x_i) \sum_{j=1}^m P(y_j/x_i) - \sum_{i=1}^k P(x_i) \sum_{j=1}^m P(y_j/x_i) \log P(y_j/x_i). \end{aligned} \quad (2.9)$$

При умові, що передачі повідомлення x_i обов'язково відповідає передача одного з повідомлень з ансамблю Y , маємо

$$\sum_{j=1}^m P(y_j/x_i) = 1,$$

і відповідно спільна ентропія джерел повідомлень $H(X, Y)$ визначиться як

$$\begin{aligned} H(X, Y) &= -\sum_{i=1}^k P(x_i) \log P(x_i) - \sum_{i=1}^k P(x_i) \sum_{j=1}^m P(y_j/x_i) \log P(y_j/x_i) = \\ &= H(X) + \sum_{i=1}^k P(x_i) H(Y/x_i), \end{aligned} \quad (2.10)$$

де $H(Y/x_i)$ – так звана *часткова умовна ентропія*, що відображає ентропію повідомлення Y за умови, що мало місце повідомлення x_i . Другий доданок в

останній формулі є усередненням $H(Y/x_i)$ по всіх повідомленнях x_i і називається *середньою умовною ентропією джерела Y за умови передачі повідомлення X* .

Таким чином *спільна ентропія двох повідомлень* визначається наступною формулою

$$H(X, Y) = H(X) + H(Y/X),$$

тобто дорівнює сумі безумовної ентропії одного з них і умовної ентропії другого.

Основні властивості ентропії складних повідомлень:

1. У випадку *статистично незалежних повідомлень X і Y спільна ентропія* дорівнює сумі ентропій джерел: $H(X, Y) = H(X) + H(Y)$, при цьому враховано, що $H(Y/X) = H(Y)$.

2. За *повної статистичної залежності повідомлень X і Y спільна ентропія* дорівнює *безумовній ентропії* одного з повідомлень. Друге повідомлення при цьому інформації не додає. Дійсно, при повній статистичній залежності повідомлень умовні ймовірності $P(y_j/x_i)$ і $P(x_i/y_j)$ рівні або нулю, або 1, тоді

$$P(x_i/y_j) \log P(x_i/y_j) = P(y_j/x_i) \log P(y_j/x_i) = 0,$$

і відповідно $H(X, Y) = H(X) = H(Y)$.

3. Умовна ентропія змінюється в межах $0 < H(Y/X) < H(Y)$.

4. Для спільної ентропії двох джерел завжди справедливе співвідношення $H(X, Y) \leq H(X) + H(Y)$, при цьому умова рівності виконується тільки для незалежних джерел повідомлень.

Отже, *при наявності зв'язку між елементарними повідомленнями ентропія джерела знижується*, причому тим більше, чим сильніше зв'язок між елементами повідомлення.

Ентропія систем з неперервною множиною станів обчислюється за правилами аналізу дискретних систем з попереднім квантуванням щільності ймовірності $p(x)$ з кроком Δx .

Тоді число станів в системі буде $N = (x_{\max} - x_{\min})/\Delta x$, а ймовірність станів $P(x_i) = p(x_i) \Delta x$. Ентропія суми дискретних повідомлень:

$$H_{\Delta x}(x) = \sum_{i=1}^N p(x_i) \Delta x \log \{p(x_i) \Delta x\}$$

Після перетворень і умови, що $\Delta x > 0$, отримуємо:

$$H_{\Delta x}(x) = H^*(X) - \log \Delta x.$$

Величина $H^*(X)$ називається *приведеною ентропією*:

$$H^*(X) = - \int_{-\infty}^{\infty} p(x) \log p(x) dx.$$

Таким чином, можна зробити наступні висновки щодо ступеня інформативності джерел повідомлень:

Розділ 2. Кількісні оцінки вимірювальної інформації

1. Ентропія джерела й кількість інформації тим більше, чим більше розмір алфавіту джерела.

2. Ентропія джерела залежить від статистичних властивостей повідомлень. Ентропія максимальна, якщо повідомлення джерела рівноймовірні й статистично незалежні.

3. Ентропія джерела, що виробляє нерівноймовірні повідомлення, завжди менше максимально досяжної.

4. За наявності статистичних зв'язків між елементарними повідомленнями (пам'яті джерела) його ентропія зменшується.

2.5. Продуктивність та надмірність джерела інформації

Розглянемо джерело двох подій s_1 і s_2 . Якщо подія s_1 зустрічається рідко, а подія s_2 - часто, то кількість інформації про реалізацію події s_1 буде значно більше, ніж події s_2 : $I(s_1) \gg I(s_2)$, де $I(s)$ - кількість інформації. Очевидно також, що кількість інформації джерела 2-х подій значно менше, ніж наприклад при кількості подій 10 або 20, тобто чим більше різних подій характеризує те чи інше явище, тим більше необхідно інформації для його опису. Таким чином, інформація є характеристикою такої загальної властивості матеріального світу, як його різноманіття.

Інформація як правило передається сигналами, фізична природа яких залежить від типу джерела повідомлень. Так наприклад сигнали – носії інформації, є механічними коливаннями у твердих матеріалах, рідинах, газах (інфразвуковими, звуковими, ультразвуковими), електричними й електромагнітними коливаннями чи хвилями (радіо, оптичними). Сигнали сприймаються приймальними пристроями, зокрема, органами почуттів живих організмів і людини, із них отримується інформація, перетворюється, чи запам'ятовується та передається по лініях зв'язку, чи перетворюється й відображається у вигляді, зручному для сприйняття людиною, осмислення й використання при прийнятті рішень.

У джерелах інформації формуються повідомлення. Повідомлення – це послідовність знаків (символів) чи неперервних сигналів, що містять ті чи інші дані, результати вимірів та інші. Множина різних знаків, що використовуються для формування повідомлень, називають *алфавітом джерела повідомлень*, а число знаків – *обсягом алфавіту*.

Неперервні повідомлення у більшості випадків є функціями часу. Типовими прикладами можуть служити мовні сигнали з виходу мікрофона, неперервно вимірювані дані про температуру, тиск, напрямок і швидкість вітру і т.д. Останнім часом вони, як правило, перетворюються в цифрові з метою підвищення якості передавання, збереження й захисту інформації.

Продуктивність джерел інформації. Швидкість передачі повідомлення визначається ресурсами її джерела на певному інтервалі часу T .

Продуктивністю джерела $H'(X)$ називається сумарна ентропія повідомлень переданих за одиницю часу: $H'(X) = H(X)/T$.

Продуктивність вимірюється в бітах на секунду. Позначивши величину $1/T$ як швидкість $\nu_c = 1/T$ (елементів в секунду), маємо

$$H'(X) = \nu_c \cdot H(X).$$

Аналогічним чином, розділивши значення ентропії та кількості інформації на T і позначаючи $H'(X/Y) = H(X/Y)/T$, $I'(X,Y) = I(X,Y)/T$, одержимо відповідну рівність для умовних ентропії і кількості інформації, розрахованих на одне повідомлення в одиницю часу.

Швидкістю передачі інформації називається величина кількості інформації $I'(X,Y)$ алфавітного ансамблю сигналів на вході системи, віднесена до одиниці часу. Якщо, наприклад X ансамбль сигналів на вході дискретного каналу, а Y ансамбль сигналів на його виході, то швидкість передачі інформації по каналу

$$I'(X,Y) = H'(X) - H'(X/Y) = H'(Y) - H'(Y/X). \quad (2.11)$$

Це співвідношення наочно ілюструється рис. 2.5.

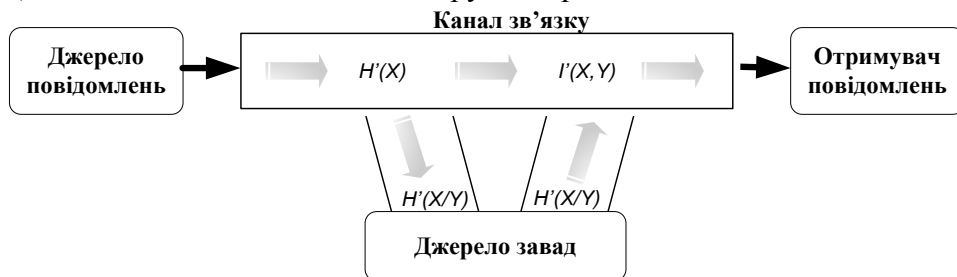


Рис. 2.5. Ілюстрація співвідношення між характеристиками каналу та швидкістю передачі інформації

На рис. 2.5 позначені: $H'(X)$ – продуктивність джерела переданого сигналу X , а $H'(Y)$ – “продуктивність” каналу, тобто повна власна інформація в прийнятому сигналі за одиницю часу. Величина $H'(X/Y)$ є втратою інформації або ненадійністю каналу в одиницю часу, а $H'(Y/X)$ – швидкість створення помилкової, сторонньої інформації в каналі, яка не має відношення до X і обумовлена присутніми в каналі завадами. За визначенням

Розділ 2. Кількісні оцінки вимірювальної інформації

К. Шеннона ненадійність каналу є ентропією входу, коли вихід відомий, тобто її можна вважати мірою середньої невизначеності прийнятого сигналу. Величина ж $H'(Y/X)$ – є ентропія виходу, коли вхід відомий і служить мірою середньої невизначеності переданого сигналу.

Співвідношення між $H'(X/Y)$ і $H'(Y/X)$ залежить від властивостей каналу. Так, наприклад, при передаванні звукового сигналу по каналу з вузькою смугою пропускання, яка недостатня для якісного відтворення сигналу і з низьким рівнем завад, втрачається частина корисної інформації, але майже не отримується даремна інформація, в цьому випадку $H'(X/Y) \ll H'(Y/X)$. Якщо ж сигнал відтворюється якісно, але при цьому прослуховуються наведення від сусіднього радіоканалу, то це означає, що, майже не втрачаючи корисної інформації, ми одержали багато зайвої - надлишкової інформації, що заважає опрацюванню, і буде дійсним співвідношення $H'(X/Y) \ll H'(Y/X)$.

Ефективність та надмірність джерел інформації. Завдання підвищення ефективності передачі зводиться до зменшення надмірності повідомлення й сигналу, точніше до раціонального її використання. Надмірність повідомлення обумовлена тим, що елементи повідомлення не є рівноймовірними й між ними є статистичний зв'язок. При кодуванні можна перерозподілити ймовірності вихідного повідомлення так, щоб розподіл ймовірностей символів коду наближався до оптимального (до рівномірного - у дискретному випадку або до нормального - при передаванні безперервних повідомлень). Такий перерозподіл дозволяє усунути надмірність, що залежить від розподілу ймовірностей елементів повідомлення. Прикладом подібного кодування є код Шенона – Фано. Що стосується сигналу, то його надмірність залежить від способу модуляції й від виду носія. Процес модуляції звичайно супроводжується розширенням смуги частот сигналу в порівнянні зі смугою частот переданого повідомлення. Це розширення смуги і є надлишковим. Частотна надмірність також збільшується при переході від синусоїдального носія до носія імпульсного або шумоподібного. З погляду підвищення ефективності передачі варто було б вибирати способи модуляції з малою надмірністю. До таких систем, зокрема, належить односмугова передача, у якій передані сигнали не містять частотної надмірності – вони є просто копіями переданих повідомлень.

Однак, говорячи про ефективність системи зв'язку, не можна забувати про її завадостійкість. Усунення надмірності підвищує ефективність передачі, але знижує при цьому вірогідність (завадостійкість) і, навпаки, збереження або введення надмірності дозволяє забезпечити високу вірогідність передавання. Отже, ефективність і завадостійкість систем передавання інформації не можна розглядати окремо один від одного. Питання підвищення ефективності й завадостійкості систем передавання інформації становлять

єдину проблему. Завдання полягає в пошуку розумного компромісу при її вирішенні.

Визначення інформації джерел двосимвольних повідомлень.

Проведемо аналіз джерела повідомлень, що складається з послідовності двох випадкових незалежних подій. Позначимо їх символами s_1 і s_2 . Розглянемо повідомлення виду $s_1 s_2 s_2 s_1 s_1 s_2 \dots s_1$, у якому n_1 випадкових подій s_1 і n_2 випадкових подій s_2 , так що загальна кількість подій $n = n_1 + n_2$. Елементи множини повідомлень цього джерела будуть відрізнятися один від одного, тому що число подій n_1 і n_2 є випадковими величинами. Кількість варіантів повідомлення можна визначити числом комбінацій із n по n_1 , тобто

$$C_n^{n_1} = \binom{n}{n_1},$$

Розглянемо таке ж джерело повідомлень, що складається з послідовності подій $D_1, D_2, D_3, \dots, D_m$, у якій число символів D_1 і D_2 дорівнює m_1 і m_2 , відповідно $m = m_1 + m_2$. Для другого джерела

$$C_m^{m_1} = \binom{m}{m_1}.$$

Кількість інформації кожного з джерел визначається варіантами $C_n^{n_1}$ і $C_m^{m_1}$. Позначимо кількість інформації розглянутих джерел символами: першого $I(S)$ і другого $I(D)$. Об'єднаємо ці два незалежних джерела в один (третій) із символами – подіями $Y=S+D$. Очевидно, що кількість інформації третього джерела $I(Y)$ повинна бути рівною сумі $I(S) + I(D)$, а число варіантів – добутку $C_n^{n_1} \cdot C_m^{m_1}$. Оскільки логарифм добутку дорівнює сумі логарифмів, то

$$I(Y) = \ln(C_n^{n_1} C_m^{m_1}) = \ln(C_n^{n_1}) + \ln(C_m^{m_1}) = I(S) + I(D).$$

Виберемо як міру кількості інформації джерел двосимвольних повідомлень число, пропорційне логарифму очікуваного числа варіантів повідомлень, тобто логарифму їх різноманіття

$$I = k \ln(C_n^{n_1}) = k \ln(n!/n_1!n_2!),$$

де k - постійна, залежна від вибору одиниці виміру інформації.

Розглянемо повідомлення при умові, що $n \rightarrow \infty$ і представимо його як

$$I = k \ln(n!) - k \ln(n_1!) - k \ln(n_2!). \quad (2.12)$$

При умові, якщо m – велике число, то має місце формула Стірлінга у виді: $\ln(m!) \approx m \ln m - m$.

Тоді вираз (2.12), з використанням формули Стірлінга приймає вид

Розділ 2. Кількісні оцінки вимірювальної інформації

$$I = k(n \ln n - n - n_1 \ln n_1 + n_1 - n_2 \ln n_2 + n_2) = -kn \left(\frac{n_1}{n} \ln \frac{n_1}{n} + \frac{n_2}{n} \ln \frac{n_2}{n} - \ln n \right),$$

де вираз у дужках можна подати як

$$\frac{n_1}{n} \left(\ln \frac{n_1}{n} + \ln n \right) + \frac{n_2}{n} \left(\ln \frac{n_2}{n} + \ln n \right) - \ln n = \frac{n_1}{n} \ln \frac{n_1}{n} + \frac{n_2}{n} \ln \frac{n_2}{n} + \frac{n_1 + n_2}{n} \ln n - \ln n.$$

У результаті одержимо
$$I = -kn \left(\frac{n_1}{n} \ln \frac{n_1}{n} + \frac{n_2}{n} \ln \frac{n_2}{n} \right). \quad (2.13)$$

Вираз (2.52) дозволяє оцінити кількість інформації джерела двосимвольних повідомлень по експериментальним даним: по числу символів n_1 і n_2 у даному повідомленні. Якщо події s_1 і s_2 є випадковими, то при $n \rightarrow \infty$ відносини n_1/n і n_2/n дорівнюють ймовірностям появи подій $P(s_1)$ і $P(s_2)$. Тоді вираз (2.13) приймає вид

$$I = -nk [P(s_1) \ln P(s_1) + P(s_2) \ln P(s_2)]. \quad (2.14)$$

Кількість інформації джерела, що припадає на один символ (подію), дорівнює його продуктивності

$$i = I/n = -k [p_1 \ln p_1 + p_2 \ln p_2], \quad (2.15)$$

де $p_1 = P(s_1)$; $p_2 = P(s_2)$.

Вираз (2.15) для обчислення інформації двосимвольного джерела випадкових подій називають *формулою К. Шеннона*.

Розглянемо її детальніше. Оскільки $p_1 + p_2 = 1$, то, позначивши $p = p_1$ і $p_2 = 1 - p$, подамо її у вигляді:
$$i = -k [p \ln p + (1-p) \ln (1-p)] = i(p).$$

Інформація, яка визначається одним символом, залежить від імовірності однієї з подій і при $p = 0$, а також $p = 1$ дорівнює нулю, тобто $i(0) = i(1) = 0$. Визначимо умови максимуму функції $i(p)$:

$$di/dp = -k [1 + \ln p - \ln (1-p) - 1] = \ln p - \ln (1-p) = 0.$$

Отже: $\ln p / (1-p) = 0$, $p / (1-p) = 1$, $p_{\max} = 0,5$.

Максимальна кількість інформації, яку видає джерело, досягається з ймовірностями появи подій $p_1 = p_2 = 0,5$, тобто $i_{\max} = k \ln 2$. Виберемо таке джерело, як еталон, вважаючи, що його інформація на один символ дорівнює одиниці ($i_{\max} = 1$). Тоді $k = (\ln 2)^{-1}$ і згідно формули Шеннона маємо:

$$i = -(p_1 \ln p_1 + p_2 \ln p_2) / \ln 2. \quad (2.16)$$

Запис формули Шеннона можна спростити, якщо змінити основу логарифма. Позначимо $x = \ln p / \ln 2$. Тоді $x \ln 2 = \ln p$, $\ln 2^x = \ln p$, $2^x = p$.

Отже, $x = \lg_2 p = \ln p / \ln 2$.

Для формули Шеннона отримаємо новий запис: $i = -p_1 \log_2 p_1 - p_2 \log_2 p_2$, де $\log_2 p$ - двійковий логарифм числа p .

Приклад. Якщо $p_1 = 0.125 = 2^{-3}$ і $p_2 = 0.875$, то $i = 0,54356$, що значно менше одиниці. Це означає наступне. Якщо два джерела видали однаково кількість інформації $I_1 = i_1 N_1$, $I_2 = i_2 N_2$ і $I_1 = I_2$, то відношення розмірів повідомлень обернено пропорційно їхній продуктивності, таким чином одні і ту ж інформацію можна передавати за менший час чи займати меншу пам'ять при її збереженні, якщо перетворити повідомлення так, щоб продуктивність джерела була максимальною (чи близькою до неї).

Продовжимо дослідження формули Шеннона, записавши її у виді:

$$i = -[P(s_1)\log_2 P(s_1) + P(s_2)\log_2 P(s_2)] = [P(s_1)i(s_1) + P(s_2)i(s_2)]. \quad (2.17)$$

Тут $i(s_1)$ і $i(s_2)$ – кількість інформації при реалізації випадкових подій s_1 і s_2 , а $P(s_1)$ і $P(s_2)$ – їх ймовірності.

Ймовірнісно-статистичний опис джерела двосимвольних повідомлень, де i – це математичне сподівання кількості інформації джерела на один символ (одну подію), а $i(s_1)$ і $i(s_2)$ – кількість інформації при реалізації випадкових подій s_1 і s_2 відповідно: $i(s_1) = -\lg_2 P(s_1)$, $i(s_2) = -\lg_2 P(s_2)$.

Формула (2.17) є ймовірнісно-статистичним описом джерела двосимвольних повідомлень при реалізації випадкових подій s_1 і s_2 .

Визначення інформації джерел багатосимвольних повідомлень. Розглянемо джерело інформації, повідомлення якого складаються з m різних символів $s_1, s_2, s_3, \dots, s_m$ і відомо, що вони можуть з'являтися в повідомленні незалежно один від одного з ймовірностями $p_1, p_2, p_3, \dots, p_m$...

Загальне число можливих повідомлень розміром n із числом символів $n_1, n_2, n_3, \dots, n_m$, тобто число варіантів такого джерела, можна обчислити за формулою

$$N = n! / \prod_{i=1}^m (n_i)!, \text{ де } n = \sum_{i=1}^m n_i.$$

Визначимо кількість такого повідомлення у виді

$$I = K \ln N = K \left(\ln (n!) - \sum_{i=1}^m \ln (n_i!) \right). \quad (2.18)$$

із використанням формули Стірлінга, отримаємо

$$I \approx -Kn \sum_{i=1}^m (n_i/n) \ln (n_i/n).$$

Формула Шеннона для обчислення інформації джерел багатосимвольних повідомлень, при $n \rightarrow \infty$ має вигляд

$$i = -(\ln 2)^{-1} \sum_{j=1}^m P_j \ln P_j. \quad (2.19)$$

Приклад. Розглянемо повідомлення українською мовою при використанні 32 символів (букв). Ймовірності їхньої появи в текстах відомі й

Розділ 2. Кількісні оцінки вимірювальної інформації

приведені в табл. 2.1. Припускаючи, що ймовірності появи букв у тексті відповідають незалежним випадковим подіям, одержимо за формулою (2.19) кількість інформації на одну букву тексту $i=4.42$ біт.

Таблиця 2.1

Продуктивність джерела багатосимвольних повідомлень буде максимальною, якщо $p_j = 1/m$, $j=1, 2, \dots, m$, тобто коли поява символів рівноймовірна. Оскільки $p_m = 1 - \sum_{j=1}^{m-1} p_j$, то функція (2.18) залежить від $(m-1)$ змінних:

$$i(p_1, p_2, \dots, p_{m-1}) = -\frac{1}{\ln 2} \sum_{j=1}^{m-1} p_j \ln p_j - \frac{1}{\ln 2} p_m \ln p_m.$$

Буква	P	Буква	P	Буква	P
А	0,062	І, Ії	0,062	Т	0,053
Б	0,014	Й	0,010	У	0,021
В	0,038	К	0,028	Ф	0,002
Г	0,013	Л	0,035	Х	0,009
Д	0,025	М	0,026	Ц	0,004
Е	0,072	Н	0,053	Ч	0,012
Є	0,003	О	0,090	Ш	0,006
Ж	0,007	П	0,023	Щ	0,003
З	0,016	Р	0,040	Ь	0,014
И	0,016	С	0,045	Ю	0,006
				Я	0,018

Необхідна умова існування внутрішнього максимуму багатовимірних функцій формується в такий спосіб: усі частинні похідні в точці максимуму повинні дорівнювати 0; тобто $\partial i / \partial p_1 = 0$; $\partial i / \partial p_2 = 0, \dots, \partial i / \partial p_{m-1} = 0$.

для даного випадку часткові похідні визначені так

$$\partial i / \partial p_j = -\left(\ln p_j + 1 + \left(\partial p_m / \partial p_j \right) \ln p_m + \partial p_m / \partial p_j \right) / \ln 2,$$

а з урахуванням $\partial p_m / \partial p_j = -\partial \left(\sum_{j=1}^{m-1} p_j - 1 \right) / \partial p_j = -1$,

у результаті отримуємо

$$\partial i / \partial p_j = -\ln(p_j / p_m) / \ln 2 = 0, \quad j=1, 2, \dots, m-1. \quad (2.20)$$

Необхідна умова виконується тільки для випадку, якщо $p_j = 1/m$, $j=1, 2, \dots, m$.

Таким чином, необхідною умовою максимуму продуктивності джерела повідомлень є рівноймовірність появи символів.

Достатньою умовою максимуму багатовимірних функцій є виконання нерівності

$$\sum_{i=1}^{m-1} \sum_{j=1}^{m-1} \left(\partial^2 i / \partial p_i \partial p_j \right) < 0 \quad (2.21)$$

у точці $p_j = 1/m$, $j = 1, 2, \dots, m-1$.

Другі похідні визначимо з формули (2.20): $\partial^2 i / \partial p_j^2 = -1 / \ln 2 (1/p_j + 1/p_m)$,

$$\partial^2 i / \partial p_i \partial p_j = -1 / p_m \ln 2.$$

У точці екстремуму:

$$\partial^2 i / \partial p_j^2 = -2m / \ln 2, \quad \partial^2 i / \partial p_i \partial p_j = -1 / m \ln 2.$$

Таким чином достатня умова максимуму виконується, тому що всі похідні менше нуля. Якби букви української мови були рівноймовірні, то продуктивність джерела повідомлення становила

$$i_{\max} = -\sum_{i=1}^{32} (1/32) \lg_2 (1/32) = \lg_2 2^5 = 5.$$

Так як поява букв української мови нерівноймовірна, то інформація зменшується і розраховується за формулою (2.18). У дійсності між парами букв існує статистичний зв'язок (кореляція), тобто якщо яка-небудь буква у повідомленні з'являється, то ймовірності появи інших букв після неї сильно відрізняються.

Визначення інформації джерела двох залежних випадкових подій. Розглянемо випадок, коли математичну модель двосимвольної корельованої послідовності використаємо ланцюг Маркова (сигнал Маркова). Для ланцюгів Маркова характерна залежність ймовірностей випадкових подій s_1 і s_2 від того, яка з подій була на попередньому кроці. Якщо попередня подія була s_1 , то на наступному кроці можлива поява події s_1 з імовірністю $P(s_1/s_1)$ і подія s_2 з імовірністю $P(s_2/s_1)$. Якщо попередня подія була s_2 , то умовні ймовірності записуються у виді $P(s_1/s_2)$, $P(s_2/s_2)$. Позначимо ймовірності появи подій s_1 і s_2 на k -ому кроці $P_1(k)$ і $P_2(k)$. Припустимо, що задано початкові ймовірності $P_1(0)$ і $P_2(0)$, причому $P_1(0) + P_2(0) = 1$, $P_1(k) + P_2(k) = 1$ і відомі умовні ймовірності переходу $P(s_1/s_1)$, $P(s_2/s_1)$, $P(s_1/s_2)$, $P(s_2/s_2)$. Тоді справедливі рекурентні співвідношення для повної ймовірності

$$\begin{aligned} P_1(k) &= P_1(k-1)P(s_1/s_1) + P_2(k-1)P(s_1/s_2), \\ P_2(k) &= P_1(k-1)P(s_2/s_1) + P_2(k-1)P(s_2/s_2). \end{aligned} \quad (2.22)$$

Для марківських дискретних випадкових послідовностей характерно, що після декількох переходів ймовірності появи подій практично стають незалежними від початкових умов і, отже, від номера кроку.

Приклад. Припустимо, що $P_1(0)=1$, $P_2(0) = 0$, $P(s_1/s_1) = 0.5$, $P(s_2/s_1) = 0.5$, $P(s_1/s_2) = 0.6$, $P(s_2/s_2) = 0.4$. Результати розрахунків за формулою (2.22) наведені в табл. 2.2. За змінених початкових умов $P_1(0) = 0$, $P_2(0)=1$ результати приведені в табл. 2.3.

Таблиця 2.2

K	0	1	2	3	4	5
$P_1(k)$	1	0,5	0,55	0,545	0,5454	0,54545
$P_2(k)$	0	0,5	0,455	0,455	0,4545	0,45455

Таблиця 2.3

K	0	1	2	3	4	5
$P_1(k)$	0	0,6	0,54	0,546	0,5454	0,54546
$P_2(k)$	1	0,4	0,46	0,454	0,4546	0,45454

З аналізу табл. 2.2 і 2.3 випливає, що незалежно від початкових умов уже на

Розділ 2. Кількісні оцінки вимірювальної інформації

s -ому кроці встановлюються значення $P_1 = 0,5454$ і $P_2 = 0,4546$.

Оскільки при великих k ймовірності не залежать від k , то вважаючи $P(k)=P(k-1)$ із виразу (2.62) отримують формули для оцінки граничних значень ймовірностей P_1 і P_2 :

$$P_1 = \frac{P(s_1/s_2)}{P(s_1/s_2) + P(s_2/s_1)} = P(s_1), \quad P_2 = \frac{P(s_2/s_1)}{P(s_1/s_2) + P(s_2/s_1)} = P(s_2).$$

Використовуючи перехідні ймовірності, можна формувати моделі двосимвольних повідомлень із різними статистичними властивостями і досліджувати вплив кореляції на кількість інформації джерел повідомлень.

Коефіцієнт кореляції між випадковими подіями A і B визначається як

$$r = (P(AB) - P(A)P(B)) / \sqrt{P(A)[1 - P(A)]P(B)[1 - P(B)]}$$

Для марківських послідовностей $P(s_1s_2) = P(s_1)P(s_2/s_1) = P(s_2)P(s_1/s_2)$, $1 - P(s_1) = P(s_2)$, $1 - P(s_2) = P(s_1)$ і для коефіцієнта кореляції отримаємо:

$$r = (P(s_1/s_2) - P(s_1)) / P(s_1) = (P(s_2/s_1) - P(s_2)) / P(s_2).$$

Виведемо формули для визначення кількості інформації джерела марківських повідомлень. Якщо мали місце події s_1 чи s_2 , то умовні математичні сподівання $i(s_1)$ і $i(s_2)$ визначимо за формулами

$$i(s_1) = P(s_1/s_1) i(s_1/s_1) + P(s_2/s_1) i(s_2/s_1), \tag{2.23}$$

$$i(s_1) = P(s_1/s_2) i(s_1/s_2) + P(s_2/s_2) i(s_2/s_2).$$

Очікувана кількість інформації джерела на один символ

$$i = P(s_1) i(s_1) + P(s_2) i(s_2). \tag{2.24}$$

Підставивши формулу (2.23) в (2.24), одержимо

$$i = P(s_1) P(s_1/s_1) i(s_1/s_1) + P(s_1) P(s_2/s_1) i(s_2/s_1) + \\ + P(s_2) P(s_1/s_2) i(s_1/s_2) + P(s_2) P(s_2/s_2) i(s_2/s_2).$$

При умові, що $i(s_1/s_1) = -\log_2 P(s_1/s_1)$, $i(s_2/s_1) = -\log_2 P(s_2/s_1)$, $i(s_2/s_2) = -\log_2 P(s_2/s_2)$, отримаємо $i = -\sum_{i=1}^2 \sum_{j=1}^2 P(s_i s_j) \log_2 P(s_i/s_j)$.

Для сигналу Маркова ймовірності добутку двох подій:

$$P(s_1s_1) = P(s_1)P(s_1/s_1) = P(s_1/s_1)P(s_1/s_2) / [P(s_1/s_2) + P(s_2/s_1)],$$

$$P(s_1s_2) = P(s_2s_1) = P(s_1/s_2)P(s_2/s_1) / [P(s_1/s_2) + P(s_2/s_1)],$$

$$P(s_2s_2) = P(s_2)P(s_2/s_2) = P(s_2/s_2)P(s_2/s_1) / [P(s_1/s_2) + P(s_2/s_1)].$$

Приклад. Розглянемо наступний сигнал Маркова. Припустимо, що $P(s_1/s_1)=3/4$, $P(s_2/s_1) = 1/4$, $P(s_1/s_2) = 1/8$, $P(s_2/s_2) = 7/8$. У такому повідомленні $P(s_1)=1/3$, $P(s_2)=2/3$. Якби символи були незалежні, то

$$i = (\log_2 3)/3 + (2/3)\log_2 (3/2) = 0,91183.$$

Умовна інформація на символи s_1 і s_2 : $i(s_1) = (3/4)\log_2(3/4) + (1/4)\log_2 4 = 0,8113$, $i(s_2) = (1/8)\log_2 8 + (7/8)\log_2(8/7) = 0,5428$. Очікувана інформація джерела на один символ дорівнює їхньому середньому значенню $i = i(s_1)/3 + 2i(s_2)/3 = 0,6323$. Вона менша, ніж у джерела повідомлень із незалежними символами.

Визначення інформації у випадку об'єднання декількох джерел повідомлень. Розглянемо два джерела повідомлень з алфавітом символів v_1, v_2, \dots, v_n і u_1, u_2, \dots, u_m і об'єднане джерело з парами символів $v_j u_j, i=1, 2, \dots, n; j=1, 2, \dots, m$. Відомий багатомірний закон розподілу ймовірностей $P(v_1, v_2, \dots, v_n, u_1, u_2, \dots, u_m) = P(|v|, |u|)$, де позначено $|v| = |v_1, v_2, \dots, v_n|^T$ – вектор символів 1-го джерела, $|u| = |u_1, u_2, \dots, u_m|^T$ – вектор символів другого джерела, T – знак транспонування.

Кількість інформації кожного з джерел і об'єданого джерела:

$$i(|v|) = -\sum_{i=1}^n P(v_i) \log_2 P(v_i), \quad (2.25a)$$

$$i(|u|) = -\sum_{j=1}^m P(u_j) \log_2 P(u_j), \quad (2.25б)$$

$$i(|v|, |u|) = -\sum_{i=1}^n \sum_{j=1}^m P(v_i, u_j) \log_2 P(v_i, u_j). \quad (2.25в)$$

Якщо джерела статистично незалежні, то $P(v_i, u_j) = P(v_i)P(u_j)$ і тоді

$$\begin{aligned} i(|v|, |u|) &= -\sum_{i=1}^n \sum_{j=1}^m P(v_i)P(u_j) \log_2 [P(v_i)P(u_j)] = \\ &= -\sum_{i=1}^n P(v_i) \log_2 P(v_i) \sum_{j=1}^m P(u_j) - \sum_{j=1}^m P(u_j) \log_2 P(u_j) \sum_{i=1}^n P(v_i). \end{aligned}$$

З огляду на, що $\sum_{i=1}^n P(v_i) = 1$ і $\sum_{j=1}^m P(u_j) = 1$, одержимо співвідношення

$$i(|v|, |u|) = i(|v|) + i(|u|), \quad (2.26)$$

тобто інформація об'єданого джерела при незалежних символах дорівнює сумі інформації джерел, що входять в об'єднання.

Якщо джерела повідомлень статистично залежні, то двовимірний закон розподілу записується у вигляді $P(v_i, u_j) = P(v_i)P(u_j/v_i) = P(u_j)P(v_i/u_j)$.

Тоді вираз (2.25в) для обчислення кількості інформації об'єданого джерела має наступний вид:

$$i(|v|, |u|) = -\sum_{i=1}^n \sum_{j=1}^m P(v_i)P(u_j/v_i) \left[\log_2 P(v_i) + \log_2 P(u_j/v_i) \right].$$

Розділ 2. Кількісні оцінки вимірювальної інформації

З огляду на те, що $\sum_{j=1}^m P(u_j/v_i) = 1$, одержимо:

$$i(|v|, |u|) = -\sum_{i=1}^n P(v_i) \log_2 P(v_i) + \sum_{i=1}^n P(v_i) i(|u|/v_i),$$

де $i(|u|, v_i) = -\sum_{j=1}^m P(u_j/v_i) \log_2 P(u_j/v_i)$ - умовна кількість інформації другого джерела, якщо мав місце символ v_i першого джерела. Очевидно, що $\sum_{i=1}^n P(v_i) i(|u|/v_i) = i(|u|/|v|)$ - умовна кількість інформації другого джерела щодо першого. Отже,

$$i(|v||u|) = i(v) + i(|u|/|v|). \quad (2.27)$$

Можна показати також, що справедливо співвідношення

$$i(|v||u|) = i(|v|) + i(|u|/|v|). \quad (2.28)$$

Оскільки $P(v_i) < 1$, то справедлива нерівність

$$\sum_{i=1}^n P(v_i) i(|u|/v_i) < i(|u|).$$

Отже, завжди має місце: $i(|v|, |u|) < i(|v|) + i(u)$, тобто *кількість інформації об'єднаного джерела менше суми інформації вихідних статистично залежних джерел повідомлень.*

Визначення інформації джерел неперервних повідомлень.

Розділимо діапазон значень заданої неперервної випадкової величини сигналу X на скінченне число n малих інтервалів ΔX так, що $X_{i+1} = X_i + \Delta X$. Будемо вважати, що реалізовано значення X_i , якщо X задовольняє нерівності $X_i - \Delta X/2 \leq x \leq X_i + \Delta X/2$. Ймовірність того, що це буде X_i

$$P(X_i) = \int_{X_i - 0,5\Delta X}^{X_i + 0,5\Delta X} p(x) dx \approx p(X_i) \Delta X,$$

де $p(x)$ - щільність розподілу випадкової величини.

Кількість інформації, що міститься у множині випадкових величин X_1, X_2, \dots, X_n з ймовірностями появи $P(X_1), P(X_2), \dots, P(X_n)$, за формулою Шеннона дорівнює

$$I_1 = -n \sum_{i=1}^n P(X_i) \log_2 P(X_i) = -n \sum_{i=1}^n p(X_i) \Delta X \log_2 [p(X_i) \Delta X].$$

Кількість інформації, що приходить у середньому на один символ

$$i_1 = -\sum_{i=1}^n p(X_i) \log_2 [p(X_i) \Delta X] - \log_2 (\Delta X),$$

причому $\sum_{i=1}^n p(X_i) \Delta X = 1$. Замінивши суму інтегралом маємо

$$i_1 = - \int_{-\infty}^{\infty} p(x) \log_2 p(x) dx - \log_2 (\Delta X)$$

Перший член у цьому виразі має скінченне значення і залежить тільки від закону розподілу сигналу, що несе інформацію, другий залежить від вибору інтервалу квантування ΔX . Якщо вважати, що при порівнянні кількості інформації різних сигналів інтервал квантування для усіх буде той самий, то:

$$i = - \int_{-\infty}^{\infty} p(x) \log_2 (p|x|) dx.$$

Середня кількість інформації на один відлік дорівнює математичному сподіванню двійкового логарифма закону розподілу для неперервних сигналів.

Основні висновки

Існують три підходи до аналізу інформації: синтаксичний, семантичний і прагматичний. На синтаксичному рівні враховуються тип носія і спосіб представлення інформації, швидкість передачі і обробки, розміри кодів представлення інформації, надійність і точність перетворення цих кодів і т.п. Семантичний аспект припускає врахування змісту інформації і базується на представленні відповідних модулів інформації, кожний з яких містить певний допустимий обсяг даних. Такі модулі даних прийнято називати інформаційними об'єктами чи повідомленнями. Прагматичний підхід пов'язаний з оцінюванням цінності отриманої інформації.

Інформація є характеристикою співвідношення між джерелом інформації (об'єктом дослідження), повідомленням і його споживачем. Інформаційні об'єкти - предмети, процеси, явища матеріальної або нематеріальної властивості, що розглядаються з погляду їх інформаційних властивостей.

Джерело інформації або повідомлення - це фізичний об'єкт, система або явище, що формують передане повідомлення. Саме повідомлення - це значення або зміна деякої величини, що відбивають стан інформаційного об'єкта (системи або явища). Засоби і процеси, що забезпечують передачу повідомлень від джерела інформації до її споживача, називаються інформаційно-комунікаційними системами.

Ступінь зміни невизначеності стану інформаційного об'єкта покладений в основу кількісної міри інформації. При введенні кількісної міри інформації прийнято не враховувати зміст повідомлень (семантику), а обмежуватися тільки формальними ознаками, важливими з погляду передачі повідомлень по каналах зв'язку.

У теорії інформації під кількістю інформації розуміють міру зменшення невизначеності стану об'єкта. Знаходження такої міри вимагає оцінки

Розділ 2. Кількісні оцінки вимірювальної інформації

кількості переданої інформації. У теорії інформації кількістю інформації називають таку числову характеристику сигналу, яка не залежить від його форми і змісту, а характеризує невизначеність, яка зникає після отримання повідомлення у вигляді даного сигналу.

Кількість інформації на синтаксичному рівні визначається через поняття ентропії системи. Ентропія визначає міру невизначеності всієї множини повідомлень на вході системи і обчислюється як середня кількість власної інформації у всіх повідомленнях. Ентропія джерела й кількість інформації тим більше, чим більше розмір алфавіту (кількість символів) джерела. Ентропія джерела залежить від статистичних властивостей повідомлень. Ентропія максимальна, якщо повідомлення джерела рівномірні й статистично незалежні. Ентропія джерела нерівномірних повідомлень завжди менше максимально можливої.

При визначенні ентропії й кількості інформації в повідомленнях, елементи яких статистично зв'язані, не можна обмежуватися тільки безумовними ймовірностями - необхідно обов'язково враховувати також умовні ймовірності появи окремих повідомлень.

Спільна ентропія двох повідомлень дорівнює сумі безумовної ентропії одного з них і умовної ентропії другого. При наявності зв'язку між елементарними повідомленнями ентропія джерела зменшується.

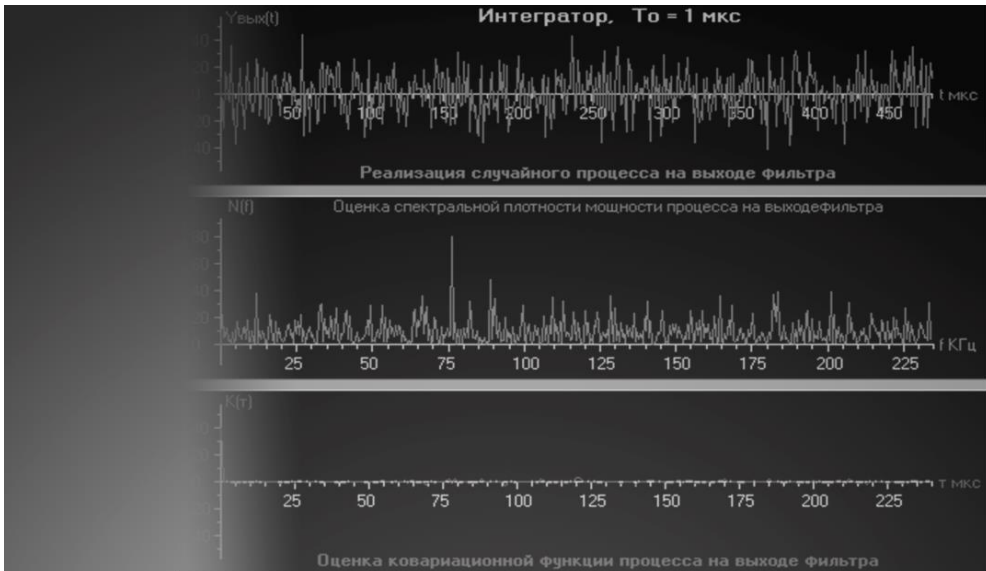
Універсальний носій інформації - електричний сигнал (або матеріали, що володіють електромагнітними властивостями, що дозволяють просто знімати з них інформацію у вигляді електричних сигналів).

Питання для самоконтролю

1. З проявами яких функцій інформації найчастіше ми зустрічаємося?
2. Наскільки актуальні в питанні інформації її динамічні властивості?
3. У чому полягає відмінність в поняттях "Дані" та "Інформація"?
4. Наведіть приклади складних інформаційних об'єктів.
5. Запишіть формули для обчислення кількості інформації що містить повідомлення.
6. Розкрийте властивість симетрії для кількості інформації.
7. Розкрийте властивість адитивності для кількості інформації.
8. Запишіть формулу для середньої кількості інформації у повідомленні.
9. Які міри інформації використовують для кількісної її оцінки?
10. Назвіть одиниці вимірювання інформації.
11. Наведіть класифікацію мір інформації та їх властивостей.
12. Дайте визначення ентропії.
13. Виведіть формулу для визначення спільної ентропії двох джерел X і Y .
14. Відзначте основні властивості ентропії складних повідомлень.
15. Визначте межі в яких може змінюватись умовна ентропія.
16. Як змінюється ентропія при наявності статистичних зв'язків між елементарними повідомленнями?

17. Які параметри сигналу можуть бути функціонально залежними від інформаційного повідомлення?

Розділ 2. Кількісні оцінки вимірювальної інформації



СИГНАЛИ ТА ЇХ МАТЕМАТИЧНІ МОДЕЛІ

3

- 3.1. Види сигналів та їх математичні моделі
- 3.2. Випадкові сигнали та завади
- 3.3. Числові характеристики сигналів і завад
- 3.4. Математичні моделі сигналів з обмеженим спектром

3.1. Види сигналів та їх математичні моделі

Згідно з загальноприйнятою термінологією **сигналом** називають процес, що характеризує зміну в просторі і часі фізичного стану деякого об'єкта і використовується для відображення, реєстрації, передавання, приймання і опрацювання повідомлень.

Сигнали можуть мати різну фізичну природу. Широко відомі електричні сигнали (струм чи напруга), конкретну реалізацію яких можна спостерігати на екрані осцилографу чи в інших пристроях відображення сигналів; акустичні сигнали, сприйняті органами слуху, тощо.

Досліджуємі в завданнях вимірювані сигнали описуються відповідними математичними моделями, аргументами яких у загальному випадку є просторові змінні (x, y, z) і час t .

Знання математичних моделей сигналів дає змогу порівнювати їх між собою, визначити їх характеристики (параметри), оцінити результати вимірювань, установлювати тотожність і розбіжності, а зрештою – класифікувати. У більшості випадків термінологія того чи іншого виду сигналів визначається видом його математичної моделі, наприклад детерміновані сигнали описуються детермінованими функціями.

Спочатку розглянемо основні види одновимірного сигналу часу $s(t)$, який має область визначення $t \in T$ і область значень $S \subseteq R$.

Неперервні і дискретні сигнали. Якщо взяти за ознаку виду сигналів характер зміни його значень і аргументів, то можливі наступні чотири види сигналів (рис. 3.1):

- 1) сигнали неперервні за значенням і часом;
- 2) сигнали дискретні в часі і неперервні за значеннями;
- 3) сигнали дискретні (квантовані) за значеннями і неперервні в часі;
- 4) сигнали дискретні за значеннями і дискретні в часі – цифрові сигнали.

Сигнали першого виду (рис. 3.1,а) називають *аналоговими* або *неперервними*.

Сигнали другого виду (рис. 3.1,б) називають *дискретними* або *дискретизованими* сигналами. Термін *дискретний* характеризує не сам сигнал, а спосіб задання його в часовій області.

Часовий інтервал $\Delta t_n = t_{n+1} - t_n$, $n = 0, 1, \dots$ називають n -им кроком часової дискретизації сигналу. Якщо крок часової дискретизації Δt_n не залежить від n , тобто є сталим, то область задання сигналу є

Розділ 2. Кількісні оцінки вимірювальної інформації

еквідистантною ґраткою часу. Інтервал Δt називають *періодом дискретизації*.

Тоді, дискретизований у часі сигнал записується у виді $\{S(n\Delta t), n = 0, 1, 2, \dots\}$ або просто $\{S(n), n = 0, 1, 2, \dots\}$

Позначимо *математичну модель дискретного сигналу*: $s(nT), x(nT), u(nT), n = 0, 1, \dots$, або, відкидаючи для спрощення T , у вигляді: $s(n), x(n), u(n), n = 0, 1, \dots$

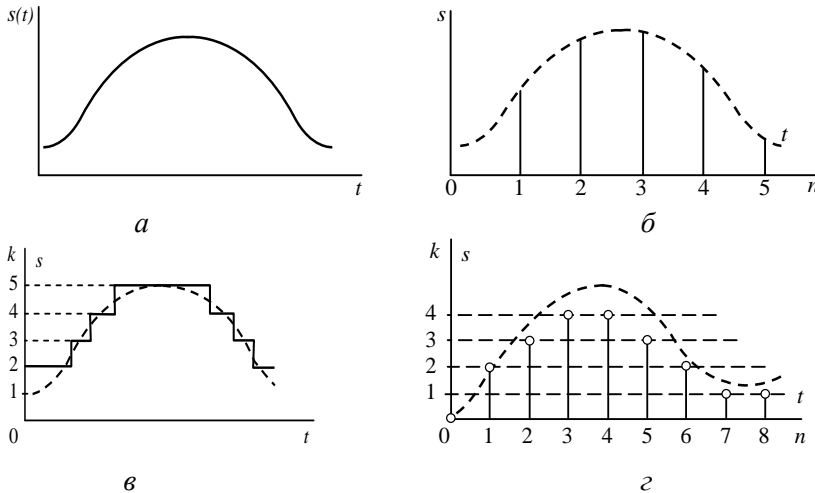


Рис. 3.1. Основні просторово-часові моделі сигналів

Сигнали третього класу (рис. 3.1,в) називають *квантованими*. Для цього класу сигналів здійснюється дискретизація за рівнем, тобто *квантований сигнал може набувати тільки дискретних значень*. Математичну модель квантованого сигналу можна позначити як $\hat{s}(t)$, при цьому час t неперервний (у загальному випадку в інтервалі від мінус нескінченності до плюс нескінченності), а величина \hat{s} може набувати одне з множини значень $s_k, k = 0, m-1$, тобто $\hat{s} \in [s_0, s_1, \dots, s_{m-1}]$. Крок квантування сигналу за рівнем, як правило, вибирається сталим.

Сигнали четвертого класу (рис. 3.1,г) називають *цифровими*. Такі сигнали утворюються з аналогових сигналів у результаті їх дискретизації в часі та квантування за величиною.

Критерій передбачуваності сигналів. Якщо за класифікаційну ознаку взяти *передбачуваність миттєвого значення сигналу в будь-*

який момент часу, то всю сукупність сигналів можна поділити на *детерміновані та випадкові сигнали*.

Детермінованими називаються сигнали, значення яких можуть бути обчислені в будь-який момент часу, тобто вони передбачувані з імовірністю, що дорівнює одиниці.

Найпростішим **прикладом** математичної моделі детермінованого сигналу може бути гармонічне коливання
$$u(t) = A \cos(\omega_0 t + \phi_0),$$
 (3.1)

для якого амплітуда A , кутова частота ω_0 і початкова фаза ϕ_0 коливань задані.

Випадковими називають сигнали, значення яких у будь-який момент часу непередбачувані, тобто в заданий момент часу t їх неможливо визначити з імовірністю, що дорівнює одиниці.

Прикладами випадкових сигналів можуть бути акустичні коливання, відтворювані акустичною апаратурою; електромагнітні завади, створювані атмосферними явищами, тощо.

Осцилограму детермінованого сигналу (3.1) із початковою фазою $\phi_0 = 0$ зображено на рис. 3.2,а, а випадкового — на рис. 3.2,б.

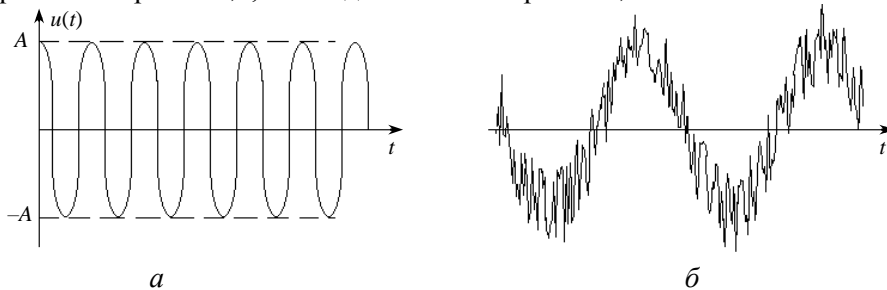


Рис. 3.2. Осцилограми типових сигналів:
а — детермінованого; б — випадкового

Усі сигнали, що несуть інформацію, є випадковими, оскільки детермінований (тобто цілком відомий) сигнал інформації не містить. Він може бути повністю відтворений у місці приймання без передавання по каналу зв'язку.

Між детермінованими і випадковими сигналами немає непереборної межі. Часто в умовах, коли рівень завад значно менший за рівень корисного сигналу з відомою формою, проста детермінована модель виявляється цілком адекватною поставленому завданню.

Критерій області існування сигналів. Якщо за класифікаційну ознаку взяти довжину часового інтервалу, у межах якого існує сигнал, то можна виділити такі класи сигналів: *нескінченні в часі сигнали* та

Розділ 2. Кількісні оцінки вимірювальної інформації

імпульсні сигнали, тобто сигнали, що існують у межах скінченного інтервалу часу.

Загальна форма запису нескінченних у часі сигналів: $s(t)$, $t \in (-\infty, \infty)$, а імпульсних сигналів: $s(t)$, $t \in [a, b]$, де a і b — межі часового інтервалу, в якому існує сигнал. В електро- і радіотехніці широко використовується поняття відеоімпульсів (рис. 3.3,*а*) і радіоімпульсів (рис. 3.3,*б*). Їх характерні особливості легко встановити за часовими діаграмами. Відмінність між цими двома основними імпульсними сигналами полягає у наступному. Якщо $A(t)$ — відеоімпульс, то відповідний йому радіоімпульс

$$u_p(t) = A(t) \cos(\omega_0 t + \phi_0), \quad (3.2)$$

причому сигнал $A(t)$ називається *обвідною* радіоімпульсу $u_p(t)$, а функція $\cos(\omega_0 t + \phi_0)$ — його *високочастотним заповненням*.

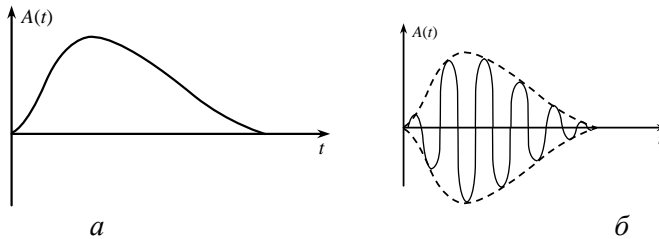


Рис. 3.3. Імпульсні сигнали

Із порівняння виразів (3.1) і (3.2) випливає, що математичну модель радіоімпульсу $u_p(t)$ можна отримати з моделі гармонічного коливання $u(t)$, заданого формулою (3.1), якщо в ній виконати заміну амплітуди A (сталой величини) на функцію часу $A(t)$, що описує відеоімпульс.

Критерій розмірності сигналів. Припустимо, що спостерігається сигнал у вигляді напруги $u(t)$ на затискачах будь-якого елемента електричного кола або струму $i(t)$, що протікає в заданій вітці кола. Такий сигнал $u(t)$ або $i(t)$, що описується однією функцією часу, називають *одновимірним сигналом*. Так звані *багатовимірні сигнали* виду $\vec{V}(t) = \{v_1(t), v_2(t), \dots, v_N(t)\}$ складаються з деякої множини одновимірних сигналів. Ціле число N називають *розмірністю* сигналу. Як приклад багатовимірного сигналу можна навести систему напруг на затискачах багатополюсника.

Критерій динаміки сигналів. За характером зміни в часі розрізняють: *статичні* та *динамічні сигнали*. У *статичній моделі сигналу немає часового параметру*. Такі сигнали використовуються передусім для зберігання інформації у вигляді цифрових кодів,

наприклад у комірках пам'яті цифрових обчислювальних машин, у програмованих логічних матрицях тощо. *Динамічні сигнали залежать від часу. Їх математичні моделі містять часовий аргумент.*

Критерій дійсності сигналів. Часові функції, за допомогою яких задаються моделі сигналів, можуть набувати як дійсних, так і комплексних значень. Вибір тієї чи іншої моделі сигналу зумовлюється лише простотою математичного аналізу. За так званим *критерієм дійсності* всі сигнали можна поділити на два класи: *дійсні та комплексні сигнали.*

Усі розглянуті раніше моделі сигналів належать до класу дійсних. Математичну модель комплексних сигналів можна подати у загальному вигляді

$$\dot{z}(t) = x(t) + jy(t),$$

де $x(t)$ і $y(t)$ — дійсні сигнали; $j = \sqrt{-1}$ — уявна одиниця.

Частковим випадком моделі (3.3) є так званий комплексно-експоненціальний сигнал $\dot{x}(t) = e^{j\omega t} = \cos \omega t + j \sin \omega t$, застосування якого продуктивне при розв'язуванні задач спектрального аналізу.

Критерій повторюваності сигналів. *Періодичність* також може бути класифікаційною ознакою сигналів. За цим критерієм розрізняють два основні класи сигналів: *періодичні та аперіодичні (поодинокі) сигнали.*

Загальна форма математичної моделі періодичного сигналу має вигляд:

$$x(t) = x(t + nT), \quad n = \pm 1, \pm 2, \dots,$$

де T - період сигналу. Деякі періодичні сигнали зображено на рис. 3.4:

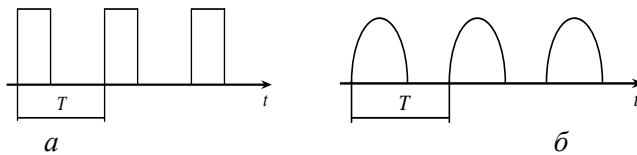


Рис. 3.4. Приклади періодичних сигналів: *а* — прямокутних; *б* — зрізаних косинус них

Аперіодичний (імпульсний чи поодинокий) сигнал $x(t)$ є частковим випадком періодичного сигналу, коли період проходження імпульсів T прямує до нескінченності: $x(t) = \lim_{T \rightarrow \infty} x(t + nT)$. Сукупність розглянутих видів сигналів і їх класифікаційних ознак зведено в табл. 3.1.

Таблиця 3.1

Класифікаційні ознаки сигналів	Типи сигналів
Характер зміни за величиною і в часі	Аналогові

Розділ 2. Кількісні оцінки вимірювальної інформації

	Дискретні
	Квантовані
	Цифрові
Передбачуваність миттєвих значень	Детерміновані
	Випадкові
Довжина інтервалу існування	Імпульсні
	Нескінченні
Розмірність	Одновимірні
	Багатовимірні
Динамічність	Статичні
	Динамічні
Дійсність	Дійсні
	Комплексні
Повторюваність	Періодичні
	Поодинокі

Гармонічні сигнали. Гармонічні сигнали називають ще тригонометричними сигналами. Математична модель таких сигналів визначається формулою (3.1), а типову осцилограму зображено на рис. 3.2, а.

Розглянемо **математичні моделі найпростіших типових сигналів:**

Комплексно-експоненціальні сигнали. Математична модель комплексно-експоненціальних сигналів має вигляд:

$$\dot{z}(t) = Ae^{j(\omega_0 t + \phi_0)}. \quad (3.4)$$

Скориставшись формулою Ейлера, подамо модель (3.4) у такий спосіб:

$$\dot{z}(t) = A \cos(\omega_0 t + \phi_0) + jA \sin(\omega_0 t + \phi_0). \quad (3.5)$$

Доданки у правій частині рівняння (3.5) — відповідно дійсна і уявна складові комплексно-експоненціального сигналу (3.4).

Надамо виразу (3.5) дещо іншої форми: $\dot{z}(t) = x(t) + jy(t)$, де $x(t)$ і $y(t)$ — дійсні сигнали; $j = \sqrt{-1}$ — уявна одиниця.

На рис. 3.5 зображено структурну схему пристрою фізичного моделювання пари дійсних сигналів $x(t)$ і $y(t)$ які ще називаються *квадратурними сигналами*. Комплексність сигналу $\dot{z}(t)$, що складається з дійсних сигналів $x(t)$ і $y(t)$, можна реалізувати при програмній чи апаратурній реалізації наступних алгоритмів перетворення сигналу $\dot{z}(t)$.

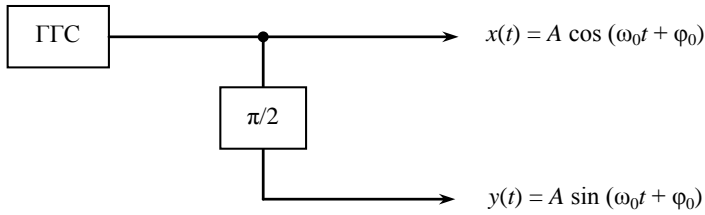
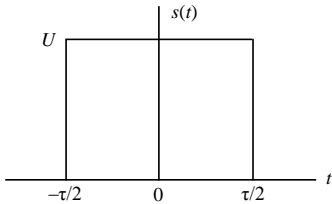


Рис. 3.5. Схема формування квадратурних сигналів:
ГГС — генератор гармонічних сигналів; $\pi/2$ — фазообертач

Прямокутні відеоімпульси. Математична модель $s(t)$ поодинокого прямокутного відеоімпульсу, симетрично розміщеного відносно початку відліку часу (рис. 3.6), задається співвідношенням



$$s(t) = \begin{cases} U, & t \in [-\tau/2, \tau/2] \\ 0, & t \notin [-\tau/2, \tau/2] \end{cases},$$

Рис. 3.6. Прямокутний відеоімпульс

де U — амплітуда; τ — тривалість

прямокутного відеоімпульсу.

У загальному випадку відеоімпульс може бути зміщений відносно початку відліку часу праворуч (*затриманий сигнал*; рис. 3.7, а) чи ліворуч (*випереджувальний сигнал*; рис. 3.7, б).

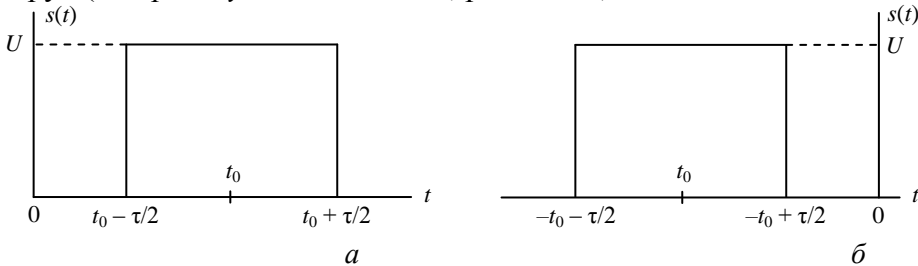


Рис. 3.7. Зміщені в часі прямокутні відео імпульси

Математичні моделі зміщених у часі прямокутних відеоімпульсів можна записати у вигляді систем рівнянь:

$$s(t) = \begin{cases} U, & t \in [t_0 - \tau/2, t_0 + \tau/2] \\ 0, & t \notin [\cdot] \end{cases} \quad \text{- для затриманого сигналу;}$$

$$s(t) = \begin{cases} U, & t \in [-t_0 - \tau/2, -t_0 + \tau/2] \\ 0, & t \notin [\cdot] \end{cases} \quad \text{- для випереджувального сигналу.}$$

Розділ 2. Кількісні оцінки вимірювальної інформації

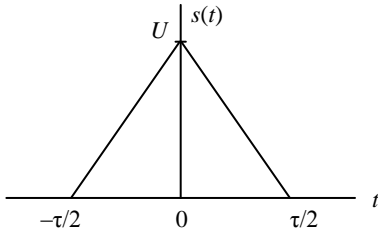


Рис. 3.8. Трикутний відеоімпульс

У загальному вигляді запис $s(t-t_0)$ стосується моделі затриманого сигналу, а $s(t+t_0)$ – моделі випереджувального сигналу, причому t_0 – інтервал тимчасового зсуву сигналу.

Трикутні відеоімпульси.

Математична модель поодинокого трикутного відеоімпульсу (рис. 3.8), симетрично розміщеного відносно початку відрітку часу, задається системою рівностей:

$$s(t) = \begin{cases} U(1+t/(\tau/2)), & t \in [-\tau/2, 0]; \\ U(1-t/(\tau/2)), & t \in [0, \tau/2]; \\ 0, & t \notin [-\tau/2, \tau/2]. \end{cases}$$

Трикутний відеоімпульс у загальному випадку також може бути зміщеним відносно початку відрітку часу (рис. 3.9).

Математична модель трикутного сигналу, затриманого на проміжок часу t_0 (рис. 3.9, а), має вигляд:

$$s(t-t_0) = \begin{cases} U(1+(t-t_0)/(\tau/2)), & t \in [t_0 - \tau/2, t_0]; \\ U(1-(t-t_0)/(\tau/2)), & t \in [t_0, t_0 + \tau/2]; \\ 0, & t \notin [t_0 - \tau/2, t_0 + \tau/2]. \end{cases}$$

Для випереджувального трикутного відеоімпульсу (рис. 3.9, б)

$$s(t+t_0) = \begin{cases} U(1+(t+t_0)/(\tau/2)), & t \in [-t_0 - \tau/2, -t_0]; \\ U(1-(t+t_0)/(\tau/2)), & t \in [-t_0, -t_0 + \tau/2]; \\ 0, & t \notin [-t_0 - \tau/2, -t_0 + \tau/2]. \end{cases}$$

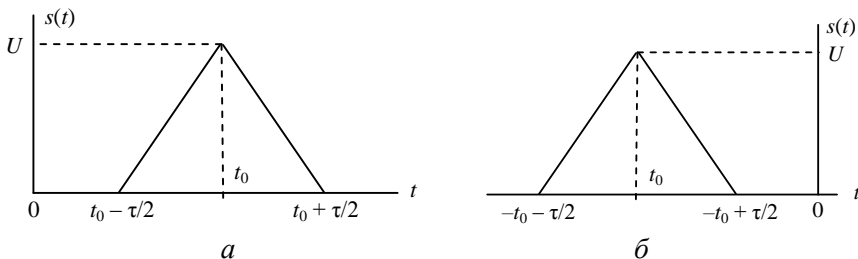


Рис. 3.9. Зміщені в часі трикутні відеоімпульси

Східчасті сигнали. Східчасті сигнали (рис. 3.10) ще називають *функціями ввімкнення*, *сигма-функціями* або *функціями Хевісайда*. Математична модель східчастого сигналу аналітично описується як:

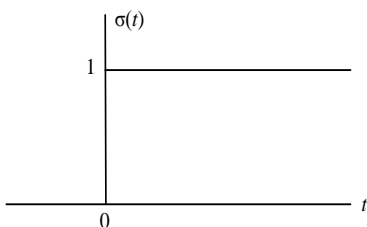


Рис. 3.10. Східчастий сигнал

$$\sigma(t) = \begin{cases} 1, & t \geq 0; \\ 0, & t < 0. \end{cases} \quad (3.6)$$

У загальному випадку східчаста функція може бути зміщена відносно початку відліку часу (рис. 3.11) на величину t_0 (затримки або випередження). Математична модель східчастого сигналу, затриманого на проміжок часу t_0 , має вигляд: $\sigma(t - t_0) = \begin{cases} 1, & t \geq t_0; \\ 0, & t < t_0. \end{cases}$

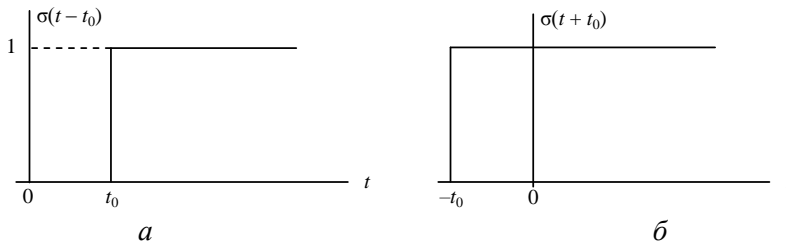


Рис. 3.11. Змішені східчасті сигнали: *a*— затриманий; *б* — випереджувальний

Математична модель випереджувальної на проміжок часу t_0 східчастої функції визначається системою: $\sigma(t + t_0) = \begin{cases} 1, & t \geq -t_0; \\ 0, & t < -t_0. \end{cases}$

Графіки на рис. 3.11 відображують характер зсуву затриманого і випереджувального сигналів відносно початку відліку часу.

Дельта-функція. Розглянемо імпульсний сигнал $v(t; \tau)$ прямокутної форми, основа якого дорівнює τ , а висота h — величина, обернена до τ : $h = 1/\tau$. Очевидно, що при будь-якому τ площа P_v , обмежена таким сигналом, дорівнює одиниці: $P_v = \tau h = 1$. Графіки сигналів $v(t; \tau)$, симетрично розміщених відносно початку відліку часу, для двох значень τ зображено на рис. 3.12.

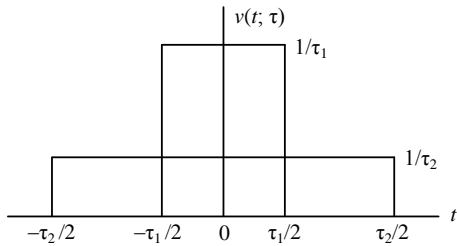


Рис. 3.12. Зображення функцій $v(t; \tau)$

Цей імпульс характерний тим, що при будь-якому виборі параметра τ його площа дорівнює одиниці: $P_v = \int_{-\infty}^{\infty} v dt \equiv 1$. (3.7)

Нехай тепер величина τ прямує до нуля. Імпульс, скорочуючись за тривалістю, зберігає свою площу, тому його висота має необмежено зростати. Границя такої функції при $\tau \rightarrow 0$ називається *дельта-функцією* або *функцією Дірака*: $\delta(t) = \lim_{\tau \rightarrow 0} v(t; \tau)$. (3.8)

Розділ 2. Кількісні оцінки вимірювальної інформації

Дельта-функція дорівнює нулю всюди, за винятком точки $t = 0$, в якій вона набуває нескінченного значення, тобто
$$\delta(t) = \begin{cases} \infty, & t = 0; \\ 0, & t \neq 0. \end{cases} \quad (3.9)$$

Щодо дельта-функції, заданої виразом (3.9), говорять, що вона зосереджена в точці $t = 0$. Однак, як випливає зі співвідношень (3.7) і (3.8), площа, обмежена дельта-функцією, скінченна:
$$\int_{-\infty}^{\infty} \delta(t) dt = 1. \quad (3.10)$$

Вираз (3.10) визначає умову нормування дельта-функції. Для стислості замість дельта-функції записують іноді δ -функція. Так само, як і східчаста, дельта-функція може бути зміщена на деякий проміжок часу t_0 (рис. 3.13). Зміщену праворуч на осі часу дельта-функцію називають затриманою дельта-функцією. Її модель записують системою рівностей:

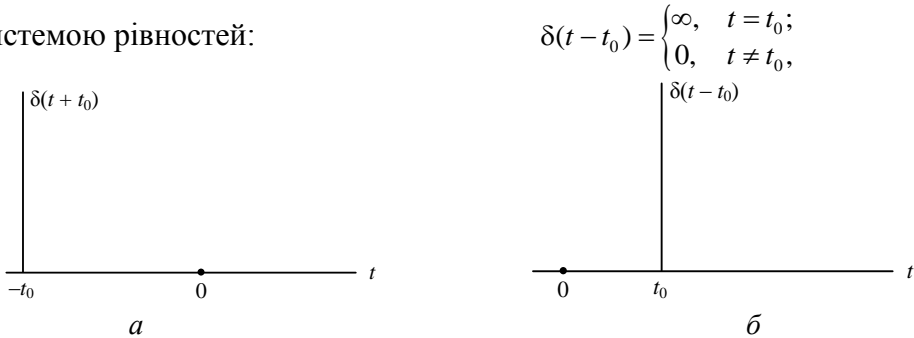


Рис. 3.13. Зображення дельта-функцій: *a* - затриманої; *b* - випереджувальної

У разі зміщення дельта-функції ліворуч відносно початку відліку часу її називають випереджувальною дельта-функцією. Її модель така:

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

Між дельта-функцією і східчастою функцією існує взаємно однозначна відповідність. Цю відповідність можна легко встановити за допомогою допоміжної функції $u(t; \tau)$, математична модель якої

задається системою рівностей:
$$u(t; \tau) = \begin{cases} 0, & t < -\tau/2; \\ t/\tau + 1/2, & -\tau/2 \leq t \leq \tau/2; \\ 1, & t > \tau/2. \end{cases}$$

(3.11)

Графік цієї функції наведений на (рис. 3.14). Зі співвідношення (3.11) і рис. 3.14 безпосередньо випливає, що, по-перше, при $\tau \rightarrow 0$ функція $u(t; \tau)$ переходить у східчасту функцію, тобто

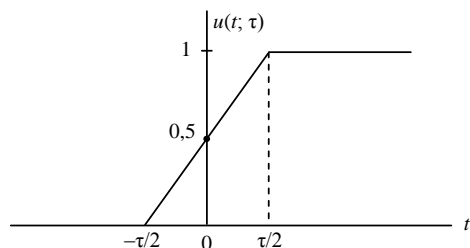


Рис. 3.14. Зображення функції $u(t; \tau)$

$$\lim_{\tau \rightarrow 0} u(t; \tau) = \sigma(t), \quad (3.12)$$

і, по-друге, похідна за часом від функції $u(t; \tau)$ збігається з функцією $v(t; \tau)$, заданої спів-відношенням (3.11), тобто $u'(t; \tau) = v(t; \tau)$, а при $\tau \rightarrow 0$:

$$\lim_{\tau \rightarrow 0} u'(t; \tau) = \delta(t). \quad (3.13)$$

Із граничних співвідношень (3.12) і (3.13) безпосередньо випливає, що *дельта-функція — це похідна від східчастої функції*: $\delta(t) = \sigma'(t)$.

Водночас правильне й обернене твердження: *східчаста функція є інтеграл від дельта-функції*:
$$\sigma(t) = \int_{-\infty}^t \delta(\tau) d\tau.$$

Якщо деяку неперервну функцію $x(t)$ помножити на дельта-функцію $\delta(t - t_0)$ і добуток зінтегрувати за часом t , то результат дорівнюватиме значенню неперервної функції в тій точці t_0 , де зосереджено δ -імпульс.

Справді, у співвідношенні
$$\int_{-\infty}^{\infty} x(t) \delta(t - t_0) dt$$
 підінтегральний вираз відрізняється від нуля лише для $t = t_0$. У цій точці функція $x(t)$ набуває конкретного значення $x(t_0)$, яке можна винести з-під знака інтеграла. Частина інтеграла, що залишилася, згідно з умовою нормування дельта-функції (3.10), дорівнює одиниці. Таким чином, дістаємо вираз:

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

У цьому виявляється *фільтрувальна властивість дельта-функції*,

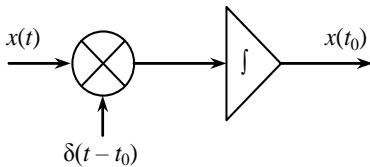


Рис. 3.15. Структурна схема вимірювання миттєвих значень сигналу

яку можна використати в пристроях вимірювання миттєвих значень деякого сигналу $x(t)$. Структурна схема вимірника (рис. 3.15) складається з двох ланок: помножувача та інтегратора.

Значення $x(t_0)$ буде виміряно тим точніше, чим коротший реальний сигнал, що наближено подає дельта-функцію.

Сигнум-функція. Функція $\text{sign}(t)$,

математична модель якої описується системою рівностей:

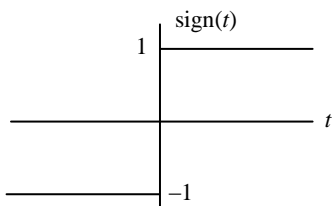


Рис. 3.16. Сигнум-функція

$$\text{sign}(t) = \begin{cases} 1, & t > 0; \\ 0, & t = 0; \\ -1, & t < 0, \end{cases}$$

називається *сигнум-функцією* (рис. 3.16).

Східчаста функція $\sigma(t)$ і сигнум-функція $\text{sign}(t)$ пов'язані співвідношеннями:
$$\sigma(t) = (1 + \text{sign}(t))/2;$$

Розділ 2. Кількісні оцінки вимірювальної інформації

$$\text{sign}(t) = 2\sigma(t) - 1.$$

Енергетичні характеристики сигналу. Відомі з фізики поняття потужності та енергії електричних сигналів можна легко узагальнити й перенести на відповідні поняття потужності та енергії довільних сигналів. Розглянемо найпростіше електричне коло (рис. 3.17), що складається з генератора напруги $u(t)$, до клем якого підімкнено резистор з опором R .

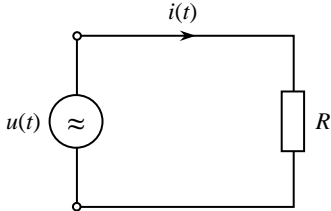


Рис. 3.17. До визначення потужності сигналу

У колі протікає струм $i(t)$, значення якого визначається за законом Ома:

$$i(t) = u(t)/R.$$

Миттєве значення потужності генератора $p(t)$ дорівнює добутку напруги $u(t)$ і струму $i(t)$:

$$p(t) = u(t)i(t).$$

Енергія E , що її генератор віддає в мережу (енергія, що виділяється на резисторі R), в інтервалі часу $[a, b]$ визначається співвідношенням:

$$E = \int_a^b p(t) dt = \int_a^b u(t)i(t) dt. \quad (3.14)$$

У випадку, коли $R=1$, струм $i(t)$ у колі дорівнює напрузі генератора $u(t)$, а значення енергії E_u обчислюється як $E_u = \int_a^b u^2(t) dt$.

Узагальнимо поняття потужності та енергії в електричному колі, в якому протікає струм $i(t)$ під впливом напруги $u(t)$, і перенесемо їх на відповідні поняття потужності та енергії довільних сигналів $u(t)$ і $v(t)$.

Отже, нехай $u(t)$ і $v(t)$ — деякі сигнали (точніше — математичні моделі сигналів). За аналогією зі співвідношенням (3.14) добуток

$$p_{u,v}(t) = u(t)v(t)$$

називатимемо *взаємною потужністю сигналів*, а інтеграл

$$E_{u,v} = \int_{-\infty}^{\infty} u(t)v(t) dt \quad (3.15)$$

— *взаємною енергією цих сигналів*. Тоді повна енергія сигналу (при $v(t)$, що тотожно дорівнює $u(t)$), визначиться виразом:

$$E_u = \int_{-\infty}^{\infty} u^2(t) dt. \quad (3.16)$$

У математиці величина $\int_{-\infty}^{\infty} u(t)v(t) dt$ називається *скалярним*

добутком функцій $u(t)$ і $v(t)$ і позначається як: $(u, v) = \int_{-\infty}^{\infty} u(t)v(t) dt$.

(3.17)

З порівняння виразів (3.17) і (3.15) випливає, що *скалярний добуток двох сигналів (функцій) — це взаємна енергія цих сигналів, тобто*

$$(u, v) = E_{u,v}. \quad (3.18)$$

Як частковий випадок із формули (3.18) дістаємо: $(u, u) = E_u$.

Таким чином, *повна енергія сигналу визначається як скалярний добуток сигналу із самим собою.*

Перенесемо поняття взаємної і повної енергії на комплексні сигнали. Якщо за формулою (3.16) обчислити повну енергію комплексного сигналу, то прийдемо до того, що й енергія такого сигналу також виявиться комплексною (оскільки квадрат комплексної функції — функція комплексна), а це неприпустимо. Тому взаємну енергію двох комплексних сигналів $\dot{u}(t)$ і $\dot{v}(t)$

задамо співвідношенням

$$E_{\dot{u}, \dot{v}} = (\dot{u}, \dot{v}^*) = \int_{-\infty}^{\infty} \dot{u}(t) \dot{v}^*(t) dt, \quad (3.19)$$

де $\dot{v}^*(t)$ — сигнал, комплексно-спряжений із сигналом $v(t)$.

Зі співвідношення (3.19) як окремий випадок дістанемо повну енергію комплексного сигналу $\dot{u}(t)$:

$$E_{\dot{u}} = (\dot{u}, \dot{u}^*). \quad (3.20)$$

Добуток комплексної та комплексно-спряженої з нею величини (функції) дорівнює квадрату модуля цієї величини: $\dot{u} \dot{u}^* = |\dot{u}|^2$. Оскільки квадрат модуля комплексної функції – функція дійсна, то повна енергія комплексного сигналу також стає дійсною величиною, тобто

$$E_{\dot{u}} = \int_{-\infty}^{\infty} |\dot{u}(t)|^2 dt.$$

Розділ 2. Кількісні оцінки вимірювальної інформації



Леонард Ейлер
(Leonhard Euler,
1707 - 1783),

видатний математик, який вніс величезний внесок у розвиток математики, а також механіки, фізики, астрономії та ряду прикладних наук. Ейлер - автор більше 800 робіт з математичного аналізу, диференціальної геометрії, теорії чисел, наближених обчислень, небесної механіки, математичної фізики, оптики, балістики, кораблебудування, теорії музики і багатьох інших, що мали значний вплив на розвиток науки. Протягом 1731 - 1741 рр. і починаючи з 1766 р. працював у Петербургській академії наук.

Корінь квадратний з повної енергії сигналу (функції) називають нормою сигналу (функції):

$$\| \dot{u} \| = \sqrt{\int_{-\infty}^{\infty} |\dot{u}(t)|^2 dt}.$$

Функцію $u(t)$, для якої виконується умова $\|u\|^2 = 1$, називають *нормованою*.

Отже, *нормованими називаються функції, повна енергія яких дорівнює одиниці*.

Ортогональні сигнали та ортогональні базиси. Два сигнали $u(t)$ і $v(t)$ називаються *ортогональними*, якщо їхній скалярний добуток, а отже, і взаємна енергія, дорівнюють нулю:

$$(u, v) = \int_{-\infty}^{\infty} u(t)v(t) dt = 0.$$

Нескінченна система дійсних функцій (сигналів) $\{\varphi_k(t)\} = \{\varphi_0(t), \varphi_1(t), \dots, \varphi_n(t), \dots\}$ називається *ортогональною*, якщо скалярний добуток двох різних сигналів, а отже, і їхня взаємна енергія, дорівнює нулю:

$$(\varphi_n, \varphi_m) = \int_{-\infty}^{\infty} \varphi_n(t)\varphi_m(t) dt = 0 \text{ при } n \neq m.$$

Припускають, що енергія кожного сигналу із системи не дорівнює нулю, тобто $\int_{-\infty}^{\infty} \varphi^2(t) dt \neq 0$. Це, зокрема, означає, що жодна з функцій системи не дорівнює тотожно нулю.

Нескінченну систему функцій $\{\varphi_k(t)\}$, попарно ортогональних одна одній і таких, що мають одиничні норми:

$$(\varphi_n, \varphi_m) = \begin{cases} 1, & n = m; \\ 0, & n \neq m, \end{cases} \quad (3.21)$$

називають *системою ортонормованих функцій* або *ортонормованим базисом*.

Зовсім не обов'язково, щоб ортонормованість базисних функцій забезпечувалася лише на нескінченному інтервалі часу. Існують численні базиси, в яких системи функцій $\{\varphi_k(t)\}$

ортонормовані на скінченному інтервалі часу.



**Олівер Хевісайд
(Oliver Heaviside,
1850-1925),**

англійський вчений самоук, інженер, математик і фізик. Створив теорію передавання сигналів на далекі відстані. Вперше застосував комплексні числа для вивчення електричних ланцюгів, розробив техніку застосування перетворення Лапласа для розв'язування диференціальних рівнянь, сформулював рівняння Максвела в термінах електричної та магнітної сил і потоку, а також незалежно від інших математиків створив векторний аналіз. Вперше розробив операційне зчислення, яке широко використовується в

Розділ 2. Кількісні оцінки вимірювальної інформації

Приклад. Розглянемо один із найважливіших і поширених базисів, що утворюється ортонормованою системою гармонічних функцій фізиці та інших науках.

На інтервалі часу $[-T/2, T/2]$ система тригонометричних функцій із кратними частотами, доповнена сталим у часі сигналом φ_0 :

$$\begin{aligned} \varphi_0(t) &= 1/\sqrt{T}; \\ \varphi_1(t) &= \sqrt{2/T} \sin 2\pi t / T; \\ \varphi_2(t) &= \sqrt{2/T} \cos 2\pi t / T; \\ &\dots\dots\dots \\ \varphi_{2k-1}(t) &= \sqrt{2/T} \sin 2\pi kt / T; \\ \varphi_{2k}(t) &= \sqrt{2/T} \cos 2\pi kt / T, \dots \end{aligned}$$

утворює ортонормований базис.

В ортогональності цих функцій легко переконатися

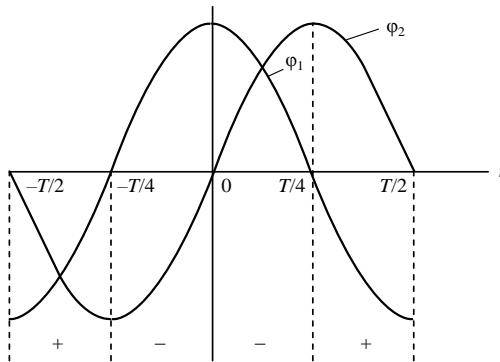


Рис. 3.18. Графіки гармонічних функцій

ся геометричною побудовою (рис. 3.18) на прикладі функцій φ_1 і φ_2 . Визначимо енергію базисної функції $\varphi_k(t)$ на інтервалі ортогональності:

$$E_k = \int_{-T/2}^{T/2} \varphi_k^2(t) dt. \tag{3.22}$$

Очевидно, що енергія, а отже, і норма функції $\varphi_0(t)$, дорівнює одиниці. При обчисленні енергії функцій $\varphi_k(t)$, $k \geq 1$, за формулою (3.22) візьмемо до уваги, що

$$\int \sin^2(pt) dt = t/2 - \sin(2pt)/4p,$$

а $\int \cos^2(pt) dt = t/2 + \sin(2pt)/4p.$



**Поль Адріан
Моріс Дірак
(Paul Adrien
Maurice Dirac,
1902 - 1984),**

англійський фізик,
один із творців
квантової механіки,
лауреат
Нобелівської премії
з фізики 1928 р.
(разом з Ервіном
Шредінгером).
Роботи Дірака
присвячені
квантовій механіці,
квантовій
електродинаміці,
теорії поля, теорії
елементарних
часток,
статистичній фізиці.
В 1926 - 1927 рр.
розробив
математичний
апарат квантової
механіки – теорію
перетворень, ввів
так звану дельта-
функцію. В 1927 р.
застосував
принципи квантової
механіки до
електромагнітного
поля і побудував
модель квантової
електродинаміки.

Розділ 2. Кількісні оцінки вимірювальної інформації

Нехай $\varphi_k(t) = \sqrt{2/T} \cos(2\pi kt/T)$, тоді для цієї функції

$$E_{\varphi_k} = \int_{-T/2}^{T/2} \varphi_k^2(t) dt = \frac{2}{T} \left(\frac{t}{2} + \frac{\sin 4\pi kt/T}{8\pi k/T} \right) \Big|_{-T/2}^{T/2}. \quad (3.23)$$

Підставивши межі інтегрування в правій частині виразу (3.23), дістанемо, що енергія (так само, як і норма) гармонічних базисних функцій (як синусних, так і косинусних) дорівнює одиниці.

Розглянутий базис гармонічних функцій є однією з багатьох систем ортонормованих функцій. Вибір тієї чи іншої системи ортонормованих функцій (базису) залежить як від властивостей сигналу, так і від конкретних умов застосування базису.

Загальні рекомендації щодо вибору базису:

1) базисні функції мають бути по зможі максимально погодженими за формою з фільтрованим сигналом.

Наприклад, коли апріорі відомо, що деякий складний сигнал містить гармонічну складову, інформацію про параметри якої дістаємо, розкладаючи сигнал в узагальнений ряд Фур'є, то як базисні функції доцільно використовувати саме систему ортонормованих гармонічних функцій, погоджених за формою з фільтрованим сигналом;

2) потрібно намагатися вибирати якомога простіший з погляду апаратної чи програмної реалізації базис.

Ці рекомендації не завжди вдається виконати, оскільки їх реалізація може призводити до конфліктних ситуацій. Тому іноді доводиться шукати компроміс між погодженістю форми базисних функцій з формою фільтрованого сигналу та простотою програмної чи апаратної реалізації базису.

Узагальнені ряди Фур'є. З математики відомо, що будь-який сигнал $s(t)$ зі скінченною енергією, тобто такий, для якого виконується умова

$$\int_{-\infty}^{\infty} s^2(t) dt < \infty, \text{ можна подати у вигляді ряду:}$$

$$s(t) = \sum_{k=0}^{\infty} c_k \varphi_k(t), \quad (3.24)$$

де c_k – коефіцієнти розкладу, які називають

спектром сигналу; φ_k – система ортонормованих дійсних функцій (базис).

Представлення (3.24) називається узагальненим рядом Фур'є сигналу $s(t)$ у вибраному базисі $\{\varphi_k\}$.

До розкладання (3.24) деякого сигналу $s(t)$ потрібно ставитися як до такого, котрий вводиться аксіоматично, тобто можна намагатися апроксимувати сигнал $s(t)$ нескінченною сумою добутків заздалегідь вибраних дійсних функцій $\varphi_k(t)$ (які зовсім не обов'язково мають бути ортонормованими) і поки не відомих коефіцієнтів розкладу c_k . Завдання полягає в тому, щоб вибрати таку систему функцій $\{\varphi_k\}$, яка, по-перше, забезпечувала б простоту обчислення коефіцієнтів розкладу $\{c_k\}$ і, по-друге, мінімізувала похибку апроксимації сигналу $s(t)$ скінченновимірним рядом Фур'є.

Якщо за систему $\{\varphi_k\}$ узяти сукупність ортонормованих функцій, то зазначену задачу можна успішно розв'язати. Отже, нехай у (3.24) $\{\varphi_k\}$ — нескінченна система дійсних функцій, ортонормованих на інтервалі часу $[-\infty, \infty]$, що задовольняють умову (3.21), тобто

$$\int_{-\infty}^{\infty} \varphi_k(t) \varphi_n(t) dt = \begin{cases} 1, & n = k; \\ 0, & n \neq k. \end{cases} \quad (3.25)$$

Помножимо обидві частини розкладу (3.24) на $\varphi_n(t)$ і проінтегруємо в межах області ортогональності: $\int_{-\infty}^{\infty} s(t) \varphi_n(t) dt = \sum_{k=0}^{\infty} c_k \int_{-\infty}^{\infty} \varphi_k(t) \varphi_n(t) dt$. (3.26)

Згідно з обмеженням (3.25) інтеграл у правій частині виразу (3.26) відмінний від нуля і дорівнює одиниці лише при $k = n$. Отже,

$$c_n = \int_{-\infty}^{\infty} s(t) \varphi_n(t) dt = (s, \varphi_n). \quad (3.27)$$

Співвідношення (3.27) є фундаментальним у теорії рядів Фур'є і визначає алгоритм обчислення коефіцієнтів розкладу сигналу $s(t)$ для заданого ортонормованого базису, а саме: *коефіцієнти розкладу узагальненого ряду Фур'є (спектр) часової функції (сигналу) $s(t)$ визначаються скалярним добутком цього сигналу та відповідних базисних функцій:*

(3.28)

$$c_k = (s(t), \varphi_k(t)), \quad k = 0, 1, \dots$$

Можливість подання сигналів за допомогою узагальнених рядів Фур'є є фактом принципового значення. Зручність застосування таких рядів при аналізі сигналів полягає ось у чому. Замість того, щоб вивчати функціональну залежність у континуальній (незчисленній) множині, достатньо подати ці сигнали зчисленною системою коефіцієнтів узагальненого ряду Фур'є c_k .

Розділ 2. Кількісні оцінки вимірювальної інформації

Зауважимо ще раз, що *спектр сигналу завжди визначається щодо конкретного базису. Зі зміною базису змінюється і спектр, хоча сигнал залишається незмінним.*

Для комплексного сигналу $\dot{s}(t)$ його розклад у ряд Фур'є записується у вигляді:

$$\dot{s}(t) = \sum_{k=0}^{\infty} \dot{c}_k \varphi_k(t), \quad (3.29)$$

тобто оскільки система базисних функцій $\{\varphi_k\}$ є системою дійсних функцій, то для комплексного сигналу $\dot{s}(t)$ його спектр $\{\dot{c}_k\}$ також є комплексним і визначається співвідношенням: $\dot{c}_k = (\dot{s}(t), \varphi_k(t))$. (3.30)

Обчислимо енергію в загальному випадку комплексного сигналу, поданого у формі узагальненого ряду Фур'є (3.29). Скориставшись співвідношеннями (3.29) і (3.30), на підставі виразу (3.20), дістанемо:

$$E_s = (\dot{s}(t), \dot{s}^*(t)) = \int_{-\infty}^{\infty} \dot{s}(t) \dot{s}^*(t) dt = \int_{-\infty}^{\infty} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} (\dot{c}_k \dot{c}_n^*) \varphi_k(t) \varphi_n(t) dt.$$

Помінявши місцями операції підсумовування та інтегрування (з огляду на їхню лінійність), отримаємо: $E_s = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} (c_k c_n^*) \int_{-\infty}^{\infty} \varphi_k(t) \varphi_n(t) dt$.

Оскільки базисні функції ортонормовані, то в сумі правої частини останнього виразу буде відмінним від нуля тільки доданок при $n = k$.

Отже, приходимо до надзвичайно важливого результату:

$$E_s = \sum_{k=0}^{\infty} \dot{c}_k \dot{c}_k^* = \sum_{k=0}^{\infty} c_k^2,$$

де $c_k = |\dot{c}_k|$ — модуль k -го коефіцієнта розкладу узагальненого ряду Фур'є.

Зміст отриманого виразу полягає в тому, що *енергія сигналу дорівнює сумі енергій усіх компонентів (гармонік), з яких складається узагальнений ряд Фур'є.*

Іншими словами, *енергія сигналу дорівнює сумі енергій спектральних складових сигналу, а квадрат модуля коефіцієнтів узагальненого ряду Фур'є чисельно дорівнює тій частці енергії сигналу, що міститься у відповідній спектральній складовій сигналу.*

А тепер доведемо оптимальність розкладення сигналу за ортонормованим базисом. Для дійсного сигналу $s(t)$ введемо

скінченновимірну апроксимацію: $\tilde{s}(t) = \sum_{n=0}^N c_n \varphi_n(t)$ з невідомими

коефіцієнтами c_n і вимагатимемо, щоб ці коефіцієнти було вибрано з умови мінімальності енергії помилки апроксимації, яка інакше називається умовою

мінімальності квадратичної помилки апроксимації:

$$\mu = \|s - \tilde{s}\|^2 = \int_{-\infty}^{\infty} \left[s(t) - \sum_{n=0}^N c_n \varphi_n(t) \right]^2 dt = \min.$$

Необхідна умова мінімуму полягає в тому, що коефіцієнти c_k мають задовольняти систему лінійних рівнянь: $\partial \mu / \partial c_k = 0, \quad k = \overline{0, N}$.

(3.31)

У розгорнутому вигляді помилка апроксимації

$$\mu = \int_{-\infty}^{\infty} \left(s^2 - 2s \sum_{k=0}^N c_k \varphi_k + \sum_{k=0}^N \sum_{n=0}^N c_k c_n \varphi_k \varphi_n \right) dt. \quad (3.32)$$

Для спрощення запису аргумент t базисних функцій φ у (3.32) опущено. Розглянемо останній доданок у виразі (4.32):

$$A_{kn} = \int_{-\infty}^{\infty} \left(\sum_{k=0}^N \sum_{n=0}^N c_k c_n \varphi_k \varphi_n \right) dt.$$

Згідно властивості лінійності операції інтегрування та підсумовування можна поміняти місцями, а доданок A_{kn} записати у вигляді:

$$A_{kn} = \sum_{k=0}^N \sum_{n=0}^N c_k c_n \int_{-\infty}^{\infty} \varphi_k \varphi_n dt. \quad \text{Оскільки базис } \{\varphi_k\} \text{ ортонормований, то,}$$

скориставшись системою (3.25), дістанемо: $A_{kn} = \sum_{k=0}^N c_k^2$. Тоді співвідношення

$$(3.32) \text{ можна подати у вигляді: } \mu = \int_{-\infty}^{\infty} \left(s^2 - 2s \sum_{k=0}^N c_k \varphi_k \right) dt + \sum_{k=0}^N c_k^2.$$

(3.33)

Беручи до уваги, що у формулі (4.33) інтеграл $\int_{-\infty}^{\infty} s^2(t) dt$ не залежить від коефіцієнта c_k , після диференціювання правої частини виразу (4.33) за c_k

дістаємо: $-2 \int_{-\infty}^{\infty} s(t) \varphi_k(t) dt + 2c_k = 0$, а це безпосередньо приводить до висновку про те, що рівності (4.31) будуть виконуватися за такого вибору

коефіцієнтів розкладення:

$$(3.34) \quad c_k = \int_{-\infty}^{\infty} s(t) \varphi_k(t) dt = (s, \varphi_k).$$

Це цілком співпадає з виразом (3.27) для коефіцієнтів узагальненого ряду Фур'є.

Більш ретельний аналіз, коли розглядається не тільки перша, а й друга похідна енергії помилки, показує, що ряд Фур'є

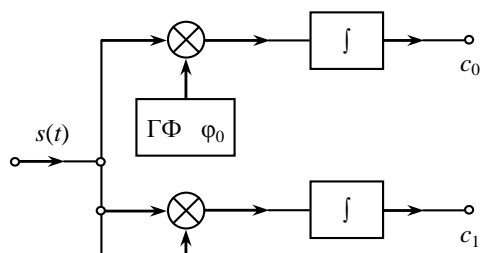


Рис. 3.19. Структурна схема пристрою спектрального аналізу сигналів

Розділ 2. Кількісні оцінки вимірювальної інформації

забезпечує не просто екстремум, а саме мінімум помилки апроксимації. Ця помилка тим менша, чим більше N . У границі, коли N прямує до нескінченності, помилка стає такою, що дорівнює нулю, і розклад типу (4.24) точно описує сигнал $s(t)$.

Насамкінець розглянемо структурну схему пристрою (рис. 3.19) для експериментального визначення коефіцієнтів розкладення довільного сигналу $s(t)$ в узагальнений ряд Фур'є (3.24) за заданою системою ортонормованих функцій $\varphi_k(t)$, $k = 0, 1, \dots$. Основні елементи пристрою — це генератори базисних функцій (ГФ) $\varphi_k(t)$, $k = 0, 1, \dots$, за якими розкладається сигнал $s(t)$ в узагальнений ряд Фур'є (тобто обчислюються коефіцієнти розкладення, сукупність яких утворює спектр c_k , $k = 0, 1, \dots$). Аналізований сигнал $s(t)$ одночасно подається на перші входи всіх помножувальних ланок, на другі входи яких підводяться базисні функції $\varphi_k(t)$, $k = 0, 1, \dots$. З виходу помножувачів сигнали надходять на інтегратори, відгуком яких є коефіцієнти c_k . За такого методу опрацювання сигналу наприкінці проміжку часу інтегрування (t_1, t_2) на виході кожного інтегратора виникає незмінний у часі сигнал, значення якого згідно формули (3.33) точно дорівнює тому чи іншому коефіцієнту розкладення узагальненого ряду Фур'є.

Отже, працездатність системи в цілому залежатиме від того, наскільки точно вдається відтворити базисні функції, а також від досконалості функціонування помножувачів та інтеграторів.

Зовсім не обов'язково обчислювати спектр сигналу апаратним способом. Алгоритм визначення коефіцієнтів розкладення узагальненого ряду Фур'є можна реалізувати й на програмному рівні.

Аналізуючи пристрій (рис. 3.19), переконуємося в тому, що всю інформацію, яка міститься в сигналі $s(t)$, можна подати у вигляді хоча й нескінченної, але все-таки зчисленної сукупності чисел c_k , $k = 0, 1, \dots$, що утворюють спектр сигналу. Цей висновок цілком очевидний, оскільки є наслідком подання сигналу у вигляді узагальненого ряду Фур'є (див. (3.24)).

Отже, будь-яку функцію (сигнал), аналітично описану в часовому просторі (таку функцію називають оригіналом), можна подати зчисленною сукупністю комплексних амплітуд \dot{c}_k , $k = 0, 1, \dots$, що утворюють спектр у частотному просторі (сукупність \dot{c}_k називають зображенням).

Між оригіналом і його зображенням існує взаємно однозначна відповідність, тобто, знаючи оригінал, можна знайти зображення, і навпаки.

Це вказує на дуальність процесу в сенсі рівносильності його подання у вигляді функції часу (у просторі оригіналів), або у вигляді сукупності елементарних частотних гармонік (у просторі зображень).

3.2 Випадкові сигнали та завади.

Основні поняття. При виконанні завдань вимірювань для опису досліджуємих реальних сигналів використовуються як детерміновані так і випадкові математичні моделі. Відповідно детерміновані сигнали описуються детермінованими функціями, а випадкові сигнали – випадковими функціями (процесами).

До детермінованих відносять сигнали, які повністю можна описати аналітично, або графічно, або у вигляді таблиць. Такі сигнали у певній мірі є ідеалізацією, яку необхідно обґрунтувати, що практично неможливо при вимірюваннях значень і характеристик реальних сигналів. Випадкові сигнали є більш адаптованими до таких вимірювань, але непрогнозованість їх значень і характеристик у значній мірі ускладнює отримання результату вимірювань.

У загальному випадку прояви випадковості (стохастичності) досліджуємих сигналів можна описати трьома наступними варіантами:

- а) динаміка зміни у просторі і часі значень і характеристик сигналів є випадковою ;
- б) початкові умови формування сигналу є випадковими, після їх реалізації динаміка сигналу є детермінованою;
- в) сигнал є комбінацією, в основному, адитивною або мультиплікативною, детермінованих і випадкових сигналів.

Варіант в) має місце практично для всіх процесів вимірювань.

Більш детально розглянемо випадкові події, випадкові функції, а саме випадкову величину і випадковий процес, які на сьогодні є основними моделями досліджуємих сигналів при вимірюваннях.

Випадкові події. Для опису випадкових подій використовується *стохастичний експеримент* проведення наступної послідовності дослідів, для яких має місце:

- а) кожний дослід проводиться незалежно один від одного, за однаковими умовами, їх кількість можна повторювати необмежено;

Розділ 2. Кількісні оцінки вимірювальної інформації

- б) перед проведенням експерименту всі його наслідки (події) відомі, але невідомо який саме з них відбудеться у наступному досліді, тобто наслідки в експерименті з'являються *випадково*;
- в) усім подіям, які можуть чи не можуть з'явитись при досліді, притаманна певна об'єктивна характеристика, яку іменують *частотою появи* відповідної події при проведенні певного числа дослідів.

Стохастичний експеримент є вимірювальним експериментом, тому що кожній появі конкретної події при експерименті ставиться у відповідність одиниця натурального ряду чисел. Саме числові результати стохастичного експерименту є основою практичної інтерпретації імовірності, яка використовується у подальшому при побудові ймовірнісної міри статистичних вимірювань.

Поставимо у відповідність стохастичному експерименту множину Ω - простір елементарних наслідків або елементарних подій. Елементами множини Ω є елементарні події $\omega \in \Omega$. При аналізі стохастичного експерименту розглядаються не тільки елементарні події, а й більш складні, як відповідна комбінація елементарних. Такі події іменують *випадковими подіями*, їх прийнято позначати початковими великими буквами латинської абетки: А, В, С, ...

Числовою характеристикою появи події А в стохастичному експерименті з n дослідідами є частота появи А: $\alpha = m/n$, де m – кількість появ події А. При необмеженому збільшенні числа дослідів $n \rightarrow \infty$, α прямує до межі P(A), яка іменується *ймовірністю* події А. Це один з фундаментальних результатів теорії ймовірності, який вперше був опублікований в працях *швейцарського математика Я. Бернуллі*. Так, з достатньою для практичних розрахунків точністю можна вважати, що якщо подія А з'явилась двічі при ста дослідідах, то частота

випадкової події А збігається з її ймовірністю $P(A) = 2 \cdot 10^{-2}$.

Для характеристики залежних випадкових подій А і В вводиться умовна ймовірність $P(A/B)$, що означає ймовірність прогнозу появи події А за умови, що подія В уже відбулась.

Випадкова величина. Така величина (random or stochastic variable) є функцією $\xi(\omega)$, що ставить у відповідність кожній елементарній події $\omega \in \Omega$ певне число з множини дійсних чисел R . Як кожна функція, випадкова

величина $\xi(\omega)$ має область визначення Ω з ймовірнісного простору (Ω, \mathcal{F}, P) , і область значень $X \subseteq \mathbb{R}$ з ймовірнісного простору

ймовірностей і теорії чисел, де його іменем названі числа з деякими важливими властивостями. Якобу Бернуллі належать також роботи з фізики, арифметики, алгебри і геометрії.



Якоб Бернуллі
(Jakob I (James) Bernoulli,
1655-1705),

швейцарський математик, брат Йогана Бернуллі, старший із відомої династії вчених. Якобу Бернуллі належать значні досягнення в теорії рядів, диференціальному численні, варіацій-ному численні, теорії

(X, \mathcal{B}, P_ξ) , де Ω - простір елементарних подій, \mathcal{F} - алгебра (σ - алгебра) підмножин Ω , P - ймовірність випадкових подій $A \in \mathcal{F}$, $X \subseteq \mathbb{R}$, \mathcal{B} - алгебра підмножин X , а, P_ξ - ймовірність значень $\xi(\omega)$.

Випадкові величини поділяються на дискретні та неперервні.

Дискретна випадкова величина $\xi(\omega)$ має скінченну або нескінченну дискретну множину значень $\{x_i, i = \overline{1, n}$ або $n \rightarrow \infty\}$. *Неперервна* випадкова величина $\xi(\omega)$ набуває будь-які значення з деякого інтервалу $X \subseteq \mathbb{R}$.

Для випадкових величин введені наступні не випадкові статистичні характеристики.

Функція розподілу ймовірності

$$F(x) = P\{\omega \in \Omega : \xi(\omega) < x\}, \quad (3.36)$$

визначає ймовірність тих $\omega \in \Omega$, для яких має місце $\xi(\omega) < x$, $x \in \mathbb{R}$. Функція $F(x)$ є монотонно зростаючою в інтервалі $0 \leq F(x) \leq 1$, при цьому $F(-\infty) = 0$ та $F(\infty) = 1$. Функція розподілу $F(x)$ – безрозмірна.

Розділ 2. Кількісні оцінки вимірювальної інформації

Щільність розподілу ймовірності $p(x)$ або щільність ймовірності визначається як похідна від функції розподілу $p(x) = dF(x)/dx$.

Одиниця виміру $p(x)$ - обернена одиниці виміру $\zeta(\omega)$.

Якщо задана $p(x)$, то $F(x)$ визначається за формулою:

$$F(x) = \int_{-\infty}^x p(x)dx. \quad (3.37)$$

Математичне сподівання

$$\mathbf{M}\zeta(\omega) = \int_{-\infty}^{\infty} xp(x)dx = \sum_{i=1}^n x_i p(x_i) = a_1, \quad (3.38)$$

у даному виразі інтегрування використовується при обчисленні математичного сподівання неперервної $\zeta(\omega)$ з відповідною щільністю ймовірності $p(x)$, а сума - при обчисленні математичного сподівання дискретної $\zeta(\omega)$ з відповідними ймовірностями $p(x_i)$ значень x_i .

Математичне сподівання $\mathbf{M}\zeta(\omega)$ є першим початковим моментом a_1 випадкової величини $\zeta(\omega)$. Одиниця виміру $\mathbf{M}\zeta(\omega)$ збігається з одиницею виміру $\zeta(\omega)$. Так, якщо $\zeta(\omega)$ - випадкова напруга чи струм, то $\mathbf{M}\zeta(\omega)$ - постійна складова або середнє значення напруги чи струму.

Дисперсія

$$\mathbf{D}\zeta(\omega) = \mathbf{M}\{[\zeta(\omega) - a_1]^2\} = \mathbf{M}\{\zeta^2(\omega)\} - a_1^2 = a_2 - a_1^2 = \sigma^2 \quad (3.39)$$

є характеристикою розсіювання (розкиду) значень $\zeta(\omega)$ відносно її середнього значення a_1 , а a_2 є другим початковим моментом $\zeta(\omega)$.

Одиниця виміру $\mathbf{D}\zeta(\omega)$ - квадрат одиниці виміру $\zeta(\omega)$.

Величина $\sigma = \sqrt{\mathbf{D}\zeta(\omega)}$ називається *середнім квадратичним відхиленням* і одиницею її виміру є одиниця виміру $\zeta(\omega)$.

Випадкові процеси. Випадковим процесомна $\zeta(\omega, t), \omega \in \Omega, t \in T$ називається двовимірна випадкова функція, значення якої при будь-якому фіксованому аргументу часу t є випадковою величиною.

Напруга шуму на виході лінії зв'язку, температура повітря, струм через мікрофон при розмові і т. ін., якщо проводити спостереження за миттєвими значеннями перелічених величин за часом, є прикладами реалізації випадкових процесів.

Результати окремих спостережень випадкового процесу іменуються його *реалізаціями*, кожна з них є детермінованою функцією і

позначаються $x(t)$. Сукупність можливих реалізацій $\{x_k(t), k = \overline{1, n}\}$ випадкового процесу називається ансамблем. Таким є ансамбль, що спостерігаються водночас на виходах різних каналів системи багатоканального зв'язку (рис. 3.20).

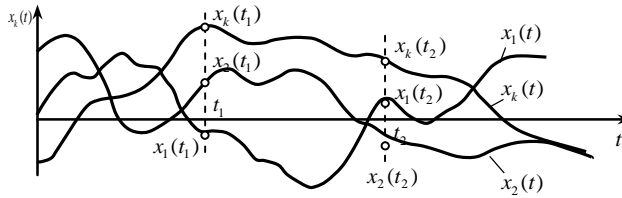


Рис. 3.20. Графіки ансамблю реалізацій випадкового процесу $\xi(\omega, t)$.

Реалізаціями випадкових процесів можуть бути детермінованими функціями часу різної складності. Так, наприклад, гармонічний сигнал $u(t) = A \cos(2\pi ft + \theta)$ може бути випадковим процесом, в якого хоча б один із параметрів A , f і θ є випадковою величиною, тобто коли або $A(\omega)$, або $f(\omega)$, або $\theta(\omega)$, а також коли має місце будь-яка їх комбінація.

При виконанні завдань вимірювань в основному використовуються стаціонарні випадкові процеси.

Відмітимо, що *стаціонарність* випадкових процесів вводиться у широкому і вузькому сенсах. У більшості практичних вимірюваннях використовуються стаціонарні процеси у широкому сенсі або так звані *слабостаціонарні*. Аналіз цих процесів проводиться в рамках кореляційної (енергетичної) теорії і при цьому визначається дві перші моментні функції процесу, а саме, середнє значення – перший початковий момент, дисперсія – другий центральний момент, кореляційна функція і спектральна щільність потужності. Такі характеристики процесу мають чітку фізичну інтерпретацію, що природно для завдань вимірювань.

Для аналізу стаціонарних процесів у вузькому сенсі необхідно використання послідовності $n(n > 2)$ вимірних функцій розподілу, тому такі процеси в подальших матеріалах розглядатись не будуть.

Більш детально зупинимось на аналізі слабостаціонарних процесів.

Випадковий процес $\xi(\omega), t \in$ слабостаціонарним, якщо задана його двовимірна функція розподілу у виді:

$$F(x_1, x_2; t, t + \tau) = P\{\omega \in \Omega : \xi(\omega, t) < x_1 \text{ і } \xi(\omega, t + \tau) < x_2\} = F(x_1, x_2; \tau)$$

Розділ 2. Кількісні оцінки вимірювальної інформації

Відповідно для слабостационарного (далі просто стационарного) випадкового процесу $\xi(\omega, t)$ маємо:

- двовимірну щільність розподілу ймовірностей

$$p(x_1, x_2; \tau) = \frac{\partial^2 F(x_1, x_2; \tau)}{\partial x_1 \partial x_2};$$

- маргінальну – одновимірну функцію розподілу

$$F(x_1, \tau) = \int_{-\infty}^{\infty} F(x_1, x_2; \tau) dx_2 = \int_{-\infty}^{\infty} F(x_1, x_2; 0) dx_2 = F(x), x \in R;$$

- одновимірну щільність розподілу ймовірності $p(x) = dF(x)/dx$;

- математичне сподівання $M\xi(\omega, t) = \int_{-\infty}^{\infty} xp(x) dx = a_1$;

- дисперсію $D\xi(\omega, t) = \int_{-\infty}^{\infty} (x - a_1)^2 p(x) dx = \sigma^2$;

- автокореляційну функцію

$$\begin{aligned} R(\tau) &= M\{[\xi(\omega, t) - a_1][\xi(\omega, t + \tau) - a_1]\} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - a_1)(x_2 - a_1)p(x_1, x_2; \tau) dx_1 dx_2; \end{aligned}$$

- спектральну щільність потужності

$$G(f) = \int_{-\infty}^{\infty} e^{j2\pi f \tau} R(\tau) d\tau.$$

У відповідності з теоремою Вінера – Хінчина для парних функцій $R(\tau)$ і $G(f)$ маємо:

$$\begin{aligned} R(\tau) &= \int_0^{\infty} G(f) \cos 2\pi f \tau df \\ G(\tau) &= 4 \int_0^{\infty} R(\tau) \cos 2\pi f \tau df. \end{aligned}$$

Наведені вирази використовуються для аналізу енергетичних характеристик стационарних випадкових процесів.

Таким чином, для стационарного $\xi(\omega, t)$: одновимірні $F(x)$ і $p(x)$, математичне сподівання a_1 і дисперсія σ^2 не залежать від часу; двовимірні $F(x_1, x_2; \tau)$ і $p(x_1, x_2; \tau)$, автокореляційна функція $R(\tau)$

залежать тільки від взаємного зсуву $\tau = t_2 - t_1$, $\tau \in \mathbb{R}$. Спектральна щільність потужності $G(f)$ є одновимірною дійсною функцією частоти ω .

Розмірність наведених характеристик стаціонарного випадкового $\xi(\omega, t)$ є наступною:

- $F(x_1, x_2; \tau)$, $F(x)$ - безрозмірні;
- $p(x_1, x_2; \tau)$ - обернена квадрату одиниці виміру $\xi(\omega, t)$;
- $p(x)$ - обернена одиниці виміру $\xi(\omega, t)$;
- a_1 - одиниці виміру $\xi(\omega, t)$;
- σ^2 , $R(\tau)$ - квадрат одиниці виміру $\xi(\omega, t)$;
- $G(f)$ - одиниця виміру Вт/Гц

Наведенні характеристики стаціонарного випадкового процесу $\xi(\omega)$ використовуються при вимірюваннях характеристик вимірювань випадкових сигналів.

Наведемо приклад одного з відомих стаціонарних випадкових процесів, які використовуються при опису шумових сигналів, таких як тепловий, дробовий та інші флуктуаційні шуми в електронних ланках технічних систем, в тому числі ІВС.

Флуктуаційний шумовий сигнал (далі просто шумовий сигнал) в залежності від постановки завдань досліджень використовується або в якості інформаційного вимірювального сигналу (корисного сигналу), або в якості випадкового сигналу завади. Традиційно фізичний механізм формування шумового сигналу розглядають, як механізм формування суми значної кількості незалежних імпульсних випадкових сигналів, які виникають в різних місцях простору і в не прогнозовані моменти часу. У відповідності з граничною теоремою теорії ймовірності такий шумовий сигнал описується гауссовим (нормальним) стаціонарним процесом.

Одновимірна щільність імовірності стаціонарного гаусового процесу $\xi(\omega, t)$ має вид

$$p(x) = \left(1/\sqrt{2\pi\sigma^2}\right) \exp\left[-(x-a_1)^2\right] \quad (3.45)$$

з відповідними математичним сподіванням

$M\xi(\omega, t) = a_1$ і дисперсією $D\xi(\omega, t) = \sigma^2$.

Відповідні графіки одновимірних законів розподілу Гаусса стаціонарного випадкового



Карл Фрідріх Гаусс (Carl Friedrich)

Розділ 2. Кількісні оцінки вимірювальної інформації

**Gauss,
1777 - 1855),**

німецький математик, астроном і фізик. Учився (1795 - 1798) в Геттінгенському університеті, з 1807 р. - професор цього університету та директор астрономічної обсерваторії. Характерними особливостями досліджень Гаусса є надзвичайна різноманітність і органічний зв'язок в них між теоретичною й прикладною математикою. Роботи

Гаусса мали великий вплив на весь подальший розвиток вищої алгебри, теорії чисел і багатьох інших наук.

процесу $\xi(\omega, t)$, а саме одновимірні щільності розподілу $p(x)$ і функції розподілу $F(x)$ наведені на рисунку 3.21

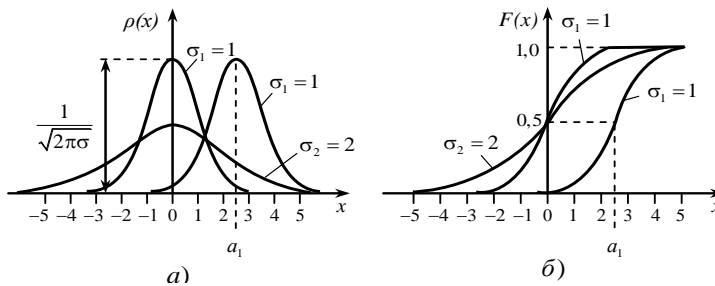


Рис. 3.21. Графіки одновимірних законів розподілу Гаусса:

а) - щільність ймовірності; б) - функція розподілу

Одновимірна функція розподілу гауссового стаціонарного випадкового процесу $\xi(\omega, t)$ згідно з (3.37) визначається так:

$$F(x) = \left(1/\sqrt{2\pi\sigma^2}\right) \int_{-\infty}^x \exp\left[-(x-a_1)^2/2\sigma^2\right] dx, \quad (3.46)$$

і після введення нової змінної $y = (x - a_1)/\sigma$ зводиться до виду

$$F(x) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{(x-a_1)/\sigma} \exp(-y^2/2) dy = 0,5 + \Phi_0\left[(x-a_1)/\sigma\right], \quad (3.47)$$

де $\Phi_0(z) = \left(1/\sqrt{2\pi}\right) \int_0^z \exp(-y^2/2) dy$.

Графік функції $F(x)$ (рис. 3.21, б) має вигляд монотонної зростаючої кривої від нуля до одиниці. Функція $\Phi_0(z)$, що входить у вираз (3.47), називається *інтегралом ймовірності* і її значення табульовані у математичних довідниках.

Спектральна щільність потужності шумового сигналу $G(f)$ визначається фізичною природою утворення сигналу. Відома ідеалізація такого шумового сигналу, так званого *білого шуму*, коли $G(f) = \sigma^2 \epsilon$ рівномірною для всього діапазону частот $f \in (0, \infty)$. Така ідеалізація є фізично необґрунтованою, тому що генератор білого шуму повинен мати нескінченну потужність. На практиці всі шумові сигнали є небілими – кольоровими, тому що їх спектральна щільність може бути апроксимована рівномірно постійною тільки в обмеженій (скінченній) смузі частот. Але при цьому модель ідеалізованого білого шуму є

Розділ 3. Сигнали та їх математичні моделі

загальною для всіх видів кольорових шумових процесів як модель породжуючого випадкового процесу.

Важливу роль при виконанні завдань вимірювань є обґрунтування конкретного виду моделі випадкових досліджуваних сигналів. Наведемо наступну таблицю (таблиця 3.1) видів випадкових функцій, їх відповідну класифікацію з метою подальшого визначення їх характеристик.

Види функцій Ознака класифікації	Види випадкових функцій (ВФ): випадкова величина (ВВ) $\xi(\omega)$, $\omega \in \Omega$; випадковий процес (ВП) $\xi(\omega, t)$, $\omega \in \Omega$, $t \in T$, де - T -інтервал часу
Кількість і вид аргументів	Одновимірні ВФ - ВВ; двовимірні ВФ - ВП
Вид законів розподілу	ВВ, ВП: гауссовий (нормальний), пуассонівський, Коші, рівномірно розподілений та інші (на сьогодні відомі більше 200 дискретних і неперервних законів розподілу)
Характер зміни значень	ВП: стаціонарні у широкому або вузькому сенсі, нестаціонарні, циклостаціонарні (періодично корельовані), з незалежними (некорельованими) приростами; з незалежними (некорельованими) значеннями.
Характер областей визначення і значень	ВВ: неперервна, дискретна, неперервно-дискретна; ВП: комбінації неперервності та дискретності аргументів із області визначення $\Omega \times X$ і $\Omega \times T$ і неперервності та дискретності області значень $X \subseteq R$

Коротко зупинимось на характеристиці, яка у значній мірі визначає ефективність результатів вимірювань значень і характеристик випадкових сигналів.

Співвідношення сигнал/завада. В залежності від постановки завдань вимірювання одні сигнали відносять до інформаційних (корисних), а інших до завад. У різних завданнях вимірювань це співвідношення є

числом, а в інших – функцією, у загальному випадку простору і часу. Розглянемо конкретний приклад. Визначимо співвідношення сигнал/завада для адитивної суміші детермінованого гармонічного сигналу $u(t)$ і стаціонарного шумового сигнал $\xi(\omega, t)$, тобто

$$\eta(\omega, t) = A \cos 2\pi ft + \xi(\omega, t), t \in T$$

Тоді в якості співвідношення сигнал/завада можна використати наступні вирази

$$\frac{c}{z} = \frac{A}{\sqrt{D\xi(\omega, t)}} = \frac{A}{\sigma} \quad \text{або} \quad \frac{c}{z} = \frac{A^2}{2\sigma^2}$$

Якщо говорити, про практику вимірювань, то можна відмітити наступне. Раніше (приблизно період до половини ХХ століття) завдання вимірювань умовно вирішувались при співвідношеннях сигнал/завада порядку 10, пізніше – порядку 3. Сучасні інформаційні вимірювальні технології дають можливість вирішувати завдання вимірювань при співвідношеннях менше 1 з використанням спеціальних методів опрацювання даних вимірювань.

3.3. Числові характеристики сигналів і завад

Енергетичні характеристики. Основними енергетичними характеристиками дійсного сигналу $s(t)$ є його потужність та енергія. Якщо $s(t)$ – напруга $u(t)$ чи струм $i(t)$, то *миттєва потужність*, що виділяється на опорі R , знаходиться через квадрат миттєвого значення: $p(t) = u^2(t)R = i^2(t)/R$. Одиниця виміру – Вт (Ват). У теорії сигналів, як правило, приймають для розрахунків $R = 1$ Ом (крім деяких випадків), і тоді в загальному вигляді: $p(t) = s^2(t)$ (3.48)

Прийняття такої умови пов'язано з тим, що в багатьох задачах теорії сигналів використовують у розрахунках не конкретні значення потужності, а відношення потужності сигналу до потужності завади. При розрахунках відношення опір R скорочується і для спрощення розрахунків його приймають за одиничний. Щоб відрізнити розрахунки потужності за таких умов ($R = 1$ Ом) від потужності на якомусь опорі $R \neq 1$ Ом, у формулі (3.41) та інших, куди входить потужність, приймається за одиницю виміру потужності V^2 (Вольти в квадраті), а не Вт.

Розділ 3. Сигнали та їх математичні моделі

Енергія сигналу на інтервалі (t_1, t_2) визначається як інтеграл його миттєвої потужності:

$$E_s = \int_{t_1}^{t_2} p(t) dt = \int_{t_1}^{t_2} s^2(t) dt. \quad (3.49)$$

Енергію сигналу можна обчислити також і за його спектральною щільністю за формулою Релея: ??

$$E_s = \frac{1}{2\pi} \int_{-\infty}^{\infty} |S(j\omega)|^2 d\omega = \int_0^{\infty} S^2(f) df. \quad (3.50)??$$

Величину $|S(j\omega)|^2$ називають *спектральною щільністю енергії*, або *енергетичним спектром* сигналу. З формули (3.50) випливає, що

$$S(f) = \sqrt{2}S(\omega), \quad ??? \quad (3.51)$$

оскільки $S(\omega)$ – двостороння спектральна щільність потужності $(-\infty < \omega < \infty)$, а $S(f)$ – одностороння $(0 \leq f \leq \infty)$.

Відношення

$$E_s / (t_2 - t_1) = (t_2 - t_1)^{-1} \int_{t_1}^{t_2} S^2(t) dt \quad (3.52)$$

має означення *середньої потужності* $P_s = \overline{s^2(t)}$ на інтервалі (t_1, t_2) .



**Норберт Вінер
(Norbert Wiener,
1894-1964),**

американський вчений – видатний математик і філософ, засновник кібернетики і теорії штучного інтелекту. Зацікавившись автоматичними розрахунками і теорією зворотнього зв'язку, сформулював у своїй фундаментальній праці "Кібернетика" (1948) положення цієї науки, предметом якої стали управління, зв'язок і опрацювання інформації в техніці, живих організмах і людському суспільстві. Автор робіт з математичного аналізу, теорії ймовірностей, електричних мереж і обчислювальної техніки.

Розрахунки середньої потужності за спектром. Середня потужність періодичного сигналу, яка розраховується на всій осі часу ($-\infty < t < \infty$), збігається з середньою потужністю за період. Для гармонічного сигналу $u(t) = U_m \cos(\omega t + \phi_0)$, згідно (3.50), середня потужність (на $R = 1$ Ом):

$$\overline{u^2(t)} = P_u = \left(U_m^2 / T \right) \int_0^T \cos^2(\omega t + \phi_0) dt = U_m^2 / 2 \quad (3.53)$$

і не залежить ні від частоти, ні від початкової фази.

Оскільки періодичний сигнал $s(t)$ можна подати у вигляді тригонометричного ряду Фур'є, а інтеграл суми дорівнює сумі інтегралів, то повна середня потужність періодичного сигналу дорівнює сумі середніх потужностей, що виділяються окремо постійною складовою $a_0/2$ і гармоніками з амплітудами A_{m1}, A_{m2}, \dots , при цьому вона не залежить від частот і фаз окремих гармонік.

Для випадкових сигналів (завад) середню потужність можна розрахувати за спектральною щільністю потужності $G_x(f)$ чи $G_x(\omega)$. Оскільки функції $G_x(f)$ та $G_x(\omega)$ показують розподіл потужності за частотами (див. (3.52)), то середня потужність

$$P_x = \int_0^{\infty} G_x(f) df = 2 \int_0^{\infty} G_x(\omega) d\omega. \quad (3.54)$$

У виразі (3.54), як і у формулі (4.50), $G_x(f) = 2G_x(\omega)$. Ця умова прийнята тому, що $G_x(\omega)$ - двостороння спектральна щільність потужності ($-\infty < \omega < \infty$), а $G_x(f)$ - одностороння ($0 \leq f < \infty$).

Щоб знайти, **наприклад**, потужність випадкового сигналу (завади) в деякій смузі частот $f_1 - f_2$, необхідно здійснити інтегрування згідно з виразом

$$(4.54) \text{ у цій смузі: } P_{x1,2} = \int_{f_1}^{f_2} G_x(f) df.$$

Приклад. Знайти середню потужність $P_{ш}$ білого шуму зі спектральною щільністю потужності $G_x(f) = N_0 = 10^{-6}$ Вт/Гц у смузі $\Delta f = 3\ 100$ Гц.

Середню потужність шуму в смузі Δf знаходимо згідно (3.54), якщо границі інтегрування вибрати від f_1 до $f_1 + \Delta f$:

$$P_{ш} = \int_{f_1}^{f_1 + \Delta f} N_0 df = N_0 \Delta f = 10^{-6} \cdot 3100 = 3,1 \cdot 10^{-3} \text{ Вт.}$$

Рівні сигналив (завад). Під *рівнем* розуміють відношення значення потужності P_x чи напруги U_x у деякій точці x електричного кола до значення потужності P_0 чи напруги U_0 , що вибрані для порівняння. Оскільки значення потужності та напруги можуть змінюватись у досить великих межах (сотні і тисячі разів), для вимірювання рівнів уведена логарифмічна одиниця рівня - децибел (дБ), що дорівнює $10 \lg(P_x/P_0)$ - за потужністю і $20 \lg(U_x/U_0)$ - за напругою. Як абсолютний нульовий рівень для порівняння, наприклад, у техніці зв'язку вибрана потужність $P_0 = 1$ мВт на опорі $R = 600$ Ом. Тоді $U_0 = \sqrt{P_0 R} = 0,7748 \approx 0,775$ В.

Децибелі, що визначені відносно потужності $P_0 = 1$ мВт, називаються *децибелами* відносно 1 мВт і скорочено позначаються дБп або дБ(мВт).

За логарифмічною одиницею виміру рівнів така характеристика якості, як відношення сигнал-завада, буде дорівнювати різниці рівнів сигналу L_s , і завади L_n , через те що $\rho = 10 \lg(P_s/P_n) = 10 \lg(P_s/P_0) - 10 \lg(P_n/P_0) = L_s - L_n$.

Динамічний діапазон і коефіцієнт амплітуди. Динамічний діапазон D_s , дБ сигналу $s(t)$, характеризує межі зміни миттєвої потужності і визначається виразом: $D_s = 10 \lg(p_{\max}/p_{\min})$, (3.55)

де p_{\max} - максимальне, а p_{\min} - мінімальне значення миттєвої потужності, які визначені будь-яким способом. Наприклад, мінімальна потужність, якщо її важко знайти, дорівнює потужності завади чи середньої квадратичної похибки.

Коефіцієнтом амплітуди сигналу K_A називається відношення його максимальної потужності до середньої. У логарифмічних одиницях, дБ: $K_A^2 = 10 \lg(p_{\max}/P_x)$. (3.56)

Розділ 3. Сигнали та їх математичні моделі

У деяких випадках динамічний діапазон і коефіцієнт амплітуди визначаються не в логарифмічних, а в абсолютних одиницях ("в разях").

Тривалість і ширина спектра сигналу (завади). Під тривалістю сигналу розуміють інтервал часу його існування. Обчислюється як різниця між часом закінчення сигналу t_k і часом його початку t_n : $T_s = t_k - t_n$.

Ширина спектра – це інтервал частот, який займає спектр. Обчислюється як різниця між максимальною частотою спектра f_{\max} і мінімальною f_{\min} : $F_s = f_{\max} - f_{\min}$. Розрахунки тривалості і ширини спектра сигналу (завади) не завдають труднощів, якщо вони мають чітко визначені початок та кінець, а спектр – граничні частоти. Але з перетворення Фур'є впливає, що коли сигнал має скінченну тривалість, то його спектр – нескінченний. І навпаки. Тому необхідно домовитись про визначення тривалості та ширини. На практиці використовуються різні визначення T_s та F_s ,

вибір яких залежить від призначення сигналу, його форми та структури. Застосовують такі методи визначення T_s і F_s :

1. Відлік на заданому значенні від максимального. Звичайно тривалість імпульсного сигналу і ширину його спектра визначають на рівні $1/\sqrt{2}$ від максимального значення відповідно $s(t)$ чи $S(f)$. Але можна вибрати для розрахунків будь-яке інше значення, наприклад, 5% від максимального, як це показано на рис. 3.22. У цьому невизначеність методу.

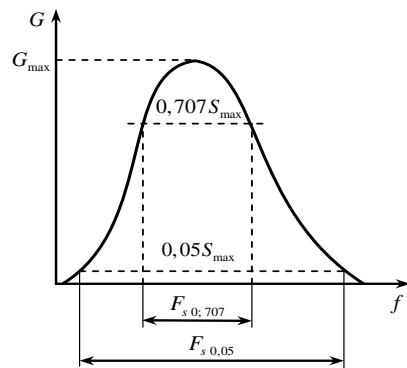


Рис. 3.22. Визначення ширини спектра на заданому рівні

Заміна реального сигналу (спектру) рівновеликим прямокутним. Така процедура частіше за інші застосовується для обчислення спектральної щільності потужності і наведена на рис. 3.23, де зображено спектральну щільність завади $G_s(f)$. Площі прямокутника і фігури, що обмежена кривою $G_s(f)$ і осями координат, рівновеликі. Ширина спектра,

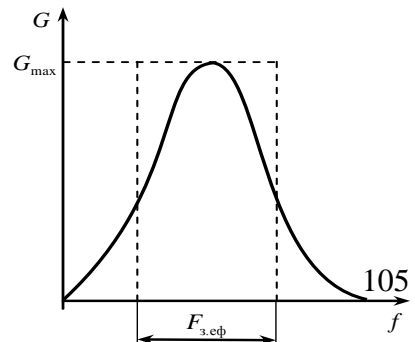


Рис. 3.23. Визначення ефективної ширини спектра

яку називають ефективною: $F_{з.еф} = \frac{1}{G_{\max}} \int_0^{\infty} G_z(f) df$.

Числові характеристики сигналів і завад широко використовуються в телекомунікаційних системах. За енергетичними характеристиками визначається необхідне відношення сигнал-завада, за шириною спектра сигналу встановлюється смуга пропускання каналу зв'язку, що необхідна для неспотвореного передавання. Для неперервних первинних сигналів ширина спектру визначається, як правило, експериментально. Для імпульсних сигналів при визначенні ширини їх спектру можна користуватись одним із найважливіших положень теорії сигналів і спектрів: якщо F_s , означає ширину спектру деякого сигналу тривалістю T_s , то завжди має місце співвідношення:

$$T_s F_s \approx \mu, \quad (3.57)$$

де μ - постійна величина, що близька до одиниці ($\mu \approx 1$) для відеоімпульсів і до двох ($\mu \approx 2$) для радіоімпульсів. Суть цього співвідношення в тому, що ширина спектра сигналу обернено пропорційна його тривалості.

3.4. Математичні моделі сигналів з обмеженим спектром

Всі розглянуті раніше сигнали належать до таких, які теоретично мають нескінченно широкий спектр. Це означає, що при спробі відновлення вихідного сигналу методом підсумовування його гармонік необхідно враховувати нескінченну множину спектральних компонентів. Утрата кожного з них, а тим більш деякої їх підмножини, супроводжується спотворенням форми сигналу. Спотворення буде тим більшим, чим більшу кількість гармонік утрачено при відновленні сигналу за його спектральними складовими. З фізичного погляду процедура відновлення сигналу, що ґрунтується на врахуванні всіх спектральних складових із нескінченно широкого спектра, нездійсненна. Не слід забувати також про те, що внесок, зроблений спектральними компонентами при $\omega \rightarrow \infty$, стає як завгодно малим згідно з властивостями самих сигналів, енергія яких скінченна. Крім того, будь-який реальний пристрій, призначений для передавання й обробки сигналів, має скінченну ширину смуги пропускання. Найбільш характерно це для пристроїв типу частотних фільтрів.

Ідеальний низькочастотний сигнал. Розглянемо особливий клас сигналів, спектральна щільність яких відмінна від нуля лише в межах деякого інтервалу частот скінченної довжини. **Приклад** таких сигналів - радіоімпульс із лінійною частотною модуляцією при значенні бази B , що прямує до нескінченності, має скінченну ширину спектра.

Нехай D — частотний інтервал, у межах якого спектральна щільність $\dot{U}(\omega)$ деякого сигналу $u(t)$ не дорівнює нулю, тобто $\dot{U}(\omega) \neq 0$, якщо

Розділ 3. Сигнали та їх математичні моделі

$\omega \in D$; $\dot{U}(\omega) = 0$, якщо $\omega \notin D$. У загальному вигляді математична модель сигналу з обмеженим спектром визначається формулою оберненого перетворення Фур'є: $u(t) = \frac{1}{2\pi} \int_D \dot{U}(\omega) e^{j\omega t} d\omega$. Залежно від вибору інтервалу D і функції

$\dot{U}(\omega)$ можна отримати різноманітні сигнали з обмеженим спектром. Розглянемо коливання, спектральна щільність якого стала і набуває дійсного значення в межах частотного інтервалу, обмеженого деякою верхньою частотою ω_b . Поза цим інтервалом спектральна щільність перетворюється на нуль:

$$\dot{U}(\omega) = \begin{cases} U_0, & \omega \in [-\omega_b, \omega_b]; \\ 0, & \omega \notin [-\omega_b, \omega_b]. \end{cases} \quad (3.58)$$

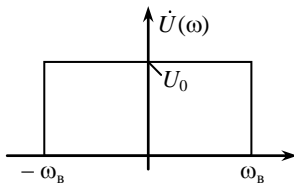


Рис. 3.24. Спектральна щільність ідеального низькочастотного сигналу

Графік спектральної щільності (3.58) зображено на рис. 3.24. Миттєве значення цього сигналу знаходимо за формулою оберненого перетворення Фур'є. Після інтегрування та елементарних перетворень дістанемо:

$$u(t) = \frac{U_0 \omega_b}{\pi} \cdot \frac{\sin \omega_b t}{\omega_b t}. \quad (3.59)$$

Це коливання називається *ідеальним низькочастотним сигналом* (ІНС), графік якого, побудований за формулою (3.59), має вигляд осцилюючої кривої, парної відносно початку відліку часу (рис. 3.25).

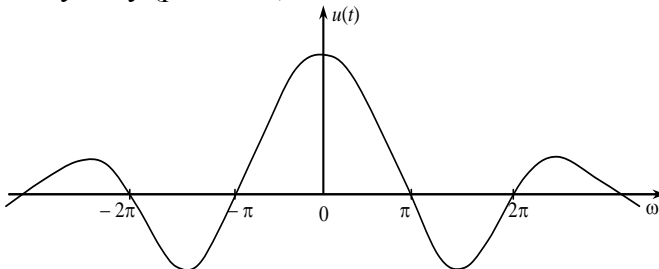


Рис. 3.25. Ідеальний низькочастотний сигнал

Зі збільшенням верхньої граничної частоти ω_b зростають як значення центрального максимуму, так і частота осциляцій. При $\omega_b \rightarrow \infty$, що прямує до нескінченності, сигнал $u(t)$ переходить у дельта-функцію, тобто $\lim_{\omega_b \rightarrow \infty} u(t) = \delta(t)$.



Олександр Михайлович Ляпунов (1857—1918),

російський математик і механік, академік Петербурської академії наук. Учень П.Л. Чебишева. Основні роботи присвячені теорії стійкості рівноваги і руху механічних систем, теорії фігур рівноваги рідини, яка рівномірно обертається, і математичній фізиці. Найважливішим досягненням є створення сучасної теорії стійкості рівноваги і руху механічних систем, визначених кінцевою кількістю параметрів. Отримав ряд вагомих результатів у теорії лінійних і нелінійних диференціальних рівнянь.

Розділ 3. Сигнали та їх математичні моделі

Ідеальний низькочастотний сигнал (ІНС) можна отримати, подавши на вхід ідеального фільтра нижніх частот (ФНЧ) сигнал $s(t)$ із рівномірною на всій осі частот спектральною щільністю (рис. 3.26). Як відомо, рівномірну в нескінченному інтервалі частот спектральну щільність має сигнал $s(t)$ типу дельта-функції $\delta(t)$ (рис. 3.27).

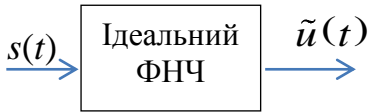


Рис. 3.26. Схема моделювання ІНС

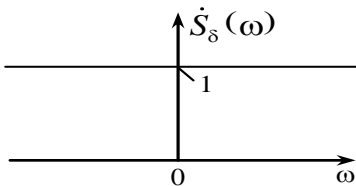


Рис. 3.27. Спектральна щільність дельта-функції

Для забезпечення на виході фільтра формування сигналу $\tilde{y}(t)$ з рівномірною в інтервалі частот $\omega \in [0, \omega_b]$ (для фізично реалізованого фільтра), частотна передавальна функція фільтра $K(\omega)$ має бути такою, як на рис. 3.28. Очевидно, що ідеальний ФНЧ (рис. 3.26) у разі подання на його вхід дельта-функції вирізає з її спектра (рис. 3.27) ділянку частот від 0 до ω_b . Сигналу $\tilde{y}(t)$ на виході фільтра відповідатиме спектральна щільність

$$\tilde{U}(\omega) = \begin{cases} U_0, & \omega \in [0, \omega_b]; \\ 0, & \omega \notin [0, \omega_b], \end{cases}$$

яка відрізняється від щільності (3.58)

відсутністю компонентів на від'ємних частотах. Зауважимо, що фізично реалізований ФНЧ формує відгук (вихідний сигнал) не раніше, ніж з моменту появи вхідного сигналу. Якщо короткий вхідний сигнал прямокутної форми (аналог дельта-імпульсу) з'являється в момент часу $t = 0$, то реакція ФНЧ на

такий сигнал матиме вигляд функції $\tilde{U}(t)$, зображеної на рис. 3.29.

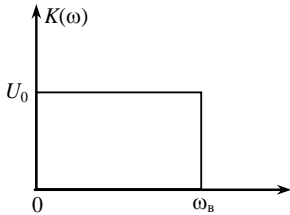


Рис. 3.28. Частотна передавальна функція ідеального ФНЧ

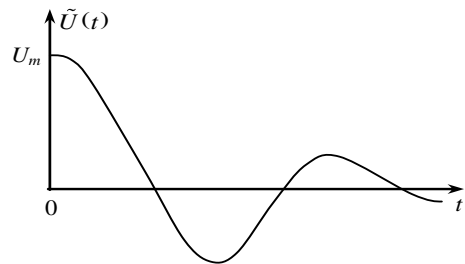


Рис. 3.29. Відгук ідеального ФНЧ на вхідний сигнал типу дельта-функції

Максимальне значення відгуку U_m в момент часу $t = 0$ визначається співвідношенням $U_m = U_0 \omega_b / \pi$.

Ідеальний смуговий сигнал. Побудова математичної моделі смугового сигналу базується на припущенні, що його спектр обмежений смугою частот шириною $\Pi = 2\Delta\omega$ з центром на частотах

$\pm\omega_0$. Якщо в межах цієї смуги спектральна щільність сигналу стала (рис. 3.30):

$$\dot{U}(\omega) = \begin{cases} U_0, & \left[\begin{array}{l} -\omega_0 - \Delta\omega < \omega < -\omega_0 + \Delta\omega \\ \omega_0 - \Delta\omega < \omega < \omega_0 + \Delta\omega \end{array} \right]; \\ 0, & \text{поза смугою пропускання,} \end{cases}$$

то за аналогією з ІНС його називають *ідеальним смуговим сигналом* (ІСС).

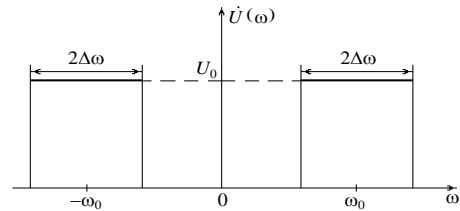


Рис. 3.30. Спектральна щільність ІСС

Миттєві значення ІСС можна знайти за формулою оберненого перетворення Фур'є. Спектральна щільність ІСС є парною функцією відносно початку осі частот і в загальному випадку

$$u(t) = \frac{1}{\pi} \int_0^{\infty} \dot{U}(\omega) \cos \omega t d\omega.$$

Оскільки $\dot{U}(\omega)$ задано як дійсну функцію, що дорівнює U_0 в межах

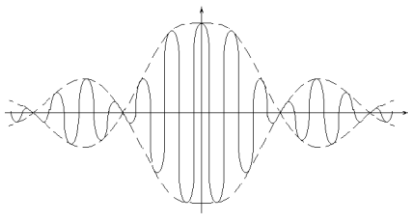


Рис. 3.31. Реалізація ідеального смугового сигналу

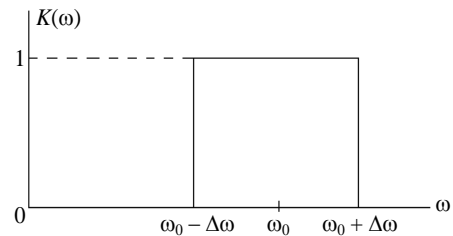


Рис. 3.32. Частотна характеристика ідеального смугового фільтра

смуги частот $\omega \in (\omega_0 - \Delta\omega, \omega_0 + \Delta\omega)$, то останній інтеграл можна записати у вигляді $u(t) = \frac{U_0}{\pi} \int_{\omega_0 - \Delta\omega}^{\omega_0 + \Delta\omega} \cos \omega t d\omega$. Після інтегрування та

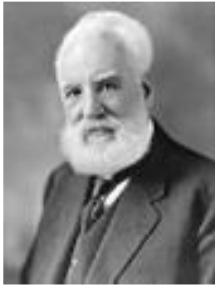
елементарного перетворення отримаємо:

$$u(t) = (2U_0\Delta\omega/\pi) \cdot (\sin \Delta\omega t / \Delta\omega t) \cos \omega_0 t.$$

Графік на рис. 3.31, демонструє структуру ІСС. Функція $\sin(\Delta\omega t)/(\Delta\omega t)$ з точністю до масштабного коефіцієнта $2U_0\Delta\omega/\pi$ описує закон зміни обвідної ІСС. Спосіб утворення ІСС цілком очевидний: на вхід ідеального смугового фільтра (рис. 3.32), що пропускає коливання з частотами в межах смуг $(\omega_0 - \Delta\omega, \omega_0 + \Delta\omega)$, потрібно подати широкосмуговий вплив типу дельта-імпульсу. На виході ідеального смугового фільтра буде спостерігатися сигнал, осцилограма якого відповідає правій половині графіка, наведеного на рис. 3.31. Це впливає з умови фізичної

Розділ 3. Сигнали та їх математичні моделі

реалізованості сигналів: відгук кола (ланцюга, системи тощо) починає формуватися не раніше появи впливу на вході.



**Олександр Грехем
Бел
(Alexander Graham
Bell,
1847-1922),**

американський фізик шотландського походження, творець телефонного апарату з металічною мембраною, один із винахідників телефону. В 1865 р. задумав передати мову електричними хвилями. Втілення ідеї зайняло наступні 10 років. В 1876 р. перше повідомлення було успішно передано по дротам. В 1877 р. була створена телефонна компанія Bell. Виконував роботи по використанню в телекомунікаціях світлового променя – напрям, який з часом привів до створення волоконно-оптичних технологій.

Базис Котельникова – у загальному випадку утворює нескінченну сукупність ортогональних функцій, норма кожної з яких дорівнює одиниці.

Нагадаємо, що ортогональними називаються такі сигнали $u(t)$ і $v(t)$, скалярний добуток яких дорівнює

$$\text{нулю: } (u, v) = \int_{-\infty}^{\infty} u(t)v(t)dt = 0.$$

Якщо до того ж енергія кожного сигналу дорівнює одиниці, то сигнали $u(t)$ і $v(t)$ називаються ортонормованими. Для сигналу $u(t)$ його енергія E_u

$$\text{визначається співвідношенням: } E_u = (u, u) = \int_{-\infty}^{\infty} u^2(t) dt.$$

Відповідно узагальненої формули Релея сигнали $u(t)$ і $v(t)$ будуть ортогональними, якщо виконується умова

$$(u, v) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{U}(\omega) \dot{V}^*(\omega) d\omega = 0. \quad (3.60)$$

Обмеження, що накладаються на смугу частот сигналу, дають змогу знаходити цікаві й важливі класи ортогональних сигналів. Як простий приклад ортогональних сигналів можна навести таку пару смугових сигналів, спектри яких не перетинаються. Рівність нулю скалярного добутку цих сигналів безпосередньо впливає з (3.60).

Менш очевидний спосіб ортогоналізації сигналів з обмеженим спектром полягає в їх зсуві в часі.

Розглянемо два ІНС $u(t)$ і $v(t)$. Обидва вони мають однакові параметри U_0 і ω_b , але відрізняються тим, що сигнал $v(t)$ запізнюється відносно сигналу $u(t)$ на час t_0 , тобто

$$v(t) = u(t - t_0). \quad (3.61)$$

Спектральна щільність сигналу $v(t)$ визначиться співвідношенням $\dot{V}(\omega) = \dot{U}(\omega) e^{-j\omega t_0}$, або:

$$\dot{V}(\omega) = U_0 e^{-j\omega t_0}, \quad \omega \in [-\omega_b, \omega_b]. \quad (3.62)$$

Припустимо, що спектральну щільність сигналу $u(t)$ задано співвідношенням (4.58). Підставивши вирази (3.58) і (3.62) у (3.60), дістанемо формулу для скалярного добутку цих сигналів

$$(u, v) = \frac{U_0^2}{2\pi} \int_{-\omega_b}^{\omega_b} e^{j\omega t_0} d\omega. \quad (3.63)$$

У результаті інтегрування отримаємо:

$$(u, v) = (U_0^2 \omega_b / \pi) (\sin(\omega_b t_0) / \omega_b t_0). \quad (3.64)$$

Зі співвідношення (3.64) випливає, що два однакові за формою ІНС виявляються ортогональними, якщо зсув між ними в часі задовольняє умову:

$$\omega_b t_0 = k\pi, \quad k = \pm 1, \pm 2, \dots \quad (3.65)$$

Мінімально можливий зсув Δt (назвемо його кроком часового зсуву), що приводить до ортогоналізації, маємо при $k = \pm 1$, тобто:

$$\Delta t = \pm \pi / \omega_b = \pm 1 / 2F_b, \quad (3.66)$$

де через F_b позначено верхню граничну частоту зрізу (у герцах) ідеального ФНЧ, що відповідає верхній гармоніці коливань у ІНС.

У співвідношенні (3.66) узято до уваги, що $\omega_b = 2\pi F_b$.

Принципово важливо, що умовою (3.65) вдається не тільки домогтися ортогоналізації двох ІНС, а й побудувати ортогональний базис для сигналів, у спектрі яких відсутні частоти, вищі за ω_b .

Покажемо це на **прикладі** ІНС. Скориставшись співвідношеннями (3.61), (4.65) і (3.66), утворимо сукупність сигналів

$$v_k(t) = u(t - k\Delta t) = u(t - k\pi / \omega_b). \quad (3.67)$$

Сигналам $v_k(t)$ відповідає спектральна щільність:

$$\dot{V}_k(\omega) = U_0 e^{-j\omega k\pi / \omega_b}, \quad \omega \in [-\omega_b, \omega_b]. \quad (3.68)$$

При $k=0$ спектральна щільність $\dot{V}_0(\omega)$ сигналу $v_0(t)$ збігається зі спектральною щільністю (3.58) ідеального низькочастотного сигналу $u(t)$, часову функцію якого задано співвідношенням (3.59). Покажемо, що сукупність $v_k(t)$ сигналів (3.68) утворює зчисленну множину ортогональних функцій. З цією метою достатньо переконатися в тому, що взаємна енергія E_l сигналів $v_k(t)$ і $v_{k+l}(t)$, рознесених на l інтервалів часу Δt , дорівнює нулю.

Справді, взаємна енергія E_l двох сигналів чисельно дорівнює скалярному добутку цих сигналів: $E_l = (v_k, v_{k+l})$.

У свою чергу, скалярний добуток можна визначити узагальненою формулою Релея, тобто

$$E_l = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{V}_k(\omega) \dot{V}_{k+l}^*(\omega) d\omega.$$

Ця рівність з урахуванням виразу (3.68) зводиться до вигляду:

$$E_l = \frac{U_0^2}{2\pi} \int_{-\omega_b}^{\omega_b} e^{j\omega l\pi / \omega_b} d\omega = \frac{U_0^2}{\pi} \int_0^{\omega_b} \cos(\omega l\pi / \omega_b) d\omega. \quad (3.69)$$

Обчислюючи інтеграл (3.69), отримаємо:

$$E_l = \frac{U_0^2 \omega_b}{\pi} \frac{\sin(l\pi)}{l\pi}, \quad l = 0, 1, \dots \quad (3.70)$$

Розділ 3. Сигнали та їх математичні моделі

Отже,

$$E_l = \begin{cases} U_0^2 \omega_b / \pi, & l = 0; \\ 0, & l \neq 0. \end{cases} \quad (3.71)$$

Із системи (3.71) випливає, що взаємна енергія E_l сигналів $v_k(t)$, заданих співвідношенням (3.67) і рознесених на l інтервалів Δt , що визначаються за формулою (3.66), дорівнює нулю, тобто функції $v_k(t)$ утворюють ансамбль ортогональних функцій. До аналітичної форми ансамблю ортогональних функцій $v_k(t)$ прийдемо, скориставшись виразами (3.58) і (3.67). Отримаємо:

$$v_k(t) = \frac{U_0 \omega_b}{\pi} \frac{\sin \omega_b (t - k\pi / \omega_b)}{\omega_b (t - k\pi / \omega_b)}, \quad k = 0, \pm 1, \pm 2, \dots \quad (3.72)$$

Таким чином, ансамбль ортогональних функцій $v_k(t)$ утворено за рахунок часового зсуву ідеального низькочастотного сигналу (3.59) на k інтервалів $\Delta t = \pi / \omega_b$. На рис. 3.33 наведено графіки сигналів (3.72) для $k = 0$ і $k = 2$.

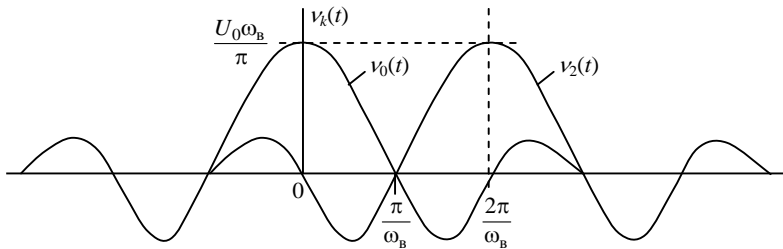


Рис. 3.33. Графіки ортогональних ІНС

Побудуємо на основі функцій $v_k(t)$, заданих співвідношенням (4.72), систему ортонормованих функцій (базис), які позначимо через $\varphi_k(t)$. Для цього необхідно вимагати, щоб енергія кожної функції $v_k(t)$ дорівнювала одиниці. Згідно зі співвідношенням (3.71) енергія E кожного сигналу (4.72):

$$E_0 = U_0^2 \omega_b / \pi. \quad (3.73)$$

Прирівнявши праву частину (3.73) до 1, дістанемо, що система ортогональних функцій $v_k(t)$ стає нормованою за умови: $U_0 = \sqrt{\pi / \omega_b}$. (3.74)

Підставивши значення (3.74) у формулу (3.72), приходимо до системи ортонормованих функцій:

$$\varphi_k(t) = \sqrt{\frac{\omega_b}{\pi}} \cdot \frac{\sin \omega_b (t - k\pi / \omega_b)}{\omega_b (t - k\pi / \omega_b)}, \quad k = 0, \pm 1, \pm 2, \dots, \quad (3.75)$$

яка утворює так званий *базис Котельникова* в просторі ІНС із частотами, обмеженими зверху величиною ω_b . Окрема функція $\varphi_k(t)$ називається *k-ю відліковою функцією*.

Таким чином, *базис Котельникова є сукупністю ортонормованих функцій $\varphi_k(t)$, утворених з ідеального низькочастотного сигналу (з верхньою граничною частотою ω_b і спектральною щільністю в смузі*

частот $[-\omega_b, \omega_b]$, що дорівнює $\sqrt{\pi/\omega_b}$) за рахунок його часового зсуву на проміжок часу $t_k = k\Delta t = k\pi/\omega$, $k = 0, \pm 1, \pm 2, \dots$.

Теорема Котельникова (теорема відліків). Теорема, яку довів В. О. Котельников у 1933 році, є одним із фундаментальних положень теоретичної радіотехніки. Теорема встановлює можливість як завгодно точного відновлення сигналу з обмеженим спектром за його дискретними значеннями, узяними через однакові проміжки часу.

Нехай $s(t)$ – довільний сигнал, спектральна щільність якого відмінна від нуля лише в інтервалі частот $-\omega_b < \omega < \omega_b$. Його можна розкласти в узагальнений ряд Фур'є за базисом Котельникова: $s(t) = \sum_{k=-\infty}^{\infty} c_k \phi_k(t)$. (3.76)

Коефіцієнтами c_k ряду (4.76) є скалярні добутки сигналу $s(t)$ і k -ої відлікової функції: $c_k = (s, \phi_k)$. Обчислення цих коефіцієнтів полягає в застосуванні узагальненої формули Релея, тобто $c_k = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}(\omega) \dot{\Phi}_k^*(\omega) d\omega$, (3.77)

де $\dot{S}(\omega)$ — спектральна щільність сигналу $s(t)$; $\dot{\Phi}_k(\omega)$ — спектральна щільність k -ої відлікової функції базису Котельникова $\phi_k(t)$, тобто $\dot{\Phi}_k(\omega) = F\{\phi_k(t)\}$, де F — оператор прямого перетворення Фур'є.

Приступимо до обчислення спектральної щільності $\dot{\Phi}_k(\omega)$. З порівняння виразів (3.74) і (3.77) випливає: $\phi_k(t) = \sqrt{\pi/\omega_b} v_k(t) \Big|_{U_0=1}$. (3.78)

Таким чином, згідно з рівністю (3.78) базисна функція Котельникова $\phi_k(t)$ з точністю до коефіцієнта $\sqrt{\pi/\omega_b}$ збігається з (3.75) для функції $v_k(t)$, спектральну щільність якої задано співвідношенням (4.67). Отже,

$$\dot{\Phi}_k(\omega) = \sqrt{\pi/\omega_b} e^{-j\omega k \frac{\pi}{\omega_b}}, \quad (3.79)$$

причому в (3.79) враховано умову (3.78), відповідно до якої $U_0 \equiv 1$. Підставивши значення (3.79) у (3.77), дістанемо

$$c_k = \sqrt{\frac{\pi}{\omega_b}} \left\{ \frac{1}{2\pi} \int_{-\omega_b}^{\omega_b} \dot{S}(\omega) e^{j\omega \frac{k\pi}{\omega_b}} d\omega \right\}. \quad (3.80)$$

Вираз у фігурних дужках правої частини формули (3.80) є не що інше, як миттєве значення сигналу $s(t_k) = s_k$ в k -ій відліковій точці: $t_k = k\pi/\omega_b = k/2F_b$. Таким чином, $c_k = \sqrt{\pi/\omega_b} s_k$, звідки випливає остаточна форма ряду

Розділ 3. Сигнали та їх математичні моделі

Котельникова:

$$s(t) = \sum_{k=-\infty}^{\infty} s_k \frac{\sin \omega_B (t - k\pi/\omega_B)}{\omega_B (t - k\pi/\omega_B)}.$$

(3.81)

Формула (3.81) є суттю теореми Котельникова (теореми відліків): довільний сигнал, спектр якого не містить частот, вищих за F_B , можна подати послідовністю дискретних відліків цього сигналу, узятих через однакові проміжки часу $1/(2F_B)$.

Позначивши розглянутий проміжок часу добору дискретних відліків сигналу (назвемо його періодом дискретизації) T , тобто взявши

$$T = 1/2F_B \tag{3.82}$$

і ввівши поняття частоти дискретизації $F = 1/T$, можна сформулювати **теорему Котельникова** в такий спосіб: для неспотвореного подання сигналу з обмеженим спектром послідовністю його дискретних відліків частота дискретизації F має дорівнювати подвоєній верхній частоті спектра сигналу F_B , тобто: $F = 2F_B$.

Особливість теореми Котельникова полягає в її конструктивному характері. Ця теорема не лише вказує на можливість розкладання сигналу у відповідний ряд, а й визначає спосіб відновлення неперервного сигналу, заданого своїми відліковими значеннями (рис. 3.34).

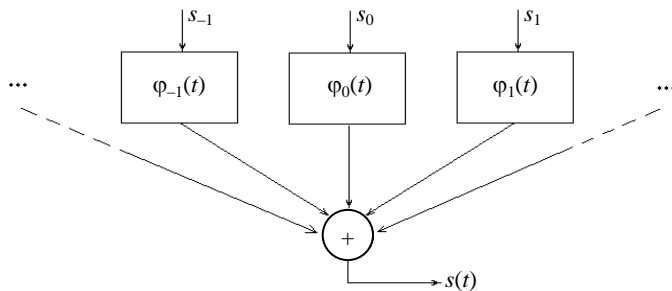


Рис. 3.34. Алгоритм синтезу сигналу

Нехай маємо сукупність генераторів, що створюють на своїх виходах відлікові функції: $\varphi_k(t) = \sin \omega_B (t - k\pi/\omega_B) / [\omega_B (t - k\pi/\omega_B)]$.

(3.83)

Генератори управляються – амплітуда їх сигналів пропорційна до відлікових значень s_k . Якщо тепер об'єднати коливання на виходах, подавши їх на суматор, то на виході суматора згідно формули (3.81) з'явиться миттєве значення синтезованого сигналу:

$$s(t) = \sum_{k=-\infty}^{\infty} s_k \varphi_k(t). \tag{3.84}$$

Фізично відновлення сигналу з обмеженим спектром з послідовності його дискретних відліків реалізується за допомогою ідеального ФНЧ, частота зрізу якого вибирається рівною верхній частоті спектра сигналу.

Припустимо, що дискретні відліки сигналу формуються, як зображено на рис. 3.35. Послідовність s_k утворюється в результаті перемножування сигналу $s(t)$ і затриманої на k періодів дискретизації дельта-функції, тобто $s_k = s(t)\delta(t - kT)$, причому T вибирається з умови (3.82). Реакцією ідеального ФНЧ на вхідний вплив типу дельта-імпульсу $\delta(t - kT)$ є функція вигляду (3.83) або відлікова функція ряду Котельникова. З урахуванням того, що ФНЧ є інтегрувальною ланкою (суматором), неважко зрозуміти, що цей ФНЧ саме й реалізує формулу (3.84), тобто відновлює сигнал $s(t)$.

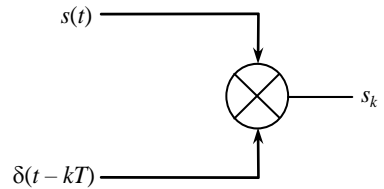


Рис. 3.35. Формування дискретних відліків сигналу

Таким чином, для відновлення сигналу $s(t)$, що має обмежений спектр F_b , за заданою послідовністю його дискретних відліків s_k , узятих із частотою дискретизації $F = 2F_b$, необхідно пропустити цю послідовність через ідеальний ФНЧ, частота зрізу якого вибирається такою, що дорівнює верхній частоті спектра сигналу F_b .

Апроксимація прямокутного сигналу рядом Котельникова. Ряд Котельникова часто використовують для наближеного описання сигналів із необмеженим спектром, значну частину енергії якого зосереджено в низькочастотній області.

Як **приклад** таких сигналів розглянемо прямокутний відеоімпульс. Прямокутний відеоімпульс не належить до сигналів з обмеженим спектром, проте модуль його спектральної щільності доволі швидко (за законом $1/\omega$) зменшується зі зростанням частоти. Щоб спростити подання прямокутного відеоімпульсу рядом Котельникова, розмістимо імпульс на осі часу так, як це показано на рис. 3.36. Математична модель $s(t)$ такого сигналу доволі проста:

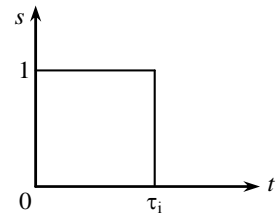


Рис. 3.36. Прямокутний відеоімпульс

$$s(t) = \begin{cases} 1, & t \in [0, \tau_1]; \\ 0, & t \notin [0, \tau_1]. \end{cases} \quad (3.85)$$

Оскільки згідно із системою (3.85) в області від'ємних значень часу сигнал відсутній, подамо ряд Котельникова такого сигналу (див. (3.81)) виразом

$$\tilde{s}_n(t) = \sum_{k=0}^{n-1} s_k \frac{\sin \omega_b (t - k\pi/\omega_b)}{\omega_b (t - k\pi/\omega_b)}, \quad (3.86)$$

в якому n - кількість відліків s_k прямокутного відеоімпульсу, які еквідистантно добираються на всьому інтервалі його існування від нуля до τ_1 .

Розділ 3. Сигнали та їх математичні моделі

Оскільки за умовою (3.85) усі відліки s_k імпульсу в межах його тривалості τ , дорівнюють одиниці, ряду (3.86) можна надати вигляду:

$$\tilde{s}_n(t) = \sum_{k=0}^{n-1} \frac{\sin \omega_b (t - k\pi/\omega_b)}{\omega_b (t - k\pi/\omega_b)}, \quad (3.87)$$

де залишилося ще недовизначеним значення верхньої частоти ω_b сигналу $\tilde{s}(t)$, що апроксимує функцію $s(t)$.

Нехай задано кількість n відліків прямокутного відеоімпульсу, які еквідистантно добираються по всій області існування $t \in (0, \tau_i)$. Очевидно, що період дискретизації:

$$T = \tau_i / n - 1. \quad (3.88)$$

Це означає, що коли $n = 2$, період дискретизації дорівнює тривалості імпульсу τ_i , а два відліки сигналу s_0 і s_1 беруться відповідно на початку і наприкінці імпульсу. При $n = 3$ відліки $0, s_1$ і s_2 беруться з інтервалом $T = \tau_i / 2$ відповідно на початку, у середині і наприкінці імпульсу. І, нарешті, при $n = 5$ відлік s_0 береться на самому початку імпульсу, а всі наступні — через чверть тривалості імпульсу, тобто $T = \tau_i / 4$ і т. д. Відповідно до виразу (3.66) ІНС ортогоналізується при зсуві на проміжок часу $\Delta t = T$, якщо його верхня частота ω_b задовольняє умову:

$$\omega_b = \pi / T. \quad (3.89)$$

Підставивши вираз (3.88) у формулу (3.89), дістанемо значення верхньої частоти ω_b спектра сигналу $\tilde{s}(t)$ за умови, що прямокутний відеоімпульс подається послідовністю n еквідистантно розміщених по всій тривалості τ_i відліків імпульсу $s(t)$, тобто:

$$\omega_b = \pi(n-1) / \tau_i. \quad (3.90)$$

На підставі співвідношень (3.87) і (3.90) приходимо до остаточної формули ряду Котельникова, що апроксимує прямокутний відеоімпульс:

$$\tilde{s}_n(t) = \sum_{k=0}^{n-1} \frac{\sin(\pi(n-1)/\tau_i)(t - k\tau_i/(n-1))}{(\pi(n-1)/\tau_i)(t - k\tau_i/(n-1))}. \quad (3.91)$$

Тобто, якщо подати прямокутний відеоімпульс усього двома його відліками, узятими на початку і в кінці імпульсу, то в спектрі цього імпульсу буде враховано складові, обмежені частотою $\omega_b = \pi / \tau_i$. За формулою (3.91) знаходимо наближений вираз математичної моделі сигналу апроксимації:

$$\tilde{s}_2(t) = \left(\frac{\sin \pi t / \tau_i}{(\pi t / \tau_i)} \right) + \left(\frac{\sin \pi / \tau_i (t - \tau_i)}{(\pi / \tau_i (t - \tau_i))} \right).$$

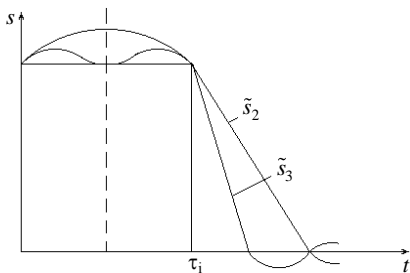


Рис. 3.37. Апроксимація прямокутного відеоімпульсу рядами Котельникова

А якщо подати цей імпульс трьома рівновіддаленими відліками (тобто взяти

$n = 3$), то, як впливає зі співвідношення (3.90), у спектрі сигналу буде враховано всі частоти, аж до $\omega_b = 2\pi / \tau_i$, і тому:

$$\tilde{s}_3(t) = \frac{\sin(2\pi t / \tau_i)}{2\pi t / \tau_i} + \frac{\sin(2\pi t / \tau_i)(t - \tau_i / 2)}{(2\pi t / \tau_i)(t - \tau_i / 2)} + \frac{\sin(2\pi / \tau_i)(t - \tau_i)}{(2\pi / \tau_i)(t - \tau_i)}.$$

Відповідні графіки зображено на рис. 3.37.

Природно, що зі зростанням числа n , тобто зі зменшенням часового інтервалу T між вибірковими відліками, точність апроксимації підвищуватиметься.

Дискретизація смугових сигналів. Попередній аналіз за замовчуванням базувався на тому припущенні, що аналогові сигнали (імпульси) з необмеженим спектром належать до класу так званих низькочастотних сигналів, для яких велика частина енергії сигналу зосереджується в низькочастотній області. Існує й інший клас сигналів, енергія яких зосереджується у високочастотній області і практично відсутня в області низьких частот. Назвемо такий клас сигналів *високочастотними*. Із цього класу сигналів широкого застосування набули так звані смугові сигнали, прикладом яких є модульовані сигнали.

Дамо деякі рекомендації щодо алгоритмів дискретизації аналогових смугових сигналів та їх відновлення з послідовності дискретних відліків.

Отже, нехай маємо деякий сигнал $s(t)$, спектр якого міститься в інтервалі $(\omega_0 - \Delta\omega, \omega_0 + \Delta\omega)$, в якому ω_0 — центральна несуча частота, а $\Delta\omega$ — девіація частоти. Вважатимемо, що поза зазначеним інтервалом гармонік спектра смугового сигналу немає. Для випадку амплітудної модуляції $\Delta\omega$ являє собою верхню гармоніку ω_a спектра модульовального сигналу. У разі кутової модуляції $\Delta\omega$ — це безпосередньо девіація частоти ФМ або ЧМ сигналів.

Теорема Котельникова і для таких сигналів дає змогу вибирати значення частоти дискретизації F смугового сигналу $s(t)$. Верхній кутовій частоті $\omega_b = \omega_0 + \Delta\omega$ сигналу $s(t)$ відповідає верхня циклічна частота F_a , пов'язана з ω_b виразом $\omega_b = 2\pi F_b$. Частота дискретизації згідно з теоремою Котельникова має задовольняти умову $F \geq 2F_b$.

Для відновлення аналогового смугового сигналу $s(t)$ послідовність дискретних відліків $\{s_k\}$ потрібно подати на вхід ідеального смугового фільтра з частотами зрізу: верхньою — ω_b , яку визначено раніше, і нижньою — $\omega_n = \omega_0 - \Delta\omega$. Відгук такого фільтра і буде доволі прийнятною апроксимацією вихідного смугового сигналу.

Основні висновки

Для теоретичного дослідження сигналів необхідно побудувати їхні математичні моделі. Математична модель сигналу є функціональною залежністю, що адекватно описує зміну в часі фізичного стану деякого об'єкта.

Класифікація сигналів виконується на підставі істотних ознак відповідних математичних моделей. Розрізняють детерміновані і випадкові, дійсні і комплексні, одновимірні і багатовимірні, статичні і динамічні, нескінченні в часі та імпульсні, періодичні та поодинокі, аналогові і дискретні сигнали. Різновидом останніх є цифрові сигнали - дискретні не тільки в часі, а й у просторі (за величиною).

До найпростіших типових елементарних сигналів належать: гармонічні та комплексно-експоненціальні сигнали, прямокутні та трикутні відеоімпульси, східчасті, сигнум- і дельта-функції тощо.

Скалярний добуток двох сигналів називається взаємною енергією цих сигналів. Повна енергія сигналу визначається скалярним добутком сигналу самого на себе. Іншими словами, повна енергія сигналу дорівнює інтегралу від квадрата функції, що описує математичну модель сигналу. Корінь квадратний з енергії сигналу називається нормою сигналу. Нормованими сигналами (функціями) називаються сигнали (функції), повна енергія яких дорівнює одиниці.

Два сигнали називаються ортогональними, якщо їхній скалярний добуток дорівнює нулю. Нескінченна система дійсних функцій (сигналів) називається ортогональною, якщо скалярний добуток двох різних функцій (сигналів) дорівнює нулю. При цьому передбачається, що енергія кожного сигналу із системи не дорівнює нулю. Це означає, що жодна з функцій, які розглядаються, не дорівнює нулю.

Нескінченну систему функцій, попарно ортогональних одна одній і таких, що мають одиничні норми, називають системою ортонормованих функцій або ортонормованим базисом.

Будь-який сигнал зі скінченною енергією можна подати у вигляді узагальненого ряду Фур'є. Узагальненим рядом Фур'є називається нескінченна сума добутків базисних функцій і відповідних їм коефіцієнтів. Сукупність коефіцієнтів розкладу утворює спектр сигналу.

Коефіцієнти розкладу узагальненого ряду Фур'є (спектр) часової функції (сигналу) визначаються скалярним добутком цього сигналу з відповідними базисними функціями. Такий спосіб визначення коефіцієнтів розкладу забезпечує мінімум квадратичної помилки апроксимації сигналу скінченновимірним рядом Фур'є.

Енергія сигналу дорівнює сумі енергій усіх компонентів (гармонік), з яких складається узагальнений ряд Фур'є. Це означає, що енергія сигналу

дорівнює сумі енергій спектральних складових, а квадрат модуля коефіцієнтів узагальненого ряду Фур'є чисельно дорівнює тій частці енергії сигналу, що міститься у відповідній складовій (гармоніці) сигналу.

Процес добування корисної інформації, що міститься в сигналі, можна подати як апаратне (чи програмне) визначення числових значень коефіцієнтів узагальненого ряду Фур'є цього сигналу.

Всі випадкові сигнали та завади є непередбачуваними. Для випадкових сигналів неможливо знайти математичну формулу, за якою можна було б розрахувати їх миттєві значення. Всі випадкові явища, що вивчаються в теорії ймовірностей, можна розділити на три типи: випадкові події; випадкові величини; випадкові процеси. Кожний з цих типів випадкових явищ має свої особливості і характеристики.

Випадкова подія - це будь-який факт, який в результаті спроби може з'явитись чи не з'явитись. Величина, значення якої змінюється від спроби до спроби випадковим чином, називається випадковою. Для такої величини неможливо певно передбачити, якого значення вона набуде за конкретних умов спроби. Функція розподілу ймовірності показує ймовірність того, що випадкова величина не перевищує конкретного значення. Математичне сподівання є середнім значенням випадкової величини. Дисперсія кількісно характеризує міру розкиду результатів окремих спроб відносно середнього значення.

Випадкові процеси бувають різних типів: нестационарні, стаціонарні, квазістаціонарні, ергодичні. Але в техніці більшість випадкових сигналів і завад відносяться до стаціонарних ергодичних випадкових процесів

Флуктуаційний шум найбільш характерний для більшості телекомунікаційних каналів і є стаціонарним ергодичним випадковим процесом із гаусовим (нормальним) розподілом імовірності. Спектральна щільність потужності флуктуаційного шуму залежить від фізичної природи його утворення, а також від точки, де він спостерігається.

Основними енергетичними характеристиками дійсного сигналу є його потужність та енергія. Для випадкових сигналів (завад) середню потужність можна розрахувати за спектральною щільністю потужності

За логарифмічною одиницею виміру рівнів така характеристика якості, як відношення сигнал-завада, буде дорівнювати різниці рівнів сигналу і завади.

Динамічний діапазон сигналу характеризує межі зміни миттєвої потужності. Коефіцієнтом амплітуди сигналу називається відношення його максимальної потужності до середньої.

Під тривалістю сигналу розуміють інтервал часу його існування. Обчислюється як різниця між часом закінчення сигналу і часом його початку. За енергетичними характеристиками визначається необхідне відношення сигнал-завада, за шириною спектра сигналу встановлюється смуга пропускання каналу зв'язку, як необхідна для неспотвореного передавання інформації.

Розділ 3. Сигнали та їх математичні моделі

Теоретично будь-які сигнали скінченної протяжності в часі мають нескінченно широкий спектр. Реальні технічні пристрої мають скінченну ширину смуги пропускання. При проходженні сигналів із нескінченно широким спектром через технічні пристрої відбувається спотворення спектра, що неминуче призводить до спотворення форми вхідного сигналу.

Для неспотвореного подання сигналу з обмеженим спектром послідовністю його дискретних відліків частота дискретизації F має дорівнювати подвоєній верхній частоті спектра сигналу.

Питання для самоконтролю

1. *Зобразіть основні просторово-часові моделі сигналу.*
2. *Які з двох типів сигналів — детерміновані чи випадкові — є носіями повідомлень?*
3. *У який спосіб гармонічне коливання можна перетворити в пару квадратурних сигналів?*
4. *Доведіть взаємозв'язок східчастої функції і дельта-функції.*
5. *У чому полягає умова нормування і фільтрувальна властивість дельта-функції?*
6. *Запишіть загальну форму для скалярного добутку двох сигналів.*
7. *Запишіть вираз для коефіцієнтів розкладу узагальненого ряду Фур'є.*
8. *Чому дорівнює сума енергій усіх гармонік сигналу?*
9. *Доведіть оптимальність за мінімумом квадратичної помилки апроксимації сигналу скінченновимірним рядом Фур'є.*
10. *У чому полягає відмінність між випадковими і детермінованими сигналами?*
11. *Яким чином формуються математичні моделі випадкових сигналів і завод?*
12. *Дайте визначення випадкових подій та їх числових характеристик.*
13. *Що таке випадкові величини та їх характеристики?*
14. *Дайте визначення стаціонарних ергодичних процесів.*
15. *Запишіть основні співвідношення для знаходження кореляційної функції та спектральної щільності потужності.*
16. *Що таке білий шум?*
17. *Назвіть основні числові характеристики сигналів і завод.*
18. *Дайте визначення миттєвої потужності.*
19. *Як знайти енергетичний спектр сигналу та середню потужність за спектром сигналу?*
20. *Як визначаються рівні сигналів і завод?*
21. *Які труднощі виникають при визначенні тривалості і ширини спектра сигналу чи заводи?*

22. Наведіть математичні моделі ідеальних низькочастотних і смугових сигналів.
23. Назвіть способи ортогоналізації сигналів у часовій і в частотній областях.
24. Як має бути вибрано частоту дискретизації сигналу з обмеженим спектром для його неспотвореного подання послідовністю дискретних відліків?
25. Запишіть формулу ряду Котельникова і дайте тлумачення його компонентів.
26. Від яких факторів залежить помилка апроксимації сигналів ряду Котельникова?

Conclusions

For theoretical study of signals it is necessary to build their mathematical models. Mathematical model of signal is a functional relationship that adequately describes temporal variation of the physical state of an object.

Signals classification is performed on the basis of essential attributes of appropriate mathematical models.

Distinction is made between deterministic and random, valid and complex, multivariate and univariate, static and dynamic, infinite in time and impulse, periodic and solitary, analog and discrete signals. A variation of the latter is digital signals, discrete not only in time but also in space.

The most typical elementary signals are harmonic and complex exponential signals, rectangular and triangular video impulses, stepped, signum- and delta- functions etc. Scalar product of two signals is called mutual energy of these signals. Total energy of the signal is determined by scalar product of signal with itself. In other words, total energy of signal is equal to integral of the square of function that describes mathematical model of signal. Square root of signal energy is called the norm of signal. Normalized signals (functions) are signals (functions), which total energy equal to one.

Two signals are called orthogonal if their scalar product is equal to zero. Infinite system of valid functions (signals) is called orthogonal if scalar product of two different functions (signals) is equal to zero. It is envisaged that energy of each signal from system is not equal to zero. This means that none of the features that are considered is equal to zero.

Infinite system of functions, pairwise orthogonal to each other, and functions that have single norm is called a system of orthonormal functions or an orthonormal basis. Any signal with finite energy could be presented in the form of generalized Fourier series. Generalized Fourier series is an infinite sum of products of basic functions and corresponding coefficients. Set of expansion coefficients forms signal spectrum. Expansion coefficients of generalized Fourier series (spectrum) time function (signal) are defined by scalar product of signal with corresponding basic functions. Such method of expansion coefficients determination provides minimum

square error of signal approximation by finite Fourier series.

Signal energy is the sum of energies of all components (harmonics) that form generalized Fourier series. It means that signal energy is the sum of spectral components energies, and squared modulus of generalized Fourier series coefficients is numerically equal to signal energy particle that is contained in corresponding component (harmonic) of a signal.

Process of obtaining useful information, contained in signal, could be represented as hardware (or software) defining the numerical values of generalized Fourier series coefficients of signal.

All random signals and noises are unpredictable. Thus, for random signals it is impossible to find a mathematical formula to calculate their instantaneous values.

All random phenomena studied in theory of probability could be divided into three types: random events; random values; random processes. Each of these types of random phenomena has its own characteristics and features. An event is called random when it may or may not occur as a result of any attempt.

Quantity which value changes from an attempt to attempt randomly is called random. For this quantity it is impossible to predict what value it will obtain under specific conditions of attempt.

Probability distribution function shows probability that random quantity does not exceed a specific value. Mathematical expectation is the average value of a random quantity. Dispersion qualitatively characterizes a measure of results variability of individual attempts relative to the average value.

Random processes are of various kinds: non-stationary, stationary, ergodic, Gaussian, Markov etc. But in technics most of random signals and noises belong to stationary ergodic random processes.

Fluctuation noise is the most typical for the majority of telecommunication channels and it is a stationary ergodic random process with Gaussian (normal) probability distribution.

Spectral power density of fluctuation noise depends on physical nature of its formation, and also from the point where it is observed. The main energy characteristics of a valid signal are its power and energy. For random signals (noises), the average power can be calculated from spectral power density. Dynamic signal range characterizes the limits of instantaneous power variation. Crest factor of signal is the ratio of its power peak to average. Signal duration is a time interval of its existence. Duration is calculated as a difference between end time and start time of signal.

Necessary signal-to-noise ratio is determined by energy characteristics, based on communication channel bandwidth the signal spectrum width is set as a must for undistorted information transmission.

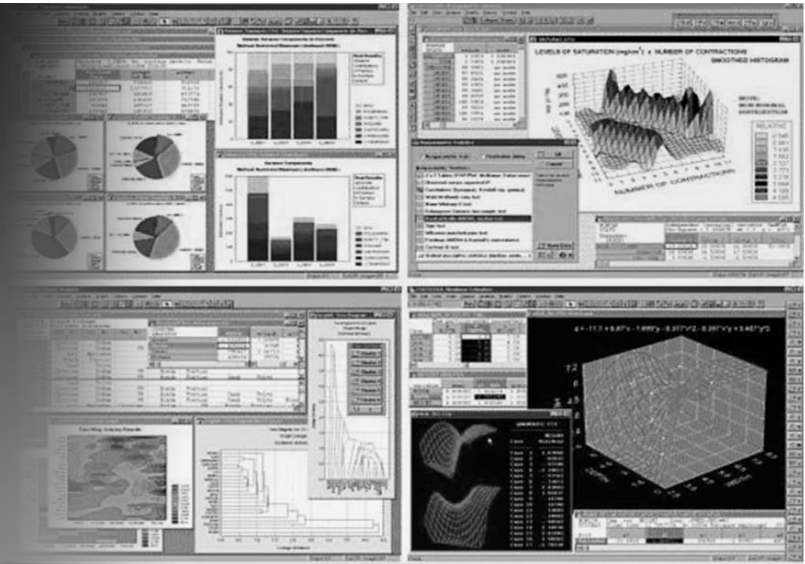
Theoretically, any signals with finite time length have infinitely wide spectrum. Real technical devices have a finite bandwidth. When signals with infinitely wide spectrum are passing through technical devices, spectrum distortion occurs, that inevitably leads to the distortion of input signal form.

For undistorted representation of signal with bounded spectrum as a sequence of its discrete samples, sampling frequency F should be equal to doubled upper fre-

quency of signal spectrum.

Ключові слова

Українською	Англійською
модель сигналу	model of signal
випадковий	casual
завади	hindrances
спектральна щільність	pectral closeness



ПЕРЕТВОРЕННЯ СИГНАЛІВ

4

- 4.1. Дискретні сигнали та їх перетворення
- 4.2. Аналого-цифрове перетворення
- 4.3. Цифро-аналогове перетворення

4.1. Дискретні сигнали та їх перетворення

Нагадаємо основні властивості дискретних сигналів. Значення дискретних сигналів визначено не для всіх моментів часу, а лише на зліченній множині точок $(\dots, t_0, t_1, t_2, \dots)$. Тому, якщо математична модель аналогового сигналу $x(t)$ має звичайні властивості гладкої функції, то дискретний сигнал $x_a(t)$ описується послідовністю $(\dots, x_0, x_1, x_2, \dots)$ своїх відлікових значень у моменти часу відповідно $(\dots, t_0, t_1, t_2, \dots)$. Дискретні сигнали набувають особливого значення в останні десятиліття під впливом як розвитку цифрової техніки зв'язку, так і способів обробки інформації в швидкодіючих ПК. Намітилася

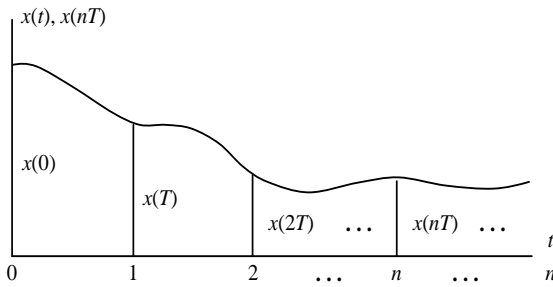


Рис. 4.1. Формування дискретних сигналів

тенденція створення спеціалізованих обчислювальних пристроїв, так званих *цифрових фільтрів*, що використовуються для опрацювання дискретних сигналів. Дискретні сигнали $x(nT)$, $n = 0, 1, 2, \dots$, утворюються в результаті дискретизації (добору

миттєвих значень) неперервних (аналогових) сигналів $x(t)$ у моменти часу t_n , що утворюють зліченну послідовність: $t = nT$, $n = 0, 1, 2, \dots$, де T - крок часової дискретизації (рис. 4.1). Дискретні сигнали $x(nT)$, $n = 0, 1, 2, \dots$, набувають значень із континуальної множини в усьому інтервалі значень (x_{\min}, x_{\max}) неперервної функції $x(t)$. При спектральному аналізі дискретних сигналів зручно застосовувати так зване Z -перетворення, яке для дискретних сигналів відіграє таку саму роль, як і інтегральне перетворення Фур'є для неперервних сигналів. Основу Z -перетворення становить інтегральне перетворення Лапласа.

Інтегральне перетворення Лапласа. Поряд із перетворенням Фур'є для розв'язування найрізноманітніших задач, пов'язаних із вивченням сигналів, широко використовується ще один вид інтегральних перетворень – *перетворення Лапласа*. Розрізняють *однобічне* і *двобічне* перетворення Лапласа.

Однобічне перетворення Лапласа вводиться для сигналів $x(t)$, заданих (тобто таких, що існують) лише на додатній півосі часу. Нехай

Розділ 4. Перетворення сигналів

$x(t)$ – деякий сигнал, дійсний або комплексний, визначений при $t > 0$ і такий, що тотожно дорівнює нулю при від’ємному значенні часу. Перетворення Лапласа $X(p)$ цього сигналу задається інтегралом

$$X(p) = \int_0^{\infty} x(t) e^{-pt} dt, \quad (4.1)$$

де p — параметр перетворення, у загальному випадку — комплексне число: $p = \sigma + j\omega$, яке дістало назву комплексної частоти.

Для сигналів, визначених на всій осі часу від мінус нескінченності до плюс нескінченності, вводять *двобічне перетворення Лапласа*:

$$X(p) = \int_{-\infty}^{\infty} x(t) e^{-pt} dt. \quad (4.2)$$

Сигнал $x(t)$ називається *оригіналом*, а функція $X(p)$ – його *зображенням* за Лапласом (для стислості – просто зображенням).

Якщо у (4.1) чи (4.2) візьмемо $p = j\omega$, то прийдемо до перетворення Фур’є. Отже, якщо для сигналу $x(t)$ відоме його перетворення Лапласа $X(p)$, то спектральна щільність $\dot{X}(\omega)$ сигналу $x(t)$ визначиться заміною в зображенні $X(p)$ параметра p на $j\omega$, тобто

$$\dot{X}(\omega) = X(p)|_{p=j\omega}. \quad (4.3)$$

Розглянемо **приклад** визначення спектральної щільності найпростіших сигналів за їхніми зображеннями за Лапласом:

1. Сигнал типу одnobічної експоненти: $x(t) = U e^{-\alpha t}$, $t \geq 0$.

Скориставшись співвідношенням (4.1), отримаємо $X(p) = U \int_0^{\infty} e^{-\alpha t} e^{-\alpha t} dt = U \int_0^{\infty} e^{-(\alpha+p)t} dt$. Обчисливши інтеграл, маємо $X(p) = U/(\alpha + p)$. Згідно з

(4.3) спектральна щільність $\dot{X}(\omega)$ цього сигналу $\dot{X}(\omega) = U/(\alpha + j\omega)$ співпадає з обчисленою в розд. 3 за допомогою інтегрального перетворення Фур’є спектральною щільністю.

2. Сигнал типу двобічної експоненти: $x(t) = U e^{-\alpha|t|}$, $t \in (-\infty, \infty)$.

Відповідно до (4.2) маємо: $X(p) = U \int_{-\infty}^{\infty} e^{-\alpha|t|} e^{-pt} dt = X_1(p) + X_2(p)$, де $X_1(p) = U \int_0^{\infty} e^{-(\alpha+p)t} dt = U/(\alpha + p)$, а $X_2(p) = U \int_{-\infty}^0 e^{(\alpha-p)t} dt = [U/(\alpha - p)] e^{(\alpha-p)t} \Big|_{-\infty}^0 = U/(\alpha - p)$. Таким чином, зображення за Лапласом двобічної експоненти

$X(p) = 2\alpha U / (\alpha^2 - p^2)$, а з урахуванням співвідношення (4.3) її спектральна щільність $\dot{X}(\omega) = 2\alpha U / (\alpha^2 + \omega^2)$ співпадає з раніше отриманим виразом.

Перетворення Лапласа має властивості, аналогічні властивостям перетворення Фур'є. Наводимо найважливіші з них:

$$\begin{aligned} \sum_k \alpha_k x_k(t) &\leftrightarrow \sum_k \alpha_k X(p); \\ x(t \pm t_0) &\leftrightarrow X(p) e^{\pm pt_0}; \\ x(\gamma t) &\leftrightarrow X(p / \gamma) / \gamma; \\ dx/dt &\leftrightarrow pX(p); \\ \int_{-\infty}^t x(\tau) d\tau &\leftrightarrow X(p) / p; \\ x_1(t) * x_2(t) &\leftrightarrow X_1(p)X_2(p). \end{aligned} \quad (4.4)$$

Множник p називають *оператором диференціювання*, а множник $1/p$ — *оператором інтегрування* в просторі зображень за Лапласом.

Елементи теорії Z-перетворення. Найпростіше до Z-перетворення можна прийти в результаті дискретизації всіх функцій часу в перетворенні Лапласа. Нагадаємо, що перетворенням Лапласа аналогової функції $x(t)$ є функція $X(p) = \int_0^{\infty} x(t) e^{-pt} dt$.

Обчислимо перетворення Лапласа для дискретизованої функції $x(nT)$, формально скориставшись співвідношенням (4.4). У результаті часової дискретизації неперервних функцій $x(t)$ і e^{-pt} , що зводиться до заміни неперервного часу t на дискретний nT , приходимо відповідно до дискретних функцій $x(nT)$ і e^{-pnT} . Отже, підінтегральна функція виявляється відмінною від нуля лише в дискретні моменти часу nT , $n=0, 1, 2, \dots$, що утворюють зліченну множину. Така операція дискретизації дає змогу перейти від інтеграла до суми добутків дискретних функцій. Позначивши нескінченну суму таких добутків через $\hat{X}(p)$, отримаємо

$$\hat{X}(p) = \sum_{n=0}^{\infty} x(nT) e^{-pnT}. \quad (4.5)$$

Скориставшись заміною $z = e^{pT}$ і позначивши ліву частину формули (3.98) через $X(z)$, остаточно знайдемо

$$X(z) = \sum_{n=0}^{\infty} x(n) z^{-n}. \quad (4.7)$$

У виразі (4.7) аргумент T для спрощення подальших перетворень опущено.

Співвідношення (4.7) називається *Z-перетворенням послідовності дискретних сигналів* $x(n)$ і часто позначається як $Z\{x(n)\}$.

Відзначимо деякі *властивості Z-перетворення*, що багато в чому аналогічні властивостям перетворень Фур'є і Лапласа.

Розділ 4. Перетворення сигналів

1. *Лінійність.* Якщо $\{x_k\}$ і $\{y_k\}$ — дві числові послідовності, що відображають деякі дискретні сигнали, причому відомі відповідні Z-перетворення $X(z)$ і $Y(z)$, то сигналу:

$$u(n) = \alpha x(n) + \beta y(n)$$

(4.8)

відповідає перетворення $U(z) = \alpha X(z) + \beta Y(z)$ (4.9)

при будь-яких сталих α і β . Доведення (4.9) виконується підставленням суми (4.8) у (4.7).

2. *Z*-перетворення зміщеного сигналу. Розглянемо дискретний сигнал $\{y_n\}$, утворений з дискретного сигналу $\{x_n\}$ зсувом останнього на одну позицію в бік запізнювання, тобто $y(n) = x(n-1)$. Безпосереднє обчислення *Z*-перетворення приводить до такого результату: $Y(z) = \sum_{n=0}^{\infty} x(n-1) z^{-n}$. Перетворимо праву частину останнього виразу до вигляду:

$z^{-1} \sum_{n=0}^{\infty} x(n-1) z^{-(n-1)}$. Після заміни $k = n-1$ маємо

$z^{-1} \sum_{k=-1}^{\infty} x(k) z^{-k} = x(-1) + z^{-1} \sum_{k=0}^{\infty} x(k) z^{-k}$. Беручи для фізично реалізовуваних сигналів $x(-1) = 0$, дістаємо

$$Z\{x(n-1)\} = z^{-1} X(z). \quad (4.10)$$

Вираз (4.10) можна узагальнити для будь-якого значення затримки k послідовності $x(n)$, а саме:

$$Z\{x(n-1)\} = z^{-k} X(z) z^{-k}. \quad (4.11)$$

Таким чином, *Z*-перетворення послідовності, затриманої на k періодів дискретизації, дорівнює добутку *Z*-перетворення вихідної (не затриманої) послідовності на z^{-k} .

Символ z^{-1} у співвідношенні (4.10) відіграє роль оператора одиничної затримки сигналу на один період дискретизації в *z*-області. Формулі (4.10) можна дати графічну інтерпретацію (рис. 4.2).

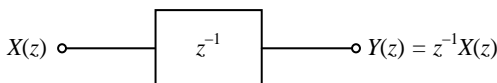
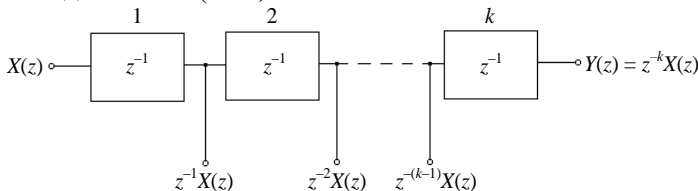


Рис. 4.2. Зображення лінії затримки дискретного сигналу на один період дискретизації в *Z*-області

На рис. 4.3 представлено інтерпретацію співвідношення (4.11).



П'єр Сімон Лаплас (Pierre-Simon Laplace, 1749 - 1827)

Французький математик і астроном, відомий своїми працями в галузі диференціальних рівнянь, один з творців теорії ймовірностей. Основні астрономічні роботи Лапласа присвячені небесній механіці. Вирішив складні задачі руху планет і їх супутників. Лаплас розширив та систематизував математичний фундамент теорії ймовірностей, запровадив похідні функції. Лаплас розвинув також теорію помилок та наближень завдяки своєму методу найменших квадратів.

Розділ 4. Перетворення сигналів

Рис. 4.3. Лінія затримки дискретних сигналів
на k періодів дискретизації

На рис. 4.2 і 4.3 позначено: $X(z)$ – вхідний вплив, а $Y(z)$ – відгук кола, що складається з послідовно з'єднаних одиничних ліній затримки (на один період дискретизації) для Z -області.

Отже, якщо вихідну послідовність $y(n)$ утворено із затриманої на k періодів дискретизації вхідної послідовності $x(n)$, тобто $y(n) = x(n - k)$, то Z -зображення відгуку кола $Y(z)$ пов'язане із Z -зображенням вхідного сигналу $X(z)$ співвідношенням $Y(z) = z^{-k} X(z)$, що відповідає виразу (4.11).

3. *Z-перетворення згортки.* Дискретним сигналам $x(n)$ і $y(n)$ можна поставити у відповідність дискретну згортку $f(k) = \sum_{n=0}^{\infty} x(n) y(k - n) = \sum_{n=0}^{\infty} y(n) x(k - n)$, якій, як і згортці аналогових сигналів, відповідає добуток їхніх Z -зображень, тобто $Z\{x(n) * y(n)\} = X(z)Y(z)$, де $*$ — знак оператора згортки.

Іншими словами, Z -зображення згортки двох дискретних сигналів дорівнює добутку Z -зображень цих сигналів.

Спектр дискретних сигналів. Як відомо, спектральна щільність $\dot{X}(\omega)$ аналогових сигналів $x(t)$ визначається на підставі прямого перетворення Фур'є часової функції $x(t)$, тобто

$$?? \quad \dot{X}_F(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt . \quad ?? \quad (4.12)$$

Тут (і далі) нижній індекс при \dot{X} відповідає типу зображення (зокрема, F — указує на Фур'є-зображення).

Спектральну щільність аналогового сигналу $x(t)$ можна дістати також на підставі перетворення Лапласа цього сигналу. Справді, якщо відоме перетворення Лапласа $X(p)$ сигналу $x(t)$

$$X(p) = \int_0^{\infty} x(t) e^{-pt} dt , \quad (4.13)$$

то, як це випливає з порівняння формул (4.12) і (4.13), спектральна щільність $\dot{X}(\omega)$ сигналу $x(t)$ можна визначити на підставі перетворення Лапласа $X(p)$ заміною параметра p на $j\omega$, тобто $\dot{X}_F(\omega) = X(p)|_{p=j\omega}$.

Часова дискретизація функцій, що входять у праву частину зображення Лапласа, приводить до Z -зображення дискретних сигналів $x(nT)$

$$X(z) = \sum_{n=0}^{\infty} x(nT) z^{-n}, \quad (4.14)$$

де z задано співвідношенням (4.6).

Вважаючи, як і при визначенні спектра сигналу $x(t)$ за його перетворенням Лапласа $X(p)$, параметр p таким, що дорівнює $j\omega$, приходимо до такого способу обчислення спектральної щільності дискретних сигналів за їхнім Z -зображенням: $\dot{X}_z(\omega) = X(z)|_{z=e^{j\omega T}}$. (4.15)

Об'єднуючи формули (4.14) і (4.15), отримаємо вираз для спектральної щільності послідовності дискретних сигналів

$$\dot{X}_z(\omega) = \sum_{n=0}^{\infty} x(n) e^{-j\omega n T}. \quad (4.16)$$

Як видно зі співвідношень (4.15) і (4.16), спектральна щільність дискретного сигналу є функцією комплексної експоненти і згідно формули Ейлера $e^{j\omega T} = \cos\omega T + j\sin\omega T$ є періодичною функцією. З огляду на це *спектр дискретного сигналу також є періодичним*.

Покажемо це на конкретному **прикладі**:

Розглянемо аналоговий сигнал $x(t)$ типу одnobічної експоненти:

$$x(t) = e^{-\alpha t}, \quad t \geq 0. \quad (4.17)$$

Піддавши його часовій дискретизації з періодом T , дістанемо послідовність дискретних сигналів

$$x(nT) = e^{-\alpha n T} = a^n, \quad \text{де } a = e^{-\alpha T} < 1. \quad (4.18)$$

Нагадаємо, що спектральна щільність сигналу (4.17) має вигляд $\dot{X}_F(\omega) = 1/(\alpha + j\omega)$. Його спектр амплітуд $A(\omega)$ є аперіодичною функцією $A_F(\omega) = 1/\sqrt{\alpha^2 + \omega^2}$. Обчислимо Z -зображення сигналу (4.18) за формулою

$X(z) = \sum_{n=0}^{\infty} a^n z^{-n} = \sum_{n=0}^{\infty} (a/z)^n$. Оскільки $a < 1$, а $|z|=1$, то ряд у правій частині останньої формули згортається, набираючи вигляду

$$X(z) = 1/(1 - a/z) = z/(z - a). \quad (4.19)$$

Скориставшись співвідношеннями (4.15) і (4.19), приходимо до такого виразу для спектральної щільності дискретизованої експоненти:

$$\dot{X}_z(\omega) = e^{j\omega T} / (e^{j\omega T} - a). \quad (4.20)$$

Згідно з формулою (4.20) амплітудний спектр $A_z(\omega)$ дорівнює відношенню модуля чисельника цього виразу до модуля знаменника. Модуль чисельника дорівнює одиниці. При визначенні модуля знаменника скористаємося формулою Ейлера для комплексної експоненти. Тоді

$$|e^{j\omega T} - a| = |(\cos\omega T - a) + j\sin\omega T| = \sqrt{(\cos\omega T - a)^2 + \sin^2\omega T}.$$

Розділ 4. Перетворення сигналів

Після елементарних перетворень дістанемо $A_z(\omega) = 1/\sqrt{1+a^2-2a\cos\omega T}$.

Графіки спектрів амплітуд аналогової $A_F(\omega)$ і дискретизованої $A_z(\omega)$ експонент зображено відповідно на рис. 4.4, а і б. З порівняння графіків (рис. 4.4) випливає, що навіть у тому разі, коли спектр аналогового сигналу аперіодичний, спектр дискретизованого сигналу стає періодичним, що є наслідком дискретизації аналогового сигналу.

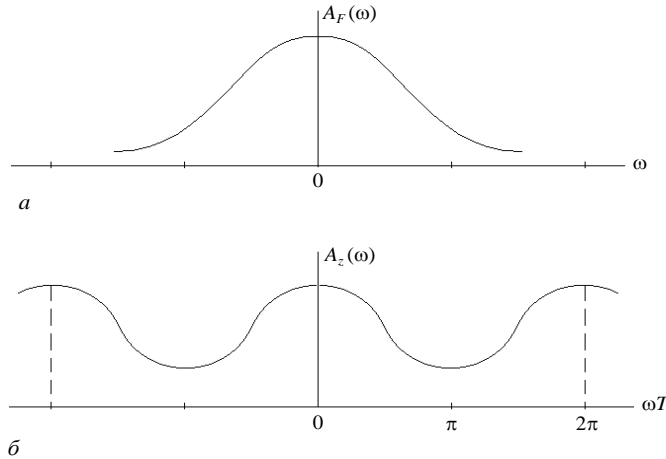


Рис. 4.4. Спектри амплітуд однієї експоненти для сигналів:
а — аналогового; б — дискретного

Отже, можна сформулювати важливий висновок: *спектр дискретних сигналів є періодичною функцією частоти навіть у тому разі, коли спектр його аналогового прототипу – аперіодична функція.*

Ефект накладання спектрів при дискретизації сигналів. Розглянемо деякий гіпотетичний сигнал, амплітудний спектр якого має вигляд рівностороннього трикутника з основою $\pm\omega_B$ (рис. 4.5).

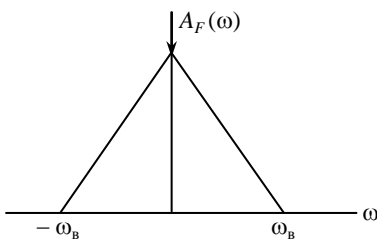


Рис. 4.5. Спектр амплітуд гіпотетичного сигналу

Як відомо, при дискретизації будь-якого аналогового сигналу $x(t)$ спектр дискретних сигналів $x(n)$ стає періодичним, причому інтервал періодичності $\omega T = 2\pi$, тобто модуль і фаза спектральних складових сигналу на частотах $\omega_k = \omega \pm k 2\pi/T$, $k = 1, 2, \dots$, (4.21) збігаються за модулем і фазою з гармонікою сигналу на частоті ω . Справді, спектральна складова дискретного сигналу $x(n)$ на деякій частоті ω визначається загальним співвідношенням (4.15) і є функцією комплексного аргументу $e^{j\omega T}$ (4.22). Замінивши кутову частоту ω в експоненті (4.22) значенням ω_k ,

заданим правою частиною рівності (4.21), отримаємо: $e^{j\omega_k T} = e^{j(\omega \pm k 2\pi/T)T} = e^{j\omega T} e^{\pm j 2k\pi}$, тобто $e^{j\omega_k T} = e^{j\omega T}$, чим і підтверджується періодичність значень спектральних складових дискретного спектру: $\Delta_\omega = 2\pi/T$. (4.22) Від інтервалу Δ_ω періодичності спектра

дискретних сигналів за кутовою частотою (4.22) можна перейти до інтервалу Δ_f періодичності спектра за циклічною частотою $\Delta_\omega = 2\pi\Delta_f$: $\Delta_f = 1/T = F$, (4.23)

де F — частота часової дискретизації сигналу.

Таким чином, згідно з (4.23): *спектр дискретизованого сигналу періодичний, причому інтервал періодичності Δ_f збігається з частотою дискретизації $F = 1/T$.*

Припустимо, що форма спектра амплітуд дискретизованого сигналу відповідає формі спектра аналогового сигналу і з урахуванням періодичності має такий вигляд, як показано на рис. 3.43.

Графіки на рис. 4.6 унаочнюють випадок, коли $\omega_b T_1 < \pi$, де ω_b — максимальна (верхня) гармоніка аналогового сигналу, а T_1 — період часової дискретизації аналогового сигналу.

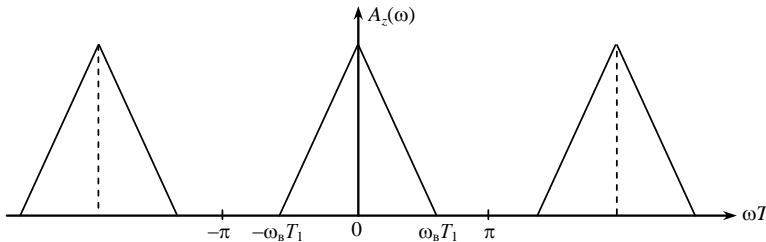


Рис. 4.6. Спектр амплітуд дискретного сигналу

Збільшимо тепер період часової дискретизації аналогового сигналу і візьмемо його таким, що дорівнює значенню T , при якому виконується рівність: $\omega_b T = \pi$, (4.24)

згідно з якою основи «трикутників» амплітудних спектрів будуть примикати один до одного, як це зображено на рис. 4.7. З подальшим збільшенням періоду дискретизації $T_2 > T$ трикутники амплітудного спектра починають накладатися один на одного (рис. 4.8). У результаті накладання високочастотних гармонік спектра на низькочастотні амплітудний спектр дискретного сигналу

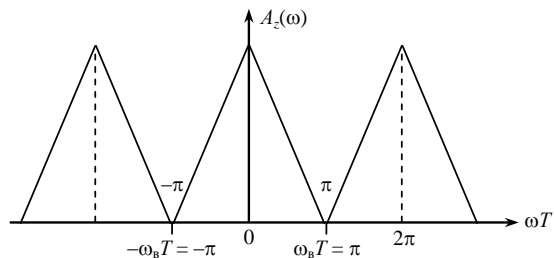


Рис. 4.7. Спектр амплітуд дискретного сигналу, для якого виконується умова (4.24)

Розділ 4. Перетворення сигналів

(верхні ламані прямі) виявляється істотно відмінним від спектра аналогового сигналу (див. рис. 4.5). Причина спотворення амплітудного спектра дискретного сигналу (рис. 4.8) полягає в тому, що частоту дискретизації $F_2 = 1/T_2$ взято недостатньо великою, у результаті чого високочастотні складові періодичного спектра потрапляють в область нижчих частот. Такий зсув спектральних складових з одного діапазону частот в інший називають *накладанням спектрів*.

Ефект накладання спектрів дискретних сигналів можна усунути за рахунок відповідного вибору частоти дискретизації аналогових сигналів.

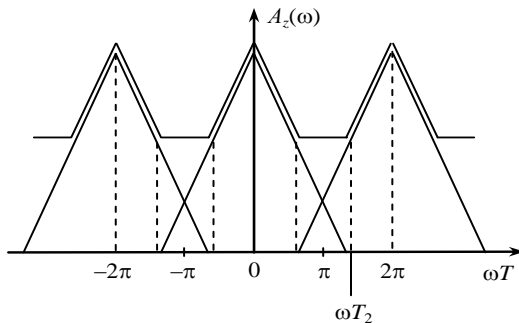


Рис. 4.8. Ефект накладання спектрів дискретизованого сигналу

умова (4.24). Переходячи у співвідношенні (4.24) від кутової частоти ω_B до циклічної $\omega_B = 2\pi F_B$ (F_B — верхня циклічна гармоніка аналогового сигналу) і замінюючи період часової дискретизації T частотою дискретизації сигналу, тобто $F = 1/T$, дістаємо співвідношення $F = 2F_B$.

Це уже відоме положення теореми Котельникова (у зарубіжній літературі її називають теоремою Найквіста або теоремою відліків): *якщо найвища частота в спектрі сигналу $x(t)$ не перевищує F_B , то функція $x(t)$ цілком визначається послідовністю своїх значень у моменти часу, віддалені один від одного не більш ніж на $1/2F_B$ секунд.*

Або іншими словами: будь-який аналоговий сигнал без спотворення можна подати послідовністю його дискретних відліків $x(nT)$ за умови, що частота дискретизації не менш як удвічі перевищує найвищу гармоніку спектра аналогового сигналу.

На практиці цю умову виконати важко, оскільки спектр аналогових сигналів надзвичайно широкий. Тому в разі скінченної

З'ясуємо умови, за яких накладання спектрів у дискретизованих сигналах відсутнє. Нехай спектр аналогового сигналу обмежено верхньою граничною частотою ω_B , як це показано на рис. 4.5. Очевидно, що спотворення спектра при дискретизації сигналу не виникає (див. рис. 4.7), якщо виконується

частоти дискретизації спектр дискретизованого сигналу відрізняється від спектра аналогового сигналу.

Дискретне перетворення Фур'є. Як було встановлено (див. (4.16)), спектральна щільність $\dot{X}_z(\omega)$ дискретного сигналу $x(nT)$

визначається виразом
$$\dot{X}_z(\omega) = \sum_{n=0}^{\infty} x(nT) e^{-j\omega nT}, \quad (4.25)$$

де n — номер дискретного відліку неперервної функції; T — період дискретизації неперервної функції $x(t)$.

Згідно з формулою (4.25) спектр дискретного сигналу — суцільний. Але таким він буває лише за умови, що обсяг вибірки дискретного сигналу нескінченний. На **практиці** вибірка відліків сигналу завжди скінченновимірна. Крім того, з багатьох причин бажано обчислювати перетворення Фур'є на ПК. Це означає, що скінченновимірною є не тільки вибірка дискретних відліків сигналу, а й відповідне цій вибірці число гармонік спектра дискретного сигналу.

Припустимо, що деяку неперервну функцію (аналоговий сигнал) $x(t)$ подано послідовністю N відліків цієї функції $x(n) = x(nT)$, де T — період часової дискретизації аналогового сигналу.

Поставимо у відповідність скінченновибірковим відлікам сигналу (4.26) скінченновибіркову послідовність спектральних складових $\dot{X}(k)$, узявши $k = \overline{0, N-1}$. Для обчислення N спектральних складових (гармонік спектра) діятимемо в такий спосіб. Спочатку замінимо у формулі (4.25) кутову частоту ω циклічною f :

$$\omega = 2\pi f, \quad (4.27)$$

а потім перейдемо від неперервних частот до дискретних:

$$\left. \begin{array}{l} \omega \rightarrow \omega_k \\ f \rightarrow f_k \end{array} \right\}, \quad k = \overline{0, N-1}. \quad (4.28)$$

Розставимо N спектральних складових дискретного сигналу еквідистантно на всьому частотному інтервалі періодичності спектра, що, як показано в попередньому підрозділі, дорівнює F (рис. 4.9).

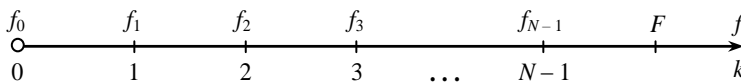


Рис. 4.9. Розподіл гармонік спектра на осі частот

Позначимо $f_k = k f_1, \quad k = \overline{0, N-1}, \quad (4.29)$

де $f_1 = F/N$

(4.30)

- інтервал частотної дискретизації спектра.

Розділ 4. Перетворення сигналів

Скориставшись (4.27) - (4.30) перетворимо нескінченний ряд (4.25) до скінченного, записавши його у вигляді: $\dot{X}(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi nTF/N}$. Оскільки

$$TF = 1, \text{ то маємо } \dot{X}(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N}. \text{ Позначивши } W = e^{-j2\pi/N}, \quad (4.31)$$

$$\text{Отримаємо } \dot{X}(k) = \sum_{n=0}^{N-1} x(n) W^{kn}, \quad k, n = \overline{0, N-1}. \quad (4.32)$$

Співвідношення (4.32) має назву *дискретного перетворення Фур'є (ДПФ)*, а комплексний множник W , заданий формулою (4.31), називається *фазовим* (або *повертальним*) *множником* ($\Phi\text{М}$).

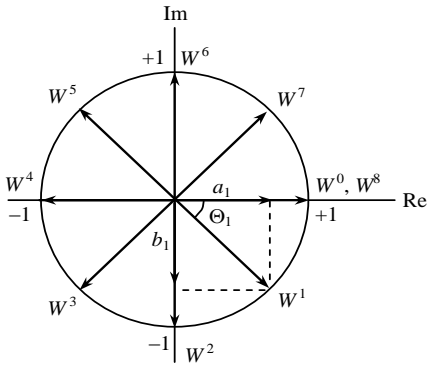


Рис. 4.10. Степені фазового множника для восьмиточкового ДПФ

тотожність (4.33) є доведеною. Періодичність ФМ можна проілюструвати графічно (рис. 4.10) на прикладі восьмиточкового ($N = 8$) ДПФ. З періодичності ФМ випливає періодичність спектра $\dot{X}(k)$ дискретних сигналів $x(n)$, $n = \overline{0, N-1}$, тобто якщо $k = mN + k_0$, (4.35)

$$\text{де } m \geq 1, \text{ а } k_0 < N, \text{ то } \dot{X}(k) \equiv \dot{X}(k_0). \quad (4.36)$$

Справді, підставивши значення k , задане виразом (4.35), у формулу (4.32), отримаємо:

$$\dot{X}(k) = \sum_{n=0}^{N-1} \dot{x}(n) e^{-j2\pi n(mN+k_0)/N}. \quad (4.37)$$

Виконуючи елементарні перетворення в показнику степеня експоненти, маємо:

$$e^{-j\frac{2\pi}{N}n(mN+k_0)} = e^{-j\frac{2\pi}{N}nk_0} e^{-j2\pi nm}.$$

Оскільки для будь-якого цілого m друга експонента в правій частині останнього виразу дорівнює одиниці, формула (4.37) приводить

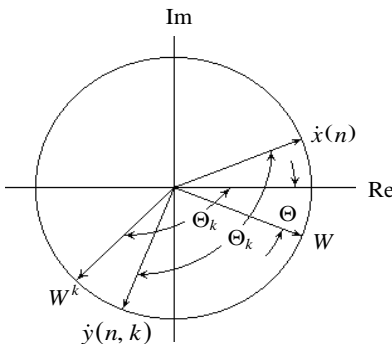


Рис. 4.11. Формування вектора $\dot{y}(n, k)$

до співвідношення $\dot{X}(k) = \sum_{n=0}^{N-1} \dot{x}(n) e^{-j2\pi nk_0/N}$, тобто $\dot{X}(k) = \dot{X}(k_0)$, що й потрібно було довести.

Таким чином, N -вибірковій сукупності дискретних відліків сигналу, еквідистантно розміщених на осі часу, відповідає N -вибіркова сукупність гармонік сигналу, еквідистантно розміщених на осі частот.

Інтервал між сусідніми гармоніками (у герцах) дорівнює: $f_1 = F/N$, де $F = 1/T$ — частота дискретизації, а N — обсяг вибірки сигналу.

При множенні вектора (у загальному випадку комплексного) сигналу $\dot{x}(n)$ на k -й степінь множника W у співвідношенні (4.32) утворюється вектор $\dot{y}(n, k) = \dot{x}(n)W^k$, повернутий щодо вектора $\dot{x}(n)$ за

годинниковою стрілкою на кут $\Theta_k = k\Theta$ (рис. 4.11). Цим

пояснюється назва комплексного множника W як «фазового», або «обертового».

Пристрій (апаратний чи програмний), що реалізує алгоритм ДПФ, називається процесором ДПФ (рис. 4.12).

Якщо на вхід процесора ДПФ подати відліки сигналу $\overline{x(n)}$, $n = 0, N-1$, то на його виході формуються гармоніки $\overline{\dot{X}(k)}$, $k = 0, N-1$, що відповідають N -вибірковій сукупності вхідних сигналів.

Обернене дискретне перетворення Фур'є. Звернемося до виразу для ДПФ:

$$\dot{X}(k) = \sum_{n=0}^{N-1} \dot{x}(n)W^{nk}. \quad (4.38)$$

Помножимо обидві частини співвідношення (4.38) на W^{-mk} і підсумуємо за k від 0 до $N-1$:

$$\sum_{k=0}^{N-1} \dot{X}(k)W^{-mk} = \sum_{k=0}^{N-1} \sum_{n=0}^{N-1} \dot{x}(n)W^{nk}W^{-mk} = \sum_{k=0}^{N-1} \sum_{n=0}^{N-1} \dot{x}(n)W^{(n-m)k}.$$

Змінивши порядок підсумовування в правій частині останнього виразу, дістанемо: $\sum_{n=0}^{N-1} \dot{x}(n) \sum_{k=0}^{N-1} W^{(n-m)k}$. Внутрішня сума тут відмінна від нуля і

дорівнює N лише при $n = m$: $N \dot{x}(m) = \sum_{k=0}^{N-1} \dot{X}(k)W^{-mk}$. Замінімо m на n . У

результаті дістанемо остаточно:

$$\dot{x}(n) = \frac{1}{N} \sum_{k=0}^{N-1} \dot{X}(k)W^{-nk}, \quad n = \overline{0, N-1}.$$

(4.39)

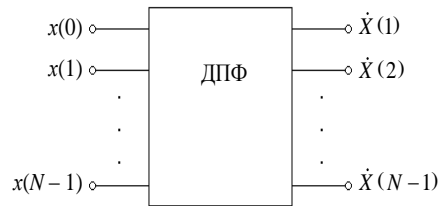


Рис. 4.12. Умовне позначення процесора ДПФ

Розділ 4. Перетворення сигналів

Вираз (4.39) називається *оберненим дискретним перетворенням Фур'є*. Його можна подати також у вигляді $\dot{x}(n) = \left(\sum_{k=0}^N \dot{X}^*(k) W^{nk} \right)^* / N$, тобто, з точністю до коефіцієнта $1/N$ *обернене ДПФ збігається з комплексно-спряженим ДПФ послідовності комплексно-спряжених гармонік сигналу*.

Співвідношення (4.38) і (4.39) утворюють пару ДПФ (відповідно прямого і оберненого).

Дискретизація періодичних сигналів. При дослідженні за допомогою ПК неперервний сигнал $\dot{x}(t)$ на інтервалі часу $(0, T)$ замінюється своїми відліковими значеннями $(\dot{x}_0, \dot{x}_1, \dots, \dot{x}_{N-1})$, узятими відповідно в моменти часу $(0, \Delta, 2\Delta, \dots, (N-1)\Delta)$. Повна кількість відліків

$$N = T / \Delta, \tag{4.40}$$

де Δ — крок (період) часової дискретизації неперервного (аналогового) сигналу. Масив чисел $\{x_n\}$, дійсних або комплексних, є тією єдиною інформацією, з якої можна судити про спектральні властивості сигналу $\dot{x}(t)$. Дослідження таких дискретних сигналів можна істотно спростити, якщо здобути вибірку відлікових значень сигналу повторити нескінченну кількість разів ліворуч і праворуч по осі абсцис (рис. 4.13). У результаті такої операції приходимо до періодичного дискретного сигналу.

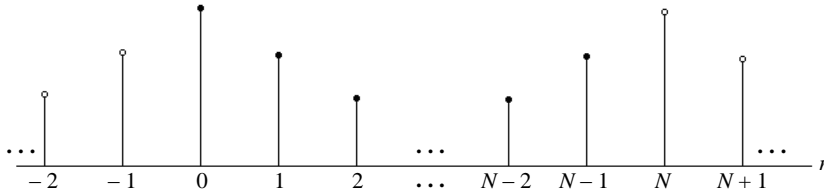


Рис. 4.13. Дискретне подання періодичного сигналу

На рис. 4.13 чорними кружечками позначено вихідну N -вимірну послідовність дискретних сигналів, а незатемненими кружечками — періодичне продовження послідовності.

Підібравши для сигналу деяку математичну модель, можна скористатися розкладом у ряд Фур'є і знайти відповідні амплітудні коефіцієнти. *Сукупність цих коефіцієнтів утворює спектр дискретного періодичного сигналу.*

Скористаємося моделлю у вигляді послідовності дельта-імпульсів і поставимо у відповідність вихідному коливанию $\dot{x}(t)$ його дискретне подання $\dot{x}_a(t)$ на інтервалі $(0, T)$: $\dot{x}_a(t) = \sum_{n=0}^{N-1} \dot{x}_n \delta(t - n\Delta)$, (4.41)

де $\dot{x}_n = \dot{x}(n\Delta)$ — відлікові значення $\dot{x}(t)$ в n -й точці.

Подамо дискретну модель (4.41) комплексним рядом Фур'є:

$$\dot{x}_a(t) = \sum_{k=-\infty}^{\infty} \dot{C}_k e^{j2\pi kt/T}$$

з коефіцієнтами
$$\dot{C}_k = \frac{1}{T} \int_0^T \dot{x}_a(t) e^{-j2\pi kt/T} dt. \quad (4.42)$$

Підставляючи співвідношення (4.41) у (4.42), з урахуванням формули (4.40) отримаємо:
$$\dot{C}_k = \frac{1}{N\Delta} \int_0^{N\Delta} \sum_{n=0}^{N-1} \dot{x}_n \delta(t - n\Delta) e^{-j2\pi kt/T} dt. \quad (4.43)$$

Введемо безрозмірну змінну: $\xi = t/\Delta, \quad dt = \Delta d\xi$.

Це означає, що необхідно поділити на період дискретизації Δ усі змінні, що мають розмірність часу в секундах, а саме: аргумент δ -функції $t - k\Delta$, змінні t і T у показнику експоненти, а також верхня межа інтегрування (4.43). У результаті такої заміни змінних вираз (4.43) набере вигляду:
$$\dot{C}_k = \frac{1}{N} \int_0^{N-1} \sum_{n=0}^{N-1} \dot{x}_n \delta(\xi - n) e^{-j2\pi k\xi/N} d\xi,$$
 а після

переставлення місцями операцій інтегрування і підсумовування —
$$\dot{C}_k = \frac{1}{N} \sum_{n=0}^{N-1} \dot{x}_n \int_0^N \delta(\xi - n) e^{-j2\pi k\xi/N} d\xi.$$
 Нарешті, використовуючи фільтрувальні властивості дельта-функції, маємо:
$$\dot{C}_k = \frac{1}{N} \sum_{n=0}^{N-1} \dot{x}_n e^{-j2\pi nk/N}. \quad (4.44)$$

Формула (4.44) визначає послідовність коефіцієнтів, що утворюють ДПФ розглядуваного сигналу. З точністю до коефіцієнта $1/N$ вона збігається з раніше отриманою формулою (4.29), але на відміну від неї — математично доволі коректна. Зі співвідношення (4.44) чітко випливає, що N -вибірковій сукупності відліків сигналу \dot{x}_n , узятих з інтервалом часової дискретизації Δ , відповідає N -вибіркова сукупність коефіцієнтів розкладу \dot{C}_k , що утворюють дискретний спектр цього дискретного сигналу. Якщо, як і раніше, позначимо через $W = e^{-j\frac{2\pi}{N}}$ фазовий множник, то вираз (4.44) набере більш звичного вигляду:

Розділ 4. Перетворення сигналів

$$\dot{C}_k = \frac{1}{N} \sum_{n=0}^{N-1} \dot{x}_n W^{nk}, \quad k = \overline{0, N-1}, \quad (4.45)$$

– прямого ДПФ, якому відповідає обернене ДПФ:

$$\dot{x}_n = \sum_{k=0}^{N-1} \dot{C}_k W^{-nk}, \quad n = \overline{0, N-1}. \quad (4.46)$$

Взаємно доповнювальні формули (4.45) і (4.46) є дискретними аналогами звичайної пари перетворень Фур'є для дискретних сигналів.

Властивості дискретного перетворення Фур'є. 1. ДПФ є лінійним перетворенням, тобто сумі зважених сигналів відповідає зважена сума їхніх

ДПФ: $\sum_{m=1}^M \alpha_m \{\dot{x}_N^{(m)}\} \leftrightarrow \sum_{m=1}^M \alpha_m \{\dot{C}_N^{(m)}\}$, де α_m – вагові коефіцієнти N -вимірної m -ї

послідовності дискретних сигналів $\dot{x}_N^{(m)} = \dot{x}_0^{(m)}, \dot{x}_1^{(m)}, \dots, \dot{x}_{N-1}^{(m)}$, $m = \overline{1, M}$,

а $\dot{C}_N^{(m)} = \dot{C}_0^{(m)}, \dot{C}_1^{(m)}, \dots, \dot{C}_{N-1}^{(m)}$ – коефіцієнт розкладу m -ї послідовності сигналів.

2. Кількість різних коефіцієнтів $\dot{C}_0, \dot{C}_1, \dots, \dot{C}_{N-1}$, що обчислюються за формулою прямого ДПФ (4.46), дорівнює кількості N відліків у вибірці.

3. Коефіцієнт \dot{C}_0 (стала складова) є середнім значенням усіх відліків:

$$C_0 = \sum_{n=0}^{N-1} \dot{x}_n / N.$$

4. Якщо N — парне число, то $C_{N/2} = \sum_{n=0}^{N-1} \dot{x}_n (-1)^n / N$. (4.47)

Справді, згідно з формулою (4.45) $W^k = e^{-j\frac{2\pi}{N}k}$, тобто при $k = N/2$: $e^{-j\frac{2\pi}{N}\frac{N}{2}} = e^{-j\pi} = -1$, що й підтверджує співвідношення (4.47).

5. Нехай відлікові значення x_n – дійсні числа. Тоді коефіцієнти ДПФ, номери яких розміщуються симетрично відносно $N/2$, утворюють комплексно-спряжені пари, тобто

$$\dot{C}_{N-k} = \dot{C}_k^*. \quad (4.48)$$

Справді, з формули (4.45) маємо: $C_{N-k} = \sum_{n=0}^{N-1} x_n e^{-2\pi(N-k)n/N} / N = \sum_{n=0}^{N-1} x_n e^{+2\pi kn/N} / N = \dot{C}_k^*$, оскільки $e^{-j2\pi m} \equiv 1$ для кожного $n = \overline{0, N-1}$, що й доводить вираз (4.48).

Тому можна вважати, що коефіцієнти $\dot{C}_{N/2+1}, \dots, \dot{C}_{N-1}$ відповідають від'ємним частотам. При вивченні амплітудного спектра сигналу вони не застосовуються, і їх можна не обчислювати.

Розглянемо **приклад** розрахунку коефіцієнтів ДПФ: Нехай дискретний сигнал (рис. 4.14) на інтервалі своєї періодичності заданий вісьмома

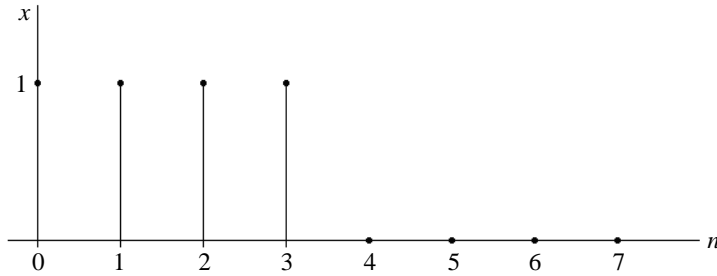


Рис. 4.14. Послідовність дискретних сигналів

рівновіддаленими відліками: $\{x_n\} = (1, 1, 1, 1, 0, 0, 0, 0)$. (4.49)

Знайдемо коефіцієнти ДПФ (спектр) цього сигналу. Використовуючи основну формулу (3.139), знаходимо безпосередньо $C_0 = 1/2$. Коефіцієнти \dot{C}_k зручно обчислювати графічно, використовуючи векторне зображення дискретних відліків сигналу на комплексній площині.

Так, для коефіцієнта $\dot{C}_1 = \sum_{n=0}^3 e^{-j\pi n/4} / 8$ (4.50)

векторну діаграму компонентів ряду (4.50) зображено на рис. 4.15.

Суму векторів, позначених цифрами 0, 1, 2, 3, запишемо у вигляді комплексного числа $\dot{C}_1 = (a, b)$, де a — дійсна, а b — уявна частини коефіцієнта \dot{C}_1 . Розглядаючи вираз (4.50) і рис. 4.15, отримаємо

$$\dot{C}_1 = (1, -(1 + \sqrt{2})) / 8. \quad (4.51)$$

Для коефіцієнта \dot{C}_2 маємо: $\dot{C}_2 = \sum_{n=0}^3 e^{-j\frac{\pi}{2} 2n} / 8$. Згідно з діаграмою (рис. 4.16)

дістанемо: $\dot{C}_2 = 0$. Виконуючи аналогічні розрахунки для інших коефіцієнтів, знаходимо $\dot{C}_3 = (1, -(\sqrt{2} - 1)) / 8, \dot{C}_4 = 0$. (4.52)

Наступні коефіцієнти визначаємо на підставі їхньої спряженості з раніше обчисленими: $\dot{C}_5 = \dot{C}_3^* = (1, \sqrt{2} - 1) / 8; \dot{C}_6 = \dot{C}_2^* = 0; \dot{C}_7 = \dot{C}_1^* = (1, 1 + \sqrt{2}) / 8$.

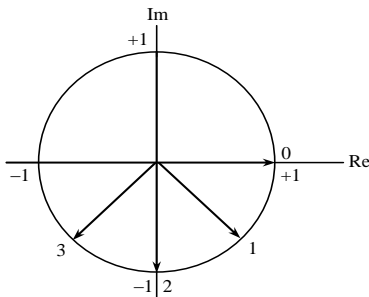


Рис. 4.15. До обчислення

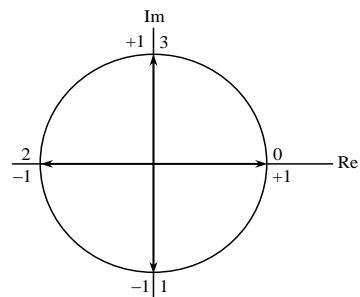


Рис. 4.16. До обчислення

Розділ 4. Перетворення сигналів

коефіцієнта \dot{C}_1

коефіцієнта \dot{C}_2

Отже, за заданими системою (4.49) значеннями дискретного сигналу $\dot{x}(n)$, $n = 0, N-1$, можна знайти сталу складову $C_0 = 1/2$, а також першу і третю гармоніки з амплітудами $A_1 = \sqrt{4 + 2\sqrt{2}}/8$; $A_3 = \sqrt{4 - 2\sqrt{2}}/8$.

Коефіцієнтам \dot{C}_7 і \dot{C}_5 відповідають комплексно-спряжені амплітуди відповідно першої і третьої гармонік вхідного сигналу, що дає змогу записати тригонометричну форму ряду Фур'є для розглянутого сигналу:

$$\dot{x}(t) = 1/3 + 2(A_1 \cos(2\pi t/T + \varphi_1) + A_3 \cos(6\pi t/T + \varphi_3)), \quad (4.53)$$

в якому початкові фази, що подаються загальним співвідношенням $\varphi = \arctg(b/a)$, на підставі виразів (4.51) і (4.52) мають такі значення: $\varphi_1 = -\arctg(1 + \sqrt{2})$; $\varphi_3 = -\arctg(\sqrt{2} - 1)$.

На рис. 4.17 зображено сигнал $\dot{x}(t)$, відновлений за своїми коефіцієнтами \dot{C}_k згідно з формулою (4.53). Ряд (4.49) утворено в результаті дискретизації періодичної послідовності прямокутних відеоімпульсів з одиничною амплітудою і шпаруватістю $q = T/\tau = 2$ (рис. 4.17). Якщо вдвічі зменшити період дискретизації Δ (вдвічі збільшити обсяг вибірки N), то якість відновлення прямокутного імпульсу тригонометричним рядом Фур'є з коефіцієнтами ДПФ $\dot{C}_0, \dot{C}_1, \dots, \dot{C}_{15}$ підвищується, що впливає з порівняння графіків на рис. 4.17 і 4.19.

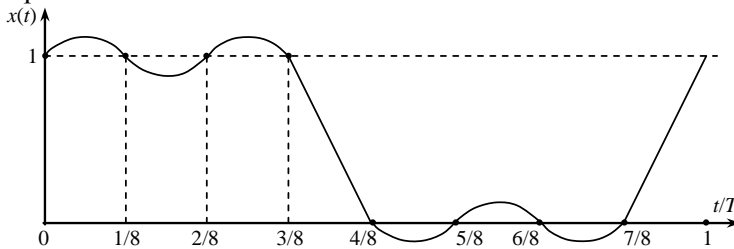


Рис. 4.17. Сигнал, відновлений за коефіцієнтами ДПФ

Слід наголосити, що відновлення неперервного сигналу рис. 4.18 за формулою (4.53) є точною операцією, цілком еквівалентною знаходженню поточних значень сигналу з обмеженим спектром за його вибірками, які утворюють ряд Котельникова.

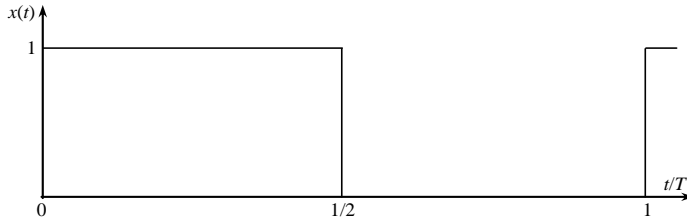


Рис. 4.18. Періодична послідовність прямокутних відеоімпульсів

Процедура, що використовує ДПФ, має переваги, оскільки приводить до скінченних сум гармонік, тоді як ряд Котельникова для періодичного сигналу принципово має містити нескінченну кількість членів.

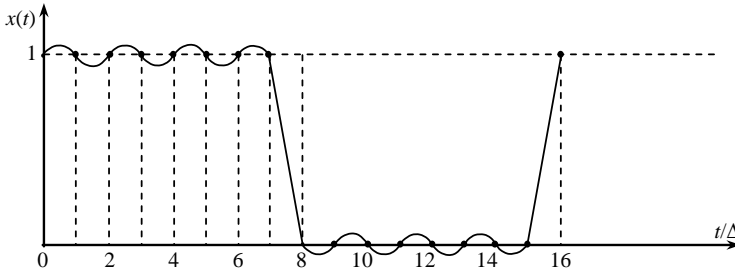


Рис. 4.19. Ілюстрація ефекту подвоєння частоти дискретизації аналогового сигналу

Справді, нехай спостерігається деяка періодична функція $x(t) = x(t \pm mT)$, де T — її період, а $m = 1, 2, \dots$. Раніше було показано, що коли відома парна кількість N відліків цієї функції на всьому її періоді, можна обчислити її сталу складову і перші $N/2$ гармоніки сигналу. Решта гармонік містяться в дзеркальній від'ємній області частот (з відповідними комплексно-спряженими коефіцієнтами ДПФ) і не впливають на амплітудний спектр сигналу.

Отже, при будь-якому парному N кількість необхідних гармонік ДПФ становить половину кількості відліків. Це положення безпосередньо впливає з теореми Котельникова. Справді, верхню граничну частоту в спектрі дискретного сигналу потрібно (згідно з теоремою Котельникова) знаходити зі співвідношення: $f_a = 1/2\Delta$, де Δ - період дискретизації періодичного сигналу. І оскільки $1/\Delta = F$ — частота дискретизації, що, у свою чергу, пов'язана з частотою першої гармоніки f_1 співвідношенням $F = N f_1$, то безпосередньо дістанемо: $f_a = (N/2)f_1$, причому $f_1 = 1/T$, де T — період повторення періодичного сигналу.

Таким чином, підтверджено доцільність застосування ДПФ при апроксимації періодичних сигналів і його переваги порівняно з наближенням періодичних функцій рядами Котельникова.

4.2. Аналого-цифрове перетворення

У вимірювальних каналах ІВС досліджувані неперервні в часі фізичні величини переутворюються сенсорами в аналогові електричні вимірювальні сигнали $s(t)$. Останні за допомогою аналого-цифрового перетворення відображаються цифровими кодами, які можна опрацьовувати цифровими методами. В комп'ютеризованих ІВС широко використовується і зворотне перетворення – цифрового коду в аналоговий сигнал (найчастіше електричний струм або напругу). Такі операції виконуються відповідно аналого-цифровими і цифро-аналоговими перетворювачами (АЦП і ЦАП).

Основними операціями, що здійснюються під час аналого-цифрового перетворення, є дискретизація сигналів в часі, їх квантування за рівнем і цифрове кодування. Дискретизація сигналів була розглянута в п. 4.1

Квантування сигналів – це процес поділу діапазону його континуальних значень (s_{\min}, s_{\max}) на кінцеве число суміжних інтервалів без їх перекриття та подання будь-якого значення з певного інтервалу одним, яке належить цьому ж інтервалу. В результаті квантування нескінченному числу значень вимірюваного сигналу з певного обмеженого діапазона ставиться у відповідність кінцева множина дозволених значень – *рівнів квантування*. Операції квантування за рівнем передують дискретизація сигналів у часі з періодом T_d .

Динамічний діапазон дискретизованого сигналу $s(nT_d)$ поділяється на окремі інтервали – кроки (ступені) квантування величиною δ . Якщо значення сигналу належить тому чи іншому кроку, здійснюється його заокруглення до рівня квантування, наприклад до середини цього кроку квантування. У моменти часу nT_d за виконання умови $s(nT_d) \in [c_{\text{кв}i}, c_{\text{кв}(i+1)})$ (рис. 4.20), первинний сигнал відображається значенням $c_{\text{кв}i} + 0,5\delta$ з похибкою квантування

$$\Delta_{\text{кв}}(kT_d) = s(kT_d) - c_{\text{кв}i} - 0,5\delta. \quad (4.53)$$

За характером похибка $\Delta_{\text{кв}}(kT_d)$ є випадковою величиною.

За рівномірного квантування кроки квантування обирають однаковими, тобто $\delta = c_{\text{кв}i} - c_{\text{кв}(i-1)} = \text{const}$. Амплітудна характеристика рівномірного квантувача зображена на рис. 4.21.

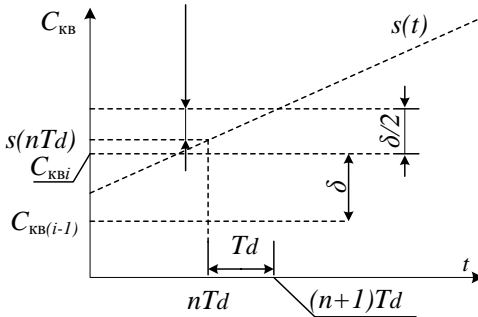


Рис. 4.20. Утворення похибки квантування

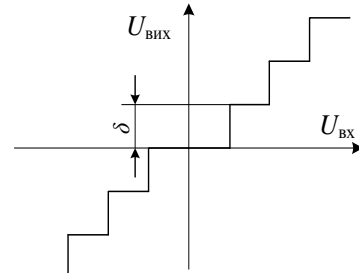


Рис. 4.21. Амплітудна характеристика квантувача

Заміна нескінченної множини значень сигналу кінцевою призводить до втрати певної частини інформації і появи похибки (4.53), яка розглядається як прояв специфічної завади – шуму квантування. Його вплив на якість вимірювання оцінюється відношенням середньої потужності сигналу P_c до середньої потужності шуму квантування $P_{шкв}$

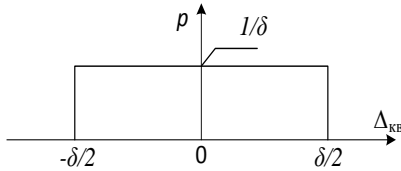


Рис. 4.22. Графічне представлення щільності ймовірності похибки квантування

$$h_{кв} = P_c / P_{шкв} \quad (4.54)$$

Визначимо потужність шуму квантування. Похибка квантування змінюється в межах $\pm 0,5\delta$ і має рівномірний розподіл ймовірностей з щільністю $p(\Delta_{кв}) = 1/\delta$ (рис. 4.22).

Потужність шуму квантування дорівнює середньому квадрату (дисперсії) похибки квантування

$$P_{шкв} = \overline{\Delta_{кв}^2} = \frac{1}{\delta} \int_{-\delta/2}^{\delta/2} \Delta_{кв}^2 d\Delta_{кв} = \delta^2/12. \quad (4.55)$$

З (4.55) випливає, що $P_{шкв}$ тим більше, чим більше крок квантування δ . Однак зі збільшенням δ зменшується вплив адитивних завад. Тому крок квантування визначають з умов досягнення компромісу між впливом зовнішніх завад та шуму квантування. Загальна кількість рівнів квантування

$$M = (s_{\max} - s_{\min}) / \delta + 1. \quad (4.56)$$

Визначимо середню потужність відліків сигналу на виході квантувача. Значення максимального відліку (позитивного чи негативного) дорівнює $s_{\max} = 0,5\delta(M - 1)$. Для великих значень M маємо $s_{\max} \approx 0,5\delta M$, а потужність максимального відліку – $P_{\max} \approx 0,25\delta^2 M^2$. З урахуванням пік-фактора сигналу

Розділ 4. Перетворення сигналів

$\Pi = 10 \lg P_{\max} / P_{\text{cp}}$, середня потужність відліків дорівнює

$$P_{\text{cp}} = M^2 \delta^2 / 4 \cdot 10^{0,1\Pi}. \quad (4.57)$$

З урахуванням (4.55) і (4.57) значення $h_{\text{кв}}$ становить

$$h_{\text{кв}} = P_{\text{cp}} / P_{\text{шкв}} = 3M^2 / 10^{0,1\Pi}. \quad (4.58)$$

Відновлення безперервного вимірального сигналу за його відліками здійснюється за допомогою ФНЧ з частотою зрізу $F_0 = F_B$. ФНЧ обмежує і спектр шумів квантування. У смузі частот $[0, 0,5F_B]$ відношення $P_{\text{cp}} / P_{\text{шкв}}$ визначається за (4.58). На виході ФНЧ потужність сигналу не зміниться, а потужність шуму квантування зменшиться в $F_d / 2F_B$ разів, тому

$$h_{\text{кв}} = 3 \cdot M^2 F_d / 2 \cdot 10^{0,1\Pi} F_B. \quad (4.59)$$

З виразу (4.59) визначають необхідне число рівнів квантування АЦП

$$M = 0,816 \cdot 10^{0,05\Pi} \sqrt{h_{\text{кв}} F_B / F_d}. \quad (4.60)$$

Кодування сигналів. Квантовані відліки можна передавати у різні способи. Найчастіше використовують кодові комбінації, кажна з яких відповідає певному рівню квантування.

Цифрове кодування – це виконується без втрати інформації операція відображення розміру квантованого сигналу цифровим кодом. Код – це сукупність символів і правил їх застосування для подання, передачі, опрацювання і зберігання інформації. Кодування виконується з метою підвищення ефективності інформаційних процесів, захисту інформації від спотворення і пошкодження її цілісності. По суті кодування – це правило, яке встановлює відповідність між M елементами двох різних множин. Всі елементи першої множини – квантовані відліки, можна розташувати у певному порядку і кожному приписати свій номер – число з натурального ряду чисел від 0 до $M-1$. Для чого ж потрібне таке додаткове перетворення як цифрове кодування?

Справа в тому, що перешкоди і шуми в каналах передачі цифрових кодів здатні змінити значення сигналу, що передається, утворюючи в околі цього значення область невизначеності. Щоби надійно розрізнити суміжні значення необхідно, аби області невизначеності цих значень не перетиналися. Кодування власне і застосовується для виконання цієї умови. Крім того, перехід до числових множин дозволяє застосувати більш досконалі методи цифрового оброблювання сигналів, істотно розширити функціональні можливості і підвищити метрологічні характеристики ІВС.

Чим більше рознесені між собою передані значення сигналу тим з більшою ймовірністю вони будуть правильно ідентифіковані за результатами

приймання. Тому за однакової потужності імпульсів для передавання символів кодів і швидкості їх передавання використовують сигнали з мінімальним значенням m основи системи числення – двійкові і трійкові. Для передачі M чисел за допомогою m -их сигналів формується множина n -розрядних кодових комбінацій, де n визначається з нерівності $m^{(n-1)} < M \leq m^n$.

Вибір m здійснюють з урахуванням простоти, економічності і зручності реалізації кодового перетворення. Найпоширенішими є двійкові натуральні рівномірні коди ($m = 2$), в яких кожна кодова група (комбінація) складається з n кодових символів, кожен з яких приймає одне з двох значень – "1" або "0". Кодова група відповідає запису номера переданого квантованого рівня у двійковій системі числення. Структура кодової групи визначається виразом

$$A = \sum_{i=1}^N a_i 2^{i-1}, \quad (4.61)$$

де a_i – кодовий символ ($a_i = 1; 0$).

Кількість необхідних елементів кодової комбінації визначається кількістю $M = 2^N$ рівнів квантування, тоді $N \geq [\log_2 M]^+$. (4.62)

В АЦП частіше використовують двійковий натуральний код і код Грея.

Код Грея відрізняється двома особливостями, які сприяють підвищенню швидкодії кодування: будь-які дві кодові комбінації, що відповідають суміжним рівням квантування, відрізняються одна від одної тільки в одному двійковому розряді, а зміна значень елементів в кожному розряді у випадку переходу від однієї кодової комбінації до іншої відбувається вдвічі рідше.

Особливості різних способів аналого-цифрового перетворення розглянемо на прикладах їх реалізації. Властивості АЦП та можливості їх використання у вимірювальних каналах ІВС визначаються параметрами: розрядністю, часом перетворення (частотою перетворення), числом аналогових входів, типом інтерфейсу для передавання цифрових даних, характеристиками точності перетворення. За поступового зростання вхідного аналогового сигналу $U_{вх}(t)$ від 0 до номінального значення U_n , що відповідає повній шкалі АЦП, вихідний цифровий код $D(iT_d)$ утворює ступінчасту функцію (рис. 4.23).

Таку залежність називають *характеристикою перетворення АЦП*. За відсутності апаратних похибок перетворення середні точки сходинок розташовуються на ідеальній прямій (лінія 1 на рис. 4.23), яка є ідеальною характеристикою перетворення. Реальна характеристика відрізняється від ідеальної розмірами і формою сходинок та розташуванням в площині координат. Для кількісного описування відмінностей застосовують статичні і динамічні параметри.

Статичні параметри АЦП. Роздільна здатність – величина, зворотна максимальному числу кодових комбінацій на виході АЦП. Цей параметр виражають у відсотках, розрядах або децибелах. Він характеризує потенційно можливу точність перетворення АЦП.

Наприклад, 12-розрядний АЦП має роздільну здатність 1/4096, або 0,0245% від повної шкали, або 72,25 дБ.

Роздільній здатності відповідає збільшення вхідної напруги АЦП у випадку зміни вихідного коду D_j , де $j \in [0, 2^N]$ – номер ступені, на одиницю молодшого розряду (ОМР). Це збільшення є кроком квантування. Номінальне значення кроку квантування становить $\Delta U_{\text{кв}} = U_{\text{н}} / (2^N - 1)$. Чим більше розрядність АЦП, тим вище його роздільна здатність.

Похибка повної шкали – це відносна різниця між реальним та ідеальним значеннями межі інтервалу перетворення за відсутності зсуву нуля

$$(4.63)$$

Ця похибка за характером є мультипликативною. Іноді вона виражається значенням ОМР.

Похибка зсуву нуля – значення $U_{\text{вх}}$, що відповідає нульовому вихідному коду АЦП. За характером вона є адитивною і виражається формулою

$$(4.64)$$

де $U_{\text{вх.01}}$ – значення вхідної напруги, за якої відбувається перехід вихідного коду з 0 в 1. Ця похибка може наводитись і у відносних одиницях

$$(4.65)$$

Похибки повної шкали та зміщення нуля АЦП зменшують підстроюванням аналогової частини схеми або корекцією алгоритму обчислень.

Нелінійність – максимальне відхилення реальної характеристики перетворення $D(U_{\text{вх}})$ від оптимальної (лінія 2 на рис 4.24). Оптимальна характеристика знаходиться емпірично з умови мінімізації похибки нелінійності.

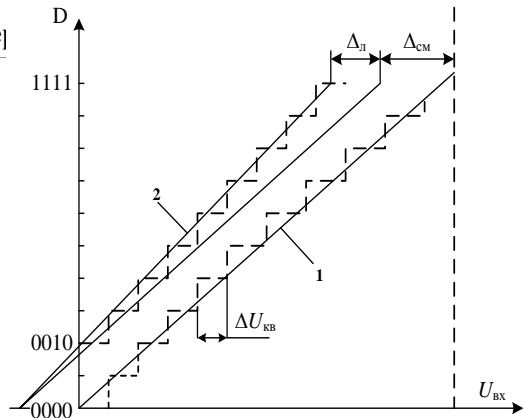


Рис.4.23. Статична характеристика перетворення АЦП

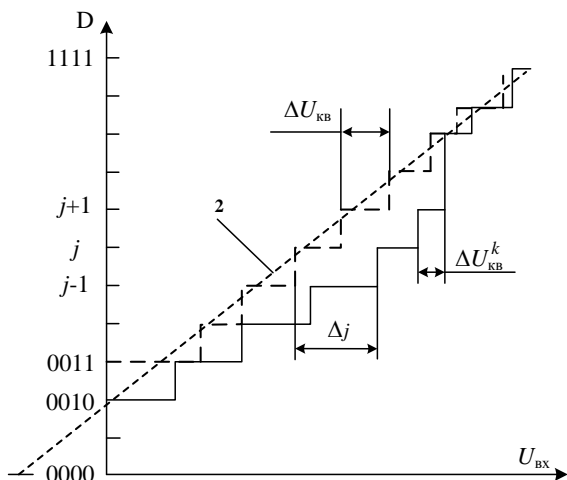


Рис. 4.24. Похибка лінійності характеристики перетворення АЦП

Нелінійність зазвичай визначається у відносних одиницях

$$\delta_{\text{л}} = (\Delta_j / U_{\text{н}}) 100 \% . \quad (4.66)$$

Диференціальною нелінійністю АЦП в певній точці k характеристики перетворення називається різниця між k -тим значенням кванта перетворення $\Delta U_{\text{кв}}^k$ і середнім значенням кванта $\Delta U_{\text{кв}}$.

Значення диференціальної нелінійності наводяться в

частках ОМР або відсотках від повної шкали

$$\delta_{\text{дл}} = \left[(\Delta U_{\text{кв}}^k - \Delta U_{\text{кв}}) / U_{\text{н}} \right] 100 \% . \quad (4.67)$$

Похибка диференційної лінійності визначає дві властивості АЦП: відсутність пропуску кодів і монотонність характеристики перетворення. Перша полягає у здатності АЦП видавати всі можливі вихідні коди за зміни вхідної напруги від 0 до $U_{\text{н}}$. Приклад прогалини j -го коду наведено на рис. 4.24.

Монотонність характеристики перетворення – це незмінність знака збільшення вихідного коду D_j за монотонної зміни вхідного сигналу.

Температурна нестабільність АЦП характеризується температурними коефіцієнтами похибки повної шкали та похибки зсуву нуля.

Динамічні параметри АЦП. Виникнення динамічних попохибок пов'язане з дискретизацією змінюваних в часі сигналів. Для характеристики динамічної точності АЦП застосовують наступні параметри.

Максимальна частота дискретизації – це найбільша частота, з якою може виконуватися перетворення вибіркового значення сигналу за збереження параметрів АЦП в заданих межах, вимірюється кількістю вибірок за секунду.

Час перетворення – це часовий інтервал, обмежений початком імпульсу дискретизації або початком перетворення і моментом появи

Розділ 4. Перетворення сигналів

на виході АЦП стійкого коду, що відповідає даній вибірці. Для одних АЦП (наприклад, АЦП послідовного рахунку) ця величина є змінною і залежить від значення вхідного сигналу. Для інших АЦП (наприклад, паралельних або послідовно-паралельних) цей час не залежить від вхідного сигналу. Для АЦП без пристрою вибірки-зберігання (ПВЗ) час перетворення є апертурним часом.

Час вибірки (стробування) – часовий інтервал, протягом якого утворюється одне вибіркове значення. Для АЦП без ПВЗ дорівнює часу перетворення АЦП.

Апертурна невизначеність аналого-цифрового перетворення. У АЦП об'єднані операції дискретизації сигналів в часі та їх квантування за рівнями. За рахунок кінцевого часу одного перетворення і невизначеності моменту його закінчення не вдається отримати однозначне відповідність між значеннями відліків і моментами часу, до яких вони приписуються. Тому під час роботи зі змінюваними в часі сигналами виникає динамічна похибка. Для її оцінювання введено поняття *апертурної невизначеності*, що характеризується апертурним часом.

Апертурним часом t_a називають часовий інтервал, протягом якого зберігається невизначеність між значенням відліку і моментом часу до якого він приписується. Ефект апертурної невизначеності проявляється або як похибка миттєвого значення сигналу у заданих моментах вимірювання, або як похибка моменту часу, в який здійснюється вимірювання за певного миттєвого значення сигналу. Ефект апертурної невизначеності можна уявити, як "тремтіння" моментів часу, в які беруться відліки сигналу відносно рівновіддалених моментів часу.

Для визначення аналітичного виразу апертурної похибки розкладемо вираз для вихідного сигналу в ряд Тейлора в околі j -ого відліку

$$s(t) = s(t_j) + t_a s'(t_j) + (t_a^2/2) s''(t_j) + \dots, \quad (4.68)$$

що дає у першому наближенні апертурну похибку

$$\Delta s_a(t_j) \approx t_a s'(t_j). \quad (4.69)$$

Для оцінювання апертурних похибок використовують синусоїдальний тестовий сигнал (рис. 4.25), для якого максимальне відносне значення апертурної похибки становить

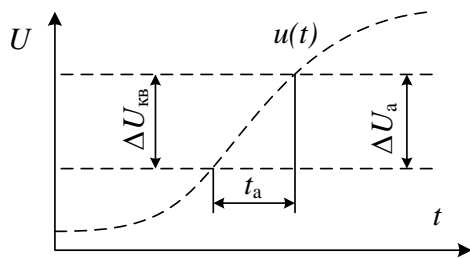


Рис. 4.25. Утворення апертурної похибки за умови її рівності кроку

$$\Delta U_a / U_m = \omega t_a. \quad (4.70)$$

Якщо для N -розрядного АЦП відносна апертурна похибка не перевищує крок квантування, то між частотою сигналу f , апертурним часом t_a та γ_a має місце співвідношення $\gamma_a = 0,5N = ft_a$.

Для дискретизації синусоїдального сигналу частотою 100 кГц з похибкою 1%, час перетворення АЦП не повинен перевищувати 25 нс. У той же час таким швидкодіючим АЦП принципово можна дискретизувати сигнали з шириною спектра ~ 20 МГц. Отже дискретизація сигналів в АЦП призводить до суперечливих вимог щодо його швидкодії і періоду дискретизації. Ця розбіжність може сягати $10^2 \dots 10^3$ разів і суттєво ускладнює і здорожує процес дискретизації, оскільки перетворення навіть порівняно вузькосмугових сигналів вимагає досить швидкодіючих АЦП. Для усунення цього протиріччя до складу АЦП включаються ПВЗ з малим апертурним часом.

Класифікація АЦП. Відомі способи перетворення і відповідні АЦП відрізняються точністю, швидкістю перетворення, складністю апаратної реалізації і інтерфейсами. Класифікація АЦП наведена на рис. 4.26.

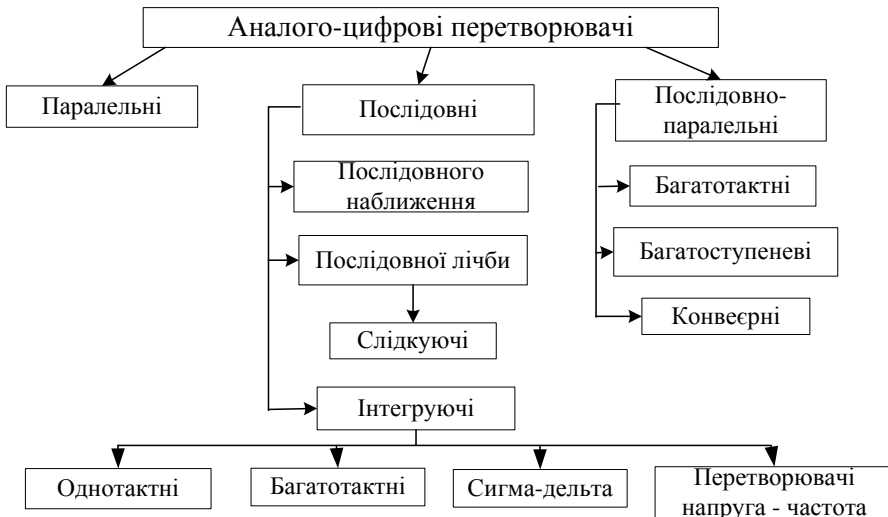


Рис. 4.26. Класифікація АЦП за способам перетворення

Операції квантування і кодування здійснюють за допомогою послідовної, паралельної, послідовно-паралельної процедур наближення цифрового еквівалента до значення аналогової перетворюваної величини. Відповідно до цього розрізняють три групи АЦП.

Розділ 4. Перетворення сигналів

Паралельні АЦП здійснюють квантування сигналу одночасно за допомогою набору компараторів, включених паралельно до джерела вхідного сигналу.

На рис. 4.27 наведено **приклад** структури паралельного АЦП для перетворення напруги в 3-розрядний вихідний двійковий код. За допомогою трьох двійкових розрядів можна представити вісім різних чисел з нулем включно. Отже, необхідно сім компараторів і сім еквідистантних опорних напруг. Останні утворюються з допомогою резистивного подільника напруги, підключеного до джерела опорного напруги $U_{оп}$.

Якщо прикладена вхідна напруга не виходить за межі діапазону від $5/2 \Delta U_{кв}$ до $7/2 \Delta U_{кв}$, де $\Delta U_{кв} = U_{оп}/7$ – квант напруги, що відповідає ОМР АЦП, то компаратори з 1-го по 3-й встановлюються в стан 1, а компаратори з 4-го по 7-й – в стан 0. Перетворення цієї групи кодів в тризначне двійкове число виконує логічний пристрій, що дістав назву «пріоритетний шифратор».

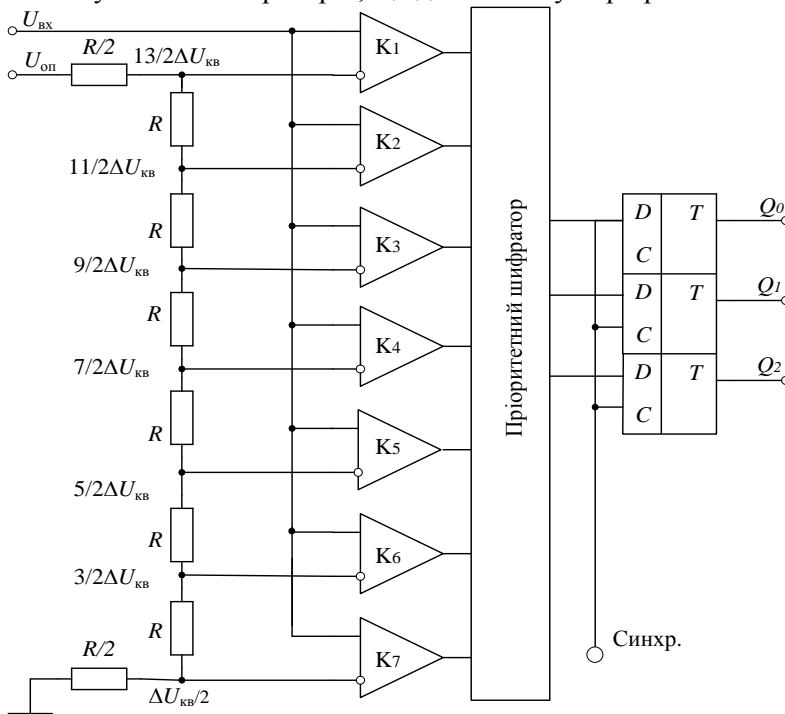


Рис. 4.27. Схема паралельного 3-розрядного АЦП

Діаграма його станів наведена в табл. 4.1.

Таблиця 4.1

Вхідна напруга	Стан компараторів							Виходи		
$U_{вх} / \Delta U_{кв}$	К ₇	К ₆	К ₅	К ₄	К ₃	К ₂	К ₁	Q ₂	Q ₁	Q ₀

0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	1	0	0	1
2	0	0	0	0	0	1	1	0	1	0
3	0	0	0	0	1	1	1	0	1	1
4	0	0	0	1	1	1	1	1	0	0
5	0	0	1	1	1	1	1	1	0	1
6	0	1	1	1	1	1	1	1	1	0
7	1	1	1	1	1	1	1	1	1	1

Якщо вхідна напруга збільшується, компаратори встановлюються в стан 1 по черзі – від низу до верху. Така черговість не гарантується у випадку швидкого наростання вхідного сигналу, оскільки через відмінності в часі затримки компаратори можуть спрацювати в іншому порядку.

Пріоритетне кодування дозволяє уникнути можливої помилки за рахунок того, що одиниці в молодших розрядах не беруться до уваги. Підключення пріоритетного шифратора безпосередньо до виходу АЦП може призвести до помилкового результату зчитування вихідного коду внаслідок різного часу затримки вихідних сигналів. Для виключення такого варіанту до складу АЦП включений вихідний регістр на *D*-тригерах.

Завдяки одночасній роботі компараторів паралельні АЦП мають найбільшу швидкодію. Для зменшення t_a на вході паралельних АЦП застосовують ПВЗ, що забезпечують час вибірки до 0,1 нс. Недоліком таких АЦП є висока складність: *N*-розрядний АЦП містить $2N - 1$ компараторів і $2N$ прецизійних резисторів. Наслідком цього є висока споживана потужність.

Послідовні АЦП перетворюють аналоговий сигнал послідовно біт за бітом. Найпоширенішими серед них є АЦП послідовного наближення (або АЦП з порозрядним зрівноважування). Структура такого АЦП наведена на рис. 4.28,а, а на рис. 4.28,б – діаграма, що пояснює принцип зрівноважування.

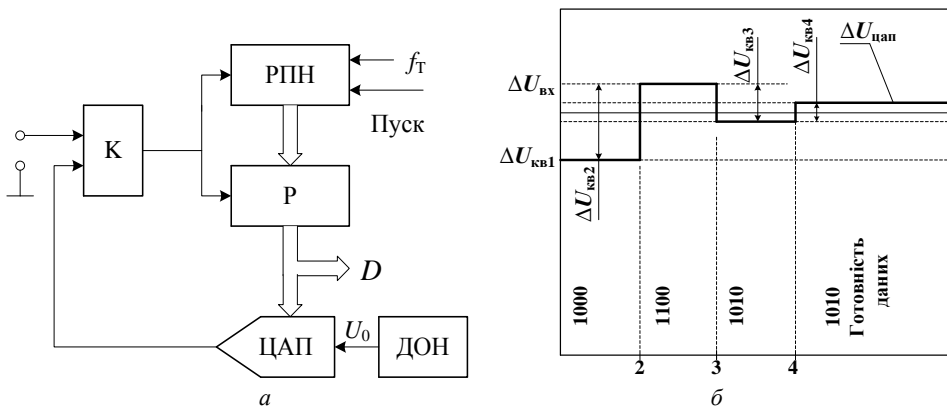


Рис. 4.28. Структурна схема АЦП послідовного наближення (а) та діаграма процесу зрівноважування 4-розрядного АЦП (б)

Розділ 4. Перетворення сигналів

В основу роботи послідовних АЦП покладено принцип дихотомії, тобто послідовного порівняння вимірюваної величини з $1/2$, $1/4$, $1/8$ і т.п. частинами можливого максимального значення. Це дозволяє в N -розрядному АЦП виконати весь процес перетворення за N послідовних кроків (ітерацій) замість $2N - 1$ крок у випадку використання послідовного рахунку.

Після подачі команди "Пуск" з приходом першого тактового імпульсу регістр послідовного наближення (РПН) примусово подає на вхід ЦАП код, що дорівнює половині його шкали (для 4-розрядного ЦАП це код $1000_2 = 8_{10}$). Вихідна напруга ЦАП на першому кроці наближення (рис. 4.28,б) становить $U_{\text{ЦАП}}(1) = \Delta U_{\text{кв}1} = 2^3 \Delta U_{\text{кв}4}$, де $\Delta U_{\text{кв}i}$ – ступінь квантування на i -му кроці наближення, $\Delta U_{\text{кв}4}$ відповідає ОМР. Величина $\Delta U_{\text{кв}1}$ становить половину діапазону перетворюваних напруг. Якщо $U_{\text{вх}} > U_{\text{ЦАП}}(1)$, на виході компаратора встановлюється 1, в іншому випадку 0. За нульового значення РПН повертає старший розряд a_3 в стан нуля. Потім залишок $U_{\text{вх}} - a_3 2^3 \Delta U_{\text{кв}4}$ в такий же спосіб порівнюється з найближчим молодшим розрядом і т.д. Після чотирьох подібних кроків РПН формує двійкове число, з якого після цифро-аналогового перетворення отримують напругу, яка з точністю до $\pm 0,5\Delta U_{\text{кв}4}$ відповідає напрузі $U_{\text{вх}}$. Отриманий результат зберігається у регістрі Р.

Другий тип послідовних АЦП – це перетворювачі *послідовної лічби*, типовим представником яких є слідкуючі АЦП (рис. 4.29).

Генератор імпульсів (ГІ) формує послідовність тактових імпульсів частотою f_T . Ці імпульси схемою управління (СУ) спрямовуються на входи реверсивного лічильника імпульсів (РСІ). Кожен імпульс змінює поточний стан РСІ на +1 або -1. Напрямок лічби визначає компаратор (К): якщо $U_{\text{вх}} > U_{\text{ЦАП}}$, то вміст РСІ збільшується, якщо $U_{\text{вх}} < U_{\text{ЦАП}}$ – зменшується.

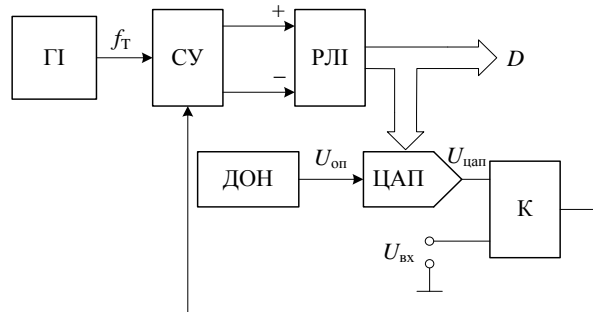


Рис. 4.29. Структурна схема слідкуючого АЦП

Вихідний код РСІ перетворюється ЦАП в напругу, яка безперервно відстежує зміну $U_{\text{вх}}$ з точністю до половини ступеня квантування. Рівняння перетворення такого АЦП задається виразом

$$U_{\text{вх}} = U_{\text{ЦАП}} = U_{\text{оп}} D / 2^N, \quad (4.71)$$

де N – розрядність ЦАП.

Недоліком послідовних АЦП є низька завадостійкість. Дійсно, вибірка миттєвого значення вхідної напруги зазвичай включає доданок у вигляді миттєвого значення завади. Цей недолік значною мірою усунений в *інтегруючих* АЦП, в яких вхідний сигнал інтегрується безперервно або на певному часовому інтервалі, тривалість якого кратна періоду завади. Це дозволяє заглушити заваду на етапі перетворення.

АЦП *багатотактного інтегрування*. Як **приклад** розглянемо АЦП двотактного інтегрування, спрощена схема якого наведена на рис. 4.30. Перетворення здійснюється в два етапи – інтегрування і лічби. На першому перемикач SW_1 замкнутий, а SW_2 розімкнений. Інтегратор (І) інтегрує вхідну напругу $U_{вх}$. Час інтегрування t_1 є постійним. Лічильник (ЛІ) працює як таймер з коефіцієнтом перерахунку $K_{ЛІ}$, тому $t_1 = K_{ЛІ} / f_T$. По завершенню інтегрування вихідна напруга інтегратора становить

$$U_i(t_1) = (1/RC) \int_0^{t_1} U_{вх}(t) dt = -U_{вх,ср} K_{ЛІ} / f_T RC, \quad (4.72)$$

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

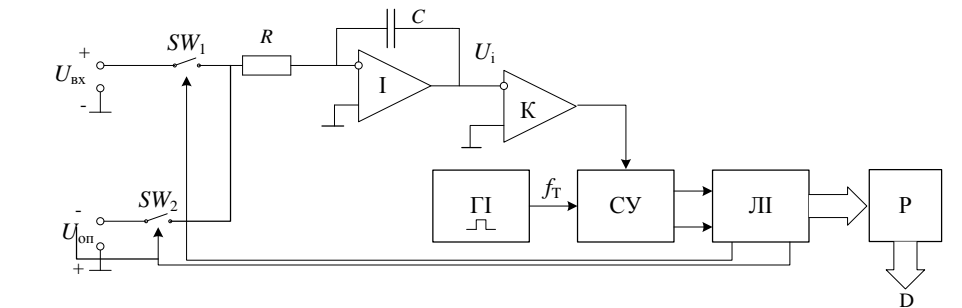


Рис. 4.30. Структурна схема АЦП двотактного інтегрування

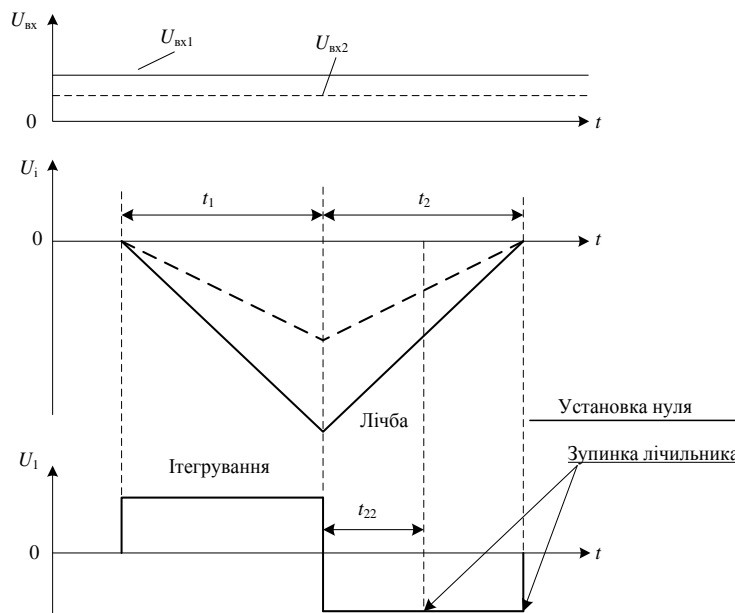


Рис. 4.31. Часові діаграми роботи АЦП двотактного інтегрування

На етапі лічби напруга $U_{оп}$ (протилежна за знаком напрузі $U_{вх}$) надходить на вхід інтегратора. Вихідна напруга інтегратора лінійно зменшується за абсолютною

Розділ 4. Перетворення сигналів

величиною (рис.4.31). Етап лічби закінчується в момент t_2 , для якого виконується умова $U_i(t_2)=0$, що фіксує компаратор К і зупиняє лічбу. Тривалість етапу лічби визначається з рівняння

$$U_i(t_1) + (1/RC) \int_{t_1}^{t_1+t_2} U_{оп} dt = 0 \quad (4.73)$$

В ЛІ накопичується код: $D = U_{вх.сер} K_{ЛІ} / U_{оп} . \quad (4.74)$

З (4.74) випливає, що ні тактова частота, ні постійна інтегрування RC не впливають на результат. Необхідно тільки забезпечити незмінність тактової частоти протягом часу $t_1 + t_2$. В (4.74) входить не миттєве напруга, а $U_{вх.сер}$. Тому завада послаблюється тим сильніше, чим вище її частота. Коефіцієнт передачі завади $K_{П}$ для такого АЦП визначається формулою

$$K_{П} = \left| \sin^2 \pi f t_1 / \pi f t_1 \right|, \quad (4.75)$$

тобто змінна напруга, період якої в ціле число разів менше t_1 , заглушується повністю (рис. 4.31, 4.32), отже тактову частоту доцільно вибирати такою, щоби добуток твір $K_{ЛІ} f$ дорівнював або був кратним періоду завади.

Сигма-дельта АЦП. Своєю назву вони отримали завдяки наявності в їх складі двох блоків: суматора (позначення операції – Σ) і інтегратора (позначення операції – Δ).

Основним принципом роботи АЦП є усереднення результатів вимірювання на великому інтервалі часу, що дозволяє зменшити внесену шумами похибку перетворення і збільшити роздільну здатність АЦП.

Основні вузли АЦП – сигма-дельта модулятор і цифровий фільтр. Структура сигма-дельта модулятора першого порядку наведена на рис. 4.33. Порядок модулятора визначається кількістю в його схемі інтеграторів і суматорів. Робота

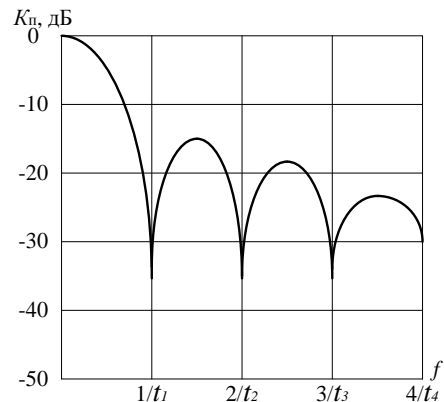


Рис. 4.32. Частотна характеристика коефіцієнта заглушення завад в АЦП двотактного інтегрування

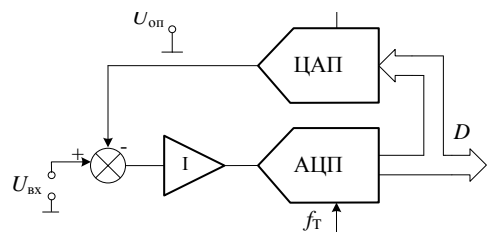


Рис. 4.33. Структура сигма-дельта модулятора

схеми ґрунтується на відніманні з $U_{вх}$ сигналу з виходу ЦАП,

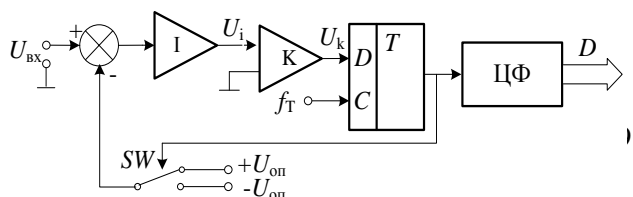


Рис. 4.34. Структурна схема однобітового сигма-дельта АЦП

отриманого в попередньому такті роботи. Ця різниця інтегрується в інтеграторі (І), а потім перетворюється в код паралельним АЦП невисокої розрядності. Послідовність кодів D надходить на цифровий ФНЧ (на рис. 4.33 відсутній). Структура однобітового сигма-дельта АЦП наведена на рис. 4.34, а принцип його функціонування пояснює табл. 4.2 на прикладі перетворення постійної напруги $U_{вх} = 0,6В$ (0В), якщо $U_{оп} = 1В$.

Нехай постійна часу інтегрування дорівнює періоду тактових імпульсів. До початку перетворення вихідна напруга І встановлюється в нуль. На виході ключа SW – однобітового ЦАП, також установлюється нульова напруга. Далі робота здійснюється як послідовність станів (табл. 4.2).

Таблиця 4.2

$U_{вх} = 0,6 В$					$U_{вх} = 0 В$				
№ такта	U_{Σ} , В	U_i , В	U_k , біт	$U_{ЦАП}$, В	№ такта	U_{Σ} , В	U_i , В	U_k , біт	$U_{ЦАП}$, В
1	2	3	4	5	6	7	8	9	10
1	0,6	0,6	1	1	1	1	1	1	1
2	-0,4	0,2	1	1	2	-1	0	0	-1
3	-0,4	-0,2	0	-1	3	1	1	1	1
4	1,6	1,4	1	1	4	-1	0	0	-1
5	-0,4	1,0	1	1	5	1	1	1	1
6	-0,4	0,6	1	1	6	-1	0	0	-1
7	-0,4	0,2	1	1	7	1	1	1	1
8	-0,4	-0,2	0	-1	8	-1	0	0	-1
9	1,6	1,4	1	1	9	1	1	1	1
1	2	3	4	5	6	7	8	9	10
10	-0,4	1,0	1	1	10	-1	0	0	-1
11	-0,4	0,6	1	1	11	1	1	1	1
12	-0,4	0,2	1	1	12	-1	0	0	-1
13	-0,4	-0,2	0	-1	13	1	1	1	1
14	1,6	1,4	1	1	14	-1	0	0	-1
15	-0,4	1,0	1	1	15	1	1	1	1
16	-0,4	0,6	1	1	16	-1	0	0	-1

Тут використано 4-розрядний лічильник. Підрахунок біт на виході компаратора за 16-тактний цикл дає число 13. Якщо $U_{вх}=1В$ на виході компаратора завжди буде присутня одиниця, що дає за повний цикл перетворення (16 тактів) число 16, тобто переповнення лічильника. Якщо $U_{вх} = -1В$ на виході компаратора завжди буде нуль, що дає нульове число в кінці циклу. Якщо $U_{вх}=0$ то результат лічби за цикл складе 8_{10} або 1000_2 . Це означає, що число на виході АЦП представлено у зміщеному коді. Верхня межа повної шкали складе 1111_2 або $+7_{10}$, а нижня – 0000_2 або -8_{10} . Для $U_{вх}=0,6В$ вміст лічильника складе 13_{10} у зміщеному коді, що відповідає числу $+5_{10}$ в прямому коді. З огляду на, що $+8_{10}$ відповідає $U_{вх}=1В$, результат перетворення – $5 \times 1/8 = 0,625В > 0,6В$.

Послідовно-паралельні АЦП є результатом компромісу між високою швидкістю і меншою ціною пристрою. Такі АЦП займають проміжне положення за роздільною здатністю і швидкістю між

Розділ 4. Перетворення сигналів

паралельними АЦП і АЦП послідовного наближення. АЦП цього класу поділяються на багатоступеневі, багатотактні і конвеєрні. У багатоступеневому АЦП процес перетворення вхідного сигналу розділений у просторі. На рис. 4.35 представлена схема двоступеневого 8-розрядного АЦП. АЦП1 здійснює грубе перетворення сигналу в чотири старших розряди вихідного коду. Цифровий код з виходу АЦП1 надходить на вихідний регістр Р і одночасно на вхід 4-розрядного швидкодіючого ЦАП. Залишок напруги після віднімання вихідної напруги ЦАП з вхідної надходить на АЦП2, опорна напруга якого в 16 разів менше, ніж у АЦП1. Як наслідок, квант АЦП2 в 16 раз менший за квант АЦП1. Перетворений АЦП2 в цифрову форму залишок напруги має чотири молодших розряди вихідного коду.

Швидкодію багатоступеневих АЦП можна підвищити застосуванням конвеєрного принципу багатоступеневого опрацювання вхідного сигналу. У звичайному багатоступеневому АЦП (рис.

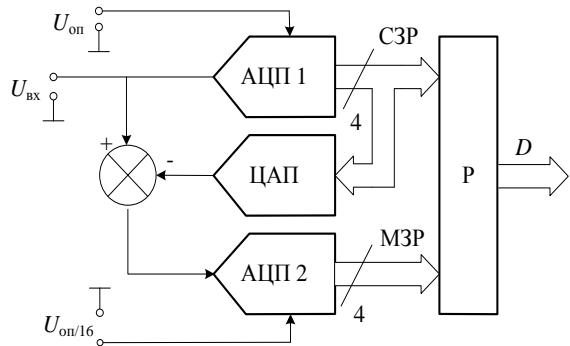


Рис. 4.35. Структурна схема двоступеневого 8-розрядного послідовно-паралельного АЦП

4.35) спочатку формуються старші розряди вихідного коду АЦП1, після чого має місце період установлення вихідного сигналу ЦАП. На цьому інтервалі АЦП2 простоє. На другому етапі, під час перетворення залишку в АЦП2, простоє АЦП1. Ввівши елементи затримки аналогового і цифрового сигналів між ступіннями перетворювача, отримують конвеєрний АЦП. Схема 8-розрядного варіанту такого АЦП наведена на рис. 4.36.

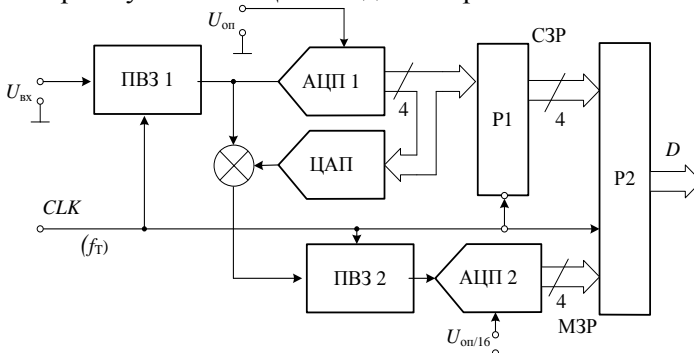


Рис. 4.36. Структурна схема конвеєрного 8-розрядного АЦП

Конвеєрна архітектура дозволяє в кілька разів підвищити максимальну частоту вибірок багатоступеневого АЦП і, без програшу у

швидкодії, збільшити число ступенів АЦП одночасно зменшивши розрядність кожного ступеня і знизити складність АЦП.

Сполучення АЦП з цифровими процесорами. Важливу частину АЦП становить цифровий інтерфейс, тобто схема, що забезпечує зв'язок АЦП з приймачем цифрових кодів. Структура цифрового інтерфейсу визначає спосіб підключення АЦП до приймача-мікропроцесора, мікроконтролера, цифрового процесора сигналів або ПК. Властивості цифрового інтерфейсу впливають на рівень верхньої межі частоти перетворення АЦП.

Найчастіше застосовують спосіб зв'язку, за якого АЦП є для процесора однією коміркою пам'яті. Для цього АЦП забезпечується необхідним числом адресних входів, дешифратором адресу і підключається безпосередньо до адресної шини і шини даних процесора. Для цього він повинен мати вихідні каскади з трьома станами. В АЦП використовують як паралельний так і послідовний інтерфейси.

Системи збирання даних. Функціональне ускладнення АЦП, необхідність опрацьовувати одночасно значну кількість сигналів призвели до розробки однокристальних систем збирання даних, які забезпечують перетворення аналогових сигналів в цифровий код від багатьох сенсорів. Структура системи збирання даних наведено на рис. 4.37.

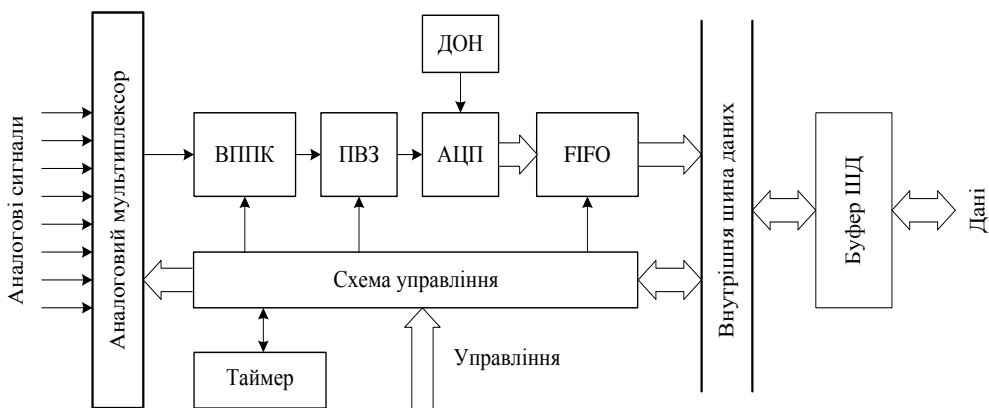


Рис. 4.37. Узагальнена структурна схема системи збирання даних

Основу системи складає АЦП. Для підключення декількох джерел вхідних аналогових сигналів використовується аналоговий мультиплексор. Для зменшення частоти переривань головного процесора в систему збирання даних включено оперативний запам'ятовуючий пристрій типу FIFO (First Input – First Output:

Розділ 4. Перетворення сигналів

першим прийшов – першим вийшов). Вимірювальний підсилювач з програмованим коефіцієнтом підсилення змінює коефіцієнт підсилення по команді зі схеми управління. Це дозволяє вирівняти діапазони аналогових сигналів з різних входів. Цифровий таймер визначає темп перетворення сигналів в АЦП.

4.3. Цифро-аналогове перетворення

Цифро-аналогове перетворення – це перетворення цифрового коду в аналогову величину (струм, напруга, частота, інтервал часу, фазовий зсув сигналів і т.п.). Найчастіше використовують перетворення цифрових кодів в електричний струм або напруга, тому далі під аналого-цифровим розумітимемо перетворення коду в напругу або струм. Пристрої, що реалізують таке перетворення, називають цифроаналоговими перетворювачами (ЦАП). У загальному вигляді ЦАП можна уявити як керований цифровим кодом потенціометр. Якщо до керованого N -розрядним двійковим кодом потенціометра підключити ДОН з напругою $U_{\text{оп}}$, то на виході ЦАП отримаємо напругу

$$U_{\text{вих}}(D) = U_{\text{оп}} D / 2^N, \quad D \in [0, 2^N - 1]. \quad (4.76)$$

З (4.76) випливає: ЦАП формує вихідну напругу $U_{\text{вих}}(D) \in [0, U_{\text{оп}})$ як частку $U_{\text{оп}}$, лінійно пов'язану з керуючим кодом D ; за монотонного збільшення D вихідна напруга ЦАП змінюється східчасто з кроком $2^{-N} U_{\text{оп}}$; ЦАП реалізує операцію множення напруги на код – $U_{\text{оп}} D$.

Параметри ЦАП. Якщо послідовно збільшувати значення вхідного цифрового коду N -розрядного ЦАП від 0 до $2^N - 1$, вихідний сигнал утворить ступінчасту криву, яка називається *характеристикою перетворення ЦАП*. За відсутності апаратних похибок середні точки сходинок розташовуються на ідеальній прямій (пряма 1 на рис. 4.38). Реальна характеристика перетворення (крива 2) істотно відрізняється від ідеальної розмірами, формою ступенів та зсувом відносно нульового коду.

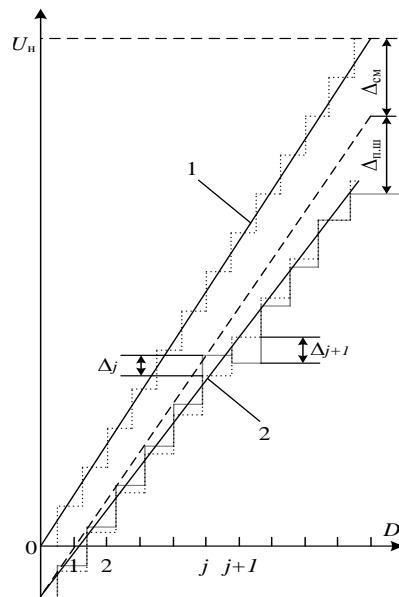


Рис. 4.38. Статична характеристика перетворення ЦАП

Кількісне описання цих розбіжностей характеризується рядом параметрів.

Статистичні параметри. Роздільна здатність – приріст напруги $U_{\text{вих}}$ внаслідок зміни вхідного коду на ОМР, який дорівнює кроку квантування. Для двійкових кодів номінальне значення кроку квантування становить $\Delta U_{\text{кв}} = U_{\text{н}} / (2^N - 1)$, (4.77)

де $U_{\text{н}}$ – номінальна (максимальна) вихідна напруга ЦАП (напруга повної шкали).

Чим більше розрядність ЦАП, тим вище його роздільна здатність.

Похибка повної шкали – відносна різниця між реальним і ідеальним значеннями меж шкали перетворення за відсутності зміщення нуля

$$\gamma_{\text{п.ш}} = (\Delta_{\text{п.ш}} / U_{\text{н}}) 100\% , \quad (4.78)$$

де $\Delta_{\text{п.ш}}$ – абсолютна різниця між реальними і ідеальними значеннями меж шкали перетворення. Ця похибка за характером є *мультиплікативною*.

Похибка зсуву нуля – відносно значення $U_{\text{вих}}$ за умови $D = 0$

$$\delta_{\text{зм}} = (\Delta_{\text{зм}} / U_{\text{н}}) 100\% , \quad (4.79)$$

де $\Delta_{\text{зм}}$ – абсолютне зміщення нуля. Це *адитивна* складова похибки ЦАП.

Нелінійність – максимальне відхилення Δ_j реальної характеристики перетворювача $U_{\text{вих}}(D)$ від ідеальної. Ідеальна характеристика знаходиться емпіричним шляхом з умови мінімізації значень похибки нелінійності. Її прийнято визначати у відносних одиницях

$$\gamma_{\text{н}} = (\Delta_j / U_{\text{н}}) 100\% . \quad (4.80)$$

Диференціальна нелінійність – це максимальна зміна (з урахуванням знаку) відхилення реальної характеристики перетворення $U_{\text{вих}}(D)$ від оптимальної при переході від одного значення вхідного коду до іншого, суміжного з ним. Цю похибку прийнято визначати у відносних одиницях

$$\delta_{\text{д.н}} = \left[(\Delta_j + \Delta_{j+1}) / U_{\text{н}} \right] 100\% . \quad (4.81)$$

Монотонність характеристики перетворення – збільшення/зменшення $U_{\text{вих}}(D)$ за збільшення/зменшення D . Якщо $\delta_{\text{д.н}} > \Delta U_{\text{кв}} / U_{\text{н}}$, характеристика ЦАП є немонотонною.

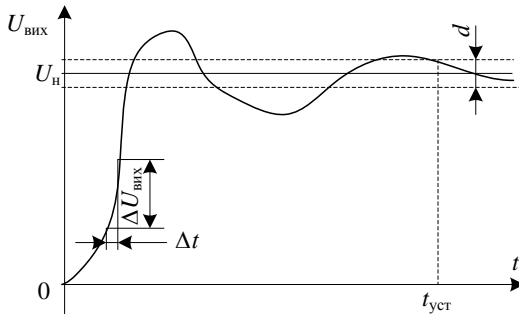


Рис. 4.39. Перехідна характеристика ЦАП

Динамічні параметри визначаються за результатами аналізу функції $U_{\text{вих}}(t)$ в разі максимального [стрибкоподібно] зміни коду D (зазвичай від 0 до $2^N - 1$). Перехідна характеристика ЦАП зображена на рис. 4.39.

Час встановлення – це інтервал часу від $t = 0$ до моменту $t_{\text{уст}}$, коли після зміни вхідного коду в останній раз виконується рівність

$$|U_{\text{вих}} - U_n| = 0,5\Delta U_{\text{кв}}.$$

Швидкість наростання – це максимальна швидкість зміни $U_{\text{вих}}(t)$ під час перехідного процесу, що визначається як відношення приросту $\Delta U_{\text{вих}}$ до інтервалу часу Δt , за який відбувався цей приріст.

Шуми ЦАП викликають різні фізичні процеси, що мають місце у напівпровідникових пристроях. Для оцінювання якості ЦАП використовують середньоквадратичне значення шуму. Цей параметр зазвичай вимірюється у $\text{нВ}/(\text{Гц})^{1/2}$ в заданій смузі частот.

Викиди (імпульсні завади) – це короткі сплески чи провали вихідної напруги, що виникають під час зміни значень вхідного коду за рахунок несинхронності розмикання/замикання аналогових ключів у різних розрядах ЦАП (наприклад, у випадку зміни коду 011...111 до значення 100...000). Викиди оцінюються за їх площею (у пВ·с). Ефективним способом заглушення викидів є використання ПВЗ у складі ЦАП.

Типи ЦАП. Класифікація ЦАП приведена на рис. 4.40.

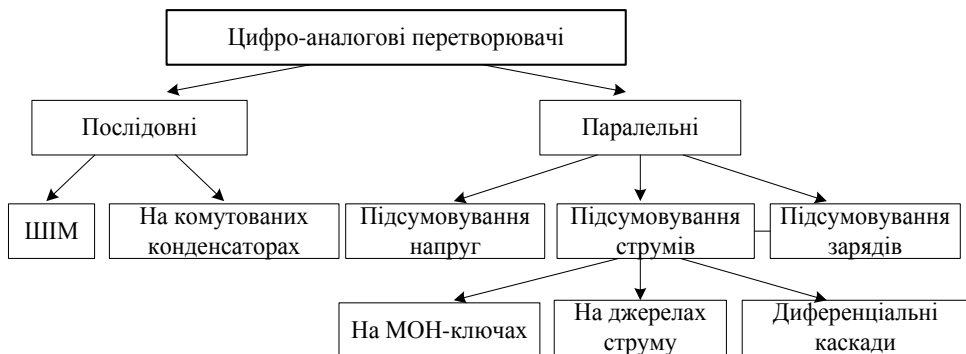


Рис. 4.40. Класифікація ЦАП за схемо технічними ознаками

Послідовні ЦАП поділяються на ЦАП з широтно-імпульсною модуляцією (ШІМ) і ЦАП на конденсаторах, що перемикаються.

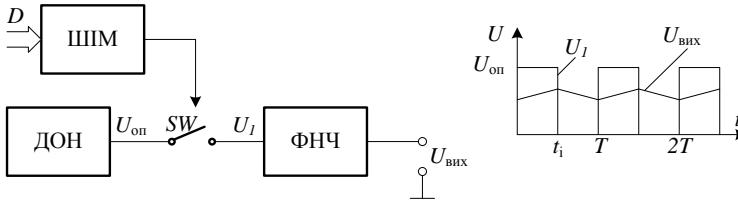


Рис. 4.41. Структура ЦАП з ШІМ (а) та графік вихідної напруги (б)

У ЦАП з ШІМ (рис. 4.41) вихід останнього управляє ключем SW , який комутує постійну напругу з ДОН. Залежно від розрядності коду D формується послідовність імпульсів з відносною тривалістю $\gamma = T/t_i = D/2^N$.

ФНЧ згладжує імпульси і виділяє середнє значення напруги

$$U_{\text{внх}} = \gamma U_{\text{оп}} = DU_{\text{оп}}/2^N. \quad (4.82)$$

Розглянута структура ЦАП проста в реалізації, має близьку до лінійної функцію перетворення і невисоку швидкодію.

У послідовному ЦАП на конденсаторах, що перемикаються, цифро-аналогове перетворення виконується за значно менший час. Принцип роботи таких ЦАП ілюструє рис. 4.42. У цій схемі $C_1 = C_2$. Перед початком циклу перетворення конденсатор C_2 розряджається ключем SW_4 . Вхідні дані задаються у вигляді послідовного коду.

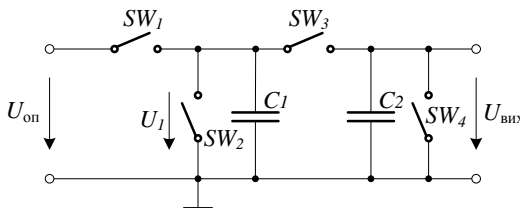


Рис. 4.42. Формування напруги $U_{\text{внх}}$ схемою на конденсаторах, що перемикаються

Перетворення коду в напругу здійснюється послідовно, починаючи з молодшого розряду d_0 . Кожен такт перетворення складається з двох напівтактів. У першому конденсатор d_0

заряджається до опорної напруги $U_{\text{оп}}$ при $d_0 = 1$ за допомогою замикання ключа SW_2 або розряджається до нуля у випадку $d_0 = 0$ шляхом замикання ключа SW_2 . У другому напівтакті за розімкнених ключів SW_1 , SW_2 і SW_4 замикається ключ SW_3 , що викликає розподілення заряду навпіл між C_1 і C_2 , тому маємо: $U_1(0) = U_{\text{внх}}(0) = d_0 U_{\text{оп}}/2$.

Поки на конденсаторі C_2 зберігається заряд, процедура заряду конденсатора C_1 повинна бути повторена для наступного двійкового розряду d_1 вхідного слова. Після нового циклу перезарядки вихідна напруга становить

$$U_{\text{внх}}(1) = U_1(1) = (d_1 + 0,5d_0)U_{\text{оп}}/2 = (2d_1 + d_0)U_{\text{оп}}/4. \quad (4.83)$$

Розділ 4. Перетворення сигналів

Аналогічно виконується перетворення для інших розрядів слова. В результаті для N -розрядного ЦАП вихідна напруга дорівнює

$$U_{\text{вих}}(N-1) = \left(U_{\text{оп}} / 2^N \right) \sum_{k=0}^{N-1} d_k 2^k = U_{\text{оп}} D / 2^N. \quad (4.84)$$

Якщо потрібно зберігати результат перетворення протягом тривалого часу, до виходу схеми підключають ПВЗ. Після закінчення циклу перетворення виконують цикл вибірки і переводять ПВЗ в режим зберігання.

Розглянута схема перетворює вхідний код за $2N$ кроків, що значно менше, ніж в ЦАП з ШІМ. Цей спосіб вимагає тільки наявності двох підібраних конденсаторів невеликої ємності. Проте у цілому послідовні ЦАП значно поступаються паралельним за швидкодією.

Паралельні ЦАП. Переважна більшість схем паралельних ЦАП реалізує ідею підсумовування вагових струмів, значення яких пропорційні вазі відповідних цифрових двійкових розрядів. Схема ЦАП, що реалізує цей принцип наведена на рис. 4.43. Опори резисторів вибирають таким чином, щоб у випадку замкнених ключів через них протікав струм, відповідний вазі розряду. Ключ замикається тоді, коли відповідний йому біт коду D дорівнює одиниці. Вихідний струм визначається рівнянням

$$I_{\text{вих}} = \left(U_{\text{оп}} / R \right) \sum_{k=0}^{N-1} d_k 2^k = U_{\text{оп}} D / R. \quad (4.85)$$

З підвищенням розрядності ЦАП зростають вимоги до точності узгодження струмозадаючих резисторів. Найбільш жорсткі вимоги висуваються до резисторів старших розрядів, оскільки розкид струмів для них не повинен перевищувати струму молодшого розряду. Тому розкид опорів в k -м розряді не повинен перевищувати величину $\Delta R / R = 2^{-k}$. З цієї умови випливає, що розкид опорів резистора, наприклад, в четвертому розряді не повинен перевищувати 3%, а в 10-му – 0,05% і т.п., що ускладнює реалізацію таких ЦАП.

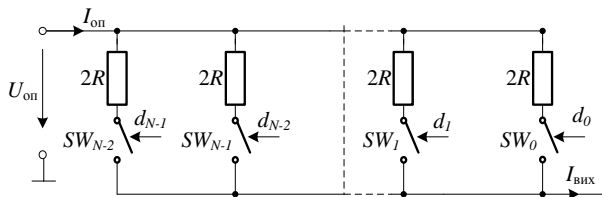


Рис. 4.43. ЦАП з підсумовуванням вагових струмів

Серед інших недоліків ЦАП з підсумовуванням вагових струмів слід вказати наступні: при різних вхідних кодах від ДОН споживається різний струм, що впливає на величину напруги $U_{\text{оп}}$; опори резисторів старших розрядів в багаторозрядних ЦАП можуть бути порівнювані з опором замкнутого ключа, що веде до додаткових похибок перетворення; до розімкнених ключів прикладається значна напруга, що ускладнює їх реалізацію.

Зазначені недоліки усунені в схемі ЦАП на матрицях резисторів типу $R - 2R$. Схема такого ЦАП представлена на рис. 4.44.

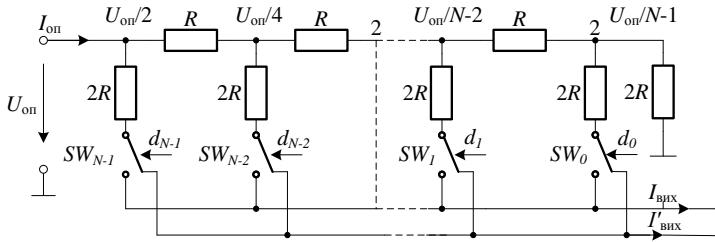


Рис. 4.44. Схема ЦАП на матриці резисторів типу $R - 2R$

Формування вагових коефіцієнтів здійснюється шляхом послідовного розподілу опорної напруги на резистивній матриці постійного імпедансу. Основним елементом матриці є подільник напруги, який задовольняє наступній умові: якщо його навантажити на опір R_n , то його вхідний опір $R_{вх} = R_n$. Коефіцієнт послаблення ланцюга $\alpha = U_2/U_1$ для такого навантаження має певне значення. Для виконання цих умов необхідно, щоб послідовні (R_p) і паралельні (R_s) опори визначались згідно з рівняннями

$$R_p = \alpha R_n / (1 - \alpha), \quad R_s = (1 - \alpha) R_n. \quad (4.86)$$

Значення опорів резисторів на схемі рис. 4.44 отримані для двійкового кодування ($\alpha = 0,5$) і $R_n = 2R$.

Оскільки в будь-якому положенні перемикачів SW_k вони з'єднують нижні виводи резисторів із загальною шиною, ДОН (на схемі відсутній) навантажений на постійний вхідний опір $R_{вх} = R$. Це гарантує незмінність опорної напруги за будь-якого вхідного коду ЦАП.

Згідно рис. 4.44 вихідні струми схеми визначаються співвідношеннями

$$I_{\text{вих}} = (U_{\text{оп}}/2^N R) \sum_{k=0}^{N-1} d_k 2^k = U_{\text{оп}} D/2^N R, \quad I'_{\text{вих}} = (U_{\text{оп}}/2^N R) \sum_{k=0}^{N-1} \bar{d}_k 2^k = U_{\text{оп}} \bar{D}/2^N R, \quad (4.87)$$

а вхідний струм
$$I_{\text{вх}} = I_{\text{вих}} + I'_{\text{вих}} + U_{\text{оп}}/2^N R = U_{\text{оп}}/R. \quad (4.88)$$

Оскільки нижні контакти резисторів $2R$ у будь-якому стані ключів під'єднані до загальної шини через низький опір замкнених ключів, напруги на останніх завжди невеликі (в межах декількох мілівольт). Це спрощує реалізацію ключів, схем управління ними і дозволяє використовувати $U_{\text{оп}}$ різної полярності і в широкому діапазоні значень. Оскільки вихідний струм ЦАП залежить лінійно від $U_{\text{оп}}$, перетворювачі такого типу можна використовувати для перемноження аналогового сигналу (подаючи його замість $U_{\text{оп}}$) на цифровий код. Такі ЦАП називають *перемножуючими*.

Підвищену точність забезпечують ЦАП на джерелах струму, в якому вагові струми задають джерела струму з високими динамічними опорами.

Отримати вихідний сигнал у вигляді напруги в ЦАП з МОН-ключами можна, використавши інверсне включення резистивної матриці (рис. 4.45).

Розділ 4. Перетворення сигналів

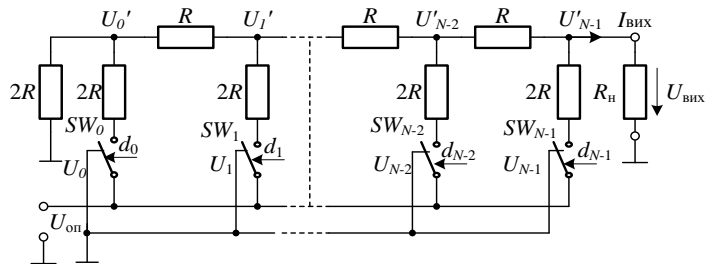


Рис. 4.45. Інверсне включення ЦАП з МОН-ключами

Для розрахунку $U_{\text{вих}}$ знайдемо зв'язок між напругою U_i на ключі SW_i і вузловою напругою U_i' . Скористаємося принципом суперпозиції. Вважатимемо рівними нулю всі напруги на ключах, крім напруги U_i . Якщо $R_H = 2R$ до кожного вузла підключені праворуч і ліворуч однакові навантаження $2R$

$$U_i' = (U_i/2R) / (1/2R + 1/2R + 1/2R) = U_i/3. \quad (4.89)$$

Знайдемо $U_{\text{вих}}$ як загальну напругу на крайньому правому вузлі, викликану сумарною дією всіх напруг U_i . З урахуванням того, що напруги вузлів сумуються з вагою, що визначається резистивною матрицею R-2R, маємо

$$U_{\text{вих}} = (2^{N-1}/3) \sum_{k=0}^{N-1} U_k 2^k = U_{\text{оп}} D/3 \cdot 2^{N-1}. \quad (4.90)$$

Паралельний ЦАП можна реалізувати на комутованих конденсаторах (рис. 4.46). Основою ЦАП такого типу є матриця конденсаторів, ємності яких відносяться як цілі ступені двійки.

Ємність k -го конденсатора матриці визначається виразом

$$C_k = 2^k C_0. \quad (4.91)$$

Цикл перетворення складається з двох фаз. В першій ключі $SW_0 \dots SW_{N-1}$ знаходяться у лівій позиції. Ключ скидання $SW_{\text{зк}}$ замкнений. Всі конденсатори розряджені. У другій фазі ключ $SW_{\text{зк}}$ розмикається. Якщо k -й біт вхідного N -розрядного слова $d_k = 1$, то відповідний ключ SW_k переключиться у праву позицію, підключаючи нижню обкладку конденсатора до ДОН, або залишається в лівій позиції, якщо $d_k = 0$. Сумарний заряд конденсаторів матриці, з урахуванням формули (4.91), становить

$$q = U_{\text{оп}} \sum_{k=0}^{N-1} C_k d_k = U_{\text{оп}} C_0 \sum_{k=0}^{N-1} 2^k d_k = U_{\text{оп}} C_0 D. \quad (4.92)$$

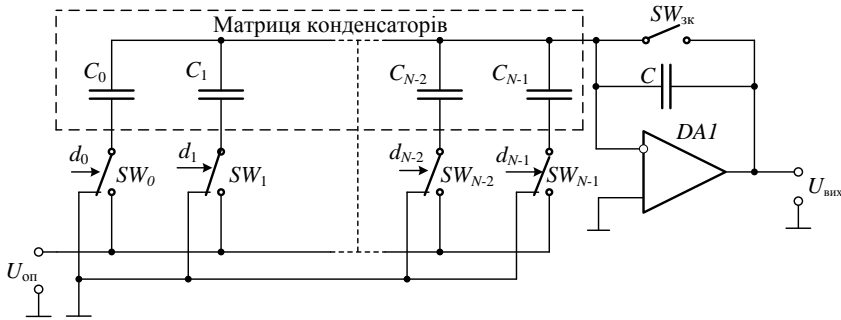


Рис. 4.46. Паралельний ЦАП на комутованих конденсаторах

Таки самий заряд отримує і конденсатор C у ланцюгу зворотного зв'язку ОП ДАІ. Вихідна напруга:
$$U_{\text{вих}} = -q/C = -U_{\text{оп}} C_0 D / C. \quad (4.93)$$

Для збереження $U_{\text{вих}}$ протягом тривалого часу до виходу ЦАП підключають ПВЗ.

Для збереження результату перетворення (постійної напруги) протягом тривалого часу до виходу ЦАП підключають ПВЗ.

Основні висновки

Дискретні сигнали утворюються в результаті часової дискретизації неперервних (аналогових) сигналів. Крок часової дискретизації вибирається, як правило, сталим. Основним математичним апаратом, що використовується при спектральному аналізі дискретних сигналів, є апарат Z -перетворень, що відіграє стосовно дискретних сигналів таку саму роль, як і інтегральне перетворення Фур'є для неперервних сигналів. Основу Z -перетворення становить інтегральне перетворення Лапласа. До Z -перетворення приходять у результаті дискретизації всіх функцій часу в перетворенні Лапласа.

Будь-який сигнал з обмеженим спектром без спотворень можна подати послідовністю його дискретних відліків за умови, що частота дискретизації не менш ніж удвічі перевищує найвищу гармоніку спектра аналогового сигналу.

Дискретне перетворення Фур'є ставить у відповідність скінченновимірній вибірці дискретних сигналів скінченновимірний спектр тієї самої розмірності, що й обсяг вибірки вхідних сигналів. Гармоніки дискретного спектра еквідистантно розміщені на інтервалі від нуля до частоти дискретизації.

В комп'ютеризованих ІВС використовується перетворення аналогового вимірювального сигналу в цифровий код і цифрового коду в аналоговий сигнал. Такі перетворення виконуються аналого-цифровими і цифро-аналоговими перетворювачами (АЦП і ЦАП).

Основними операціями аналого-цифрового перетворення є дискретизація сигналів в часі, їх квантування за рівнем і цифрове кодування. Дискретизація сигналу – це перетворення, внаслідок якого він відображається

Розділ 4. Перетворення сигналів

кінцевою сукупністю своїх значень – відліків, отриманих в дискретні моменти часу. Інтервал часу між сусідніми відліками у випадку рівномірної дискретизації сигналу називається періодом дискретизації.

Квантування сигналів – це процес поділу безперервного діапазону значень перетворюваної величини на кінцеве число суміжних інтервалів та подання будь-якого значення з певного інтервалу одним дозволеним значенням, яке належить цьому ж інтервалу. Квантування сигналу призводить до втрати частини інформації. Різниця між миттєвими значеннями сигналу і дозволеними рівнями називається шумом квантування.

Цифрове кодування – це виконувана без втрати інформації операція відображення розміру квантованого сигналу цифровим кодом. Код – це сукупність символів і правил їх застосування для подання, передачі, опрацювання і зберігання інформації. Кодування виконується з метою підвищення ефективності інформаційних процесів, а також для захисту інформації від спотворення і пошкодження її цілісності.

В АЦП найчастіше використовують двійковий код і код Грея. Код Грея відрізняється двома особливостями, які сприяють підвищенню швидкодії перетворення і заводо захищеності інформації: будь-які дві кодові комбінації для сусідніх рівнів квантування відрізняються одна від одної тільки в одному двійковому розряді; під час переходу від однієї кодової комбінації до іншої зміна значень елементів в кожному розряді відбувається вдвічі рідше.

Основними параметрами АЦП є: розрядність, час перетворення (частота перетворення), число каналів для підключення аналогових сигналів, тип інтерфейсу, характеристики точності перетворення, а також статичні і динамічні параметри. До статичних відносяться: роздільна здатність, похибка повної шкали, похибка зміщення нуля, нелінійність, диференціальна нелінійність, похибка диференціальної лінійності, монотонність характеристики перетворювача, температурна нестабільність. Динамічними параметрами АЦП є: максимальна частота дискретизації, час перетворення, час вибірки й апертурний час. Апертурним часом називають часовий інтервал, протягом якого зберігається невизначеність між значенням відліку і моментом часу, до якого він приписується. Для зменшення апертурного часу до складу АЦП включаються пристрої вибірки та зберігання.

Операції квантування й кодування здійснюються за допомогою послідовної, паралельної, послідовно-паралельної процедур.

Паралельні АЦП здійснюють перетворення сигналу одночасно за декількома рівнями квантування, отже мають найвищу швидкодію. Послідовні АЦП перетворюють сигнал послідовно біт за бітом, що потребує більшого часу, але спрощує їх апаратурну реалізацію.

Інтегруючі АЦП дозволяють підвищити заводостійкість за рахунок інтегрування вхідного сигналу на часовому інтервалі, тривалість якого кратна періоду завади.

В основу роботи сигма-дельта АЦП покладений принцип усереднення результатів вимірювання на значному інтервалі часу. Це дозволяє зменшити

внесену шумами похибку перетворення і збільшити роздільну здатність АЦП, покращити лінійність характеристики перетворення.

Послідовно-паралельні АЦП є компромісом між високою швидкістю і простотою схемотехнічного рішення. Такі АЦП займають проміжне положення по роздільній здатності і швидкодії між паралельними і послідовними АЦП і поділяються на багатоступеневі, багатотактного і конвеєрні.

Важливою частиною АЦП є цифровий інтерфейс, який забезпечує зв'язок АЦП з цифровими каналами ІВС. Найчастіше застосовують спосіб зв'язку, за якого АЦП представляє для одержувача коду як модна комірка пам'яті. Для цього АЦП має необхідну кількість адресних входів, дешифратор адресу і підключається безпосередньо до адресної шини і шини даних процесора.

Цифро-аналогове перетворення – це перетворення цифрового коду в аналогову величину, найчастіше в електричний струм або напруга. Характеристикою перетворення ЦАП називають функціональну залежність значень вихідної напруги (струму) від цифрового коду.

Основними статичними параметрами ЦАП є: роздільна здатність, похибка повної шкали, похибка зміщення нуля, нелінійність, диференціальна нелінійність, монотонність характеристики перетворення, температурна нестабільність. Динамічні властивості ЦАП характеризуються часом встановлення і швидкістю наростання вихідного сигналу.

За принципом роботи розрізняють послідовні і паралельні ЦАП. Паралельні реалізують ідею підсумовування вагових струмів (напружень, зарядів), значення кожного з яких пропорційне вазі цифрового двійкового розряду.

Запитання для контролю

1. Дайте означення інтегрального перетворення Лапласа.
2. Назвіть основні властивості перетворення Лапласа.
3. Назвіть основні властивості Z-перетворення.
4. У чому полягає «ефект накладання» спектрів при дискретизації і які способи його усунення?
5. У чому полягає подібність, а в чому відмінність прямого і оберненого ДПФ?
6. Назвіть основні властивості ДПФ.
7. Які основні операції виконуються в процесі перетворення аналогових сигналів в цифрові?
8. В чому полягає сутність процесу квантування сигналів?
9. Розкрийте механізм виникнення шумів квантування, приведіть їх основні характеристики.
10. Наведіть рівняння перетворення ЦАП.
11. У чому полягає сенс цифро-аналогового перетворення?
12. Наведіть приклади застосування ЦАП.

Розділ 4. Перетворення сигналів

13. Назвіть і розкрийте зміст основних статичних параметрів АЦП.
14. Які параметри характеризують динамічну точність АЦП?
15. Розкрийте зміст поняття «апертурна невизначеність аналого-цифрового перетворення».
16. Назвіть основні типи АЦП.
17. У чому полягає ідея паралельних АЦП?
18. У чому полягає основна перевага інтегруючих АЦП?
19. Як треба вибирати співвідношення часу інтегрування і періоду імовірної завади для найбільш ефективного її заглушення? Обґрунтуйте свій висновок.
20. У чому полягає принцип роботи сигма-дельта АЦП?
21. Порівняйте фільтруючі властивості інтегруючих і сигма-дельта АЦП.
22. У чому полягає ідея послідовно-паралельних АЦП?
23. Назвіть відомі вам варіанти сполучення АЦП з цифровими пристроями.
24. Які основні вузли входять до складу схем збору даних?
25. Назвіть і розкрийте зміст основних параметрів ЦАП.
26. Дайте порівняльну характеристику різних типів ЦАП.
27. У чому полягає основна ідея паралельних ЦАП?
28. Які функції виконує інтерфейс в схемах ЦАП?
29. Наведіть схеми ЦАП з паралельним інтерфейсом і поясніть їх роботу за часовими діаграмами.

Conclusions

Discrete signals are formed by temporal sampling of continuous (analog) signals. Temporal sampling step is selected as a constant. The main mathematical apparatus that is used by spectral analysis of discrete signals is the apparatus of Z-transformation, which plays on relation to discrete signals, the same role as an integral Fourier transformation for continuous signals. The basis of Z-transformation is integral Laplace transformation. Z-transformation is a result of all time functions in Laplace transformation.

Any signal with a limited spectrum without distortions could be represented by sequence of its discrete samples on condition that sampling frequency at least twice bigger than the highest harmonic spectrum of analog signal.

Discrete Fourier transformation associates a finite sampling of discrete signals to a finite spectrum of the same dimension as the sampling size of the input signals. Harmonics of discrete spectrum are equidistantly allocated on a range from zero to the sampling frequency.

In computerized IMS conversion from analog measurement signal into digital code and from digital code into analog signal is used. Such conversions are performed by analog-to-digital and digital-to-analog converters (ADCs and DACs).

Amongst main operations of analog-to-digital signals conversion are sampling by time, quantization by level and digital coding. Signal sampling is a signal conversion, as a result of which it is displayed as a finite set of its values - counts at discrete points in time. Time interval between adjacent counts, when uniform signal sampling, is called sampling period.

Terms of unambiguous reconstruction of analog signal by its counts are given in sampling theorem (Shannon's theorem) and recovery theorem (Shannon-Kotelnikov theorem).

Quantization of signals is the process of separating of a continuous values range of transforming quantity into a finite number of adjacent non-overlapping intervals and presenting of any value from given interval by single value belonging to the same interval. Substitution of continuous infinite set of values by finite one leads to a partial loss of information. The difference between instantaneous values of signal and permitted levels is called quantization noise.

Digital coding is, when performed without information loss, operation of quantized signal size presenting as a digital code. Code is a set of symbols and their application rules for presenting, transmitting, processing and storage of information. Coding is performed in order to increase information processes effectiveness, and also to protect information from distortion and integrity damage while transmitting, processing and storing.

ADCs more often use binary natural code and Grey code. Grey code differs by two features that facilitate the increase in performance and interference immunity of coding devices in comparison with binary natural code: any two code combinations that match adjacent quantization levels differ from each other only in one bit; elements value change in every digit while transferring from one code combination to another occur half as much.

Basic parameters of ADC are: bitness, conversion time (conversion rate), number of channels to connect analog signals, interface type, conversion accuracy characteristics, as well as static and dynamic parameters. Static parameters are: resolution, full scale error, zero offset error, nonlinearity, differential nonlinearity, differential linearity error, monotony of conversion characteristic, temperature instability. ADC dynamic parameters are: maximum sampling rate, conversion time, sampling time and aperture time. Aperture time is time interval during which there is still uncertainty between count value and time moment to which it is attributed. To reduce aperture time ADC has sampling and storage devices.

Quantization and encoding operations are carried out by serial, parallel, serial-parallel procedures. Parallel ADCs perform signal quantization simultaneously with a set of comparators connected in parallel with signal input source, which leads to limitation of bitness, however, such ADC have the highest performance. Sequential ADCs convert the analog input signal sequentially bit by bit, so conversion process takes more time, but greatly simplifies converter scheme.

Integrating ADCs can improve interference immunity due to the fact that input signal is integrated over a time interval, the duration of which is selected to be aliquot to interference period.

Basic principle put in a basis of sigma-delta ADC is measurement results averaging over large time interval. It allows to reduce conversion error introduced by noise and increase ADC resolution and provides better linearity of conversion characteristic. Integrating ADCs based on voltage-frequency converters provide relatively high conversion accuracy at low cost.

Розділ 4. Перетворення сигналів

Serial-parallel ADCs are a compromise between high performance and simplicity of scheme solution. Such ADCs are in intermediate position by resolution and performance between parallel and sequential ADCs and are divided into multi-stage, multicycle and conveyor.

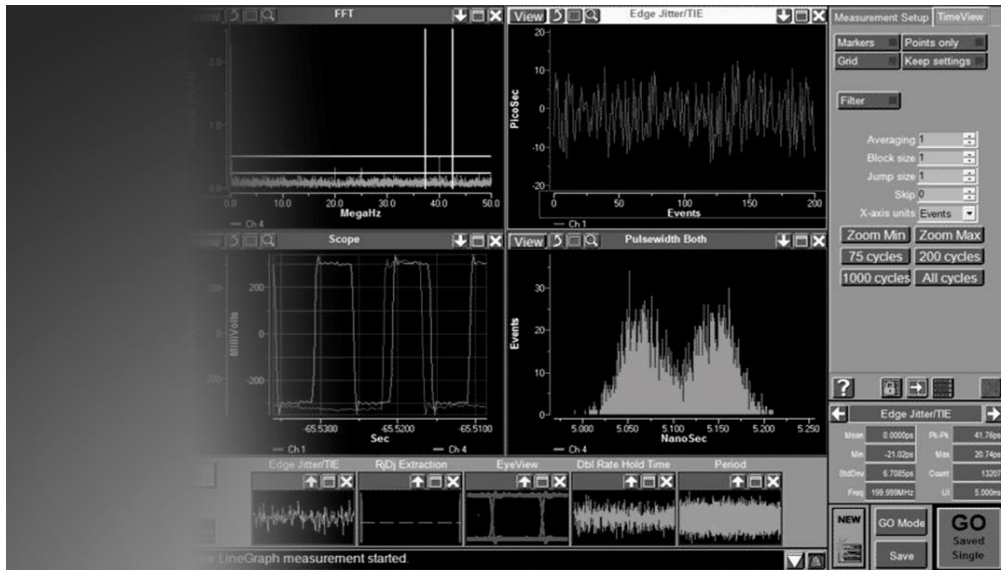
An important part of ADC is digital interface that provides connectivity between ADC and IMS digital channels. Often is used communication method, wherein ADC gives to code receiver one memory cell. To do this, ADC has required number of address inputs, address decoder and connects directly to address bus and processor data bus.

ADC with parallel output data interface is implemented using multi-bit storage register having three stable output states. ADC with sequential output data interface outputs them as a sequential code directly from comparator or successive approximation register. Digital-to-analog conversion is a conversion of digital code into analog value, usually into electric current or voltage. Functional dependence of output voltage (current) values from digital code is called DAC conversional characteristic.

Basic DAC static parameters are: resolution, full scale error, zero offset error, nonlinearity, differential nonlinearity, transformation characteristics monotony, temperature instability. DAC dynamic properties are characterized by setting time and slew rate of the output signal. By operational principle DACs are divided into serial and parallel. Parallel DACs realize the idea of summation of weighted currents (voltages or charges), value of which is proportional to the weight of digital binary bit.

Ключові слова

аналого-цифровий перетворювач	analog-to-digital converter
цифро-аналоговий перетворювач	digital-to-analog converter
квантування	quantization
дискретизація	sampling
цифрове кодування	digital coding



СТАТИСТИЧНЕ ОПРАЦЮВАННЯ ДАНИХ ВИМІРЮВАНЬ

5

- 5.1. Термінологія статистичного опрацювання даних вимірювань
- 5.2. Методи статистичного опрацювання даних вимірювань
- 5.3. Основи регресійного аналізу
- 5.4. Статистичне опрацювання випадкових кутів
- 5.5. Типові методи опрацювання результатів вимірювань

5.1. Термінологія статистичного опрацювання даних вимірювань

В ході контролю параметрів технологічних процесів, дослідження характеристик фізичних полів, моніторинга стану навколишнього середовища, контролю якості продукції, неруйнівного контролю, технічної і медичної діагностики, а також під час калібрування, валідації та верифікації ІВС та її окремих складових виникає широке коло завдань опрацювання, аналізу, зберігання і подання значних обсягів вимірювальної інформації. Розв'язання цих завдань ґрунтується на використанні методів теорії ймовірності та математичної статистики. Розглянемо основні терміни статистичного опрацювання даних вимірювань.

Вимірювання, вибірка, варіаційний ряд. Результатом вимірювання є число, вектор, функція або множина чисел як іменованих, так і неіменованих. Результат вимірювання містить всі види похибок притаманних інформаційно-вимірювальним процедурам. Похибки обумовлені недосконалістю використовуваної моделі вимірювання, небажаним взаємним впливом різних технічних засобів, впливом зовнішніх чинників і неінформативних параметрів об'єкта вимірювання, недосконалістю обчислювальних алгоритмів, втратою і спотворенням даних у вимірювальних каналах, суб'єктивними чинниками та ін.

Випадковий характер результату вимірювання обумовлюють дві причини. Перша відноситься до вимірювань постійних, незмінних в часі фізичних величин. Результати вимірювання спотворюються випадковим в часі вимірювальним шумом, тому отримані послідовно в часі результати стохастично змінюються відносно певного значення. Друга причина стосується вимірювання величин, які самі приймають випадкові значення, наприклад, значення параметра в партії виробів.

У сучасній теорії вимірювань апріорна модель результату вимірювання – це випадкова величина з невідомими характеристиками. Статистичному оцінюванню цих характеристик і присвячені методи опрацювання даних вимірювань. Результатом їх застосування є як оцінка самого результату вимірювання, так і оцінки його характеристик якості (похибки, невизначеності, вірогідність, достовірність).

Одне із завдань математичної статистики полягає в тому, щоби на підставі інформації про деяку властивість підмножини елементів

генеральної сукупності обґрунтувати цю властивість для всієї множини – генеральної сукупності.

В теорії похибок випадкові величини розглядаються зазвичай як результат прямих вимірювань певної фізичної величини невідомого, але незмінного розміру, спотворені випадковими похибками. Якщо провести n повторних вимірів випадкової величини, то отримуємо n різних числових значень x_1, x_2, \dots, x_n як реалізації цієї величини.

Випадкова вибірка – це декілька вибірових елементів, узятих з генеральної сукупності і призначених для отримання інформації про неї. Випадкова вибірка повинна досить повно відображає особливості всіх об'єктів генеральної сукупності, щоб отримувані оцінки були достовірними, тобто вона повинна бути *репрезентативною* (*представницькою*). Ця якість забезпечується випадковим і незалежним відбором її елементів. Це означає, що всі елементи генеральної сукупності мають однакову ймовірність стати елементами вибірки.

В імовірнісній моделі випадкову вибірку зазвичай розглядають як реалізацію незалежних однаково розподілених випадкових величин $\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega)$, де ω – елементарна подія з простору подій Ω . В результаті вимірювального експерименту отримують відповідні реалізації випадкової вибірки – конкретні значення x_1, x_2, \dots, x_n . Елементи вибірки $x_j, j = \overline{1, n}$ називають *варіантами*.

Функції випадкової вибірки, використовуваної для оцінки параметрів розподілів та (або) для перевірки статистичних гіпотез, називають *статистиками*. Якщо в ймовірнісній моделі дані вимірювань розглядаються як випадкові величини (або випадкові елементи), то статистики, як функції випадкових величин (елементів), самі є випадковими величинами (елементами). Статистики використовують для визначення характеристик випадкових величин (математичного сподівання, медіани, дисперсії, моментів вищих порядків і ін.), а їх реалізації – для оцінки цих характеристик, які іменуються *статистичними оцінками*.

Варіаційним рядом $x_{[1]}, x_{[2]}, \dots, x_{[n]}$ називають реалізацію випадкової вибірки, записану в порядку зростання елементів $x_{[1]} \leq x_{[2]} \leq \dots \leq x_{[n]}$, де $x_{[j]}$ – позначення елемента варіаційного ряду.

У разі аналізу вибірки значного обсягу ($n > 50$) проводиться розбиття варіаційного ряду на класи (інтервали) і отримують *інтервальний варіаційний ряд*. Ширину інтервалів Δx визначають за

Розділ 6. Статистичне опрацювання даних вимірювань

формулою $\Delta x = (x_{[n]} - x_{[1]}) / k$, де k – кількість класів, їх рекомендовані значення вказані в табл. 5.1.

Таблиця 5.1

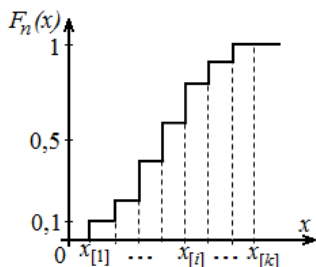
Обсяг вибірки n	25÷40	40÷60	60÷80	100÷200	200÷1000
Кількість класів k	5÷6	6÷8	7÷10	8÷12	10÷15

Припустимо, що зроблено n вимірювань випадкової величини $\xi(\omega)$ і отримано k різних реалізацій, з яких отримано варіаційний ряд $x_{[1]}, x_{[2]}, \dots, x_{[k]}$. В цьому експерименті значення $x_{[1]}$ спостерігалось m_1 разів, $x_{[2]}$ – m_2 разів, ..., $x_{[k]}$ – m_k разів. Числа m_1, m_2, \dots, m_k є частотами вимірених значень $x_{[1]}, x_{[2]}, \dots, x_{[k]}$. Величини $\tilde{m}_1 = m_1 / n, \tilde{m}_2 = m_2 / n, \dots, \tilde{m}_k = m_k / n$ називають відносними частотами варіанти $x_{[i]}$. Зрозуміло, що $\tilde{m}_1 + \tilde{m}_2 + \dots + \tilde{m}_k = 1$. Отримані результати зручно представити в табличній формі (табл. 5.2, четвертий рядок таблиці представляє значення кумулятивних (накопичених) відносних частот).

Таблиця 5.2

$x_{[i]}$	$x_{[1]}$	$x_{[2]}$...	$x_{[i]}$...	$x_{[k]}$
m_i	m_1	m_2	...	m_i	...	m_k
\tilde{m}_i	m_1 / n	m_2 / n	...	m_i / n	...	m_k / n
$F_i = \sum_{j=1}^i \tilde{m}_j$	m_1 / n	$(m_1 + m_2) / n$...	$\sum_{j=1}^i m_j / n$...	$\sum_{j=1}^k m_j / n = 1$

Якщо кількість варіант k не дуже велика, для отримання наочного уявлення про розподіл випадкової величини будують полігони частот. Для цього на осі абсцис відкладають значення варіантів, а на осі ординат відповідні значення частот m_1, \dots, m_k , або відносних частот $\tilde{m}_1, \dots, \tilde{m}_k$, або кумулятивних відносних частот F_i . Полігон відносних частот дає уявлення про розподіл ймовірностей. Графік кумулятивних відносних частот називають *емпіричною функцією розподілу*. Вона виражає залежність між значеннями кількісної ознаки і накопиченої частоти і містить всю інформацію про результати вимірювань.



Згідно закону великих чисел для кожного дійсного числа x емпірична функція розподілу $F_n(x)$ збігається до функції розподілу $F(x)$ відповідної випадкової величини якщо $n \rightarrow \infty$ (теорема В.І. Гливенка). Функція розподілу визначена на

Рис. 5.1. Приклад реалізації емпіричної функції розподілу ймовірностей

всій числовій осі. Ясно, що $F_n(x)=0$ для всіх $x < x_{[1]}$ і $F_n(x)=1$ для всіх $x > x_{[k]}$. На інтервалі $x_{[1]} < x < x_{[k]}$ реалізація емпіричної функція має вигляд східчастої монотонно зростаючої від 0 до 1 функції, такий, що $F_n(a) = P(X < a)$ (рис. 5.1). У разі інтервального варіаційного ряду частота дорівнює загальній кількості варіант в даному класі. Всі класи, крім останнього, є напіввідкритими інтервалами (наприклад, $[a_i; a_{i+1})$), а останній $[a_{k-1}; a_k]$ – закритий.

Можна скласти таку таблицю (табл. 5.3) інтервальних варіаційних рядів, де a_i – межі класових інтервалів. Якщо на осі абсцис відкласти класові інтервали і над ними побудувати прямокутники з висотами, рівними відповідним дільностям відносних частот $\tilde{m}_i / \Delta x$, то площа кожного прямокутника дорівнюватиме: $S_i = \Delta x p_i = \Delta x \tilde{m}_i / \Delta x = \tilde{m}_i$, тобто відносній частоті.

Таблиця 5.3

Номер класу	1	2	...	k
Класовий інтервал	$[a_0; a_1)$	$[a_1; a_2)$...	$[a_{k-1}; a_k]$
Частота	m_1	m_2	...	m_k
Відносна частота	$\tilde{m}_1 = m_1 / n$	\tilde{m}_2	...	\tilde{m}_k
Щільність відносної частоти	$p_1 = \tilde{m}_1 / \Delta x$	p_2	...	p_k

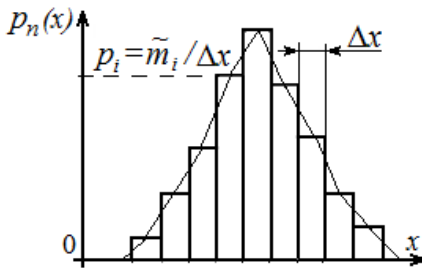


Рис. 5.1. Приклад реалізації емпіричної функції щільності розподілу ймовірностей

ймовірностей.

За нескінченно великої кількості класових інтервалів, частоти \tilde{m}_i слугують наближенням ймовірності попадання випадкової величини в i -тий інтервал: $\int_{x_i}^{x_{i+1}} p(x) dx \approx P_i$. Цю властивість використовують для порівняння теоретичного та емпіричного розподілів.

Отримана східчаста фігура називається *гістограмою* (рис. 5.2). Площа під гістограмою

$$S = \sum_{i=1}^k S_i = \sum_{i=1}^k \tilde{m}_i = 1.$$

Ламана лінія, яка йде по осі абсцис, потім обводить гістограму і знову проходить по осі абсцис є графіком реалізації емпіричної функції щільності розподілу

Розділ 6. Статистичне опрацювання даних вимірювань

Точкові оцінки характеристик ряду вимірювань.

Методологічною основою статистичного опрацювання результатів вимірювань є вибірковий метод, сутність якого полягає у перенесенні результатів дослідження вибірки на генеральну сукупність. Така можливість обґрунтовується наступним чином. Нехай генеральна сукупність має генеральний параметр Q . Проводиться серія з n незалежних вимірювань випадкової величини $\xi(\omega)$ з невідомим законом розподілу ймовірностей. Результатом вимірювання є вибірка $(x_1, \dots, x_i, \dots, x_n)$ обсягом n .

Необхідно знайти оцінку $\hat{Q} = \varphi(x_1, \dots, x_i, \dots, x_n)$ генерального параметра Q . Як параметр Q можуть виступати, наприклад, математичне сподівання, дисперсія, інші моменти випадкової величини вищих порядків, а також медіана і мода розподілів випадкової величини. Вважається, що x_i отримують за незмінних умов, і всі вони незалежні.

Кожне значення x_i можна розглядати як реалізацію випадкової величини $\xi_i(\omega)$, а всю вибірку $(x_1, \dots, x_i, \dots, x_n)$ – як реалізацію випадкового вектора (вибірки) $(\xi_1(\omega), \dots, \xi_n(\omega))$, в якому всі $\xi_i(\omega)$, $j = \overline{1, n}$ незалежні і мають той же розподіл ймовірності, що і $\xi(\omega)$.

Точкове оцінювання – спосіб статистичного оцінювання, який полягає в тому, що безпосередньо отримане значення оцінки застосовується як невідоме значення параметра розподілу. Точкова оцінка параметра визначається одним числом $\hat{Q} = \varphi(x_1, \dots, x_i, \dots, x_n)$ за вибіркою $x_1, \dots, x_i, \dots, x_n$, тобто визначає одну точку числової осі.

Наявність декількох методів оцінювання одних і тих самих параметрів призводить до необхідності вибору кращого серед цих методів. Ступінь відповідності Q його оцінці \hat{Q} залежить не тільки від обсягу вибірки, але й від виду функції φ . Вимоги найкращого наближення \hat{Q} до Q повинні відповідати таким показникам якості: незміщеність; слухність; ефективність.

Незміщеність оцінки означає, що зі зростанням кількості вимірювань середнє значення оцінки наближається до значення оцінюваного параметра $\lim_{n \rightarrow \infty} (M\hat{Q}) = Q$, де M – оператор математичного сподівання.

Якщо ця умова не виконується, оцінку називають *зміщеною*, а *зміщення* визначається різницею $M\hat{Q} - Q$.

Це зміщення може бути обумовлено як властивостями самої функції, так і похибками вимірювання, юстування, калібрування, випадковим характером отриманої вибірки або комбінацією цих

факторів. У ряді випадків воно не залежить від обсягу вибірки і називається *систематичною похибкою*. Наявність систематичних похибок не дозволяє формувати висновки щодо істинного значення вимірюваного параметру.

Оцінка називається *слухною*, якщо зі збільшенням обсягу вибірки $n \rightarrow \infty$ вона збігається з відповідним параметром генеральної сукупності, тобто для будь-якого як завгодно малого позитивного числа $\lim_{n \rightarrow \infty} P(|Q - \hat{Q}| \leq \varepsilon) = 1, \varepsilon > 0$. Це означає, що розсіювання оцінки \hat{Q} (як випадкової величини) зі збільшенням n все щільніше концентрується навколо Q .

Вимога *ефективності* оцінки: для вибірок рівного обсягу ефективна оцінка має мінімальне розсіювання (вибіркову дисперсію). Якщо виконується нерівність $D\hat{Q}_1 \leq D\hat{Q}_2$, то оцінка \hat{Q}_1 є більш ефективною. У цьому виразі D – оператор дисперсії.

Точковою оцінкою математичного сподівання генеральної сукупності є *вибіркове середнє* $\bar{x} = n^{-1} \sum_{i=1}^n x_i$.

Ця оцінка є *незміщеною, слухною і ефективною*.

Якщо вибіркові дані згруповані у варіаційний ряд, то оцінка математичного сподівання визначається за формулою $\bar{x} = n^{-1} \sum_{i=1}^k m_i x_{[i]}$. У наведеній формулі x_i – значення варіанти для варіаційного ряду або середина класового інтервалу для інтервального варіаційного ряду; m_i – частота варіанти або класова частота.

Вибіркова медіана \hat{M}_e також може служити оцінкою математичного сподівання *генеральної сукупності* (особливо в разі симетричного генерального розподілу). Для дисперсії вибіркової медіани справедливим є граничне співвідношення $\lim_{n \rightarrow \infty} (D\hat{M}_e) = 0,5\pi D\bar{x}$.

Дисперсія оцінки медіани більше дисперсії вибіркового середнього в $\pi/2$ рази, тобто оцінка вибіркового середнього є ефективнішою. Вибіркова дисперсія \hat{s}^2 є оцінкою дисперсії генеральної сукупності $s^2 = n^{-1} \sum_{i=1}^n (x_i - M\xi)^2$. Ця оцінка є незміщеною, але якщо дійсне значення математичного сподівання $M\xi$ невідоме і використовується його оцінка, то отримана оцінка дисперсії $\hat{s}^2 = n^{-1} \sum_{i=1}^n (x_i - \bar{x})^2$ є зміщеною:

$\hat{s}^2 = s^2 (n-1)n^{-1}$. Незміщену оцінку дисперсії отримують за формулою

$$s^2 = (n-1)^{-1} \sum_{i=1}^n (x_i - \bar{x})^2 .$$

У цій формулі знаменник $(n-1) = \nu$ називається

кількістю *ступенів свободи*. В загальному випадку ν дорівнює різниці між кількістю значень, на підставі яких обчислено статистичну оцінку і кількість констант, необхідних для обчислення цієї оцінки, які були визначені на підставі тих же значень. У наведеному вище рівнянні був втрачений один ступінь свободи за рахунок визначення вибіркового середнього.

Незміщеною точковою оцінкою стандартного відхилення вибіркового середнього є статистика $\hat{s}_{\bar{x}} = \sqrt{[1/n(n-1)] \sum_{i=1}^n (x_i - \bar{x})^2}$.

Ця статистика є характеристикою сукупності вимірювань, тоді як \hat{s} є характеристикою окремого вимірювання.

Інтервальні оцінки характеристик ряду вимірювань.

Розглянуті точкові оцінки дають одне числове значення параметра, яке є випадковою величиною і відрізняється від детермінованого істинного значення параметру. Ця випадковість обумовлена тим, що оцінки параметрів розподілу генеральної сукупності вираховують не по всій сукупності, а по вибірках з обмеженою кількістю елементів. Випадковість значень елементів, що потрапили у вибірку, і визначає випадковість значень оцінки. За значенням оцінки параметру не можна зробити висновок про її точність. Інформативнішим є результат оцінювання, який характеризується можливим відхиленням оцінки від параметра і ймовірністю виконання цієї умови, що дає інтервальне оцінювання.

Інтервальне оцінювання – це спосіб оцінювання, який передбачає визначення інтервалу, в який із заданою ймовірністю містить шукане (істинне) значення параметра розподілу. Цей інтервал називається *довірчим* або *надійним*, а відповідна ймовірність – *довірчою ймовірністю* або *надійністю*. Зрозуміло, що чим більше довірчий інтервал, тим більше впевненість, що цей інтервал містить оцінювану величину.

Довірча ймовірність визначається як $P=1-\alpha$, де α називається *рівнем значущості* і дорівнює ймовірності того, що надійний інтервал не накриває значення параметра (це ймовірність помилкової оцінки).

Припустимо, необхідно визначити довірчий інтервал для параметру Q генеральної сукупності, використовуючи незміщену і

слушну вибірккову оцінку \hat{Q} . Ця оцінка має вибіркковий розподіл із середнім $M\hat{Q} = Q$ і стандартним відхиленням $\sigma_{\hat{Q}} \rightarrow 0$ у випадку $n \rightarrow \infty$. Можна вказати ймовірність $P(|\hat{Q} - Q| \leq \varepsilon) = 1 - \alpha$, з якою відхилення оцінки від її математичного сподівання не перевищуватиме деякого малого позитивного числа ε . Цей вираз можна переписати наступним чином: $P(\hat{Q} - \varepsilon \leq Q \leq \hat{Q} + \varepsilon) = 1 - \alpha$, де $\hat{Q} - \varepsilon$, $\hat{Q} + \varepsilon$ відповідно нижня і верхня межі довірчого інтервалу. Ці межі є випадковими величинами, оскільки отримана за вимірами оцінка \hat{Q} є випадковою величиною.

5.2. Методи статистичного опрацювання даних вимірювань

Критерії перевірки згоди законів розподілів. Під час опрацювання результатів багаторазових вимірювань виникає завдання оцінювання їх закону розподілу, від якого залежать значення меж інтервалів випадкової похибки. Існує два підходи до вирішення цього завдання.

Перший ґрунтується на застосуванні критеріїв згоди для згрупованих даних і використовується за наявності вибірки результатів вимірювань досить великого обсягу ($n \geq 50$). Такі критерії вимагають попередньої побудови таких оцінок закону розподілу, як полігон частот, гістограма, емпірична функція розподілу, за виглядом яких власне і формулюють гіпотезу про закон розподілу. Однак групування даних призводить до втрати інформації.

Інший підхід ґрунтується на використанні спеціалізованих критеріїв згоди, які не вимагають групування даних і можуть застосовуватися на вибірках обсягом від 10 до 50 значень. Для цих критеріїв необхідна апріорна інформація про вид закону розподілу.

Розглянемо найпоширеніші критерії згоди для згрупованих даних.

Критерій χ^2 Пірсона. За досліджувану вибіркою будується гістограма (рис. 5.3) і для кожного інтервалу розраховують емпіричні частоти – кількість значень ряду, що потрапили в j -й інтервал, а також оцінки значень емпіричної щільності ймовірності $p_j = \hat{m}_j / (nh)$, де h – ширина стовпчиків гістограми. За вибіркковими даними розраховують

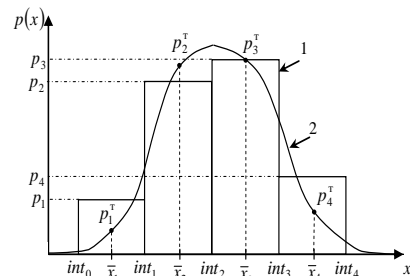


Рис. 5.3. Гістограма (1) і порівнювана теоретична щільність розподілу (2)

параметри обраної теоретичної щільності розподілу $p(x)$. Наприклад, для гауссового розподілу це оцінки математичного сподівання \bar{x} і середньоквадратичного відхилення (СКВ) s . Вираховуються значення теоретичної функції щільності ймовірності в точках, що відповідають серединам інтервалів гістограми $\bar{x}_j - p_j^T = p(\bar{x}_j)$. Для всіх інтервалів визначаються теоретичні частоти $m_j^T = p_j^T nh$, а потім і статистика χ_p^2 :

$$\chi_p^2 = \sum_{j=1}^l \left[(\hat{m}_j - m_j^T)^2 / m_j^T \right] = nh \sum_{j=1}^l \left[(p_j - p_j^T)^2 / p_j^T \right]. \quad (5.1)$$

Якщо $\chi_p^2 < \chi_\alpha^2(v)$ – приймається гіпотеза про згоду емпіричного закону розподілу з обраним теоретичним. Значення $\chi_\alpha^2(v)$ знаходять з таблиць χ^2 -розподілу для заданого рівня довіри α і числа ступенів свободи $v = l - k - 1$, де k – кількість параметрів обраного закону розподілу, l – кількість інтервалів гістограми.

Критерій Колмогорова – Смірнова.

За вибіркою будується оцінка інтегральної функції розподілу $F(x)$, за графіком якої виноситься гіпотеза про тип закону розподілу. Відповідна теоретична інтегральна функція $F^T(x)$, розрахована за оцінками параметрів, наноситься на той же графік (рис. 5.4), після чого визначається максимальне значення $D = \max |F^T(x) - F(x)|$ і статистика

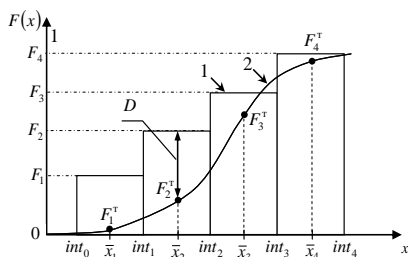


Рис. 5.4 Емпірична (1) і теоретична (2) функції розподілу

Таблиця 5.4

P	0,99	0,95	0,9
$\lambda(P)$	0,44	0,52	0,57

$\lambda = D\sqrt{n}$. Якщо для заданої ймовірності P маємо $\lambda < \lambda(P)$, то приймається гіпотеза про те, що закон

розподілу $F(x)$ відповідає $F^T(x)$. Значення $\lambda(P)$ наведені в табл. 5.4.

Критерій згоди без групування даних. Складений критерій. Цей критерій застосовують для перевірки згоди емпіричного розподілу з гауссовим за кількості результатів вимірювань $n < 50$. Він фактично складається з двох критеріїв.

Критерій 1. Обчислюють відношення $\tilde{d} = \sum_{i=1}^n |x_i - \bar{x}| / n\hat{s}$, (5.2)

де \hat{s} – зміщена оцінка СКВ $\hat{s} = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / n}$

Результати вимірювань групи можна вважати розподіленими за гауссовим законом, якщо $d(1-\alpha_1/2, n) < \tilde{d} \leq d(\alpha_1/2, n)$, де $d(1-\alpha_1/2, n)$ і $d(\alpha_1/2, n)$ – табличні квантілі розподілу для обсягу вибірки n , $\alpha_1/2$ і $(1-\alpha_1/2)$, α_1 – заздалегідь обраний рівень значущості критерія.

Критерій 2. Можна вважати, що результати вимірювань належать гауссовому розподілу, якщо не більш ніж m різниць $|x_i - \bar{x}|$ перевищили значення $z_{p/2}s$, де $z_{p/2}$ – верхній квантиль розподілу нормованої функції Лапласа для ймовірності $P/2$, s – оцінка СКВ, обчислена за формулою $s = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)}$. Ймовірності P визначають за обраним рівнем значущості α_2 і кількості вимірів n . Для рівнів значущості відмінних від наведених у таблиці, необхідні значення P знаходять шляхом лінійної інтерполяції.

Якщо для критерію 1 обрано рівень значущості α_1 , а для критерію 2 – рівень α_2 , результуючий рівень значущості складеного критерію $\alpha \leq \alpha_1 + \alpha_2$. Гіпотеза про згоду емпіричного розподілу з гаусовим приймається тільки тоді, коли виконуються обидва критерії.

Критерій Фроціні використовується для перевірки гіпотези про рівномірний розподілі експериментальних даних. Вибірка ранжується за зростанням і нормується, таким чином виходить упорядкований ряд $U = (U_{[1]} \leq \dots \leq U_{[n]})$, розподілений на відрізку $[0, 1]$. Статистика критерію

$$B_n = \left(1/\sqrt{n}\right) \sum_{i=1}^n |U_{[i]} - (i-0,5)/n|. \quad (5.3)$$

Гіпотеза про рівномірність розподілу приймається тоді, коли $B_n < B_n(P)$, де $B_n(P)$ – граничне значення статистики Фроціні для заданої довірчої ймовірності.

Перевірка гіпотез про математичне сподівання. У вимірювальній практиці зустрічаються завдання з експериментального визначення систематичних похибок вимірювальних каналів і ІВС в цілому, коли значення випадкової похибки може бути відомим заздалегідь або визначеним експериментально. В математичній статистиці такі завдання називають *перевіркою гіпотез про математичне сподівання*.

Порівняння невідомого середнього значення з заданим значенням μ_0 за відомої дисперсії. В даному випадку перевіряється гіпотеза про відсутність систематичної похибки в результатах

Розділ 6. Статистичне опрацювання даних вимірювань

вимірювань фізичної величини зі значенням μ_0 , що відтворюється за допомогою еталона, виконаних з відомим значенням СКВ σ_0 ,

Для розв'язання даного завдання необхідні наступні вихідні дані: обсяг вибірки n , сума спостережуваних значень $\sum_{i=1}^n x_i$, задане значення μ_0 , значення дисперсії генеральної сукупності σ_0^2 або стандартного відхилення σ_0 , обраний рівень значущості α . Наступним кроком є визначення квантиля $u_{1-\alpha}$ стандартного гауссового розподілу рівня $(1-\alpha)$, квантиля $u_{1-\alpha/2}$ стандартного гауссового закону розподілу рівня $(1-\alpha/2)$, а також обчислення значення $\bar{x} = \sum_{i=1}^n x_i / n$. Для отриманих результатів виконують порівняння вибіркового середнього \bar{x} із заданим значенням μ_0 наступним чином:

1) в двосторонньому випадку:

гіпотеза статистичної рівності вибіркового середнього і заданого значення μ_0 (нульова гіпотеза про відсутність систематичної похибки) відхиляється, якщо

(5.4)

2) в односторонньому випадку:

а) гіпотеза про те, що вибіркоче середнє не менше μ_0 (нульова гіпотеза) відхиляється, якщо

б) гіпотеза про те, що вибіркоче середнє не більше μ_0 (нульова гіпотеза) відхиляється, якщо

Порівняння невідомого середнього значення з заданим значенням μ_0 за невідомої дисперсії. Даний критерій у вимірвальній практиці застосовується в тих же випадках, що і попередній, з єдиною відмінністю – прецизійність ІВС або вимірвального каналу (СКВ) заздалегідь не відома і оцінюється за експериментальними даними.

Для вирішення даного завдання потрібні такі вихідні дані: обсяг вибірки n , сума значень $\sum x_i$ і сума квадратів значень спостережуваних величин $\sum x_i^2$, задане значення μ_0 , ступінь свободи $\nu = n - 1$, а також обраний рівень значущості α .

Визначаються квантиль $t_{1-\alpha}(v)$ розподілу Стьюдента рівня $(1-\alpha)$ з v ступенями свободи, квантиль $t_{1-\alpha/2}(v)$ рівня $(1-\alpha/2)$ і середнє значення $\bar{x} = \sum_{i=1}^n x_i / n$. Далі знаходять вибіркові дисперсію і СКВ:

$$D = [1/(n-1)] \sum_{i=1}^n (x_i - \bar{x})^2 = [1/(n-1)] \left[\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 / n \right] \quad \text{і} \quad s = \sqrt{D}.$$

За отриманими результатами виконують порівняння \bar{x} та μ_0 :

1) в двосторонньому випадку – гіпотеза рівності вибіркового середнього і заданого значення (нульова гіпотеза про відсутність систематичної похибки) відхиляється, якщо виконується $|\bar{x} - \mu_0| > (t_{1-\alpha/2}(v) / \sqrt{n})s$; (5.7)

2) в односторонньому випадку:

а) гіпотеза про те, що вибіркове середнє не менше μ_0 (нульова гіпотеза) відхиляється, якщо $\bar{x} < \mu_0 - (t_{1-\alpha}(v) / \sqrt{n})s$; (5.8)

б) гіпотеза про те, що вибіркове середнє не більше μ_0 (нульова гіпотеза) відхиляється якщо $\bar{x} > \mu_0 + (t_{1-\alpha}(v) / \sqrt{n})s$. (5.9)

Порівняння середніх значень двох вибірок з невідомими але рівними дисперсіями (рівноточні вимірювання). Даний критерій може бути застосований у випадках:

1) коли процес вимірювання проводять паралельно двома різними ІВС, точність кожної з них невідома, але може бути визначена за експериментальними даними; необхідно відповісти на питання, чи можна вважати, що обидва прилади дають результати рівної точності;

2) якщо необхідно порівняти результати вимірювань, отриманих за двох різних значеннях впливаючих величин.

Для розв'язання даного завдання необхідні такі вихідні дані: обсяги двох вибірок n_1 і n_2 , сума значень величин $\sum x_{1,i}$ і $\sum x_{2,i}$, сума квадратів значень величин $\sum x_{1,i}^2$ і $\sum x_{2,i}^2$, ступень свободи $v = n_1 + n_2 - 2$, обраний рівень значущості α . Визначають квантиль $t_{1-\alpha}(v)$ розподілу Стьюдента рівня $(1-\alpha)$ з v ступенями свободи і квантиль $t_{1-\alpha/2}(v)$ – для рівня $(1-\alpha/2)$ і v ступенями свободи, а також значення $\bar{x}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} x_{1,i}$ і $\bar{x}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} x_{2,i}$. Далі

обчислюють
$$s = \sum_{i=1}^{n_1} (x_{1,i} - \bar{x}_1)^2 + \sum_{i=1}^{n_2} (x_{2,i} - \bar{x}_2)^2$$
 і
$$s_d = \sqrt{[(n_1 + n_2) / n_1 \cdot n_2] \cdot s / (n_1 + n_2 - 2)}.$$

Розділ 6. Статистичне опрацювання даних вимірювань

За отриманими результатами виконують порівняння середніх значень двох сукупностей:

1) в двосторонньому випадку:

припущення про рівність середніх значень (нульова гіпотеза) відхиляється, якщо

$$|\bar{x}_1 - \bar{x}_2| > t_{1-\alpha/2}(v) s_d ;$$

(5.10)

2) в односторонньому випадку:

а) припущення про те, що перше середнє \bar{x}_1 не менше другого \bar{x}_2 (нульова гіпотеза) відхиляється, якщо

$$\bar{x}_1 < \bar{x}_2 - t_{1-\alpha}(v) s_d ;$$

(5.11)

б) припущення про те, що перше середнє \bar{x}_1 не більш другого \bar{x}_2 (нульова гіпотеза) відхиляється, якщо

$$\bar{x}_1 > \bar{x}_2 + t_{1-\alpha}(v) s_d .$$

(5.12)

Примітка. Дисперсії невідомі, але у припущенні можуть бути рівними.

Гіпотези про дисперсію. Порівняння дисперсії або стандартного відхилення із заданою величиною. Вихідні дані: обсяг вибірки n , суми значень величин $\sum x_i$ та їх квадратів $\sum x_i^2$, задане значення σ_0^2 , ступені свободи $\nu = n - 1$, обраний рівень значущості α . Спочатку визначають квантилі χ^2 -розподілу для рівнів значущості α , $(1-\alpha)$, $\alpha/2$ ($1-\alpha/2$) і відповідно: $\chi_\alpha^2(\nu)$, $\chi_{1-\alpha}^2(\nu)$, $\chi_{\alpha/2}^2(\nu)$, $\chi_{1-\alpha/2}^2(\nu)$. Далі обчислюються значення $C = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 / n$ і C / σ_0^2 .

За отриманими результатами виконується порівняння оцінки дисперсії \hat{D} з заданим значенням σ_0^2 (або порівняння стандартного відхилення σ з заданим значенням σ_0):

1) двосторонній випадок:

припущення рівності дисперсії (стандартного відхилення) і заданого значення (нульова гіпотеза) відхиляється, якщо виконується умова

$$C / \sigma_0^2 < \chi_{\alpha/2}^2(\nu) \text{ или } C / \sigma_0^2 > \chi_{1-\alpha/2}^2(\nu); \quad (5.13)$$

2) односторонній випадок:

а) припущення про те, що дисперсія (стандартне відхилення) не більше заданого значення (нульова гіпотеза) відхиляється, якщо

$$C / \sigma_0^2 > \chi_{1-\alpha}^2(\nu); \quad (5.14)$$

б) припущення про те, що дисперсія (стандартне відхилення) не менше заданого значення (нульова гіпотеза) відхиляється, якщо

$$C / \sigma_0^2 < \chi_\alpha^2(\nu); \quad (5.15)$$

Цей критерій може застосовуватися для порівняння оціненої експериментально точності досліджуваної ІВС із заданою точністю (тобто відомим параметром σ_0) або з точністю іншого обладнання.

Порівняння стандартних відхилень двох генеральних сукупностей. Для розв'язання даного завдання потрібні такі вихідні дані: обсяги двох вибірок n_1 і n_2 , суми значень величин $\sum x_{1,i}$ і $\sum x_{2,i}$ та їх квадратів $\sum x_{1,i}^2$ і $\sum x_{2,i}^2$, ступенів свободи $\nu_1 = n_1 - 1$ і $\nu_2 = n_2 - 1$, обраний рівень значущості.

Необхідно визначити квантилі розподілу Фішера $F_{1-\alpha/2}(\nu_1, \nu_2)$ і $F_{1-\alpha}(\nu_1, \nu_2)$ та обчислити значення

$$C_1 = \sum_{i=1}^{n_1} (x_{1,i} - \bar{x}_1)^2 = \sum_{i=1}^{n_1} x_{1,i}^2 - \left(\sum_{i=1}^{n_1} x_{1,i} \right)^2 / n_1, \quad s_1^2 = C_1 / n_1 - 1,$$

$$C_2 = \sum_{i=1}^{n_2} (x_{2,i} - \bar{x}_2)^2 = \sum_{i=1}^{n_2} x_{2,i}^2 - \left(\sum_{i=1}^{n_2} x_{2,i} \right)^2 / n_2, \quad s_2^2 = C_2 / n_2 - 1.$$

Отримані результати використовують для порівняння дисперсій двох сукупностей:

1) двосторонній випадок:

гіпотеза рівності дисперсій або рівності двох стандартних відхилень (нульова гіпотеза) відхиляється за виконання умов

$$s_1^2 / s_2^2 < 1 / F_{1-\alpha/2}(\nu_2, \nu_1) \quad \text{або} \quad s_1^2 / s_2^2 > F_{1-\alpha/2}(\nu_1, \nu_2); \quad (5.16)$$

2) односторонній випадок:

а) гіпотеза про те, що $D_1 \leq D_2$ ($\sigma_1 \leq \sigma_2$) (нульова гіпотеза) відхиляється якщо

$$s_1^2 / s_2^2 > 1 / F_{1-\alpha}(\nu_1, \nu_2); \quad (5.17)$$

б) гіпотеза про те, що $D_1 \geq D_2$ ($\sigma_1 \geq \sigma_2$) (нульова гіпотеза) відхиляється, якщо

$$s_1^2 / s_2^2 < 1 / F_{1-\alpha}(\nu_2, \nu_1). \quad (5.18)$$

Розглянутий критерій застосовується для наступних завдань:

1) порівняння точності двох засобів вимірювань за результатами вимірювань значення однієї фізичної величини;

2) співвідношення стабільності метрологічних характеристик в часі або у випадку зміни впливаючих факторів.

Гіпотези про математичні сподівання і дисперсії серій вимірювань. Дані критерії, дозволяють порівняти математичні сподівання або дисперсії декількох груп даних. Такими групами можуть бути, наприклад, групи результатів вимірювань, виконаних декількома вимірювальними каналами ІВС.

Критерій Кохрена. Для сукупності з p стандартних відхилень s_i , розрахованих виходячи з однієї і тієї ж кількості результатів вимірювань n в групах даних, статистика Кохрена має вид

$$C = s_{\max}^2 / \sum_{i=1}^p s_i^2, \quad (5.19)$$

де s_{\max} – найбільше значення стандартного відхилення в сукупності.

У разі, якщо значення статистики менше (або дорівнює) 5% -го критичного значення, тестову позицію (s_{\max}) визнають коректною, тобто стандартні відхилення груп даних можна вважати статистично рівними (вимірювання в групах рівноточні).

Якщо значення тестової статистики більше 5% -го критичного значення і менше (або дорівнює) 1% -го значення, то тестову позицію (s_{\max}) вважають відмінною від загальної групи. Використання результатів вимірювань даного каналу ІВС може привести до помилок.

У разі, якщо значення тестової статистики більше 1% -го критичного значення, тестова позиція суттєво відрізняється від загальної групи і дані результати вимірювань не повинні використовуватися.

Критерій Кохрена застосовують у випадках, коли всі стандартні відхилення знайдені за однаковою кількістю результатів вимірювань, отриманих в умовах повторюваності. За допомогою критерія Кохрена перевіряють найбільше значення в сукупності стандартних відхилень, і тому така перевірка є односторонньою. Розсіювання в дисперсіях може також проявлятися і в мінімальних значеннях стандартних відхилень.

Якщо найбільше значення стандартного відхилення класифіковано як таке, що містить надмірну похибку, воно повинно бути виключено, а перевірка з використанням критерію Кохрена може бути повторена на значеннях, які залишилися. Процедура повторення може призвести до надмірних виключень даних у випадках, коли гауссовий розподіл даних, прийнятий за основу, не є досить коректною апроксимацією.

Критерій Граббса. Для перевірки гіпотези про статистичну рівність груп оцінок математичних сподівань виділяють найбільшу величину з групи $\bar{x}_{[p]}$ розташованих в порядку зростання даних $\bar{x}_{[i]}$, $i = 1, 2, \dots, p$, і обчислюють статистику Граббса

$$G_p = (\bar{x}_{[p]} - \bar{x}) / s. \quad (5.20)$$

Для перевірки значущості мінімальної оцінки математичного сподівання $\bar{x}_{[1]}$ обчислюють тестову статистику $G_1 = (\bar{x} - \bar{x}_{[1]}) / s$.

Якщо значення тестової статистики менше (або дорівнює) 5% -го критичного значення, тестову позицію $\bar{x}_{[p]}$ або $\bar{x}_{[1]}$ визнають коректною. Якщо значення тестової статистики більше 5% -го і менше (або дорівнює) 1%

-го критичного значення, тестову позицію вважають такою, що відрізняється від загальної групи. Її використання призводить до помилок тому вона повинна вилучатись з подальших розрахунків.

Гіпотези про приналежність результату вимірювання до генеральної сукупності експериментальних даних. Зазвичай, першим етапом опрацювання результатів вимірювань є перевірка приналежності варіант отриманого ряду вимірювань до загальної сукупності експериментальних даних. Фактично проводиться перевірка на наявність *грубої похибки (промаху)* у вимірюваннях. Останні виникають випадково у разі повторних вимірювань і за своїми значеннями явно перевищують похибки, виправдані умовами проведення вимірювань. Причини грубих похибок – неконтрольовані зміни умов вимірювань, стрибки напруги живлення, несправність ІВС, помилки оператора тощо. Для виключення результатів вимірювань з грубими похибками застосовують апарат перевірки статистичних гіпотез і відповідні статистичні критерії.

Критерій Романовського. Результат x_i з ряду вимірювань $x_i, i = \overline{1, n}$ містить грубу похибку, якщо виконується умова

$$t = |x_i - \bar{x}| / s \geq t_\alpha(v) \sqrt{(n+1)/n}, \quad (5.21)$$

де \bar{x}, s – відповідно середнє значення і оцінка СКВ ряду без "підозрілого" значення; $v = n - 1$ – кількість ступенів свободи.

За допомогою даного критерію перевіряють максимальне і мінімальне значення в ряді вимірювань, а також декілька мінімальних або максимальних значень, граничне значення статистики Стьюдента $t_\alpha(v)$ вибирають для $v = n - k$, де k – кількість "підозрілих" результатів.

Критерій Діксона застосовують у випадку гауссова розподілу похибки і невеликого обсягу даних вимірювань $n \leq 25$ для перевірки максимального або мінімального значення з ряду вимірювань. Значення ряду упорядковують за зростанням: $x_{[1]} < x_{[2]} < \dots < x_{[n]}$. Результати $x_{[n]}$ (максимальний) і $x_{[1]}$ (мінімальний) містять надмірну похибку, якщо

$r_{\max} = (x_{[n]} - x_{[n-1]}) / (x_{[n]} - x_{[1]}) > Z_\alpha(n); \quad r_{\min} = (x_{[2]} - x_{[1]}) / (x_{[n]} - x_{[1]}) > Z_\alpha(n),$
де $Z_\alpha(n)$ – коефіцієнт, значення якого вибирають виходячи з обсягу ряду n і рівня значущості α (табл. 5.5). У разі проміжних значень n для визначення $Z_\alpha(n)$ застосовують лінійну інтерполяцію.

Таблиця 5.5

n	Рівень значущості α			n	Рівень значущості α		
	0,10	0,05	0,10		0,10	0,05	0,10

Розділ 6. Статистичне опрацювання даних вимірювань

4	0,68	0,76	0,89	16	0,28	0,33	0,43
6	0,48	0,56	0,70	18	0,26	0,31	0,41
8	0,40	0,47	0,59	20	0,26	0,30	0,39
10	0,35	0,41	0,53	30	0,22	0,26	0,34
14	0,29	0,35	0,45				

Критерій Діксона може використовуватися для перевірки двох і більше результатів на наявність промахів

Критерій "вільний" від закону розподілу похибки. У разі обмеженої кількості вимірювань n і (або) складності оцінювання параметрів або визначення закону розподілу рекомендується виключати результати з грубими похибками через наближені оцінки коефіцієнтів виду розподілу. З ряду вимірювань виключаються результати $x_i < x_{r-}$ і $x_i > x_{r+}$, де граничні значення визначають за формулами

$$x_{r-} = \bar{x} - s \left(1 + A \sqrt{1/\gamma^2 - 1} \right), \quad x_{r+} = \bar{x} + s \left(1 + A \sqrt{1/\gamma^2 - 1} \right), \quad (5.22)$$

де $A \in [0,85 \dots 1,3]$ – коефіцієнт, значення якого вибирається залежно від заданої довірчої ймовірності (зазвичай вибирається $A = 1,3$); γ – конгресес, значення якого залежить від форми закону розподілу

$$\gamma = \sqrt{s^4 / \mu_4}, \quad \mu_4 = n^{-1} \sum_{i=1}^n (x_i - \bar{x})^4.$$

Розрахунки \bar{x}, s, μ_4 проводять без "підозрілого" вимірювання.

Гіпотези про наявність тренда в ряді вимірювань. Методи виявлення тренда в ряді вимірювань залежать від його характеру: лінійне зростання/спадання математичного сподівання або дисперсії, періодична зміна математичного сподівання і т.п. За можливості нехтування випадковою похибкою, тренд виявляють шляхом аналізу чергування знаків відхилень результатів або від центру розподілу ймовірності, яким може бути математичне сподівання, медіана і т.п., або від попередніх результатів. Отже чергування знаків невиправлених випадкових відхилень з якоюсь закономірністю свідчить про наявність тренда. **Наприклад**, якщо спостерігається послідовність знаків "+" або "-" випадкових відхилень попередніх результатів від наступних, то в наявності *прогресуюча систематична похибка* (рис. 5.5,а). Чергування груп знаків "+" і "-" випадкових відхилень від центру розподілу свідчить про присутність *періодичної систематичної похибки* (рис. 5.5,б).

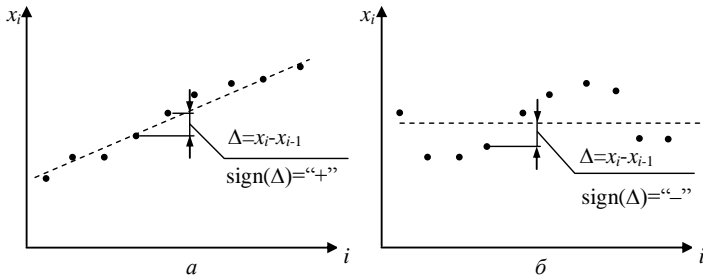


Рис. 5.5. Графічне пояснення процесу виявлення тренда

За значних випадкових похибок використовують інші критерії.

Критерій серій, заснований на використанні медіани. Цей критерій дозволяє виявити як періодичний, так і прогресуючий тренди. Послідовність його застосування наступна.

1. Визначають оцінку медіани $\mathbf{Med}(x)$ з ряду вимірювань $x_i, i = \overline{1, n}$.
2. Для ряду вимірювань замість кожного x_i ставлять плюс, якщо $x_i > \mathbf{Med}(x)$, і мінус, якщо $x_i < \mathbf{Med}(x)$, а якщо $x_i = \mathbf{Med}(x)$, то його пропускають в отриманій послідовності знаків.
3. Для отриманих послідовностей знаків визначають серії, загальне число серій $\nu(n)$ і тривалість найбільшої серії $\tau(n)$. Під серією розуміють послідовність плюсів чи мінусів.
4. Перевіряють гіпотезу H_0 про статистичну незалежності випадкових відхилень, тобто про відсутність систематичної похибки. Конкуруюча гіпотеза H_1 – випадкові відхилення статистично залежні, тобто є періодична систематична похибка. Гіпотеза H_0 для рівня значущості $0,05 < \alpha < 0,0975$ приймається, якщо одночасно виконуються нерівності

$$\nu(n) > 0,5(n+1-1,96\sqrt{n-1}), \quad \tau(n) < (3,31g(n+1)). \quad (5.23)$$

Якщо хоча б одна з нерівностей не виконується, приймають гіпотезу H_1 про наявність періодичного тренду.

Критерій "зростаючих" і "спадаючих" серій чутливий до прогресуючих трендів (систематичних похибок) реалізується поетапно.

1. Для ряду вимірювань $x_1, \dots, x_i, \dots, x_n$ на i -му місці ставлять плюс, якщо $x_{i+1} > x_i, (x_{i+1} - x_i > 0)$, і мінус, якщо $x_{i+1} < x_i, (x_{i+1} - x_i < 0)$, (якщо $x_{i+1} = x_i$, враховують тільки одне значення).
2. Для отриманої послідовності плюсів і мінусів визначають серії, загальну кількість серій $\nu(n)$ і тривалість найдовшої серії $\tau(n)$.
3. Перевіряють гіпотезу H_0 про статистичну незалежності випадкових відхилень $(x_{i+1} - x_i)$, тобто відсутність систематичної похибки. Конкуруюча

Розділ 6. Статистичне опрацювання даних вимірювань

гіпотеза H_1 – випадкові відхилення статистично залежні, тобто є прогресуюча систематична похибка. Гіпотеза H_0 для рівня значущості $0,05 < \alpha < 0,0975$ приймається, якщо одночасно виконуються дві нерівності:

$$v(n) > \left((2n-1)/3 - 1,96\sqrt{(16n-29)/90} \right), \quad \tau(n) < \tau_0(n). \quad (5.24)$$

Порогове значення $\tau_0(n)$ обирають в залежності від n :

$$\tau_0(n \leq 26) = 5; \quad \tau_0(26 < n \leq 153) = 6; \quad \tau_0(153 < n \leq 1170) = 7.$$

Якщо хоча б одна з нерівностей не виконується, H_0 відхиляють.

5.3. Основи регресійного аналізу

Регресійний аналіз набув широкого застосування для калібрування вимірювальних каналів ІВС, градування сенсорів у системах вимірювання та контролю фізико-хімічних властивостей матеріалів. В залежності від моделі шуканої функції $Y(X)$ розрізняють лінійну, нелінійну і множинну регресію, а відповідно до методу оцінювання параметрів моделі – параметричну і непараметричну.

Загальна схема регресійного аналізу містить наступні етапи.

1. Знаходження вибіркової оцінки регресії:
вибір моделі і оцінювання її параметрів;
оцінювання статистичної значущості отриманих параметрів.
2. Оцінювання адекватності обраної моделі.
3. Визначення довірчих інтервалів функції регресії, які із заданою ймовірністю містять середні або індивідуальні значення за якими визначають межі похибки (розширену невизначеність) значення Y для заданого значення X .

На сьогодні найбільш повно розроблений апарат регресійного аналізу, який передбачає, що вибіркові значення статистично незалежні і мають гауссовий закон розподілу. У випадку порушення цієї умови доцільно застосовувати методи непараметричної регресії.

Лінійна регресія. У разі лінійної регресії розглядається модель залежності двох величин X і Y в наступному загальному вигляді.

$$Y(X) = kX + b, \quad k, b \in R.$$

Початковими даними можуть бути n пар вибірових значень x_i і y_i , $i = \overline{1, n}$ (рис. 5.6,а) або в випадку багаторазових вимірювань значень $y_{i,k}$ $i = \overline{1, n}$, $k = \overline{1, m}$ n пар x_i і середніх значень \bar{y}_i (рис. 5.6,б).

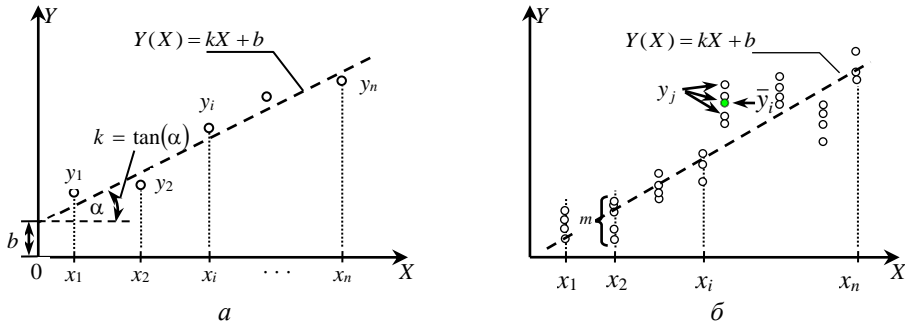


Рис. 5.6. Дані для побудови регресії: значення Y отримані як результати вимірювань з одноразовими (а) або багаторазовими (б)

Існує кілька методів оцінювання коефіцієнтів регресії – спрощені, метод найменших квадратів (МНК), метод правдоподібності.

Оцінювання коефіцієнтів регресії методом найменших квадратів. Значення коефіцієнтів регресії k і b знаходять з умови мінімуму суми квадратів відхилень початкових значень від значень отриманих на основі побудованої функції регресії $\min \sum_{i=1}^n (y_{\text{перп}} - y_i)^2$.

З цієї умови значення коефіцієнти b і k обчислюють за формулами:

$$b = \left[\sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n (x_i \cdot y_i) \right] / \left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right],$$

$$k = \left[n \sum_{i=1}^n (x_i y_i) - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i \right] / \left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right].$$

Перевірка значущості коефіцієнтів b і k . Для оцінювання значущості використовується критерій Стьюдента для перевірки гіпотези: H_0 – оцінка статистично незначуща, коефіцієнт дорівнює 0; H_1 – оцінка статистично значуща. Розраховуються t -статистики: $t_k = |k|/s_k$, $t_b = |b|/s_b$, де s_b і s_k – оцінки СКВ коефіцієнтів b ті k відповідно:

$$s_b = s_k \sqrt{\sum_{i=1}^n x_i^2 / n}, \quad s_k = s_y / \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad s_y = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (y_i - y_{\text{перп}i})^2}.$$

Якщо $t_k < t_{\alpha}(v)$ або $t_b < t_{\alpha}(v)$ приймається гіпотеза H_0 . В іншому випадку відповідний коефіцієнт є статистично значущим. Значення

Розділ 6. Статистичне опрацювання даних вимірювань

$t_{\alpha}(v)$ знаходиться як $(1-\alpha) \cdot 100\%$ одностороння довірча межа розподілу Стьюдента для $v = n - 2$.

Перевірка адекватності отриманої моделі виконується через перевірку гіпотези про те, що обрана модель регресії виду $Y = g(X)$, побудована за результатами парних вимірів $x_i, y_i, i = \overline{1, n}$, задовільно узгоджується з цими даними. Перевірку адекватності моделі регресії проводять за критерієм Фішера. Суть перевірки: значення спостережень y_i (\bar{y}_i) повинні лежати близько до прямої $Y(X)$, тобто їх відхилення від прямої регресії не повинні бути більшими відхилення значень в групі спостережень $y_{i,k}$ від їх середнього \bar{y}_i .

Статистика Фішера розраховується за формулою $F = s_y^2 / s_{\text{гр}}^2$, де s_y – СКВ розсіювання значень y_i в околі лінії регресії; $s_{\text{гр}}$ – СКВ розсіювання значень в групі спостережень в околі \bar{y}_i .

$$\text{У випадку багаторазових вимірювань } s_{\text{гр}}^2 = \frac{1}{n(m-1)} \sum_{i=1}^n \sum_{k=1}^m (y_{i,k} - \bar{y}_i)^2.$$

У випадку одноразових вимірювань значення $s_{\text{гр}}^2$ повинно бути відомим до проведення експерименту.

Визначення довірчих інтервалів (розширеної невизначеності) розрахованих за регресією значень y . Оцінюється СКВ випадкової складової похибки (стандартна невизначеність) середнього \bar{y} та індивідуальних значень y_i , визначених за отриманою залежністю, як функції x :

$$s_{\bar{y}}(x) = u_{\bar{y}}(x) = s_y \sqrt{1/n + (x - \bar{x})^2 / \sum_{i=1}^n (x_i - \bar{x})^2},$$
$$s_y(x) = u_y(x) = s_y \sqrt{1 + 1/n + (x - \bar{x})^2 / \sum_{i=1}^n (x_i - \bar{x})^2}.$$

СКВ випадкової складової похибки середнього \bar{y} сягає максимального значення на краях лінії регресії: $s_{\bar{y}\text{max}} \approx \sqrt{1 + 3w^2} s_y / \sqrt{m}$, де w – кількість напіврозмахів ($V = |x_{\text{max}} - \bar{x}| = |x_{\text{min}} - \bar{x}|$) у відхиленні аргументу x від середнього значення; $x_{\text{max}}, x_{\text{min}}$ – максимальне і мінімальне значення аргументу. За отриманими значеннями СКО оцінюють довірчий інтервал випадкової похибки (розширену невизначеність) середнього та

індивідуального значень y : $\Delta_{\bar{y}}(x) = U_{\bar{y}}(x) = t_{\alpha/2}(v) \cdot s_{\bar{y}}(x)$,
 $\Delta_y(x) = U_y(x) = t_{\alpha/2}(v) \cdot s_y(x)$, де $t_{\alpha/2}(v)$ – квантиль розподілу Стьюдента для довірчої ймовірності $P=1-\alpha/2$ і $v=n-2$ ступенів свободи.

Величини довірчих областей, які із заданою ймовірністю P містять середні або індивідуальні значення y залежать від відстані значення x (за яким визначають y) від середнього \bar{x} (рис. 5.7,а).

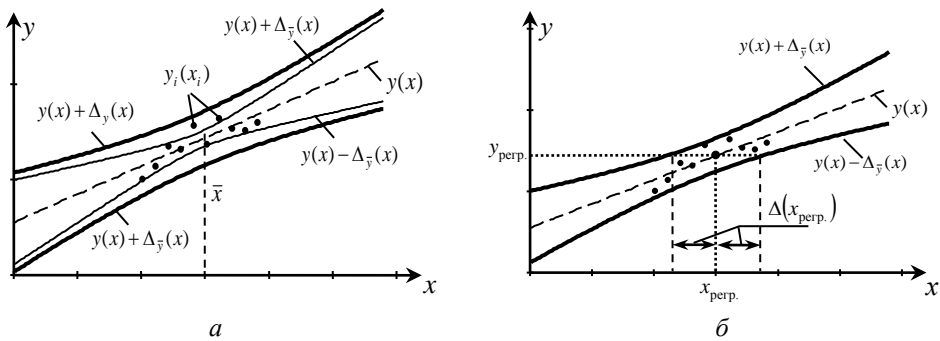


Рис. 5.7. Графіки довірчих областей регресії для значень y (а) і x (б)

Визначення довірчих інтервалів (розширеної невизначеності) значень x розрахованих за регресією. Для певних метрологічних завдань, часто застосовується градування (побудова лінійної регресії) шляхом вимірювання певного показника для різних значень іншого оцінюваного показника об'єкта чи речовини, наприклад залежність поглинання, проходження і відбиття інфрачервоного випромінювання від вологості матеріалу або інших його властивостей. Отримана градувальна залежність y подальшому використовується для обчислення показника $x_{\text{перп}}$ об'єкту, що дає відгук $x_{\text{перп}} = (y_{\text{перп}} - b) / k$. СКВ випадкової похибки значення $x_{\text{перп}}$ пов'язано з s_y

$$s(x_{\text{перп}}) = (s_y / k) \sqrt{1/m + 1/n + (x_{\text{перп}} - \bar{x})^2 / \sum_{i=1}^n (x_i - \bar{x})^2}$$

Довірчий інтервал похибки (розширена невизначеність) (рис. 5.7,б) знаходять як: $\Delta(x_{\text{перп}}) = U(x_{\text{перп}}) = t_{\alpha/2}(v) s(x_{\text{перп}})$, (5.25)
 Значення x_i у можуть мати постійні систематичні похибки θ_x або θ_y . В

цьому випадку СКВ значення $x_{\text{перп}} : s_{\Sigma}(x_{\text{перп}}) = \sqrt{(\theta_x / \sqrt{3})^2 + (\theta_y / k \sqrt{3})^2 + s^2(x_{\text{перп}})}$.
 Підставляючи $s_{\Sigma}(x_{\text{перп}})$ у формулу (5.25) замість $s(x_{\text{перп}})$ отримаємо шукані значення довірчих інтервалів похибки $\Delta(x_{\text{перп}})$.

5.4. Статистичне опрацювання випадкових кутів

Завдання аналізу випадкових кутів на площині зустрічаються в різних областях діяльності людини – в медицині, біології, географії, метеорології, економіці, фізиці, радіотехніці, неруйнівному контролю та діагностиці тощо. Для визначення статистичних характеристик кутових даних застосовують відповідні статистичні методи аналізу.

Результати кутових вимірів відображають точками чи напрямками, що задаються променями на колі одиничного радіусу (рис. 5.8).

Розбиття кола на клас-інтервали і групування даних за цими клас-інтервалами дозволяє відображати кутові дані значного обсягу M круговими гістограмами (рис. 5.9,а). Останні є колом, розділеним на t секторів (найчастіше рівної величини). Для j -го сектора підраховується кількість кутів M_j , значення яких належать цьому сектору.

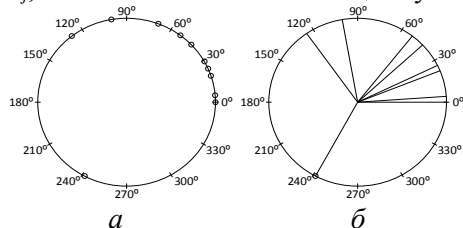


Рис. 5.8. Графічне зображення результатів кутових вимірів точками на колі (а) чи променями (б)

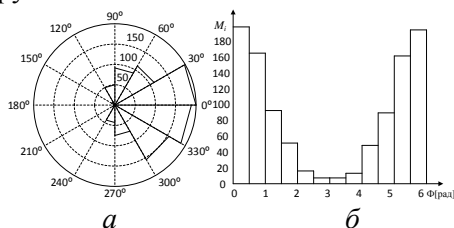


Рис. 5.9. Кругова (а) та розгорнута кругова (б) гістограми результатів кутових вимірів

Друга форма графічного представлення кутових вимірів – розгорнута на інтервалі $\theta \in [0, 2\pi)$ кутова гістограма (рис. 5.9,б). На цих гістограмах в певних секторах маємо трикутники або прямокутники, висоти яких дорівнюють значенням відносних частот – M_j/M .

У загальному випадку вимірювані кути можуть виходити за межі інтервалу $[0, 2\pi)$. В цьому випадку областю значень кута $\Phi > 2\pi$ є множина всіх дійсних чисел R , а значення таких кутів зручно інтерпретувати за допомогою гвинтової лінії (рис. 5.10), що в декартовій системі координат (x, y, Φ) задається рівняннями: $x = \cos \Phi$, $y = \sin \Phi$, $\Phi = 2\pi n + \varphi$, $n = 0, \pm 1, \dots$

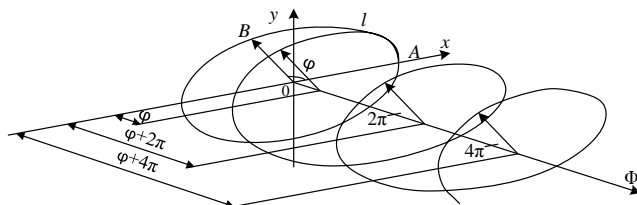


Рис. 5.10. Представлення кутів $\Phi > 2\pi$ на гвинтовій лінії

На рис. 5.10 напрямку вектора в площині xOy відповідає дуга AB довжиною l . Цій дузі через гвинтову лінію ставиться у відповідність одне з чисел осі ОФ виду $\varphi + 2\pi n$. Числове значення n задається початковими умовами або визначається умовами фізичної реалізації експерименту.

$$\text{Форма подання кута у вигляді } \Phi = [\Phi/2\pi] \cdot 2\pi + \varphi, \quad (5.27)$$

де $[\Phi]$ – ціла (кількість повних обертів) і $\varphi \equiv \Phi \pmod{2\pi} \in [0, 2\pi)$ – дробова (за модулем 2π) частини кута Φ , є базовою формулою подання довільного кута.

Розглядаючи періодичні процеси, що розгортаються в часі з періодом T (або в просторі з просторовим періодом λ), можна поставити у відповідність їх періодам величину 2π . Тоді кожному інтервалу часу t (або відстані D) відповідає кут $\varphi \equiv 2\pi(t/T) \pmod{1}$ (або $\varphi \equiv 2\pi(D/\lambda) \pmod{1}$). Отже дробова частина t (або дробова частина D) ідентифікується з кутом φ , після чого вона може досліджуватися методами кутометрії.

Статистичні кутові вимірювання. Результати аналізу даних кутових вимірів мають характерні властивості, відмінні від властивостей лінійних вимірів, аргументом яких є числова вісь. Це визначає необхідність використання відповідних статистичних оцінок для опрацювання кутових даних. Табл. 5.6 подає порівняльний аналіз таких оцінок випадкових величин і кутів.

Імовірнісна модель кута на площині. У статистичному аналізі кутових даних за основну математичну модель покладено модель випадкового кута

$$\Psi(\omega) = [\Psi(\omega)/2\pi] 2\pi + \psi(\omega), \quad \omega \in \Omega, \quad (5.28)$$

де $\psi(\omega) \equiv \Psi(\omega) \pmod{2\pi}$ – дробова частина випадкового кута.

Таблиця 5.6

Характеристика	Випадкова величина	Випадковий кут
1	2	3

Розділ 6. Статистичне опрацювання даних вимірювань

Визначення	<p>Дійсною випадковою величиною є функція $\xi(\omega)$ з областю визначення $\Omega \subset \{\omega\}$ і областю значень $X \subset R$ така, що для довільного $x \in X$ множина $\omega \in \Omega$, для яких $\xi(\omega) < x$ є подія A з множини випадкових подій \mathfrak{Z}, яка задана на фіксованому імовірнісному просторі $(\Omega, \mathfrak{Z}, P)$</p>	<p>Дійсним випадковим кутом є функція $\Psi(\omega)$ з областю визначення $\Omega \subset \{\omega\}$ і областю значень $\Theta \subset R$ така, що для довільного $\theta \in \Theta$ дробова частина кута</p> $\psi(\omega) = \Psi(\omega) - [\Psi(\omega)/2\pi]^+ 2\pi = \Psi(\omega) \bmod 2\pi$ <p>є випадковою величиною</p>
Функція розподілу ймовірностей	<p>Функція розподілу ймовірностей $F(x)$, $x \in R$ випадкової величини $\xi(\omega)$:</p> $F(x) = P\{\omega \in \Omega : \xi(\omega) < x\},$ $-\infty < x < \infty.$ <p>Властивості $F(x)$: монотонність (якщо $x_1 < x_2$, то $F(x_1) \leq F(x_2)$); непервність зліва ($\lim_{x \uparrow x_0} F(x) = F(x_0)$); $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$; $F(\infty) = \lim_{x \rightarrow \infty} F(x) = 1$</p>	<p>Функція розподілу ймовірностей $F(\theta)$, $-\infty < \theta < \infty$ випадкового кута $\Psi(\omega)$:</p> $F(\theta) = G(\theta') + [\theta/2\pi]^+ + C, \quad 0 \leq \theta' < 2\pi,$ <p>де $G(\theta')$ – функція розподілу випадкової величини $\psi(\omega)$, θ' – дробова частина θ, C – довільна постійна.</p> <p>Властивості $F(\theta)$: монотонність (якщо $\theta_1 < \theta_2$, то $F(\theta_1) \leq F(\theta_2)$); $G(\theta')$ неперервна справа ($\lim_{\theta' \downarrow \theta'_0} G(\theta') = G(\theta'_0)$), тому $G(2\pi) = 1$; $F(\theta + 2\pi n) - F(\theta) = n$; $n = 0, 1, 2, \dots$; $F(-\infty) = \lim_{\theta \rightarrow -\infty} F(\theta) = 0$; $F(\infty) = \lim_{\theta \rightarrow \infty} F(\theta) \rightarrow \infty$</p>
Квантили законів розподілу	<p>Розв'язок рівняння $F(\alpha_\gamma) = \gamma$, де $\gamma = \int_{-\infty}^{\alpha_\gamma} p(x) dx \in [0, 1]$ – задана ймовірність, відносно $\alpha_\gamma \in (-\infty, \infty)$ є квантилем розподілу $F(x)$ рівня γ</p>	<p>Розв'язок рівняння $G(\theta'_\gamma) = \gamma$, де $\gamma \in [0, 1]$ – задана ймовірність, відносно $\theta'_\gamma \in [0, 2\pi)$ є квантилем розподілу $F(\theta)$ рівня γ, де</p> $\gamma = \int_0^{\theta'_\gamma} p(\theta) d\theta$
Характеристична функція	<p>Характеристичною функцією випадкової величини $\xi(\omega)$ називається математичне сподівання функції</p> $f(u) = M e^{ju\xi} = \int_{-\infty}^{\infty} e^{(ju)x} dF(x),$ $u \in (-\infty, \infty)$	<p>Характеристичною функцією випадкового кута $\Psi(\omega)$ (послідовністю тригонометричних моментів відносно нульового напрямку) називається математичне сподівання функції $\exp(j\rho\Psi)$</p> $f_n = M e^{jn\rho} = \int_{2\pi k}^{2\pi(k+1)} \exp(jn\theta) dF(\theta) = \rho_n \exp(j\mu_n),$ <p>k – довільне ціле число; $n = 0, \pm 1, \pm 2, \dots$</p>

<p>Щільність розподілу ймовірностей</p>	<p>Якщо існує така функція $p(x)$, і для довільного $x \in R$</p> $F(x) = \int_{-\infty}^x p(y) dy,$ <p>то $p(x)$ називається щільністю ймовірностей.</p> <p>Властивості $p(x)$:</p> $p(x) \geq 0; \int_{-\infty}^{\infty} p(x) dx = 1;$ <p>3) для довільних a і b</p> $P\{a \leq \xi < b\} = F(b) - F(a) = \int_a^b p(x) dx.$	<p>Якщо існує така функція $p(\theta)$, $\theta \in [0, 2\pi]$, і для довільних $\theta: F(\theta_2) - F(\theta_1) = \int_{\theta_1}^{\theta_2} p(\theta) d\theta$,</p> <p>$(\theta_2 - \theta_1) \in [0, 2\pi]$, то $p(\theta)$ називається щільністю ймовірностей на колі.</p> <p>Властивості $p(\theta)$: $p(\theta) \geq 0$;</p> $p(\theta + 2\pi) \equiv p(\theta); \int_{\theta_1}^{\theta_1 + 2\pi} p(\theta) d\theta = 1;$ <p>для довільних θ_1, θ_2:</p> $P(\theta_1 < \Psi \leq \theta_2) = \begin{cases} 0, \theta_2 \leq \theta_1, & F(\theta_2) - F(\theta_1), \\ \theta_1 \leq \theta_2 \leq \theta_1 + 2\pi, & 1, \theta_2 > \theta_1 + 2\pi. \end{cases}$ <p>$p(\theta)$ періодична функція з періодом 2π</p>
<p>Нерівність Чебишева</p>	$P\{ \xi - M\xi \geq g\sqrt{D\xi}\} \leq \frac{1}{g^2}$	$P\{ \sin(\Psi - \mu_1/2) > \varepsilon\} < v/2\varepsilon^2,$ $0 < \varepsilon < 1, 0 < \frac{v}{2\varepsilon^2} < 1.$
<p>Перший початковий момент (середнє значення)</p>	<p>Математичним сподіванням випадкової величини $\xi(\omega)$ з функцією розподілу $F(x)$ називається число, що визначено інтегралом Стільтєса</p> $M\xi = \int_{-\infty}^{\infty} x dF(x)$	<p>Круговим середнім значенням випадкового кута $\Psi(\omega)$ з функцією розподілу $G(\theta)$, для якого $f_1 = Me^{j\Psi} \neq 0$, називається кут</p> $\mu_1 = \text{Arg} f_1$
<p>Медіана</p>	<p>Медіаною неперервного розподілу $F(x)$ випадкової величини $\xi(\omega)$ називається таке значення $x = Me$, для якого однаково ймовірно, чи виявиться випадкова величина більше чи менше Me, тобто</p> $P(\xi < Me) = P(\xi > Me)$	<p>Круговою медіаною неперервного розподілу на колі $G(\theta)$ випадкового кута $\psi(\omega)$ називається значення кута μ_m, яке є одним з розв'язків рівняння</p> $Q(\mu_m) = F(\mu_m + \pi) - F(\mu_m) - 0,5 = \int_{\mu_m}^{\mu_m + \pi} p(\theta) d\theta - 0,5 = 0,$ <p>і для якого $Q(\mu_m - 0,5\pi)$ максимальне</p>
<p>Мода</p>	<p>Мода – це значення Mod випадкової величини $\xi(\omega)$, для якого функція $p(x)$ набуває максимального значення</p>	<p>Мода – це значення Mod випадкового кута $\psi(\omega)$, для якого функція $p(\theta)$ набуває максимального значення</p>

Розділ 6. Статистичне опрацювання даних вимірювань

Модуль тригонометр. моменту	–	Модуль тригонометричного моменту n -го порядку – це $\rho_n = f_n $; для $n=1$ модуль $\rho = f_1 $ визначає довжину результуючого вектора, що уявляє математичне сподівання випадкового вектора $(\cos \Psi, \sin \Psi)$
Розкид значень випадкової величини і випадкового кута	Дисперсією випадкової величини $\xi(\omega)$ з функцією розподілу $F(x)$ називається математичне сподівання квадрата відхилення $\xi(\omega)$ від $M\xi$ $D\xi = \int_{-\infty}^{\infty} (x - M\xi)^2 dF(x).$	Круговою дисперсією випадкового кута $\Psi(\omega)$ з функцією розподілу $G(\theta')$ називається величина $\nu = 1 - \rho = 1 - f_1 ,$ яка характеризує відхилення значення випадкового кута від його середнього значення

Функція розподілу ймовірностей $\psi(\omega) \in [0, 2\pi)$ в загальному вигляді визначається як: $G(x') = P\{\omega \in \Omega : 0 < \psi(\omega) \leq x'\}$, $x' \equiv x \pmod{2\pi}$. (5.29)

Вираз (4.29) задає функцію розподілу ймовірностей $\psi(\omega)$. Її властивості приведені в табл.5.6, а приклади графіків – на рис. 5.11.

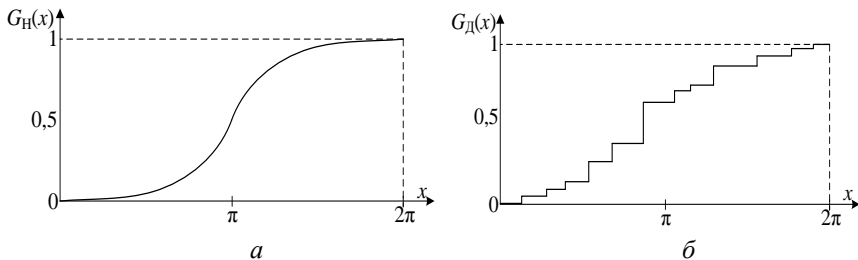


Рис. 5.11. Графіки функцій розподілу ймовірностей випадкового кута безперервної (а) та дискретної (б) функцій

Інтегральна функція розподілу ймовірностей випадкового кута $\Psi(\omega)$ в загальному вигляді визначається як

$$F(x) = G(x') + [x/2\pi] + C, \quad x \in \mathbb{R}, \quad x' \equiv x \pmod{2\pi}. \quad (5.30)$$

Крім зазначених у в табл. 5.6, вона має такі властивості:

1. $F(x') = F(x'') - F(0_-)$, $x' \in [0, 2\pi)$.
2. $F(x) - x/2\pi$ є періодичною функцією з періодом 2π .
3. Для $0 \leq x_2 - x_1 \leq 2\pi$ маємо

$$P\{\omega \in \Omega : x_1 < \Psi(\omega) \leq x_2\} = \begin{cases} 0, & x_2 \leq x_1, \\ F(x_2) - F(x_1), & (x_1 < x_2 \leq x_1 + 2\pi), \\ 1, & x_2 > x_1 + 2\pi \end{cases}$$

4. $F(x+2\pi) - F(x) = 1, x \in R.$

5. Для характерного випадку різниці кутів в межах $0 < x_2 - x_1 \leq 2\pi$

$$\begin{cases} F(x_2) - F(x_1) = P\{x_1 < x' \leq x_2\} \text{ для } [x_2/2\pi] = [x_1/2\pi]; \\ (F(x_2') - F(0)) + (F(x_1') - F(2\pi)) = \\ = P\{(0 \leq x' \leq x_2') + P(x_1' \leq x' \leq 2\pi)\} \text{ для } [x_2/2\pi] = [x_1/2\pi] + 1. \end{cases}$$

6. Постійна складова C у (4.30) частіше дорівнює нулю, але в залежності від задачі кутових вимірів може набувати і інших значень.

З наведених властивостей $F(x), x \in R$ витікає, що ця функція на інтервалах з початком в довільній точці числової осі має однакові прирости $F(x+2\pi(k+1)) - F(x+2\pi k) = 1$, що не суперечить властивостям нормованої імовірнісної міри. Приклад графіку $F(x)$ дано на рис. 5.12.

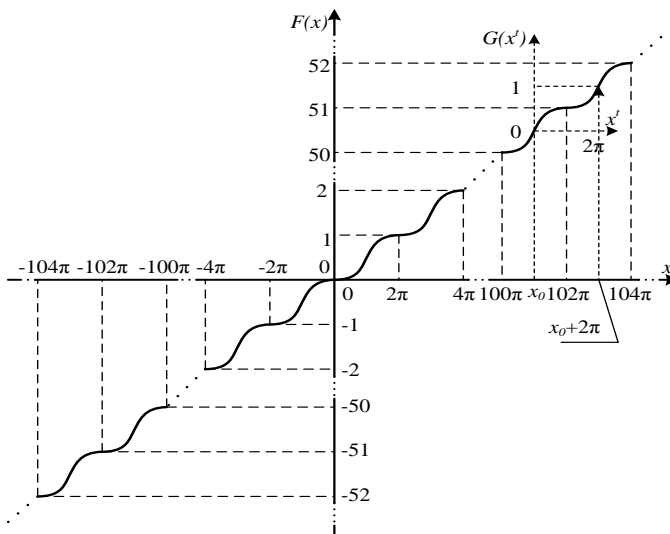
Властивості характеристичної функції випадкового кута $\Psi(\omega)$

Характеристична функція випадкового кута, визначення якої дано в табл. 5.6, є послідовністю тригонометричних моментів виду

$$f_n = a_n + jb_n, \tag{5.31}$$

де косинус і синус моменти визначаються відносно початкового нульового напрямку відповідно до формул

$$a_n = \mathbf{M}\{\cos n\Psi(\omega)\} = \int_0^{2\pi} \cos nx dG(x), \quad b_n = \mathbf{M}\{\sin n\Psi(\omega)\} = \int_0^{2\pi} \sin nx dG(x).$$



Характеристична функція випадкового кута має такі властивості.

1. Модуль характеристичної функції $|f_n| \leq 1$.

2. Для $n = 0$ маємо $f_0 = 1$.

3. Характеристична функція від'ємного аргумента дорівнює комплексноспряженій характеристичній

Рис. 5.12. Приклад графіку неперервної функції $F(x)$

Розділ 6. Статистичне опрацювання даних вимірювань

функції $f_{-n} = f_n^*$, так як $a_{-n} = a_n$, $b_{-n} = -b_n$.

4. Характеристична функція f суми незалежних випадкових кутів $\psi_1(\omega), \dots, \psi_m(\omega)$ дорівнює добутку їх характеристичних функцій $f_n^{(k)}$ (де n –

порядок функції, k – порядковий номер випадкового кута): $f = \prod_{k=1}^m f_n^{(k)}$.

5. Щільність розподілу ймовірностей випадкового кута однозначно визначається його характеристичною функцією $p(\theta) = (2\pi)^{-1} \sum_{n=-\infty}^{\infty} f_n e^{-jn\theta}$.

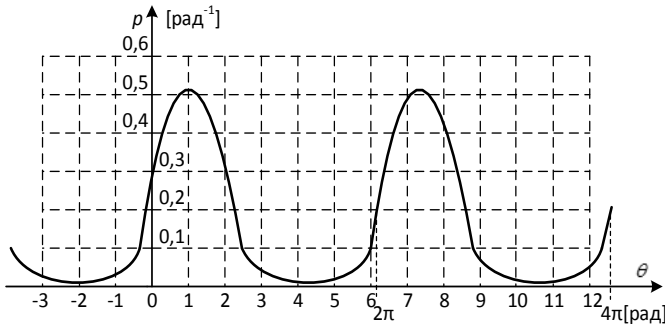
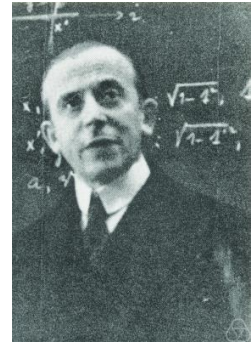


Рис. 5.13. Приклад графіку щільності розподілу ймовірностей випадкового кута

Закони розподілів ймовірностей випадкових кутів. Для кутів має місце операція додавання за модулем 2π , що визначає властивість періодичності щільності розподілу ймовірностей випадкових кутів (рис. 5.13)



**Ріхард Едлер
фон Мізес
(Richard von Mises,
1883 - 1953),**

видатний німецький математик і механік австрійського походження.

Народився в Лемберзі, Австро-Угорщина (нині Львів, Україна). Роботи в області прикладної механіки, механіки рідин, аеродинаміки, аеронавтики, статистики, теорії ймовірностей. У теорії ймовірностей запропонував частотну концепцію поняття ймовірності, ввів до загального вжитку інтеграл Стілтєса і першим роз'яснив роль теорії марківських ланцюгів у фізиці; був неймовірно динамічною і універсальною людиною, особливо добре обізнаним в області технології; визнаний

і істотно відрізняє їх від розподілів випадкових експерт в поезії.
величин.

Щільність розподілу ймовірності на рис. 5.13 періодична з періодом 2π . Такі розподілу називають *одночастотними*. *Багаточастотні* розподілу мають період менше в ціле число разів, тобто період $2\pi/j$, $j=2,3,\dots$. Розглянемо найживаніші щільності ймовірностей випадкових кутів.

Розподіл Мізеса. Щільність розподілу ймовірностей Мізеса випадкового кута $\psi(\omega)$ визначається за формулою

$$p_M(\theta|\mu, k) = (2\pi I_0(k))^{-1} e^{k \cos(\theta - \mu)}, \quad |\mu| < \infty, k > 0, \quad (5.32)$$

де $I_0(k)$ – модифікована функція Бесселя першого роду нульового порядку; μ – круговий середній напрям випадкового кута; k – параметр концентрації випадкового кута в околі μ . Приклади графіків $p_M(\theta|\mu, k)$ наведені на рис. 5.14.

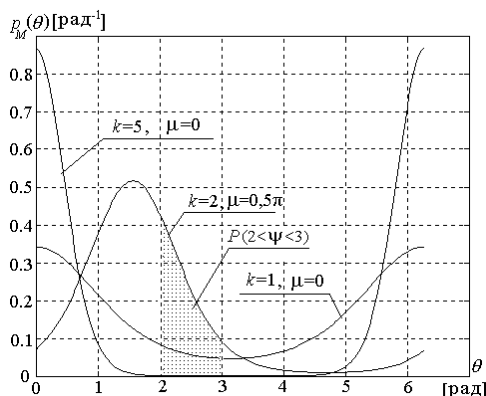


Рис. 5.14. Графіки щільності розподілу Мізеса

Характеристична функція розподілу Мізеса

$$f_n = \alpha_n = I_n(k)/I_0(k). \quad (5.33)$$

Цей розподіл має важливу властивість: найбільш правдоподібною оцінкою параметра μ є круговий середній напрямок. Розподілом Мізеса задовільно апроксимується намотаний гауссовий

розподіл при виконанні наступного співвідношення між їх параметрами розсіювання:

$$\sigma = \sqrt{-2 \ln(I_1(k)/I_0(k))} .$$

Намотаний гауссовий розподіл належить до сім'ї намотаних розподілів, які отримують за нелінійного перетворення випадкової величини $\xi(\omega)$:

$$\psi(\omega) \equiv [K\xi(\omega)] \bmod 2\pi, \quad (5.34)$$

де K – масштабний коефіцієнт.

Перетворення (5.34) призводить до трансформації законів розподілу на прямій в намотані (на коло одиничного радіусули) закони розподілу ймовірностей випадкових кутів.

Якщо $\xi(\omega)$ має щільність розподілу ймовірностей $p(x)$, то безперервний випадковий кут $\psi(\omega)$ також має безперервну щільність

$$p_{2\pi}(\theta) = \sum_{i=-\infty}^{\infty} p(\theta + 2\pi i). \quad (5.35)$$

У загальному випадку функція $p_{2\pi}(\theta)$ несиметрична відносно середини інтервалу $[0, 2\pi)$, а на його кінцях має однакові значення: $p_v(0) = \lim_{\theta \rightarrow 2\pi} p_v(\theta)$.

Щільність ймовірності намотаного гауссового розподілу рис. 5.15:

$$p_{2\pi}(\theta/\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{j=-\infty}^{\infty} \exp\left(-\frac{[(\theta - \mu)(\bmod 2\pi) + 2\pi j]^2}{2\sigma^2}\right), \quad (5.36)$$

де μ – математичне сподівання; σ – СКВ випадкової величини $\xi(\omega)$.

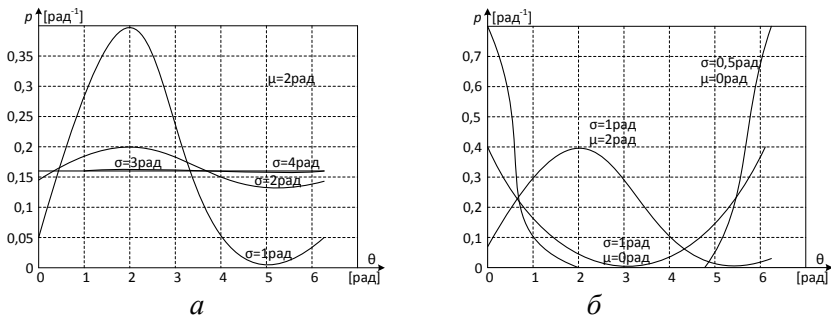


Рис. 5.15. Приклади щільності ймовірностей намотаного гауссового розподілу

Розподіл (5.36) має таку властивість: сума незалежних кутів з цим розподілом має розподіл того ж виду, але з іншими характеристиками.

Характеристична функція намотаного гауссового розподілу ймовірностей: $f_n = \alpha_n = \exp(-n^2 \sigma^2 / 2)$, $\beta_n = 0$, $n \in Z$. (5.37)

Розділ 6. Статистичне опрацювання даних вимірювань

У статистичної кутометрії має місце *центральна гранична теорема*, яка стверджує наступне: для незалежних випадкових кутів $\psi_1(\omega), \dots, \psi_n[\omega]$, що мають однакову функцію розподілу ймовірності, розподіл ймовірностей нормованої суми кутів $\left[\sum_{j=1}^n \psi_j(\omega) / \sqrt{n} \right] \bmod 2\pi$, в разі $n \rightarrow \infty$, наближається до намотаного гауссового розподілу.

Дані про інші безперервні розподіли ймовірностей випадкових кутів зведені в табл. 5.7.

Таблиця 5.7

Назва	Щільність розподілу ймовірностей $p(\theta)$	Характеристична функція f_n
Рівномірний	$\frac{1}{2\pi}$	$\frac{e^{2\pi i j} - 1}{2\pi i j} = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases}$
Кардіоїдний	$(2\pi)^{-1} [1 + 2\rho \cos(\theta - \mu)]$, $ \mu < \infty, \quad \rho < 0,5$	-
Трикутний	$[4 - \pi^2 \rho + 2\pi \rho \pi - \theta] / 8\pi$, $\rho \leq 4 / \pi^2$	$\alpha_{2n-1} = \rho / (2n-1)^2$, $\alpha_{2n} = 0, \quad \beta_n = 0$
Намотаний Коші	$\frac{1}{2\pi} \cdot \frac{1 - \rho^2}{1 - \cos 2\rho + \rho^2}$ $\rho = e^{-a} \in [0, 1]$	$\rho^{ a }$

Вибіркові кругові характеристики випадкових кутів. У випадку багаторазових вимірювань випадкового кута $\psi(\omega) \in [0, 2\pi)$ отримують вибірку його реалізацій $\{\varphi_j, j = \overline{1, M}\}$ обсягу M . Оцінювання числових характеристик $\psi(\omega)$ виконується шляхом визначення *вибіркових кругових характеристик*, зміст яких наведено в табл. 5.8. Найчастіше вживаними є вибіркове кругове середнє і вибіркова кругова дисперсія.

Вибіркове кругове середнє. Результат одного виміру – φ_j , зображається плоским кутом, якому відповідає на колі радіуса $r=1$ дуга довжиною l_j між позитивною напіввіссю абсцис і вектором $\overrightarrow{OP_j}$ з декартовими $(\cos \varphi_j, \sin \varphi_j)$ (полярними $(1, \varphi_j)$) координатами (рис. 5.16,а). Передбачається, що вимірювання виконуються з кроком $2\pi/m$. Дискретному характеру результатів вимірів відповідає розбиття кола на m клас-інтервалів (точка P_j відповідає середині j -го клас-інтервалу).

Будь-яка конструктивна характеристика кругового середнього повинна задовольняти умові адитивності

$$\{L(\varphi_1 - v, \dots, \varphi_M - v)\}(\bmod 2\pi) \equiv \{L(\varphi_1, \dots, \varphi_M) - v\}(\bmod 2\pi), \quad v \in [0, 2\pi),$$

тобто для довільного кута дробові частини (по модулю 2π) чисел $\{L(\varphi_1, \dots, \varphi_M) - v\}$ і $L(\varphi_1 - v, \dots, \varphi_M - v)$ збігаються. Тобто кут, що задається характеристикою $L(\varphi_1, \dots, \varphi_M)$, повинен адитивно залежати від початкового кута v . Цій вимозі задовольняє вибірковий круговий середній кут.

Таблиця 5.8

Найменування статистичної характеристики	Зміст, математичний запис
Вибіркове кругове середнє статистики углов	$\varphi_c = \{\arctg(S/C) + 0,5\pi\{2 - (\text{sign}S) \times [1 + \text{sign}C]\}\} \bmod 2\pi;$ $C = M^{-1} \sum_{j=1}^M \cos \varphi_j; \quad S = M^{-1} \sum_{j=1}^M \sin \varphi_j$
Вибіркова середня довжина результуючого вектора	$r = \sqrt{C^2 + S^2}$
Вибіркова кругова дисперсія статистики кутів	$V = 1 - r$
Кругове стандартне відхилення статистики кутів	$\sigma = \sqrt{-2 \ln(1 - V)} = \sqrt{-2 \ln r};$ $V = 1 - \exp(-0,5\sigma^2)$
Вибіркова кругова медіана статистики кутів	Це кут, якому відповідає точка кола P, для якої діаметр PQ ділить значення статистики навпіл, і в околі P маємо максимальну концентрацію значень
Вибіркова мода статистики кутів	Це кут, що відповідає точка кола, в околі якої спостерігається максимальна концентрація значень статистики
Вибірковий круговий розмах статистики кутів	Довжина найменшої дуги статистики, що визначається з варіаційного ряду $T_j = \varphi_{j+1} - \varphi_j, \quad j = 1, \dots, M - 1; \quad T_M = 2\pi - \varphi_M + \varphi_1;$ $W = 2\pi - \max\{T_1, \dots, T_m\}$
Вибіркова характеристика асиметрії статистики кутів	$g_1 = \frac{b_2(m)}{V^{3/2}} = \frac{r_2 \sin[m_2(0) - 2m]}{V^{3/2}}$
Вибіркова характеристика ексцесу статистики кутів	$g_2 = \frac{a_2(m) - (1 - V)^4}{V^2} = \frac{r_2 \cos m_2(0) - 2n - (1 - V)^4}{V^2}$
Вибірковий тригонометричний момент порядку u відносно напрямку α (u – ціле число)	$T_u(\alpha) = M^{-1} \sum_{j=1}^M e^{iu(\theta_j - \alpha)} = a_u(\alpha) + ib_u(\alpha) = r_u(\alpha) e^{im_u(\alpha)};$ $a_u(\alpha) = M^{-1} \sum_{j=1}^M \cos[u(\varphi_j - \alpha)] =$ $= a_u(0) \cos(u\alpha) + b_u(0) \sin(u\alpha);$

Розділ 6. Статистичне опрацювання даних вимірювань

$$\begin{aligned}
 b_u(\alpha) &= M^{-1} \sum_{j=1}^M \sin[u(\varphi_j - \alpha)] = \\
 &= -a_u(0) \sin(u\alpha) + b_u(0) \cos(u\alpha); \\
 r_u(\alpha) &= \sqrt{a_u^2(\alpha) + b_u^2(\alpha)} = \sqrt{a_u^2(0) + b_u^2(0)} = r_u(0); \\
 m_u(\alpha) &= m_u(0) - u\alpha.
 \end{aligned}$$

Для множини $\{\varphi_j, j = \overline{1, M}\}$ *вибіркове кругове середнє* – це напрямок суми всіх одиничних векторів $\overrightarrow{OP_1}, \dots, \overrightarrow{OP_M}$ (тобто аргумент сумарного вектора). На рис. 5.16,б сумарний вектор $\overrightarrow{OP} = \sum_{j=1}^3 \overrightarrow{OP_j}$ має декартові координати $(\sum_{j=1}^3 \cos \varphi_j, \sum_{j=1}^3 \sin \varphi_j)$ і вибіркове кругове середнє φ_c .

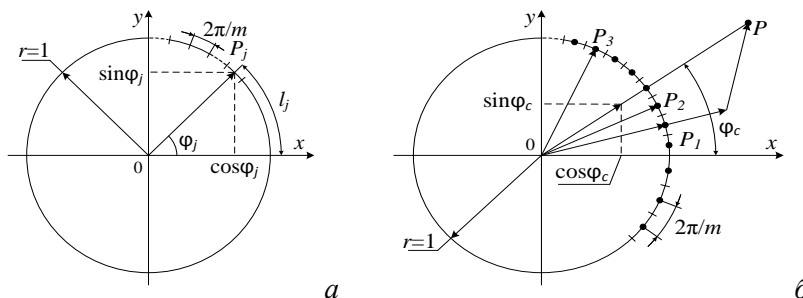


Рис. 5.16. Графічне зображення результатів одного вимірювання (а) і усереднення вибірки кутів обсягом $M = 3$ (б)

Його механістичне тлумачення полягає в наступному: якщо всім точками одиничного кола, що відповідають множині $\{\varphi_j, j = \overline{1, M}\}$, приписати однакову "масу", то координати "центру мас" цієї системи визначаються як:

$$C = M^{-1} \sum_{j=1}^M \cos \varphi_j, \quad S = M^{-1} \sum_{j=1}^M \sin \varphi_j. \quad (5.38)$$

Значення C и S дозволяють обчислити кут φ_c (див. табл. 5.8), правильність обчислення якого перевіряється виконанням умови

$$\sum_{j=1}^M \sin(\varphi_j - \varphi_c) = 0. \quad (5.39)$$

Перерахунок координат вектора \mathbf{r} з декартової системи в полярну виконується за формулами: $r = |\mathbf{r}| = \sqrt{C^2 + S^2}$, $C = r \cos \varphi_c$, $S = r \sin \varphi_c$. Величину r називають *вибірковою середньою результуючою довжиною* вектора. Якщо $r = 0$ значення $\varphi_c \in [0, 2\pi)$ однозначно не визначається.

Вибіркова кругова дисперсія випадкового кута відносно φ_c визначається як:
$$V(\varphi_c) = M^{-1} \sum_{j=1}^M \{1 - \cos(\varphi_j - \varphi_c)\} = 1 - M^{-1} \sum_{j=1}^M \cos(\varphi_j - \varphi_c) = 1 - r. \quad (5.40)$$

Характеристика $V(\varphi_c)$ має властивість інваріантності відносно початку відліка кутів.

Статистичні інтервальні оцінки випадкових кутів. У випадку усереднення на колі сукупності M статистично незалежних випадкових кутів вибірка кругова дисперсія середнього кута зменшується в \sqrt{M} разів. З урахуванням цього факту методика інтервального оцінювання середнього значення випадкового кута з намотаним гауссовим розподілом визначається такою послідовністю дій.

1. Розрахунок сум C і S (5.38).
2. Розрахунок вибіркового кругового середнього φ_c .
3. Перевірка умови (5.39) для підтвердження правильності розрахунків.
4. Визначення вибіркової кругової дисперсії $V = 1 - r$ одного кутового виміру. У разі $M < 15 \dots 20$ до отриманого значення r застосовують поправку типу *поправки Шеннарда*: $r_g = r c_g$, де $c_g = \pi/m \sin(\pi/m)$.
5. Розрахунок кругового стандартного відхилення σ .
6. Визначення кругового стандартного відхилення кругового середнього $\sigma_c = \sigma/\sqrt{M}$ вибірки обсягу M .
7. Розрахунок межі довірчого інтервалу $\alpha_{\text{дов}}$ для заданої довірчої ймовірності $P_{\text{дов}}$ і функції Лапласа.
8. Формування довірчого інтервалу:
нижня межа довірчого інтервалу – $(\varphi_c - \alpha_{\text{дов}}) \bmod 2\pi$;
верхня межа довірчого інтервалу – $(\varphi_c + \alpha_{\text{дов}}) \bmod 2\pi$

5.5. Типові методи опрацювання результатів вимірювань

Опрацювання прямих багаторазових вимірювань регламентовано нормативною документацією. Статистичне опрацювання групи результатів прямих багаторазових незалежних вимірювань передбачає:

- виключення з результатів вимірювань систематичних похибок;
- обчислення оцінки вимірюваної величини;
- обчислення СКВ результатів вимірювань;
- перевірку наявності грубих похибок і їх виключення;
- перевірку гіпотези належності результатів вимірювань гауссовому розподілу;

Розділ 6. Статистичне опрацювання даних вимірювань

обчислення довірчих меж випадкової похибки оцінки вимірюваної величини;

обчислення довірчих меж невиключеної систематичної похибки оцінки вимірюваної величини;

обчислення довірчих меж похибки оцінки вимірюваної величини.

Перевірка гіпотези про приналежність результатів вимірів гауссовому розподілу проводиться з рівнем значущості q від 10 до 2%. Конкретні значення q вибираються з умов проведення вимірювань.

Для визначення довірчих меж похибки оцінки вимірюваної величини довірна ймовірність P в більшості технічних вимірювань приймається рівною 0,95.

За оцінку вимірюваної величини x приймається середнє арифметичне значення виправлених результатів вимірювань:

$\bar{x} = \sum_{i=1}^n x_i / n$, де x_i – i -й результат вимірювань; n – число виправлених результатів.

Отриманий результат заокруглюється.

Якщо всі результати вимірювань містять постійну систематичну похибку, її допускається виключити після обчислення середнього арифметичного значення невиправлених результатів вимірювань.

СКВ s групи, що містить n результатів вимірювань, яке є оцінкою характеристики розсіювання результатів вимірювань, викликаних наявністю випадкових складових похибки, обчислюється як:

$$s = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)}.$$

СКВ середнього арифметичного (фактично результату вимірювання) $s_{\bar{x}}$ визначається як: $s_{\bar{x}} = s / \sqrt{n}$.

Довірчі межі випадкової похибки оцінки вимірюваної величини встановлюються для результатів вимірювань, що належать гауссовому розподілу. Якщо число результатів вимірювань $n \leq 15$ належність їх до гауссового розподілу не перевіряють. В цьому випадку висновок про гауссовість розподілу приймається на підставі апріорних відомостей. Якщо немає достатніх підстав вважати розподіл випадкової похибки гауссовим, методи обчислення довірчих меж випадкової похибки розробляються додатково.

У випадку $15 < n \leq 50$ для перевірки належності результатів вимірювань до нормального розподілу слід віддати перевагу розглянутому вище складеному критерію, а якщо $n > 50$ – можна застосовувати критерії χ^2 Пірсона або Колмогорова – Смірнова.

Довірчі межі ε (без урахування знаку) випадкової похибки оцінки вимірюваної величини обчислюються за формулою $\varepsilon = t s_{\bar{x}}$, де t – коефіцієнт Стьюдента, що залежить від довірчої ймовірності P і числа результатів вимірювань n (числа ступенів свободи).

Невиключена систематична похибка (НСП) оцінки вимірюваної величини має складові, які залежать від методу вимірювання, засобу вимірювальної техніки або обумовлені іншими чинниками.

Якщо випадкові складові похибки нехтовно малі, за межу складових НСП приймають, наприклад, межі допустимих основної і додаткової похибок засобів вимірювань. Межа НСП оцінки вимірюваної величини за наявності менше трьох складових ($m < 3$) НСП, кожна з яких представлена межею Θ_i , оцінюється за формулою

$$\Theta_{\Sigma} = \pm \sum_{i=1}^m |\Theta_i|.$$

Якщо $m \geq 3$ розподіли похибок всередині меж цих складових (похибки засобів вимірювань кожного типу, похибки поправок і т. і.) розглядаються як розподіли випадкових величин. За відсутності даних про вид розподілів складових, останні розглядаються як рівномірні. Тоді довірчі межі допускається обчислювати як

$$\Theta_{\Sigma}(P) = \pm k \sqrt{\sum_{i=1}^m |\Theta_i|^2},$$

де k – коефіцієнт, що визначається прийнятою довірчою ймовірністю, числом складових НСП і їх співвідношенням між собою.

Для довірчої ймовірності $P = 0,95$ значення k нехтовно мало залежить від m і співвідношення між складовими НСП, тому для зазначеної довірчої ймовірності приймається $k = 1,1$. Якщо закони розподілу складових НСП відомі і відмінні від рівномірного, то $\Theta_{\Sigma}(P)$ обчислюють шляхом побудови композиції цих законів розподілу.

У випадках, коли НСП з'являється в результаті виключення систематичної похибки, викликаної дією впливаючої величини Y на вимірювану величину X , необхідно визначити залежність $X = f(Y)$. В цьому разі в обчисленні меж НСП оцінки вимірюваної величини необхідно враховувати коефіцієнт впливу $\partial f / \partial Y$, який одержують при розкладанні даної функції в ряд Тейлора.

За наявності k складових НСП, представлених межами, і $m - \mu$ складових НСП, отриманих від дії впливаючих величин та представлених з коефіцієнтами впливу, формула для $\Theta_{\Sigma}(P)$ матиме вид

$$\Theta_{\Sigma}(P) = \pm k \sqrt{\sum_{i=1}^{\mu} \Theta_i^2 + \sum_{i=\mu+1}^m \left(\frac{\partial f}{\partial Y_i} \Theta_i\right)^2}. \quad (5.41)$$

Розділ 6. Статистичне опрацювання даних вимірювань

Довірча ймовірність для обчислення меж невиключеної систематичної похибки набуває такого ж виду, як і у випадку обчислення довірчих меж випадкової похибки результату вимірювання.

Довірчі межі похибки оцінки вимірюваної величини знаходять шляхом побудови композиції розподілів випадкових похибок і НСП, що розглядаються як випадкові величини. Межі похибки оцінки вимірюваної величини (без урахування знака) визначаються як $\Delta = K s_{\Sigma}$, де K – коефіцієнт, що залежить від співвідношення випадкової складової похибки і складової, що обумовлена НСП. Сумарне СКВ s_{Θ}

оцінки вимірюваної величини: $s_{\Sigma} = \sqrt{s_{\theta}^2 + s_x^2}$, де $s_{\theta} = \Theta_{\Sigma} / (k\sqrt{3})$ – середньоквадратичне відхилення НСП; k – коефіцієнт, який визначається значенням P , числом складових НСП і їх співвідношенням між собою. Коефіцієнт K визначається за емпіричною формулою $K = [\varepsilon + \Theta_{\Sigma}(P)] / (s_x + s_{\theta})$.

Оформлення запису оцінок вимірюваних величин проводиться згідно з правилами міждержавної стандартизації. За симетричних довірчих меж похибки оцінку вимірюваної величини представляють у формі $\bar{x} \pm \Delta, P$. Числове значення оцінки вимірюваної величини має закінчуватись цифрою того розряду, що і значення похибки.

За відсутності даних про вид функцій розподілів складових похибок оцінювання вимірюваної величини і необхідності подальшого опрацювання результатів вимірювань або аналізу похибки оцінювання вимірюваної величини результат представляють у формі $\bar{x}; s_x; n; \Theta$. Оцінки s_x і Θ можуть бути виражені в абсолютній і відносній формах.

Опрацювання результатів непрямих вимірювань. Непрямі вимірювання передбачають визначення шуканого значення фізичної величини A через результати вимірювань інших фізичних величин a_1, a_2, \dots, a_m , пов'язаних із шуканою відомою функціональною залежністю

$$A = f(a_1, \dots, a_i, \dots, a_m),$$

(5.42)

яка встановлюється або з теоретичних передумов, або експериментально з похибкою, якою можна знехтувати. Результати вимірювань аргументів і оцінки їх похибок можуть бути отримані з прямих, непрямих, сукупних або спільних вимірювань. Відомості про аргументи можуть бути взяті з довідкової літератури або технічної документації.

Методики опрацювання результатів непрямих вимірювань поділяються на три типи:

методика опрацювання за лінійної залежності і відсутності кореляції між похибками вимірювань аргументів;

методика опрацювання за нелінійної залежності і відсутності кореляції між похибками вимірювань аргументів;

методику опрацювання для корельованих похибок вимірювань аргументів за наявних рядів окремих значень вимірюваних аргументів.

Оцінювання довірчих меж похибок результату непрямого вимірювання зазвичай виконують для довірчої ймовірності 0,95.

Непрямі вимірювання за лінійної залежності. Результат вимірювання A пов'язаний з m вимірюваними аргументами a_1, a_2, \dots, a_m рівнянням з постійними коефіцієнтами b_1, b_2, \dots, b_m виду

$$A = b_1 a_1 + b_2 a_2 + \dots + b_m a_m. \quad (5.43)$$

Якщо коефіцієнти b_1, b_2, \dots, b_m визначають експериментально, то завдання визначення результату вимірювання величини вирішується поетапно: спочатку оцінюють кожний доданок $b_i a_i$ як результат непрямого вимірювання, отриманий як добуток двох вимірюваних величин, а потім знаходять оцінку вимірюваної величини A .

$$\text{Результат непрямого вимірювання } \tilde{A}: \quad \tilde{A} = \sum_{i=1}^m b_i \tilde{a}_i, \quad (5.44)$$

де \tilde{a}_i – результат вимірювання аргумента a_i . СКВ результату

$$\text{непрямого вимірювання:} \quad s(\tilde{A}) = \sqrt{\sum_{i=1}^m b_i^2 s^2(\tilde{a}_i)}, \quad (5.45)$$

де $s(\tilde{a}_i)$ – СКВ a_i .

Для непрямих вимірювань довірчі межі випадкової похибки результату, за умови узгодженості результатів вимірювань аргументів з гауссовим розподілом, обчислюється (без урахування знака) як

$$\varepsilon(p) = t_\alpha(v_{\text{эф}}) S(\tilde{A}), \quad (5.46)$$

де $t_\alpha(v_{\text{эф}})$, – коефіцієнт Стьюдента, відповідний довірчій ймовірності $P = 1 - \alpha$ і числу ефективних ступенів свободи $v_{\text{эф}}$, обчислюваному за наведеною нижче формулою і заокругленому до меншого цілого

$$v_{\text{эф}} = \left[\sum_{i=1}^m b_i^2 S^2(\tilde{a}_i) - 2 \left(\sum_{i=1}^m b_i^4 S^4(\tilde{a}_i) / (n_i + 1) \right)^2 \right] / \sum_{i=1}^m (b_i^4 S^4(\tilde{a}_i) / (n_i + 1)),$$

де n_i – число вимірювань для визначення аргумента a_i .

Розділ 6. Статистичне опрацювання даних вимірювань

Довірчі межі невиключеної систематичної похибки результату непрямого вимірювання $\Theta_{(p)}$ (без урахування знака) для ймовірності P обчислюють за заданими межами Θ_i невиключених систематичних похибок результатів вимірювань аргументів $\Theta_{(p)} = k \sqrt{\sum_{i=1}^m b_i^2 \Theta_i^2}$, де k – поправковий коефіцієнт, який визначається прийнятою довірчою ймовірністю P і числом m складових Θ_i . Для $P = 0,95$ приймають $k = 1,1$. Сумарну похибку результату непрямого вимірювання оцінюють на основі композиції законів розподілів випадкових і невиключених систематичних похибок, згідно наступних правил:

якщо $\Theta(P)/S(\tilde{A}) > 8$, то як похибку результату непрямого вимірювання приймають невиключену систематичну складову похибки вимірювання;

якщо $\Theta(P)/S(\tilde{A}) < 0,8$, то як похибку результату непрямого вимірювання приймають випадкову складову похибки вимірювання;

якщо $0,8 \leq \Theta(P)/S(\tilde{A}) \leq 8$, то довірча межа похибки результату непрямого вимірювання $\Delta(P)$ визначається як $\Delta(P) = K(\varepsilon(P) + \Theta(P))$, де K – коефіцієнт, що залежить від P і відношення $\Theta(P)/S(\tilde{A})$ (табл.5.9).

Таблиця 5.9

$\Theta_{(p)}/S(\tilde{A})$	0,5	0,75	1	2	3	4	5	6	7	8
K (для $P = 0,95$)	0,81	0,77	0,74	0,71	0,73	0,76	0,78	0,79	0,8	0,81

Непрямі вимірювання за нелінійної залежності. Для непрямих вимірювань за нелінійної залежності і некорельованих похибках вимірювань аргументів використовується метод лінеаризації, який передбачає розкладання нелінійної функції в ряд Тейлора:

$$f(a_1, \dots, a_m) = f(\tilde{a}_1, \dots, \tilde{a}_m) + \sum_{i=1}^m (\partial f / \partial a_i) \Delta a_i + R, \quad (5.47)$$

де $f(a_1, \dots, a_m)$ – нелінійна функціональна залежність вимірюваної величини \tilde{A} від вимірюваних аргументів a_i ; $\partial f / \partial a_i$ – перша похідна від функції f за аргументом a_i , обчислена в точках $\tilde{a}_1, \dots, \tilde{a}_m$; Δa_i – відхилення результату вимірювання аргументу a_i від його середньоарифметичного; $R = \frac{1}{2} \sum_{j=1}^m \frac{\partial^2 f}{\partial a_i \partial a_j} (\Delta a_i a_j)$ – залишковий член,

яким можна знехтувати за умови виконання нерівності

$R < 0.8 \sqrt{\sum_{i=1}^m (\partial f / \partial a_i)^2 S^2(\tilde{a}_i)}$, де $S^2(\tilde{a}_i)$ – СКВ випадкових похибок

результату вимірювання a_i -го аргументу.

Результат вимірювання \tilde{A} : $\tilde{A} = f(\tilde{a}_1, \dots, \tilde{a}_m)$. (5.48)

Значення СКВ випадкової похибки результату непрямого вимірювання: $S(\tilde{A}) = \sqrt{\sum_{i=1}^m (\partial f / \partial a_i)^2 S^2(\tilde{a}_i)}$.

Довірчі межі випадкової похибки, межі НСП і сумарну похибку результату непрямого вимірювання за умови, що розподіли похибок результатів вимірювань аргументів узгоджуються з гауссовим розподілом, обчислюють згідно з наведеною вище процедурою для лінійної залежності, підставляючи замість коефіцієнтів b_1, b_2, \dots, b_m перші похідні $\partial f / \partial a_1, \partial f / \partial a_2, \dots, \partial f / \partial a_m$ відповідно.

Непрямі вимірювання за наявності кореляції між похибками вимірювань аргументів. Для визначення наявності кореляційного зв'язку між похибками вимірювань аргументів розраховується оцінка

коефіцієнта парної кореляції $\tilde{r} = \sum_{i=1}^n (a_{hi} - \tilde{a}_h)(a_{ji} - \tilde{a}_j) / \sqrt{\sum_{i=1}^n (a_{hi} - \tilde{a}_h)^2 (a_{ji} - \tilde{a}_j)^2}$,

где a_j, a_{hi}, a_{ji} – результати i -го вимірювання h -го і j -го аргументів; n – число вимірювань кожного з аргументів.

Якщо вимірювана величина залежить від аргументів, необхідно перевірити відсутність кореляційних зв'язків між похибками всіх парних поєднань аргументів.

За гауссового закону розподілу випадкових похибок аргументів критерієм відсутності кореляційного зв'язку між похибками результатів вимірювань аргументів є виконання нерівності

$|\tilde{r} \sqrt{n-2} / \sqrt{1-\tilde{r}^2}| < t_\alpha$, де t_α – коефіцієнт Стьюдента, відповідний рівню значущості α і числу ступенів свободи $\nu = n - 2$.

Якщо гіпотеза про наявність кореляційного зв'язку між похибками аргументів підтверджується, то для обчислення результату непрямих вимірювань використовують метод приведення, який передбачає наявність ряду окремих значень вимірюваних аргументів, отриманих шляхом багаторазових вимірювань. Цей метод можна застосовувати за невідомих розподілів похибок вимірювань аргументів. Метод ґрунтується на приведенні ряду окремих значень вимірюваних непрямим методом величин до ряду прямих вимірювань. Отримувані поєднання окремих результатів вимірювань аргументів підставляють у формулу

Розділ 6. Статистичне опрацювання даних вимірювань

(5.43) і обчислюють окремі значення вимірюваної величини A : A_1, A_2, \dots, A_l .

Результат непрямого вимірювання визначається як: $\tilde{A} = \sum_{j=1}^l A_j / l$, де l – число значень вимірюваної величини; A_j – окреме значення вимірюваної величини, отримане в результаті підстановки j -го поєднання узгоджених результатів вимірювань аргументів у (5.43).

СКВ випадкових похибок результату непрямого вимірювання обчислюється за формулою $S(\tilde{A}) = \sqrt{\sum_{j=1}^l [(A_j - \tilde{A})^2 / l(l-1)]}$.

Довірчі межі випадкової похибки для результату вимірювання $\Delta = TS(\tilde{A})$, де T – коефіцієнт, що залежить від виду розподілу окремих значень вимірюваної величини і від вибраної довірчої ймовірності.

Межі невиключеної систематичної похибки результату непрямого вимірювання і довірчі межі похибки результату непрямого вимірювання обчислюють за наведеними вище формулами.

Результат непрямого вимірювання і його похибку можна представити або у вигляді $\tilde{A}, s(\tilde{A}), n, \Theta(P)$, якщо передбачається його дослідження і зіставлення з іншими результатами, або як $\tilde{A} \pm \Delta(P)$ у випадку симетричності меж похибки результату вимірювання.

Основні висновки

Випадковий характер результату вимірювання обумовлюють дві основні причини. Перша відноситься до вимірювань постійних і незмінних в часі фізичних величин. Результати вимірювання спотворюються випадковим вимірювальним шумом, тому отримані послідовно результати хаотично змінюються в часі відносно певного значення. Друга причина виникає через вимірювання величин, які самі по собі мають випадковий характер, наприклад значення параметра в партії виробів. У цьому випадку існують випадкові значення в просторі набору значень. Тому для опрацювання результатів вимірювань широко використовується статистичний аналіз.

Застосовуючи статистичний аналіз, на основі отриманих результатів вимірювань необхідно отримати інформацію про істине значення вимірюваної величини і оцінку ступеня її достовірності. Статистичні методи ґрунтуються на таких поняттях, як вимірювання, генеральна сукупність і вибірка.

За результатами проведення багаторазових вимірювань фізичної величини отримують вибірку результатів, які повинні представляти істинне

значення величини з необхідною (заданою) точністю або достовірністю, що фактично відповідає вимозі репрезентативності вибірки.

На основі статистичного опрацювання вибірових даних необхідно отримати оцінки значень вимірюваної величини і її похибки. Оцінки бувають точковими і інтервальними. Точкове оцінювання полягає в тому, що безпосередньо отримане значення оцінки приймається за невідоме значення параметра розподілу. Якість точкових оцінок визначають за таким показником – незміщеність, слухності, ефективності.

Незміщеність оцінки передбачає, що зі зростанням кількості вимірювань середнє значення оцінки прямує до значення оцінюваного параметра. Зсув оцінки може бути обумовлений як властивостями рівняння вимірювання, так і похибками вимірювання, юстування, калібрування, випадковим характером отриманої вибірки або комбінацією цих факторів. У супротивном випадку воно не залежить від обсягу вибірки і називається систематичною похибкою. Наявність систематичних похибок не дає можливості судити про істинне значення вимірюваного параметра.

Оцінка називається слухною якщо зі збільшенням обсягу вибірки вона збігається з відповідним параметром генеральної сукупності. Це означає, що прецизійність результату вимірювання зі збільшенням обсягу аналізованих даних повинна зростати.

Вимога ефективності оцінки: для вибірок рівного обсягу ефективна оцінка повинна мати мінімальне розсіювання.

Результат вимірювання може також бути представлений у вигляді інтервалу, якому із заданою ймовірністю належить істинне значення вимірюваної величини. У цьому випадку застосовується інтервальне оцінювання, тобто визначається інтервал, який із заданою ймовірністю містить шукане (істинне) значення параметра розподілу. Цей інтервал називається довірчим або надійним, а відповідна ймовірність – довірчою ймовірністю або надійністю. Чим більше довірчий інтервал, тим більше упевненість у тому, що оцінювана величина знаходиться в межах цього інтервалу.

У практиці вимірювань зустрічаються завдання з експериментального визначення значущості систематичних похибок вимірювальних каналів, алгоритмів опрацювання даних і ІВС в цілому. Значення випадкової похибки (прецизійність) може бути відоме заздалегідь або визначено експериментальним шляхом. Дані завдання в математичній статистиці називаються перевіркою гіпотез про математичне сподівання.

Випадкові похибки вимірювальних каналів і ІВС досліджують за допомогою критеріїв порівняння стандартних відхилень (дисперсій) – це критерій порівняння стандартного відхилення із заданою величиною, який застосовують для порівняння прецизійності досліджуваної ІВС, оціненої експериментальним шляхом або із заданою точністю іншого обладнання, і критерій порівняння стандартних відхилень двох генеральних сукупностей, застосований для порівняння точності двох засобів вимірювань за

Розділ 6. Статистичне опрацювання даних вимірювань

результатами вимірювань значення однієї фізичної величини і визначення стабільності метрологічних характеристик ІВС у часі або за зміни впливаючих факторів.

Критерії перевірки гіпотез Граббса і Кохрена про середні значення і стандартних відхилення серій вимірювань дозволяють одночасно порівняти результати вимірювань та їх прецизійність, які оцінені за кількома групами даних. Такими групами можуть бути, наприклад, групи результатів вимірювань, виконаних декількома вимірювальними каналами ІВС.

Кутові вимірювання мають місце в багатьох областях діяльності людини. Результати кутових вимірів в інтервалі $[0, 2\pi)$ відображають точками на колі одиничного радіуса або променями на площині. Вимірювані кути можуть виходити за межі інтервалу $[0, 2\pi)$. Областю значень кутів більших 2π є множина дійсних чисел. Значення таких кутів інтерпретують гвинтовою лінією.

Кутові вимірювання мають характерні властивості відмінні від властивостей лінійних вимірювань, аргументом яких є числова вісь. Це визначає необхідність застосування відповідних статистичних оцінок для опрацювання результатів кутових вимірювань.

Виконання для кутів операції додавання по модулю 2π визначає властивість періодичності законів щільності розподілу ймовірностей випадкових кутів, чим вони істотно відрізняються від розподілів ймовірностей випадкових величин. У практиці кутових вимірювань найчастіше використовують розподіли Мізеса, сім'ю намотаних розподілів (Гаусса, Коші та ін.), кардіоїдний, рівномірний, трикутний.

Найчастіше використовуваними статистичними характеристиками результатів кутових вимірів є вибіркові кругові середнє, дисперсія, медіана, результуюча середня довжина вектора, асиметрія і ексцес.

Запитання для самоконтролю

1. Поясніть терміни "вибірка", "генеральна сукупність", "порядкова статистика", "статистична оцінка".
2. Які показники якості статистичних оцінок Вам відомі?
3. Наведіть незміщену, слушну і ефективну оцінку математичного сподівання генеральної сукупності.
4. Які оцінки називають точковими, інтервальними?
5. Що таке ранг елемента вибірки і які правила його визначення?
6. Яку вибірку називають ранжированою?
7. Наведіть і поясніть відомі методи визначення закону розподілу за даними ряду вимірювань.
8. Які критерії згоди Ви знаєте, і з якою метою вони застосовуються в опрацюванні результатів вимірювань?
9. Поясніть суть і послідовність застосування спеціалізованих критеріїв

згоди з певними законами розподілу.

10. Які гіпотези про статистичні характеристики Вам відомі?
11. Які є способи перевірки рівноточності результатів багаторазових вимірювань?
12. Як підтвердити відсутність систематичної похибки в результатах вимірювань?
13. З якою метою проводиться перевірка статистичної рівності математичних сподівань або дисперсії груп даних?
14. З якою метою проводиться перевірка приналежності вимірювання до загальної сукупності експериментальних даних?
15. Які методи виявлення результатів з грубими похибками Вам відомі?
16. Що таке статистична однорідність експериментальних даних. Які існують методи перевірки цієї властивості?
17. Що таке тренд? З якою метою проводиться перевірка результатів вимірювань на наявність тренда?
18. Назвіть відомі Вам одиниці вимірювання кутів.
19. У чому полягає відмінність випадкових кутів і величин.
20. Дайте визначення і наведіть основні властивості функції розподілу ймовірностей випадкового кута.
21. Дайте визначення і наведіть властивості: щільності ймовірності випадкового кута, характеристичної функції випадкового кута.
22. Наведіть приклади розподілів ймовірностей випадкових кутів.
23. Назвіть числові характеристики випадкових кутів і поясніть їх зміст.
24. Викладіть суть і властивості вибіркової кругової дисперсії.
25. Дайте визначення і геометричну інтерпретацію вибіркового кругового розмаху статистики кутів.
26. У чому полягає особливість формування інтервальних оцінок результатів кутових вимірів?

Conclusions

The random nature of measurement result, in general, is caused by two main reasons. The first reason relates to measurement of constants, not time varying physical quantities. Measurement results are distorted by random in time measuring noise, so results, obtained consistently in time, vary randomly in time relative to a certain value. The second reason concerns measurement of quantities with random values, for example, the parameter value in production lot. In this case, there are random values in the space of value set. Therefore, when processing measurement results, statistical analysis is widely used.

Operations conducted with the purpose of timely and correct results fixation are the measuring under observation ones. Observation results include all types of errors inherent in measurement procedure. Thus, measurement results could be considered as observation.

Розділ 6. Статистичне опрацювання даних вимірювань

Quantity results sampling is obtained as a result of repeated measurements. These results should represent true value of a quantity with necessary (given) adequacy or accuracy that actually meets the requirement of sampling representativeness.

On the basis of statistical processing of sample data it is necessary to obtain estimation values of measured quantity and its errors. Estimates could be point and interval. Point estimation is an estimation method, meaning that directly obtained estimation value is used as an unknown value of distribution parameter. Point estimates should satisfy the following quality indicators - unbiasedness, consistency, efficiency.

Estimation unbiasedness suggests that while observations number increases the average estimation value tends to the value of estimated parameter. This bias may be caused by measurement equation properties and measurement errors, alignment, calibration, random nature of obtained sampling, or a combination of these factors. Otherwise, it does not depend on sampling size and is called a systematic error. Presence of systematic errors makes it impossible to judge the true value of measured parameter.

Estimation will be consistent if increase of sampling size coincides with the corresponding parameter of general totality. This means that measurement result precision should increase with an increase in volume of analyzed data.

Requirement for estimation effectiveness is as follows - for samplings of equal volume effective estimate should have a minimal dissipation.

The measurement result could also be represented as an interval where the true value of measured quantity is given with a given probability. In this case interval estimation is applied. Interval estimation is an estimation method, where the defined interval is the one in which with a given probability a required (true) value of distribution parameter exists.

In measurement practice there are tasks of experimental determination of systematic errors importance of measurement channels, processing algorithms and IMS as a whole, while the value of random error (precision) can be known in advance or determined experimentally. These tasks in mathematical statistics are called testing hypotheses about mathematical expectation.

When studying random errors of measurement channels and IMS, a criteria of standard deviations (variances) comparison is used – it is a criteria of comparison of standard deviation and given quantity, that is used to compare precision of studied IMS, estimated experimentally with a given IMS or with a given precision of other equipment; and a criteria of comparison of standard deviations of two general totalities, that is used to compare the accuracy of two measurement means using the results of value measurement of the single physical quantity and stability determination of IMS metrology characteristics in time or within changing of influencing factors.

Testing criteria for hypothesis about average values and standard deviations of the sets of observations allows simultaneously compare measurement results and their precision, estimated on several data sets. Such groups could be, for example, groups of the results of measurements, performed by several IMS measurement channels. These are Grubbs and Cochran criteria.

When processing observation results, as a rule, the first step is to check whether individual observations belong to general set of experimental data. As a matter of fact, it is check for gross errors in observations. Gross errors belong to errors that occur randomly in repeated observations. They clearly exceed value of errors, justified by measurement conditions and significantly affect the estimating accuracy of average value and standard deviation. Reasons for gross errors could be uncontrolled changes in measurement conditions, power surges, sudden failures in IMS or measuring channels, operator errors, etc. To exclude the results of observations with gross errors criteria of statistical hypothesis testing' are used, such as criteria by Romanovsky, Dixon, etc.

Systematic errors in IMS or individual channels may vary over time that does not allow amending the individual measurement results, and may also cause accuracy characteristics to overstep regulated limits. Identification of such errors can be done by using the criteria of testing the hypothesis of trend presence in a number of observations. Criteria selection depends on its nature: the linear growth or decline of the average value or standard deviation of measurement results, the periodic change of mathematical expectation, etc.

Under condition that random error can be neglected, the trend could be detected by analyzing the alternation of deviations signs or from the center of probability distribution, which may be mathematical expectation, median etc., or derived from previous results. Median criterion of series could also be used, that allows to identify periodic and progressive trend; and criterion of "growing" and "declining" series that is sensitive to the presence of progressive systematic errors.

Angular measurements are used in many areas of human activity. It is accepted that angular measurements results in interval $[0, 2\pi)$ displayed by dots on a unit circle or as a plane rays. In general, measured angles could go beyond interval $[0, 2\pi)$.

Range of angle values that are bigger than 2π is a set of all real numbers. Values of such angles are convenient to interpret as a spiral line. Angles are measured in radians and degrees. The results of angular measurement data analysis have typical properties different from those of linear measurements, i.e. measurements, the argument of which is numerical axis. It determines the necessity to build corresponding statistical estimates for processing the results of angular measurements.

Performing the addition for angles using module 2π determines the periodicity property of laws of random angles probability distribution density, that is significantly differ them from random variables probability distributions. Mises probability distribution, Gauss, Cauchy distributions, cardioid, uniform, triangular distributions are mostly used in angle measurement practice.

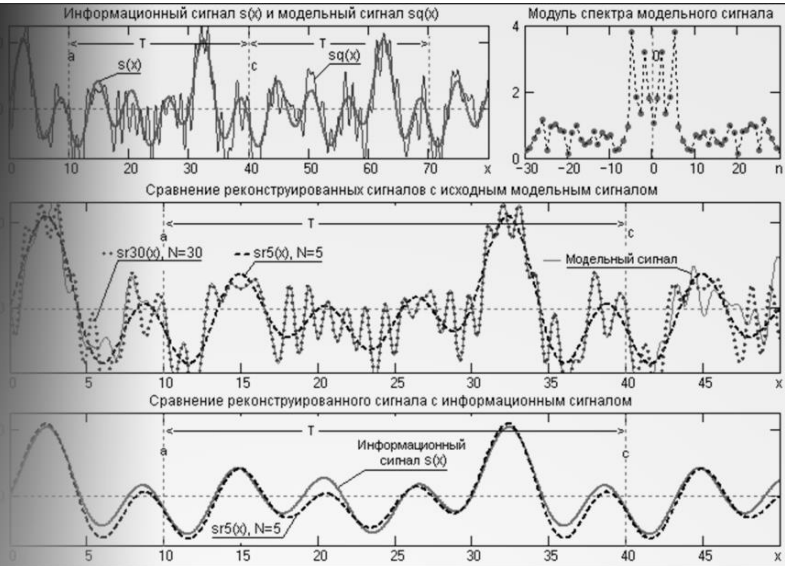
Most commonly used statistical characteristics of angular observations results are sample circular angles statistics average, variance and median, asymmetry and excess characteristics, as well as resultant vector length.

Ключові слова

систематична похибка	systematic error
----------------------	------------------

Розділ 6. Статистичне опрацювання даних вимірювань

прецизійність	precision
результат вимірювання	result of a measurement
статистичне оцінювання	statistical estimation
статистичні критерії	statistical criterions
опрацювання результатів вимірювань	processing of measurement data
кругова гістограма	circular histogram
кругове середнє	mean angle
кругова дисперсія	angular dispersion
розподіл Мізеса	the von Mises distribution
намотаний розподіл	wrapped distribution
середня результуюча довжина вектора	mean resultant length of the vector



СТАТИСТИЧНЕ ОЦІНЮВАННЯ ХАРАКТЕРИСТИК ВПАДКОВИХ ПРОЦЕСІВ

6

- 6.1. Статистичне оцінювання характеристик стаціонарних процесів
- 6.2. Статистичне оцінювання стаціонарності випадкової послідовності
- 6.3. Статистичне оцінювання часових характеристик стаціонарної послідовності
- 6.4. Статистичне оцінювання спектральних характеристик стаціонарної послідовності

6.1. Статистичне оцінювання характеристик стаціонарних процесів

Стаціонарні та ергодичні процеси. Розглянемо характерні особливості використання стаціонарного випадкового процесу для визначення його статистичних характеристик за даними вимірювань в рамках кореляційної теорії.

Відомо, що в рамках кореляційної теорії стаціонарний випадковий процес $\xi(\omega, t)$, $\omega \in \Omega$, $t \in (-\infty, \infty)$ (синоніми назви – *стаціонарний у широкому сенсі, гільбертів стаціонарний, стаціонарний другого порядку, слабостаціонарний*) повністю заданий, якщо визначені:

а) двовимірна функція розподілу

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

$$x_1, x_2 \in R, t_1, t_2, \tau \in (-\infty, \infty);$$

б) або двовимірна щільність імовірностей $p(x_1, x_2, \tau)$;

в) або двовимірна характеристична функція $f(u_1, u_2, \tau)$, $u_1, u_2 \in R$.

Числові характеристики стаціонарного процесу $\xi(\omega, t)$:

математичне сподівання $a_1 = \int_{-\infty}^{\infty} xp(x)dx$, де $p(x)$ – одновимірна

(маргінальна) щільність розподілу процесу $\xi(\omega, t)$;

дисперсія
$$\sigma^2 = \int_{-\infty}^{\infty} (x - a_1)^2 p(x)dx;$$

автоковаріаційна функція
$$B(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2, \tau) dx_1 dx_2;$$

автокореляційна функція

$$R(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - a_1)(x_2 - a_1) p(x_1, x_2, \tau) dx_1 dx_2; \quad (6.1)$$

спектральна щільність потужності, що визначається згідно з теоремою Вінера – Хінчіна

$$S(f) = \int_{-\infty}^{\infty} e^{-i2\pi f \tau} R(\tau) d\tau.$$

(6.2)

Для стаціонарно зв'язаних стаціонарних процесів $\xi_1(\omega, t)$ і $\xi_2(\omega, t)$ визначається взаємокореляційна функція

$$R_{12}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - a_1)(x_2 - a_1) p_{12}(x_1, x_2, \tau) dx_1 dx_2, \quad (6.3)$$

і відповідно взаємна спектральна щільність потужності

$$S_{12}(f) = \int_{-\infty}^{\infty} e^{-i2\pi f\tau} R_{12}(\tau) d\tau \quad (6.4)$$

та функція когерентності

$$v_{12}^2(f) = |S_{12}(f)|^2 / S_1(f)S_2(f), \quad v_{12}^2(f) \in [0,1] \quad (6.5)$$

як нормована спільна функція четвертого порядку статистичного взаємозв'язку двох стаціонарно зв'язаних стаціонарних процесів $\xi_1(\omega, t)$ і $\xi_2(\omega, t)$ в частотній області.

Надалі для опису випадкових процесів використовуватимемо такі позначення:

$\{\xi(\omega, t), \omega \in \Omega_1, t \in (-\infty, \infty)\}$ – стаціонарний процес з неперервною областю визначення $\Omega_1 \times R$ і неперервною областю значень $X \subseteq R$;

$\{\xi(\omega, t_j), \omega \in \Omega_2, t_j \in (-\infty, \infty), j \in N\}$ – стаціонарний процес з дискретним часом як стаціонарна послідовність неперервних випадкових величин;

$\{\zeta(\omega, t_j), \omega \in \Omega_3, t_j \in R, j \in N\}$ – стаціонарний процес з дискретним часом і значеннями як стаціонарна послідовність дискретних випадкових величин.

У класі стаціонарних випадкових процесів виділяється підмножина *ергодичних процесів*, які є конструктивною моделлю для статистичного опрацювання даних вимірювань. Достатні і необхідні умови ергодичності процесу визначаються окремо для математичного сподівання, кореляційної функції і одновимірної функції розподілу. Наведемо необхідні і достатні умови ергодичності стаціонарного процесу $\xi(\omega, t)$:

а) відносно математичного сподівання a_1

$$\lim_{c \rightarrow \infty} (2/c) \int_0^c R(\tau)(1 - \tau/c) d\tau = 0, \quad (6.6)$$

де $R(\tau)$ – автокореляційна функція процесу $\xi(\omega, t)$;

б) відносно автокореляційної функції $R(\tau)$

$$\lim_{c \rightarrow \infty} (2/c) \int_0^c R(s, \tau)(1 - s/c) ds = 0, \forall \tau, \quad (6.7)$$

де автокореляційна функція $R(s, \tau)$ є моментом четвертого порядку відносно процесу $\xi(\omega, t)$, оскільки сформований випадковий процес

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

другого порядку $\zeta(\omega, \tau, t) = [\xi(\omega, t) - a_1][\xi(\omega, t + \tau) - a_1]$ (6.8)

має такі характеристики: математичне сподівання процесу $\zeta(\omega, \tau, t)$ визначається як центральний момент другого порядку процесу $\xi(\omega, t)$

$$\mathbf{M}\zeta(\omega, \tau, t) = R(\tau) \quad (6.9)$$

і відповідно його автокореляційна функція як центральний момент четвертого порядку процесу $\xi(\omega, t)$

$$R(s, \tau) = \mathbf{M}[(\zeta(\omega, \tau, t) - R(\tau))(\zeta(\omega, \tau, t + s) - R(\tau))]; \quad (6.10)$$

в) відносно одновимірної функції розподілу $F(x, t) \equiv F(x)$

$$\lim_{c \rightarrow \infty} (2/c) \int_0^c (F(x_1, x_2, \tau) - F(x_1)F(x_2))(1 - \tau/c) d\tau = 0, \quad (6.11)$$

де $F(x_1, x_2, \tau)$ – двовимірна функція розподілу процесу $\xi(\omega, t)$.

Відзначимо, що всі зазначені умови ергодичності процесу $\xi(\omega, t)$ відносно своїх характеристик виражаються через інші характеристики того ж процесу, але на порядок більше. Так, для ергодичності математичного сподівання процесу $\xi(\omega, t)$ потрібні знання його кореляційної функції $R(\tau)$, для одновимірної функції розподілу $F(x)$ – знання двовимірної функції розподілу $F(x_1, x_2, \tau)$. Тому для більшості завдань вимірювань на практиці спочатку необхідно підтвердити гіпотезу про стаціонарність досліджуваного процесу з наступним підтвердженням гіпотези ергодичності процесу.

Статистичне оцінювання. Розглянемо особливості визначення статистичних характеристик ергодичного процесу за даними вимірювання – часового ряду як реалізації досліджуваного процесу. Іншими словами, для апріорно заданого ергодичного процесу з невідомими характеристиками необхідно визначити їх статистичні оцінки і тим самим підтвердити або спростувати статистичну гіпотезу про адекватність моделі.

Така постановка завдання є частинним випадком більш загального завдання побудови (розроблення, створення, обґрунтування) математичної моделі об'єкта дослідження за експериментальними даними (даними спостережень) різного роду. Рішення такого загального завдання неоднозначне і визначається обсягом, змістом експериментальних даних і апріорно відомими характеристиками моделі. Розглянемо два підходи до вирішення такого завдання: у вузькому сенсі, коли як експериментальні дані використовуються дані вимірювань з відповідною похибкою їх визначення, та в широкому

сенсі, коли як експериментальні дані використовуються всі інші дані спостережень.

По суті, така класифікація первинних експериментальних даних для побудови моделі визначає і рівень достовірності результатів вирішення завдання. Отримання даних вимірювань серед експериментальних даних є найбільш витратним як у часі, так і економічно і вимагає використання відповідних засобів вимірювальної техніки. У ряді випадків проводиться узгодження даних вимірювань, які стосуються одного і того ж об'єкту вимірювань, але самі вимірювальні експерименти проведені в різних місцях, в різний час і різними засобами вимірювань. У таких випадках проводиться їх попереднє статистичне опрацювання з метою створення єдиної бази даних вимірювань – як правило, часового ряду. Основними операціями такого опрацювання є: видалення тренду, відсіювання грубих промахів, центрування і, за необхідності, нормування ряду. Разом з цим не можна нехтувати експериментальними даними будь-якого рівня, оскільки вони дають свій внесок у вирішення поставленого завдання. На рис. 6.1 зображена послідовність операцій отримання і опрацювання експериментальних даних у вузькому і широкому сенсах для обґрунтування моделі об'єкта вимірювання.

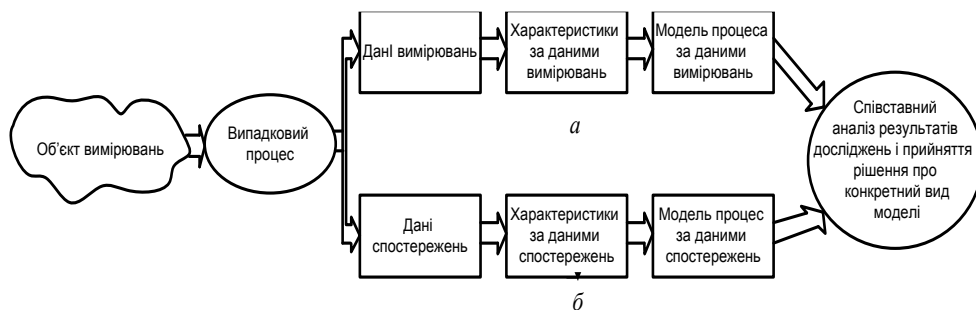


Рис. 6.1. Послідовність операцій отримання і опрацювання експериментальних даних для обґрунтування конкретного виду моделі об'єкта вимірювання: *а* – у вузькому сенсі; *б* – у широкому сенсі

Характерні особливості формування та статистичного опрацювання експериментальних даних у вузькому і широкому сенсах полягають у наступному. Спочатку розглянемо випадок експериментальних даних у вузькому сенсі, коли за умовами експерименту і застосування ІВС є можливість отримати дані вимірювань. Реалізація процесу вимірювання передбачає використання двох мір – міри досліджуваної фізичної величини, яка необхідна для отримання її числового значення, і ймовірнісна міра, яка застосовується на етапі обґрунтування моделі

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

вимірювального експерименту і в процесі статистичного опрацювання експериментальних даних і формування результату (рис. 6.2). На рис. 6.3 для ілюстрації експериментальних даних у вузькому сенсі наведені графіки вимірювання діаметра нейлонової нитки $d(l)$ вздовж її довжини (а) та коливання об'єму льоду на Землі (б).

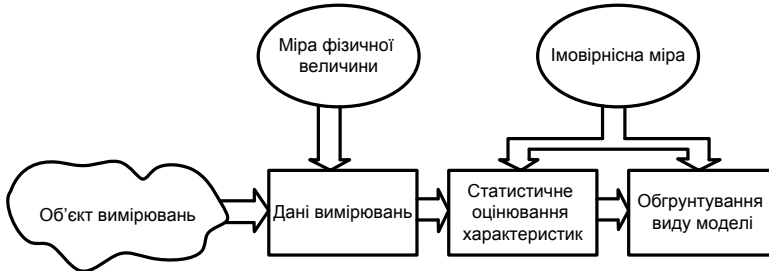


Рис. 6.2. Структура взаємозв'язків різних об'єктів в процесі отримання моделі за даними фізичних вимірювань

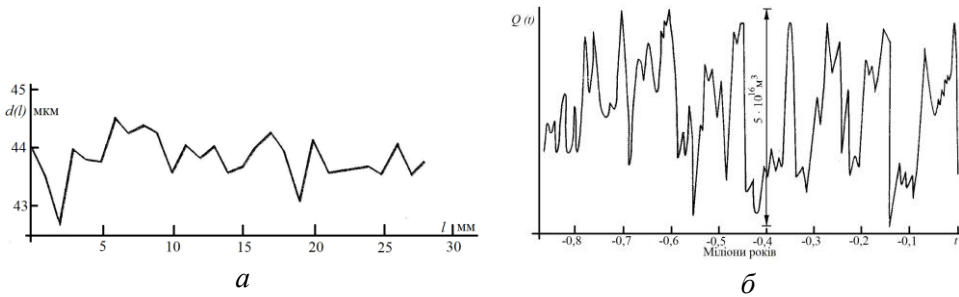


Рис. 6.3. Реалізація випадкового процесу на прикладах вимірювання діаметра нейлонової нитки $d(l)$ вздовж її довжини (а) та зміни сумарного об'єму льоду на Землі $Q(t)$ за останній мільйон років (б)

У другому випадку використовуються експериментальні дані в широком сенсі. Такі дані (наприклад, результати спостережень, випробувань тощо) подаються неіменованими числами, тобто реєструються в числовому вигляді придатному для їх подальшого статистичного опрацювання. Проведення операцій з такими даними спостережень передбачає використання тільки ймовірнісної міри (рис. 6.4). На рис. 6.5 подано приклад експериментальних даних у широкому сенсі.

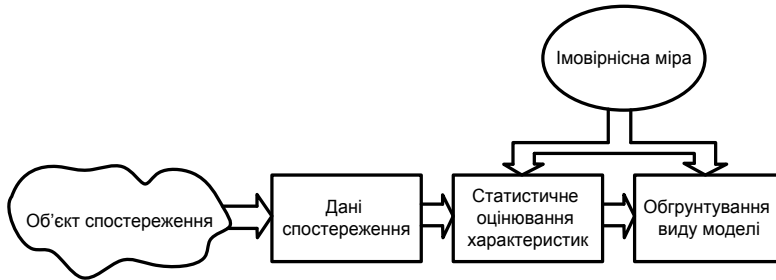


Рис. 6.4. Структура взаємозв'язків різних об'єктів в процесі отримання моделі за даними спостережень

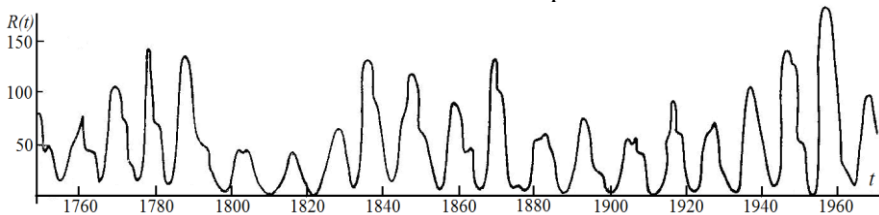


Рис. 6.5. Реалізація випадкового процесу на прикладі оцінювання середніх річних значень числа Вольфа $R(t)$ (числа сонячних плям)

Розглянемо відомі методи статистичного оцінювання характеристик випадкових процесів за експериментальними даними на конкретному прикладі.

Приклад. В результаті проведення експериментальних досліджень зареєстрована серія m реалізацій стаціонарного випадкового процесу $\xi(\omega, t)$ на кінцевому інтервалі часу T . Серія реалізацій $x_1(t), x_2(t), \dots, x_m(t)$ задана на однакових кінцевих часових інтервалах T і отримана за однакових умов проведення експерименту.

Графіки m реалізацій процесу $\xi(\omega, t)$ з різними варіантами усереднення наведені на рис. 6.6, на якому також показані стрілками два можливі методи усереднення для отримання статистичних оцінок характеристик процесу:

- 1) усереднення за множиною (ансамблем) реалізацій;
- 2) усереднення в часі за однією реалізацією.

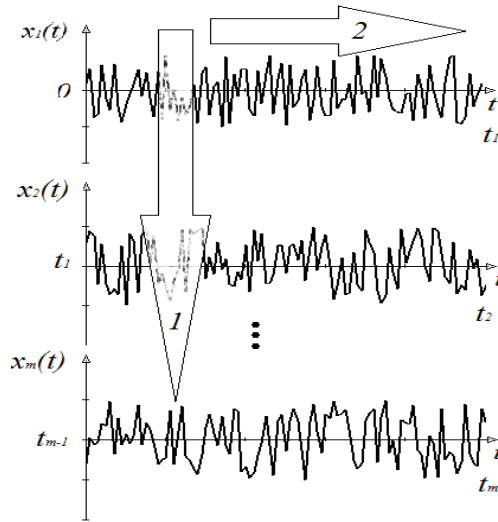


Рис. 6.6. Графіки серії m реалізацій $\{x_j(t), j = \overline{1, m}\}$ стаціонарного випадкового процесу $\xi(\omega, t)$ та умовне позначення можливих методів усереднення за ансамблем реалізацій (1) і за однією реалізацією в часі (2)

Метод усереднення за ансамблем реалізацій є загальним методом статистичного оцінювання характеристик випадкових сигналів за експериментальними даними. Такий метод використовується як для стаціонарних, так і для нестаціонарних випадкових процесів. Для нестаціонарних процесів метод використовується тільки в разі формування ансамблю синхронізованих у часі (тобто з однаковою початковою фазою розвитку процесу або його динаміки) реалізацій.

Метод усереднення в часі за однією реалізацією застосовують для ергодичних випадкових процесів. Зупинимося на цьому методі детальніше.

На практиці важливо знати, коли результати усереднення – середні – збігаються до певної величини і чи рівні середні за ансамблем реалізацій середнім у часі. Від цього залежить весь план вимірювального експерименту, його тривалість і економічні витрати. Ці питання є предметом розгляду ергодичної теорії, а досліджувані при цьому закони збіжності іменуються законами *великих чисел* або *ергодичними теоремами*. Центральне місце у вирішенні ймовірнісних завдань займає ергодична теорема Бірхгофа – Хінчина, яка була доведена як загальна теорема механіки у 1931 р. американський математиком Г. Бірхгофом. Через три роки радянський математик А. Я. Хінчин показав, що ця теорема має теоретико-імовірнісне узагальнення і для стаціонарної послідовності випадкових величин і формується таким чином.

Для стаціонарної послідовності випадкових величин $\xi_1(\omega)$, $\xi_2(\omega)$, $\xi_3(\omega)$, ..., у яких $\mathbf{M}\xi_j(\omega)$ є кінцевими з ймовірністю одиниця, послідовність середніх арифметичних $n^{-1} \sum_{j=1}^n \xi_j(\omega)$ збігається до межі.

Відзначимо наступне: якщо в умовах цієї теореми припустити існування дисперсії $\mathbf{M}\left\{\left[\xi_j(\omega) - \mathbf{M}\xi_j(\omega)\right]^2\right\} = \mathbf{D}\xi_j(\omega) < \infty$, то така теорема доводиться на основі нерівності Чебишева.

Дана теорема має не тільки теоретичне, але й велике прикладне значення: за виконання умови ергодичності під час дослідження стаціонарного випадкового процесу не обов'язково проводити багаторазові експерименти для отримання набору його реалізацій, а достатньо обмежитись однією реалізацією на тривалому інтервалі часу.

Поняття ергодичності відноситься до конкретної характеристики процесу і умови їх ергодичності були наведені вище. Це означає, що один і той же процес може бути ергодичним відносно математичного сподівання і не бути ергодичним відносно, наприклад, кореляційної функції. Тому завжди необхідно вказувати, в якому сенсі розуміється властивість ергодичність, а замість виразу *ергодичний процес* правильніше вживати вираз *процес ергодичний щодо конкретної характеристики*, наприклад, кореляційної функції.

Математичні моделі. Відмітимо ще один дуже важливий момент, пов'язаний з адаптацією математичних моделей досліджуваних стаціонарних процесів до реальних даних вимірювань їх значень. Використання вимірювальних сигналів у цифровому вигляді, цифрового опрацювання під час статистичного оцінювання характеристик процесів набуває все більшого поширення в практиці вимірювань. Використання цифрових методів обумовлює появу додаткових чинників, наприклад шумів заокруглення, шумів квантування, мінливість частоти дискретизації і зміни рівнів квантування сигналів в часі і інші, які приводять до появи відповідних похибок вимірювання. З іншої сторони, застосування цифрових методів в ІВС створює принципово нові можливості автоматизації, адаптації, інтелектуалізації вимірювань.

На рис. 6.7 наведені графіки реалізацій стаціонарного процесу на кінцевому інтервалі часу T як вимірювальних сигналів у трьох видах: неперервного сигналу (а); сигналу з дискретним часом (б); цифрового сигналу з дискретним часом і квантованими за рівнем значеннями (в).

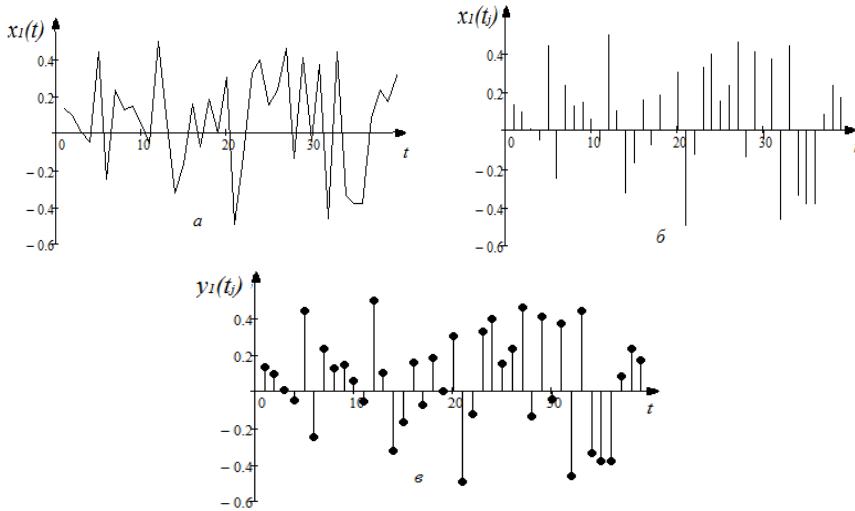


Рис. 6.7. Графіки досліджуваних вимірюваних сигналів:
a – неперервного $x(t)$; *б* – дискретизованого $x(t_j)$; *в* – цифрового $y_1(t_1)$

Випадковий процес $\xi(\omega, t)$ є функцією двох аргументів ω і t , перший з яких $\omega \in \Omega$ є неупорядкованим на просторі елементарних подій Ω . Така неупорядкованість і характеризує випадковість, стохастичність і непрогнозованість динаміки процесу $\xi(\omega, t)$ в часі. Аргумент t є впорядкованим у часі і для кожного фіксованого $t = t_1$ ми маємо перетин випадкового процесу $\xi(\omega, t)_{t=t_1}$ – випадкову величину $\xi_1(\omega)$.

Для вирішення завдань статистичного оцінювання характеристик випадкових процесів найцікавішим є дослідження динаміки і характерних особливостей простору Ω в часі, оскільки для цього використовується тільки ймовірність (ймовірнісна міра) якості їх оцінювання. Якщо простір Ω не залежить від часу, тобто залишається незмінним в часі, маємо стаціонарний, а у всіх інших – нестаціонарний випадки. Відомо також, що простір Ω змінюється у випадку різного роду перетворень вимірювальних сигналів в електронних трактах ІВС. Цей факт необхідно враховувати при статистичному оцінюванні характеристик випадкових процесів, а саме відмінність просторів елементарних подій Ω під час формування:

а) аналогових вимірювальних сигналів як множини реалізацій відповідних випадкових процесів для випадку $\Omega_1 \Rightarrow \{x_i(t), i \in N, t \in T\}$

, де Ω_1 є нескінченною безперервною множиною і відображає випадковість механізму формування реалізацій;

б) дискретизації за часом аналогових сигналів $\{x_i(t)\}$

$$\Omega_2 \Rightarrow \{x_i(t_j), i, j \in N, t_j \in T\},$$

де множина Ω_2 в загальному випадку не є еквівалентною множині Ω_1 .

Це пояснюється тим фактом, що дискретизація в часі проводиться не ідеальними перетворювачами, а реальними технічними пристроями, які вносять свою специфічну випадковість в досліджуваний сигнал $x_i(t_j)$ як відгук на виході технічного пристрою на вхідну дію $x_i(t)$.

Множина Ω_2 також є нескінченною і безперервною;

в) квантуванні за рівнем сигналів $\{x_i(t_j)\}$

$$\Omega_3 \Rightarrow \{y_i(t_j), i = \overline{1, d}, j \in N, t_j \in T\},$$

де простір Ω_3 , на відміну від Ω_1 і Ω_2 , є кінцевою дискретною числовою множиною.

Математичною моделлю безперервної реалізації $x(t)$ (рис. 6.7,а), яка належить до класу аналогових вимірювальних сигналів, є стаціонарний процес $\xi(\omega, t), \omega \in \Omega_1, t \in (-\infty, \infty)$. Безлічі значень області визначення і області значень процесу $\xi(\omega, t)$ нескінченні і мають потужність континууму. Процес $\xi(\omega, t)$ породжує безліч реалізацій $\{x_j(t), j \in N\}$, і кожна конкретна реалізація $x_j(t)$ відповідає конкретному елементарній події $\omega = \omega_1$ з нескінченного простору елементів-тарних подій Ω_1 .

Розглянемо реалізацію процесу $x_1(t_j)$ (рис. 6.7,б), яку отримують як результат дискретизації в часі аналогового сигналу $x_1(t)$. Математична модель такої реалізації – стаціонарна послідовність неперервних випадкових величин $\xi(\omega, t_1), \xi(\omega, t_2), \dots, \xi(\omega, t_j), \dots, \omega \in \Omega_2, j \in N$.

Множини значень областей визначення і областей значення стаціонарної послідовності неперервних випадкових величин $\{\xi(\omega, t_j)\}$ також нескінченні, проте мають різну потужність, а саме: область визначення – зліченна (дискретна) множина, а область значень – множина (неперервна) потужності континууму.

Після квантування за рівнем значень процесу $\xi(\omega, t_j)$ маємо випадковий

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

процес $\{\zeta(\omega, t_j), \omega \in \Omega_3, j \in N\}$ з дискретним часом і дискретними значеннями, який має кінцеву множину рівнів квантування, нескінченну зліченну множину елементів t_j і кінцеву дискретну множину значень $\zeta(\omega, t_j) \in Q$, де Q – кінцева дискретна множина чисел. Отже процес $\{\zeta(\omega, t_j), \omega \in \Omega_3, j \in N\}$ є стаціонарною послідовністю дискретних випадкових величин $\xi(\omega, t_1), \dots, \xi(\omega, t_j), \omega \in \Omega_2, j \in N$.

Процес $\zeta(\omega, t_j)$ можна сформулювати двома способами:

1) за заданого безперервного процесу $\xi(\omega, t)$ шляхом дискретизації області визначення аргументу t і одночасно квантування за рівнем його області значень;

2) за заданого процесу $\xi(\omega, t_j)$ з дискретним часом шляхом квантування за рівнем його області значень.

Графік відповідної реалізації $y_i(t_j)$ на кінцевому інтервалі часу T процесу $\zeta(\omega, t_j)$ наведено на рис. 6.7,в. Як правило, реалізації цього процесу як цифрові часові ряди задані на рівномірній часовій решітці з постійним кроком $\Delta t = h: 0, h, 2h, \dots, (n-1)h$ (6.12) де $[0, T = (n-1)h]$ – інтервал часу, який визначає область вимірювання цифрового часового ряду $y_i(t_j)$.

Таким чином, досліджуваний цифровий часовий ряд $y_i(t_j)$ можна позначати так: $y_i(j\Delta t) = y_i(jh) = y_i(t_j), i = \overline{1, m}, j = \overline{0, n-1}$, (6.13)

тобто як i -у реалізацію процесу $\zeta(\omega, t_j)$ на кінцевому інтервалі часу T .

На підставі аналізу наведених вище моделей стаціонарних випадкових процесів розглянемо математичну модель, реалізацією якої є дані вимірювань в цифровому вигляді як цифровий часовий ряд на кінцевому інтервалі часу T .

Випадкова вибірка обсягу n як скінченновимірна послідовність випадкових величин

$$\zeta(\omega, t_1), \zeta(\omega, t_2), \dots, \zeta(\omega, t_n) = \{\zeta(\omega, t_j), \omega \in \Omega_3, j = \overline{1, n}\} \quad (6.14)$$

із генеральної сукупності послідовності випадкових величин

$$\{\zeta(\omega, t_j), \omega \in \Omega_3, j \in N\} \quad (6.15)$$

є математичною моделлю цифрового часового ряду, заданого на рівномірній часовій решітці (6.12) у виді (6.13), тобто

$$y(jh) = y(t_j), j = \overline{0, n-1}. \quad (6.16)$$

Якщо проводиться m вимірювальних експериментів, то послідовність m реалізацій формує матрицю даних вимірювань в цифровому вигляді – матрицю цифрових часових рядів

$$\begin{pmatrix} y_1(0) & \cdots & y_1(t_{n-1}) \\ \vdots & \vdots & \vdots \\ y_m(0) & \cdots & y_m(t_{n-1}) \end{pmatrix}.$$

Досліджувані характеристики випадкових процесів визначаються відповідними *статистиками*

$$\mathbf{T}_k(\zeta(\omega, t_1), \dots, \zeta(\omega, t_n)), k=1, 2, \dots, g \quad (6.17)$$

як однозначно визначеними функціями кожної k -ї характеристики.

Статистична оцінка k -ї характеристики випадкової послідовності (6.15) визначається згідно $\mathbf{T}_k(y(0), \dots, y(t_{n-1}))$, де $y(0), \dots, y(t_{n-1})$ – часовий ряд даних вимірювань у вигляді (6.13) як реалізація (6.14).

Наведені характерні особливості статистичного оцінювання характеристик випадкових процесів за даними вимірів в рамках кореляційної теорії не відображають всього діапазону багатоплановості і різноманітності завдань вимірювання. Так, наприклад, не розглянуті важливі завдання вимірювань, коли результати вимірювань отримують у цифровій формі і виникають питання їх інтерполяції, апроксимації для первинних досліджуваних процесів, явищ, сигналів в аналоговому вигляді. Ця проблематика заслуговує окремого розгляду.

6.2. Статистичне оцінювання стаціонарності випадкової послідовності

Постановка завдання. Одним з перших завдань аналізу даних вимірювань і визначення характеристик випадкових процесів є перевірка даних на стаціонарність. Загальна постановка такого завдання полягає в перевірці статистичних гіпотез про стаціонарність досліджуваного випадкового процесу на підставі результатів статистичного опрацювання отриманих експериментально реалізацій процесу.

Результат статистичного оцінювання – приймається чи відхиляється гіпотеза про стаціонарність випадкового процесу – має принципове значення для подальшого статистичного опрацювання даних вимірювань. На сьогодні в значній мірі розроблена методологія

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

(методи, способи, інформаційне забезпечення) статистичного опрацювання реалізацій стаціонарних випадкових процесів. В процесі статистичного опрацювання реалізацій нестационарних випадкових процесів виникають труднощі порівняно зі стаціонарним випадком, для подолання яких необхідно проведення додаткового значного обсягу досліджень

Розглянемо завдання статистичного оцінювання стаціонарності послідовності дискретних випадкових величин (6.15) за даними вимірювань реалізацій випадкової вибірки послідовності (6.14) у вигляді двох цифрових часових рядів

$$\{y_1(ih) = y_1(t_i), i = \overline{0, n-1}\}, \{y_2(jh) = y_2(t_j), j = \overline{0, n-1}\}, \quad (6.18)$$

заданих на неперетинних інтервалах часу T_1 і T_2 , які найчастіше обирають рівними. Одним із статистичних методів вирішення такого роду завдань є метод перевірки статистичних гіпотез.

Сформулюємо дві (основну H_0 і альтернативну H_1) статистичні гіпотези: H_0 – послідовність (6.15) є стаціонарною; H_1 – послідовність (6.15) не стаціонарна. Статистична гіпотеза H_0 про стаціонарність послідовності (6.15) проти альтернативної гіпотези H_1 про нестационарність цієї послідовності перевіряється на основі реалізації випадкової вибірки (6.14) у вигляді двох цифрових часових рядів (6.18).

Вирішення завдання. Розглянемо два варіанти вирішення сформульованого завдання про стаціонарність досліджуваної послідовності:

- 1) випадкова послідовність (6.15) розподілена за законом Гаусса;
- 2) випадкова послідовність (6.15) розподілена за іншим законом.

1. Для перевірки стаціонарності послідовності (6.15) за умови нормального закону розподілу скористаємось методами перевірки статистичних гіпотез на основі F - і t - критеріїв, які іменуються відповідно статистичними критеріями t – Стьюдента і F – Фішера. Їх застосування дає змогу перевірити досліджувану послідовність за математичним сподіванням і дисперсією, що для гауссового розподілу послідовності (6.15) еквівалентно перевірці її на стаціонарність в строгому сенсі.

Гіпотеза про стаціонарність нормальної послідовності (6.15) на основі двох незалежних часових рядів (6.18) однакового обсягу n перевіряється з використанням F -критерію Фішера. За цим критерієм гіпотеза про рівність дисперсій, які оцінюються за двома реалізаціями $\{y_1(ih) = y_1(t_i), i = \overline{0, n-1}\}$ і $\{y_2(jh) = y_2(t_j), j = \overline{0, n-1}\}$ послідовності (6.15) на

неперетинних інтервалах часу T_1 і T_2 перевіряється за статистикою

$$W_{n-1,n-1} = \hat{\sigma}_1^2 / \hat{\sigma}_2^2, \quad (6.19)$$

де $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$ - незміщені оцінки дисперсій на вказаних інтервалах T_1 і T_2

$$\hat{\sigma}_1^2 = \frac{1}{n-1} \sum_{i=1}^n (y_1(t_i) - {}_1\hat{a}_1)^2, \quad \hat{\sigma}_2^2 = \frac{1}{n-1} \sum_{i=1}^n (y_2(t_i) - {}_2\hat{a}_1)^2, \quad (6.20)$$

а відповідні незсунені і ефективні оцінки математичного сподівання

$${}_1\hat{a}_1 = \frac{1}{n} \sum_{i=1}^n y_1(t_i), \quad {}_2\hat{a}_1 = \frac{1}{n} \sum_{i=1}^n y_2(t_i). \quad (6.21)$$

Статистика $W_{n-1,n-1}$ за справедливості гіпотези про сталість дисперсії в часі послідовності $\{\zeta(\omega, t_j)\}$ має F - розподіл зі степенями свободи $n-1$, $n-1$. Надалі припускаємо $\hat{\sigma}_1^2 > \hat{\sigma}_2^2$, що не обмежує спільності, оскільки у супротивному випадку треба перенумерувати реалізації, що спостерігаються, у зворотному порядку.

Критична область, яка відповідна розглянутому випадку перевірки гіпотези про стаціонарність відносно дисперсії, будується наступним чином. За таблицею для F - розподілу, які є в довідковій літературі з математичної статистики, для заданого рівня значущості α (зазвичай $\alpha \in \{0,1; 0,05; 0,01\}$), знаходимо табличне значення $F_{\alpha/2, n-1, n-1}$.

Якщо $W_{n-1,n-1} < F_{\alpha/2, n-1, n-1}$, то розбіжність між дисперсіями послідовності $\{\zeta(\omega, t_j)\}$ вважають несуттєвими або незначущим. Тоді приймається гіпотеза про стаціонарність досліджуваної послідовності (6.15). В іншому випадку, тобто коли $W_{n-1,n-1} > F_{\alpha/2, n-1, n-1}$, гіпотеза про стаціонарність відхиляється.

У розглянутому випадку перевірки на стаціонарність за допомогою F -критерію по дисперсії немає ніяких припущень щодо математичного сподівання послідовності (6.15). Якщо виникають сумніви щодо сталості математичного сподівання в часі, здійснюється ще одна додаткова перевірка з використанням t -критерію Стьюдента. Така перевірка, як правило, необхідна з метою підвищення достовірності про стаціонарність досліджуваної послідовності.

Двовибірковий t -критерій Стьюдента використовується для перевірки гіпотези про стаціонарність послідовності (6.15) за математичним сподіванням у випадку сталої дисперсії. Гіпотеза H_0 про рівність математичних сподівань двох незалежних випадкових вибірок (6.14) однакового обсягу n з генеральної сукупності (6.15) з гауссовим

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

розподілом перевіряється за t -статистикою $|t| = |{}_1\hat{a}_1 - {}_2\hat{a}_1| / \sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_2^2/n}$,
(6.22)

де ${}_1\hat{a}_1$ і ${}_2\hat{a}_1$ незміщені оцінки математичного сподівання двох цифрових часових рядів $\{y_1(ih) = y_{1i}, i = \overline{0, n-1}\}$ і $\{y_2(jh) = y_{2j}, j = \overline{0, n-1}\}$ послідовності (6.15) на неперетинних інтервалах часу T_1 і T_2 , обчислені за формулами (6.21), а $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$ - незміщені оцінки дисперсій за (6.20).

За рівних математичних сподівань процесу $\{\zeta(\omega, t_j)\}$ на неперетинних інтервалах T_1 і T_2 - справедливості статистичної гіпотези H_0 , статистика (6.18) має t -розподіл Стьюдента з $m = 2n - 2$ ступенями свободи. Критична область для перевірки статистичної гіпотези H_0 про стаціонарність за математичним сподіванням, будується наступним чином. Для заданого рівня значущості α за таблицями t -розподілу Стьюдента знаходимо табличне значення критичної області $t_{m, 1-\alpha/2}$. Якщо для заданого рівня значущості α виконується нерівність $|t| < t_{m, 1-\alpha/2}$, то розбіжність між оцінками математичного сподівання вважається незначною і має місце гіпотеза H_0 . Різниця ${}_1\hat{a}_1 - {}_2\hat{a}_1$ вважається значущою для заданого α , якщо $|t| > t_{m, 1-\alpha/2}$. У такому разі гіпотеза H_0 відхиляється, тобто вважається, що послідовність (6.15) стаціонарна.

Рівень значущості α виражає ймовірність прийняття неправильного рішення. На практиці, як правило, його подають у процентах і вибирають рівним 5%, тобто $\alpha=0,05$; в ряді випадків обирають $\alpha=0,01$ або $\alpha=0,005$, що відповідає 1% і 0,5%. Для підвищення достовірності прийняття рішення необхідно збільшувати обсяг вибірки.

2. Для статистичного оцінювання стаціонарності послідовничності дискретних випадкових величин (6.15) з відміним від гауссового розподілом за даними вимірювань в рамках кореляційної теорії можна скористатися описаними вище параметричними F - і t -критеріями. Для більш ефективної перевірки на стаціонарність необхідно провести таку перевірку різними попарними комбінаціями на більшій ($m > 2$) кількості неперетинних інтервалах часу $\{T_j, j = \overline{1, m}\}$.

В практиці вимірювань зустрічаються і інші підходи до статистичного оцінювання стаціонарності послідовності (6.15) за

даними вимірювань. Однак вони мають індивідуальний частинний характер і не реалізуються як типові методики для негауссових розподілів.

6.3. Статистичне оцінювання часових характеристик стаціонарної послідовності

Спочатку виконаємо порівняльний аналіз статистичного оцінювання характеристик випадкової величини і стаціонарної послідовності дискретних випадкових величин за даними вимірювань. Послідовність процедур такого порівняльного аналізу зображена на рис. 6.8.

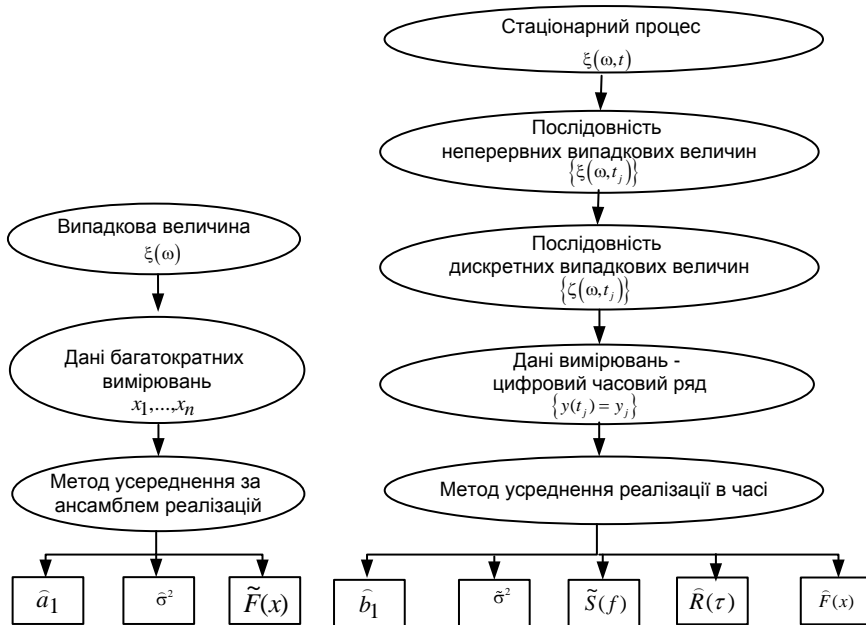


Рис. 6.8. Послідовність процедур статистичного оцінювання характеристик випадкової величини і стаціонарного процесу

Перейдемо до конкретних результатів порівняльного аналізу.

Випадкова величина. Математична модель багаторазових прямих вимірювань випадкової величини $\xi(\omega)$ виду

$$\xi_1(\omega), \xi_2(\omega), \dots, \xi_n(\omega), \omega \in \Omega \quad (6.23)$$

є випадковою вибіркою обсягу n однаково розподілених незалежних випадкових величин з генеральної сукупності $\{\xi_j(\omega), j \in Z\}$. Реалізація випадкової вибірки (6.23) обсягу n – це послідовність чисел x_1, x_2, \dots, x_n .

Формули для обчислень статистичних оцінок характеристик

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

випадкової величини відомі і мають вигляд:

$$\hat{a}_1 = \sum_{i=1}^n x_i / n, \quad \hat{\sigma}^2 = \sum_{i=1}^n (x_i - \hat{a}_1)^2 / (n-1),$$

де \hat{a}_1 , $\hat{\sigma}^2$ – незміщені статистичні оцінки відповідно математичного сподівання і дисперсії. Як основний методу визначення \hat{a}_1 , $\hat{\sigma}^2$ використано метод усереднення за ансамблем реалізацій.

Стаціонарна послідовність дискретних випадкових величин. Математична модель випадкової вибірки стаціонарної послідовності дискретних випадкових величин виду (6.14), тобто $\zeta(\omega, t_1), \zeta(\omega, t_2), \dots, \zeta(\omega, t_n)$ відрізняється від випадкової вибірки (6.23) тим, що це послідовність залежних випадкових величин.

Як зазначалося вище, реалізацією послідовності (6.14) є цифровий часовий ряд (6.13), тобто $y(jh) = y(t_j), j = \overline{0, n-1}$.

Сформулюємо більш чітко постановку завдання статистичного оцінювання часових характеристик стаціонарної послідовності дискретних випадкових величин (6.15), яка в загальному випадку є вкладеною випадковою послідовністю в стаціонарну послідовність неперервних випадкових величин $\{\xi(\omega, t_j)\}$. Послідовність $\{\xi(\omega, t_j)\}$ є дійсною, для неї виконується ергодична гіпотеза відносно перших чотирьох моментів, існування яких також передбачається.

Потрібно за даними вимірів досліджуваного процесу – цифрового часового ряду (6.13) обчислити статистичні оцінки математичного сподівання, дисперсії, автокореляційної функції і одномірної емпіричної функції розподілу. Однією з умов успішного вирішення такого завдання є значний обсяг n цифрового часового ряду.

Відомі варіанти вирішення такого завдання можна умовно розділити на дві групи:

- 1) часові ряди задаються без урахування похибок їх вимірювань (ряди в широкому сенсі);
- 2) часові ряди задаються з урахуванням похибок їх вимірювань (ряди у вузькому сенсі).

Характерним моментом статистичного оцінювання одновимірних часових характеристик стаціонарної послідовності дискретних випадкових величин $\{\xi(\omega, t_j)\}$ – середнього (математичного сподівання) a_1 і дисперсії σ^2 – є той факт, що ці ймовірнісні в теоретичному сенсі характеристики є постійними величинами і не залежать від часу.

Одновимірні функції розподілу $F(x)$, щільність розподілу $p(x)$ і характеристична функція $f(u)$ послідовності $\{\zeta(\omega, t_j)\}$ також не залежать від часу t . Двовимірні характеристики стаціонарного процесу – автоковаріаційна $B(\tau)$, автокореляційна $R(\tau)$, двовимірна функція (щільність) розподілу $F(x_1, x_2; \tau)$ – залежать лише від різниці часу $\tau = t_2 - t_1$.

Така ідеалізація моделі стаціонарного процесу не має місця у практиці статистичного оцінювання характеристик процесу за даними вимірювань. Вимірювання проводяться на кінцевому інтервалі часу, на результати вимірювання діють як контрольовані, так і неконтрольовані, як правило, випадкові чинники. В результаті проведення статистичного оцінювання теоретичні числові постійні характеристики або функції детермінованих аргументів стаціонарного процесу стають функціями додаткових аргументів, наприклад обсягу вибірки, інтервалу часового вікна, похибок вимірювань, заокруглень результатів обчислень числових даних і інших. По суті мають місце невідповідності (неадекватності) моделі реальному досліджуваному процесу. І компроміс адекватності можна досягнути на рівні можливих похибок статистичного оцінювання, які не перевищують заданих. У більшості випадків дисперсія оцінки шуканої характеристики служить показником точності її статистичного оцінювання. Якщо за умовами завдання статистичного оцінювання вдається визначити довірчий інтервал із заданою ймовірністю оцінки шуканої характеристики, можна говорити про досягнення мети – вирішення завдання статистичного оцінювання в більш повному обсязі.

У більш строгій постановці завдання відповідності моделі результатам вимірювань використовується апарат математичної статистики, на-приклад, перевірка адекватності моделі та отриманих експериментальних даних на основі перевірки відповідних статистичних гіпотез на підставі обґрунтованого статистичного критерію

Статистичне оцінювання часових характеристик. Розглянемо основні алгоритми статистичного оцінювання зазначених характеристик стаціонарної послідовності дискретних випадкових величин.

Статистичне оцінювання математичного очікування. Статистична оцінка математичного очікування (середнього) стаціонарної послідовності дискретних випадкових величин, для якої виконується ергодична гіпотеза відносно математичного сподівання, обчислюється усередненням в часі однієї реалізації стаціонарної послідовності –

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

цифрового часового ряду виду (6.13) за формулою $\hat{b}_1 = \frac{1}{n} \sum_{j=1}^n y(t_j)$.

(6.23)

Оцінка \hat{b}_1 є незміщеною і ефективною, оскільки $\mathbf{M}\hat{b}_1 = a_1$.

Для конкретного цифрового часового ряду (6.13) оцінка \hat{b}_1 є середнім значенням і статистичною оцінкою математичного сподівання стаціонарної послідовності дискретних випадкових величин (6.15).

Якщо відома кореляційна функція стаціонарної послідовності (6.19) у вигляді

(6.24)

$$R(k) = \mathbf{M}\left\{\left[\zeta(\omega, t_j) - a_1\right]\left[\zeta(\omega, t_{j+k}) - a_1\right]\right\},$$

то дисперсія оцінки математичного сподівання (6.23) визначається як

$$\mathbf{D}\hat{b}_1 = \mathbf{M}\left[\hat{b}_1 - a_1\right]^2 = \frac{1}{n}\left[R(0) + 2\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)R(k)\right], \quad (6.25)$$

де кількість значень функції $R(k)$ задана на кінцевій множині точок, яка визначається обсягом вибірки n цифрового часового ряду.

У виразі (6.25) значення a_1 не входить в явному вигляді для обчислення дисперсії, але його знання необхідно при обчисленні за формулою (6.24).

У формулу (6.25) значення a_1 не входить в явному виді, але його знання необхідно для обчислення $R(k)$ за формулою (6.24).

На підставі аналізу завдання статистичного оцінювання характеристик стаціонарних процесів відносно статистичного оцінювання математичного сподівання a_1 стаціонарної послідовності дискретних випадкових величин $\{\zeta(\omega, t_j)\}$ можна зробити такі висновки.

У більшості випадків для $\{\zeta(\omega, t_j)\}$ невідомими є як a_1 , так і кореляційна функція $R(k)$. За результатами статистичного оцінювання визначають їх оцінки – \hat{b}_1 і $\hat{R}(k)$. У загальній постановці завдання статистичного оцінювання a_1 і $R(k)$ стаціонарних процесів не має загального розв'язку і це, ймовірно, є природним через різноплановість і різноманіття завдань статистичного оцінювання. У той же час на базі використання додаткових умов, наприклад для $\{\zeta(\omega, t_j)\}$ заданий гауссовий закон розподілу або відома кореляційна характеристика $R(k)$ чи a_1 , завдання статистичного оцінювання вирішується в необхідному для практичного використання вигляді.

За своєю суттю завдання статистичного оцінювання, а це в повній мірі відноситься до завдання вимірювання, ставиться, щоби визначити ту чи іншу характеристику випадкового процесу, а для цього необхідне знання

характеристик того ж процесу на порядок вище. **Наприклад**, для оцінювання якості визначення математичного сподівання для обчислення дисперсії $\widehat{D}b_1$ необхідне знання кореляційної функції $R(k)$ (див. (6.25)).

Гіпотеза ергодичності для $\{\zeta(\omega, t_j)\}$, в даному випадку ергодичної гіпотези відносно математичного сподівання $\{\zeta(\omega, t_j)\}$ має важливе значення. В цьому випадку повинна виконуватися умова (6.6), тобто $\lim_{n \rightarrow \infty} 2n^{-1} \sum_{j=1}^n R(j)(1 - j/n) = 0$, яку для практичного застосування можна замінити іншою умовою. На основі нерівності $\sum_{j=1}^n R(j)(1 - j/n) \leq \sum_{j=1}^n R(j)$ на практиці використовується достатня і необхідна умова ергодичності $\{\zeta(\omega, t_j)\}$ відносно математичного сподівання $\lim_{n \rightarrow \infty} \sum_{j=1}^n R(j) < \infty$. Цій умові задовольняє значна частина досліджуваних стаціонарних процесів.

Статистичне оцінювання кореляційної функції. Статистична оцінка кореляційної функції часового ряду $y_i(t_j)$ (6.13) визначаються як

$$\tilde{R}(k) = \frac{1}{n-k} \sum_{j=1}^{n-k} (y(t_j) - \tilde{b}_1)(y(t_{j+k}) - \tilde{b}_1). \quad (6.26)$$

Ця оцінка асимптотично зміщена і в разі заміни випадкової величини \tilde{b}_1 на a_1 отримана за формулою (6.26) статистична оцінка кореляційної функції $\tilde{R}(k)$ стає асимптотично незміщеною, тобто $\lim_{n \rightarrow \infty} M\tilde{R}(k) = R(k)$.

Відзначимо наступний факт, який має важливе практичне значення. У випадку обчислення статистичних оцінок кореляційної функції стаціонарної послідовності дискретних випадкових величин (6.15), для якої виконується ергодична гіпотеза щодо кореляційної функції за даними вимірів – цифрового часового ряду (6.13), використовується не класична форма визначення кореляційної функції у вигляді (6.1), а так зване *кореляційне перетворення* – дія відповідного кореляційного оператора у вигляді (6.26).

Таким чином, оцінка $\tilde{R}(k)$ є відгуком кореляційного оператора, який здійснює усереднення за часом добутків вихідного центрованного часового ряду $\{y(t_j) - \tilde{b}_1\}$ на зсунутий в часі центрований часовий ряд

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

$\{y(t_{j+k}) - \tilde{b}_1\}$. Якщо $k = 0$, обчислюється статистична оцінка дисперсії $\tilde{R}(0) = \tilde{\sigma}^2$. Дисперсія статистичної оцінки $\tilde{R}(k)$ визначається за формулою $D\tilde{R}(k) = \left[R(k,0) + 2 \sum_{j=1}^{n-k-1} (1-j/(n-k)) R(k,s) \right] / (n-k-1)$, отриманою аналогічно до (6.8),..., (6.10): формується випадкова послідовність другого порядку відносно досліджуваної стаціонарної послідовності дискретних випадкових величин $\{\zeta(\omega, t_j)\}$ виду $\zeta(\omega, \tau, t) = [\zeta(\omega, t_j) - a_1][\zeta(\omega, t_{j+\tau}) - a_1]$, потім визначається її математичне сподівання $M\zeta(\omega, \tau, t) = R(\tau)$ і відповідно її автокореляційна функція як центральний момент четвертого порядку послідовності $\{\zeta(\omega, t_j)\}$:

$$R(s, \tau) = M \left[\zeta(\omega, \tau, t) - R(\tau) \right] \left[\zeta(\omega, \tau, t + s) - R(\tau) \right].$$

Отже, дисперсія статистичної оцінки $\tilde{R}(k)$ стаціонарної послідовності дискретних випадкових величин $\{\zeta(\omega, t_j)\}$, що задовольняє ергодичній гіпотезі відносно кореляційної функції визначається, якщо задана $R(s, \tau)$ – четвертий центральний момент послідовності $\{\zeta(\omega, t_j)\}$.

Статистичне оцінювання одновимірної функції розподілу. Це одне з найпоширеніших завдань статистичного оцінювання для стаціонарної послідовності дискретних випадкових величин $\{\zeta(\omega, t_j)\}$. Емпіричною або вибірковою функцією розподілу випадкової вибірки (6.14) обсягу n називається функція $F_n(x) = g_x/n$, де g_x – число тих $\zeta(\omega, t_j)$, значення яких менше x .

Ця функція є випадковою, а кожна її реалізація – кусково-постійна функція x , що змінюється від 0 до 1 тільки стрибками, число яких дорівнює n . Реалізація функції $F_n(x)$ іноді визначається безпосередньо, та в більшості практичних випадків ґрунтується на використанні гістограм як реалізацій щільності розподілу.

Результати вирішення цього завдання аналогічні результатам статистичного оцінювання функції розподілу випадкової величини, які були наведені в п'ятому розділі. Тільки інтерпретація отриманих результатів статистичного оцінювання зовсім інша, а саме:

а) статистична оцінка одновимірної функції розподілу $\tilde{F}(x)$,

обчислена з часового ряду (6.13), характеризує стаціонарну послідовність дискретних випадкових величин (6.15) і не залежить від часу t ;

б) для виконання необхідного і достатнього умови ергодичності одновимірної функції розподілу $F(x)$ послідовності $\{\zeta(\omega, t_j)\}$, необхідно виконання умови (6.11), тобто

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n [F(x_1, x_2; \tau) - F(x_1)F(x_2)] \cdot (1 - j/n) = 0, \forall \tau, \quad \text{де } F(x_1, x_2; \tau) -$$

двовимірна функція розподілу послідовності у вигляді (6.15).

У більшості випадків перевірка виконання умови (6.11) викликає на практиці певні труднощі. Тому, якщо виконані умови ергодичності щодо математичного сподівання і кореляційної функції для досліджуваної стаціонарної послідовності, то можна припустити можливість застосування ергодичної гіпотези і для одновимірної функції розподілу $F(x)$.

Статистичне оцінювання двовимірної функції розподілу $F(x_1, x_2; \tau)$. Для статистичного оцінювання такої функції як приклад розглянемо дещо іншу постановку задачі у порівнянні з розглянутими раніше. Двовимірна функція розподілу $F(x_1, x_2; \tau)$ задається аналітично з невідомими параметрами і функціями.

Приклад. Задана двовимірна щільність розподілу стаціонарної послідовності гауссових дискретних випадкових величин $\{\zeta(\omega, t_j)\}$

$$p(x_1, x_2; \tau) = \left(1/2\pi\sigma^2\sqrt{1-\rho^2(\tau)}\right) \exp\left\{-\left(1/2(1-\rho^2(\tau))\left[\frac{(x_1-a_1)^2 - 2\rho(\tau)(x_1-a_1)(x_2-a_1) + (x_2-a_1)^2}{\sigma^2}\right]\right)\right\},$$

де математичне сподівання a_1 , дисперсія σ^2 і нормована кореляційна функція $\rho(\tau)$ не задані. Необхідно визначити статистичні оцінки a_1 , σ^2 , $\rho(\tau)$ і двовимірну функцію розподілу $F(x_1, x_2; \tau)$ за даними вимірювань – цифровому часовому ряду (6.13), тобто $\{y(jh) = y(t_j), j = \overline{0, n-1}\}$.

Передбачається, що досліджувана послідовність $\{\zeta(\omega, t_j)\}$ має кінцеві перші два моменти і має місце ергодична гіпотеза для математичного сподівання, кореляційної функції і одновимірної функції розподілу процесу.

Використовуючи раніше наведені формули, обчислюємо оцінки

$$\tilde{a}_1 = \frac{1}{n} \sum_{j=1}^n y(t_j), \quad \tilde{\sigma}^2 = \frac{1}{n-1} \sum_{j=1}^n [y(t_j) - \tilde{a}_1]^2,$$

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

$$\tilde{\rho}(\tau) = \frac{1}{n-\tau} \sum_{j=1}^{n-\tau} \frac{[y(t_j) - \tilde{a}_1][y(t_{j+\tau}) - \tilde{a}_1]}{\tilde{\sigma}^2}, \quad \tau = 0, 1, 2, \dots, k, \quad k \leq 0, 1, n,$$

і визначаємо шукану статистичну оцінку двовимірної функції розподілу у вигляді

$$\tilde{F}(x_1, x_2; \tau) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{1}{2\pi\tilde{\sigma}^2 \sqrt{1-\tilde{\rho}^2(\tau)}} \times \\ \times \exp \left\{ -\frac{1}{2(1-\tilde{\rho}^2(\tau))} \cdot \left[\frac{(y_1 - \tilde{a}_1)^2 - 2\tilde{\rho}(\tau)(y_1 - \tilde{a}_1)(y_2 - \tilde{a}_1) + (y_2 - \tilde{a}_1)^2}{\tilde{\sigma}^2} \right] \right\} dy_1 dy_2.$$

В даному розділі наведені типові завдання статистичного оцінювання часових характеристик стаціонарної послідовності дискретних випадкових величин за даними вимірів у цифровому вигляді. Результати розв'язання конкретних завдань і більш детальних досліджень статистичного оцінювання стосуються різних комбінацій детермінованих і випадкових сигналів за різних відношень сигнал/шум. Більшість таких завдань полягає у виявленні сигналів і оцінюванні їх параметрів і характеристик в умовах дії шумів, статистичному оцінюванні трансформації сигналів в трактах електронних систем і ін.

6.4. Статистичне оцінювання спектральних характеристик стаціонарних послідовностей

Серед прикладних методів статистичного оцінювання спектральної щільності потужності стаціонарних процесів $S(f)$ виділяють наступні:

- 1) на основі перетворення Фур'є реалізацій стаціонарного процесу;
- 2) основі перетворення Фур'є статистичної оцінки кореляційної функції стаціонарного процесу (теорема Вінера – Хінчина);
- 3) на основі використання вузькосмугової лінійної фільтрації реалізацій стаціонарного процесу, зведення в квадрат відгуків з подальшим їх усередненням за ансамблем.

Розглянемо перший метод статистичного оцінювання $S(f)$ за даними вимірів реалізацій стаціонарної послідовності дискретних випадкових величин – цифрового часового ряду (6.13).

Обґрунтуванням вибору даного методу є наступні фактори:

а) дані вимірювань у вигляді часового ряду без додаткових перетворень і, відповідно, внесення додаткових похибок вимірювання безпосередньо використовуються для статистичного оцінювання $S(f)$;

б) інформаційне забезпечення у вигляді алгоритмів, програмного забезпечення, практичних рекомендацій вибору частоти дискретизації, числа рівнів квантування, обсягу n відліків, часових і спектральних

вікон статистичного оцінювання $S(f)$ даного напрямку найбільш розроблено і впроваджено в практику вимірювань сучасних ІВС;

в) безпосереднє використання даних вимірювань для статистичного оцінювання характеристик досліджуваних процесів мотивує застосування єдиної методології статистичного опрацювання інформації в ІВС з подальшим оцінюванням результатів і похибок вимірювань.

Слід зазначити, що раніше – в 50 - 70-ті роки ХХ ст., для статистичного оцінювання в основному використовувався метод оцінювання кореляційної функції на основі теореми Вінера – Хінчина (6.2). Спочатку обчислювалася оцінка автокореляційної функції досліджуваного процесу за даними вимірювань (6.1), потім використовувалося перетворення Фур'є обчисленої оцінки за алгоритмом Блекмана-Тьюкі.

У подальшому було розроблено інформаційне забезпечення цифрового опрацювання сигналів, що дало змогу поліпшити якість статистичного оцінювання спектральної щільності у випадку дискретно перетворення Фур'є реалізацій досліджуваних стаціонарних процесів і використанні відповідних часових і спектральних вікон.

Зупинимося на цьому більш детально. Розглянемо статистичне оцінювання таких спектральних характеристик для двох стаціонарних і стаціонарно зв'язаних послідовностей дискретних випадкових процесів $\{\zeta_1(\omega, t_j)\}$ і $\{\zeta_2(\omega, t_j)\}$:

спектральних щільностей потужності $S_{11}(f)$ и $S_{22}(f)$;

взаємної спектральної щільностей потужності $S_{12}(f)$ (6.3);

функції когерентності $v^2(f)$ як нормованої спільної функції четвертого порядку статистичного взаємозв'язку центрованих послідовностей $\{\zeta_1(\omega, t_j)\}$ і $\{\zeta_2(\omega, t_j)\}$ в частотній області, яка

визначається виразом (6.4), тобто

$$v^2(f) = |S_{12}(f)|^2 / S_{11}(f) \cdot S_{22}(f), \quad v^2(f) \in [0, 1].$$

Раніше зазначалося, що дані вимірів формуються у вигляді цифрових сигналів, тобто сигналів з дискретним часом і квантовими за рівнем значеннями. Отже цифрові вимірювальні сигнали у вигляді часових рядів мають кінцевий обсяг n елементів. Частота дискретизації безперервних вимірювальних сигналів і число рівнів квантування для кожного конкретного завдання вибирається на підставі апріорних даних сигналів.

Таким чином, на вхід підсистеми статистичного оцінювання спектральних характеристик стаціонарних послідовностей ІВС надходить два цифрових часових ряда як реалізації досліджуваних послідовностей $\{\zeta_1(\omega, t_j)\}$ і

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

$\{\xi_2(\omega, t_j)\}$, кожна з яких має вигляд (6.13), тобто

$$\left\{y_1(ih) = y_1(t_i), i = \overline{0, n-1}\right\} \text{ и } \left\{y_2(jh) = y_2(t_j), j = \overline{0, n-1}\right\}, \quad (6.27)$$

де $h = \Delta t$ – постійний крок рівномірної часової решітки завдання відліків.

Надалі розглядатимемо обидва часових ряди (6.27) з однаковими об'ємом n відліків і часом дискретизації Δt , тобто (6.27) можна записати у вигляді $\left\{y_m(ih) = y_m(t_i), m = \overline{1, 2}, i = \overline{0, n-1}\right\}$.

Розглянемо послідовність обчислень спектральних характеристик методами цифрового опрацювання часових рядів.

1. На початковому першому етапі попереднього статистичного опрацювання вихідних цифрових часових рядів виконують:

- а) видалення промахів, помилкових значень відліків ряду;
- б) цифрову фільтрацію;
- в) видалення тренду;
- г) центрування обох часових рядів:

якщо відомо середнє (математичне сподівання) досліджуваної послідовності $\left\{m a_1, m = \overline{1, 2}\right\}$, то

$$\bar{y}_m(ih) = y_m(ih) - m a_1 = \bar{y}_m(t_i), m = \overline{1, 2}, i = \overline{0, n-1}; \quad (6.28)$$

якщо $\left\{m a_1, m = \overline{1, 2}\right\}$ невідоме, то для операції центрування використовується її статистична оцінка $\left\{m \tilde{b}_1, m = \overline{1, 2}\right\}$, тоді

$$\tilde{y}_m(ih) = \tilde{y}_m(t_i) = y_m(ih) - m \tilde{b}_1, m = \overline{1, 2}, i = \overline{0, n-1}. \quad (6.29)$$

Кожен з часових рядів (6.29), що використовується для цифрового статистичного опрацювання, є центрований цифровий часовий ряд з n відліками як реалізацією випадкової вибірки центрованої стаціонарної послідовності дискретних випадкових величин $\left\{\zeta(\omega, t_i)\right\}$ (6.14).

Часовий ряд (6.29) можна розглядати і в інший спосіб: якщо припустити, що зареєстрована безперервна реалізація $x(t)$ безперервного стаціонарного процесу $\xi(\omega, t)$, то часовий ряд (6.29) можна сформуванати шляхом проведення наступних операцій:

дискретизацією реалізації $y(t)$ в часі з постійним кроком $\Delta t = h$ на рівномірній часовій решітці, тобто $x(t) \Rightarrow x(ih) = x(t_i), i = \overline{0, n-1}$;

квантуванням з числом Q рівнів значень $x(t_i)$ і формуванням цифрового сигналу $x(t_i) \Rightarrow y(ih) = y(t_i), y(t_i) \in [-Q, Q], i = \overline{0, n-1}$.

В ІВС формування цифрових сигналів виду (6.29) виконується в АЦП, які мають свої метрологічні характеристики (див. розд. 4).

2. На другому етапі попереднього статистичного опрацювання рекомендується з метою поліпшення якості (як правило, зменшення дисперсії статистичних оцінок) статистичного оцінювання спектральних характеристик

стаціонарної послідовності виконувати:

а) розбиття вихідного часового ряду (6.29) на k відрізків однаково обсягу N , створення ансамблю з K реалізацій і використання в подальшому метод усереднення за ансамблем статистичних оцінок досліджуваних спектральних характеристик;

б) множення вихідного часового ряду (6.29) на часове вагове вікно $W(t_i)$, вибір виду якого ґрунтується на постановці завдання статистичного оцінювання спектральних характеристик і апіорних даних про досліджуваний випадковий процес, а також спектральній характеристикі вагового вікна $W(t_i)$

Наведені рекомендації є типовими для статистичного оцінювання спектральних характеристик стаціонарних процесів, а їх зміст обґрунтовано результатами значного практичного досвіду застосування як аналогових, так і цифрових статистичних методів опрацювання.

Представлені для подальшого цифрового статистичного опрацювання ряди (6.29), кожен з яких містить n відліків, розбиваються на $k = \overline{1, K}$ відрізків з N відліків, тобто $N = n/k$. Значення K доцільно вибирати якомога більше, оскільки дисперсію відповідної спектральної характеристики можна потенційно зменшити в $1/k$ разів на підставі застосування статистичного методу усереднення за ансамблем реалізацій.

Таким чином, кожен з центрованих цифрових часових рядів $\tilde{y}_m(ih)$, $m = \overline{1, 2}$, $i = \overline{0, n-1}$ (6.29) розбивається на K ділянок (рис. 6.9).

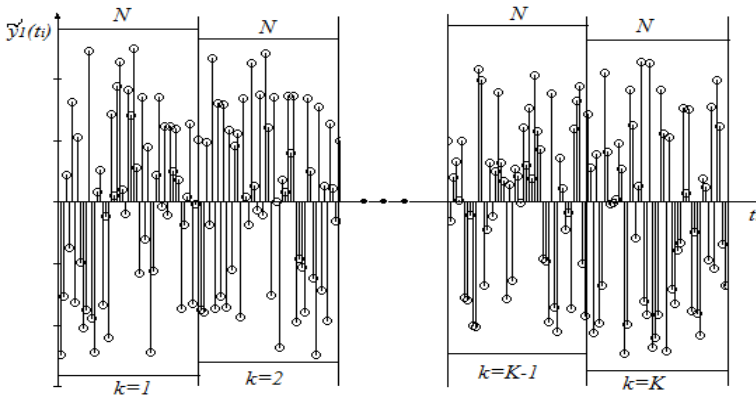


Рис. 6.9. Графічна ілюстрація розбиття досліджуваного центрованого цифрового часового ряду з n відліками на K ділянок з N відліками на кожній

Приклади деяких часових вагових вікон, що використовуються для підвищення якості статистичного оцінювання спектральних характеристик стаціонарної послідовності $\{\zeta(\omega, t_j)\}$, наведені в табл.6.1. У таблиці наведені і частотні (спектральні) характеристикі вагових вікон $W(f)$. Отже для

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

подальшого статистичного опрацювання для кожного часового ряду (6.29) сформовано ансамбль K реалізацій виду

$$Zm(t_i) = \tilde{y}_m(t_i)w(t_i) = \begin{cases} Z_{m1}(t_j) \\ \vdots \\ Z_{mk}(t_j), \quad m = \overline{1,2}, j = \overline{0, N-1}, k = \overline{1, K}, i = j + (k-1)N \\ \vdots \\ Z_{mK}(t_j) \end{cases}$$

3. Для кожної реалізації з $\{Z_{mk}(t_j), m = \overline{1,2}, k = \overline{1, K}, j = \overline{0, N-1}\}$ виконуються наступні операції:

а) дискретне перетворення Фур'є (ДПФ), на практиці в більшості випадків використовується економний з огляду на число обчислювальних операцій спосіб реалізації ДПФ – швидке перетворення Фур'є

$$F_{mk}(f_g) = \Delta t \sum_{i=0}^{N-1} Z_{mk}(t_i) \cdot \exp(-j2\pi \cdot i \cdot g/N), \text{ де } j = \sqrt{-1}; \quad (6.30)$$

б) модуль виразу (6.30) підноситься в квадрат: $|F_{mk}(f_g)|^2$;

в) використовується метод усереднення за ансамблем K компонент (6.30) і визначається статистична оцінка двосторонньої спектральної щільності стаціонарної послідовності дискретних випадкових величин (6.14), тобто $\{\zeta(\omega, t_j)\}$, виразом $S_{mm}(f_g) = \frac{1}{KN\Delta t} \sum_{k=1}^K |F_{mk}(f_g)|^2, g = \overline{0, N-1}$.

Таблиця 6.1

Назва вікна	Дискретно-часова функція $w(n), n = 0 \dots N-1$	Частотна характеристика $w(f)$ $f_k = 0, 1, \dots, N/2,$ $D_N(f) = T \exp(-j2\pi f TN) \frac{\sin[N\pi f T]}{\sin[\pi f T]}$
Прямокутне	$w(n) = 1$	$D_N(f)$
Трикутне (Бартлетта)	$w(n) = 1 - n/A - 1 ,$ $A = (N-1)/2$	$\frac{2}{N} \cdot D_N^2\left(\frac{f}{2}\right)$
Ханна	$w(n) = 0.5 - 0.5 \times$ $\times \cos(2\pi n/(N-1))$	$0.5D_N(f) + 0.25 \times (D_N(f-1/NT) + D_N(f+1/NT))$
Хеммінга	$w(n) = 0.54 - 0.46 \times$ $\times \cos(2\pi n/(N-1))$	$0.54D_N(f) + 0.23 \times (D_N(f-1/NT) + D_N(f+1/NT))$

Для розрахунку оцінки спектральної щільності $S_{mm}(f_g)$ використовуються дискретні значення частот $f_g = g/N\Delta t, g = \overline{0, N-1}$.

Підкреслимо, що оцінка $S_{mm}(f_g)$ визначена в діапазоні як негативних, так і позитивним частот. Перші $N/2+1$ значень оцінки спектральної щільності для $j=0,1,\dots,N/2$ задають оцінку $S_{mm}(f_g)$ в діапазоні від'ємних частот f_j від 0 до $N/2$, а інші $N/2-1$ для $j=N/2+1,N/2+2,\dots,N-1$ – в діапазоні негативних (симетрично 0) частот $-N/2 < f_j < 0$. Статистична оцінка односторонньої спектральної щільності $S_{mm}(f_g)$ визначається як $G_{mm}(f_g) = \frac{2}{KN\Delta t} \sum_{k=1}^K |F_{mk}(f_g)|^2, g = \overline{0, N/2}$.

4. На четвертому етапі статистичного оцінювання спектральних характеристик стаціонарної послідовності дискретних випадкових величин $\{\zeta(\omega, t_j)\}$ (6.14), після виконання операцій попередніх трьох етапів цифрового статистичного опрацювання первинних цифрових часових рядів виду (6.27), визначають наступні статистичні оцінки.

1). Односторонні спектральні щільності:

а) стаціонарної послідовності $\{\zeta_1(\omega, t_j)\}$ у вигляді

$$G_{11}(f_g) = \frac{2}{KN\Delta t} \sum_{k=1}^K |F_{1k}(f_g)|^2, g = \overline{0, N/2};$$

б) стаціонарної послідовності $\{\zeta_2(\omega, t_j)\}$ у вигляді

$$G_{22}(f_g) = \frac{2}{KN\Delta t} \sum_{k=1}^K |F_{2k}(f_g)|^2, g = \overline{0, N/2}.$$

2). Односторонню взаємну спектральну щільність стаціонарних послідовностей $\{\zeta_1(\omega, t_j)\}$ і $\{\zeta_2(\omega, t_j)\}$ у вигляді

$$G_{12}(f_g) = \frac{2}{KN\Delta t} \sum_{k=1}^K |F_{1k}(f_g) \cdot F_{2k}(f_g)|^2, g = \overline{0, N/2}.$$

3). Функції когерентності

$$v^2(f_g) = |G_{12}(f_g)|^2 / [G_{11}(f_g) \cdot G_{22}(f_g)], g = \overline{0, N/2}.$$

Слід зазначити наступне. В аналізі результатів статистичного оцінювання наведених вище спектральних характеристик стаціонарної послідовності виникають питання чисто вимірювального (метрологічного) плану. Так, не наведено вирази для оцінки якості отриманих статистичних оцінок спектральних характеристик, наприклад вирази для дисперсії оцінок. Такі результати можуть бути отримані тільки для конкретних завдань статистичного оцінювання спектральних характеристик з використанням додаткових апріорних відомостей про

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

моделі вимірюваних сигналів, частотний діапазон спектральних характеристик та ін. Тому в наведену вище типову методику статистичного оцінювання спектральних характеристик стаціонарних послідовностей результати таких досліджень не були включені.

5. На заключному етапі статистичного оцінювання спектральних характеристик стаціонарних процесів проводиться аналіз отриманих результатів; в ряді випадків з урахуванням використання конкретних вагових вікон, вводяться масштабні коефіцієнти. За результатами проведеного аналізу приймається рішення про досягнення або недосягнення поставленої мети статистичного оцінювання спектральної щільності потужності стаціонарної послідовності.

6. Результати статистичного оцінювання спектральної щільності потужності стаціонарних процесів також використовуються і в ряді інших завдань, наприклад для статистичного оцінювання аналогічних характеристик як функцій часу нестаціонарного випадкового процесу – періодично корельованого – а також для дослідження адитивної суміші двох стаціонарних незалежних випадкових процесів.

Основні висновки

Статистичне оцінювання характеристик випадкових процесів – це визначення їх статистичних оцінок за даними вимірів значень процесів.

Кореляційний теорія використовується для дослідження і статистичного оцінювання перших двох моментних функцій випадкових процесів.

У теорії і практиці статистичного оцінювання найбільш повно розроблені моделі, алгоритми і програмне забезпечення обчислення статистичних оцінок характеристик стаціонарних випадкових процесів. Характеристиками стаціонарних процесів є: математичне сподівання, дисперсія, кореляційна функція, спектральна щільність потужності і одновимірна функція розподілу.

До інформаційного забезпечення статистичного оцінювання стаціонарних процесів відносять математичні моделі, їх основні статистичні характеристики, алгоритми і програмне забезпечення розрахунків статистичних оцінок характеристик стаціонарних процесів.

Статистичні оцінки математичного сподівання, дисперсії, кореляційної функції, спектральної щільності потужності стаціонарних процесів обчислюються за даними вимірів реалізацій стаціонарних процесів.

Застосування ергодичної теореми Біркхофа - Хінчіна дає змогу використовувати усереднення за часом реалізацій досліджуваних процесів для статистичного оцінювання характеристик стаціонарних процесів.

Достатні і необхідні умови ергодичності стаціонарного процесу відносно математичного сподівання базуються на використанні відомої кореляційної функції процесу.

Достатні і необхідні умови ергодичності стаціонарного процесу відносно кореляційної функції базуються на використанні центрального четвертого моменту досліджуваного стаціонарного процесу.

До характерних особливостей завдань статистичного оцінювання як процесу вимірювання відносять специфіку формування експериментальних даних про процеси і подальшу їх цифрове статистичне опрацювання.

Експериментальні дані умовно розділені на дані вимірювань і інші експериментальні дані, що дає можливість виконувати статистичне оцінювання характеристик стаціонарних процесів у вузькому і широкому сенсах.

У цифровому статистичному оцінюванні характеристик стаціонарних процесів використовують цифрові часові ряди як реалізації досліджуваного процесу, отримані шляхом дискретизації у часі і квантування за рівнем вихідної (безперервної) реалізації.

Серед основних методів статистичного оцінювання виділено два метода, один з яких дає можливість здійснювати оцінювання шляхом усереднення за безліччю (ансамблю) реалізацій досліджуваного процесу; інший – усередненням однієї реалізації процесу в часі. Застосування останнього методу дозволяє значно спростити вирішення завдань обчислення статистичних оцінок характеристик стаціонарних процесів на практиці.

Математична модель, реалізацією якої є цифровий часовий ряд як дані вимірювання, представляється як випадкова вибірка певного обсягу стаціонарної послідовності дискретних випадкових величин.

Статистика у статистичному оцінюванні k -ї характеристики стаціонарної послідовності визначається як відповідна функція випадкової вибірки стаціонарної послідовності дискретних випадкових величин і за своїм визначенням є випадковою функцією.

Завдання статистичного оцінювання стаціонарності досліджуваної послідовності вирішується на основі застосування статистичних критеріїв t -Стюдента і F -Фішера. Статистичний критерій t -Стюдента використовується на основі визначення статистичних оцінок математичних очікувань, а статистичний критерій F -Фішера – на основі визначення статистичних оцінок дисперсії двох реалізацій на непересічних часових інтервалах.

Для обчислення за даними вимірів статистичних оцінок середнього (математичного сподівання) і дисперсії стаціонарної послідовності дискретних випадкових величин використовують ергодичну гіпотезу.

Статистична оцінка кореляційної функції стаціонарної послідовності отримана за даними вимірів у вигляді цифрового часового ряду, а дисперсія отриманої оцінки визначається як центральний момент четверте порядку досліджуваного процесу.

Методика отримання статистичної оцінки одновимірної функції розподілу стаціонарної послідовності за даними вимірів аналогічна методиці отримання статистичної оцінки функції розподілу випадкової величини, але має іншу фізичну інтерпретацію: як оцінка стаціонарної послідовності дискретних випадкових величин, заданих на часовій осі.

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

Для визначення двовимірної функції розподілу стаціонарної послідовності за даними вимірів для гаусового закону розподілу визначені статистичні оцінки математичного очікування, дисперсії і кореляційної функції.

Основною спектральною характеристикою стаціонарної послідовності випадкових величин є її спектральна щільність потужності. До основних прикладних методів її статистичного оцінювання відносяться: дискретне перетворення Фур'є реалізацій досліджуваної стаціонарної послідовності; дискретне перетворення Фур'є статистичної оцінки кореляційної функції послідовності; лінійна вузькосмугової фільтрації реалізацій послідовності; зведення в квадрат відгуків з наступним їх усередненням за ансамблем.

У сучасних вимірювальних інформаційних технологіях використовується статистичне оцінювання спектральної щільності потужності, що ґрунтується на дискретному перетворенні Фур'є, як правило реалізованому через алгоритм швидкого перетворення Фур'є, вибірок досліджуваної стаціонарної послідовності з подальшим статистичним обробленням.

Вимірювальна інформаційна технологія статистичного оцінювання спектральної щільності потужності стаціонарної послідовності базується на операціях: центрування, дискретного перетворення Фур'є, зведення в квадрат, використання часових вікон і усереднення за ансамблем результатів обчислень.

Ефективність статистичного оцінювання спектральної щільності потужності стаціонарних процесів підвищується за рахунок конкретизації додаткової інформації про досліджуваний стаціонарний процес, наприклад завдання одновимірного закону розподілу або кореляційної функції процесу.

Питання для самоконтролю

- 1. Яка інформація використовується для визначення математичного сподівання дисперсії стаціонарного випадкового процесу?*
- 2. Наведіть визначення коваріаційної і кореляційної функцій стаціонарного випадкового процесу.*
- 3. Розкрийте суть кореляційної теорії випадкових процесів.*
- 4. Назвіть основні характеристики стаціонарного в широкому сенсі випадкового процесу.*
- 5. Сформулюйте задачу статистичного оцінювання характеристик випадкових процесів в рамках кореляційної теорії?*
- 6. Розкрийте основну суть ергодичної теорем в дослідженні випадкових процесів.*
- 7. Що таке ергодичність стаціонарного процесу щодо математичного очікування, кореляційної функції?*
- 8. Які необхідні і достатні умови ергодичності стаціонарного процесу відносно кореляційної функції?*
- 9. Наведіть постановку задачі побудови математичної моделі об'єкта*

дослідження за експериментальними даними.

10. Сформулюйте два підходи у вузькому і широкому сенсах до класифікації отриманих експериментальних даних про досліджувані процеси, включаючи дані вимірів.

11. Викладіть послідовність операцій отримання і опрацювання даних вимірювань.

12. Проаналізуйте динаміку зміни характеристик досліджуваних процесів на основі аналізу реалізації на короткочасових інтервалах.

13. Наведіть приклади реалізації досліджуваних процесів на тривалих часових інтервалах і дайте стисло характеристику динаміки їх змін.

14. Викладіть суть двох основних статистичних методів усереднення даних вимірів в статистичному оцінюванні характеристик процесів.

15. Назвіть основні операції перетворення безперервного (аналогового) сигналу в цифровий часовий ряд.

16. Розкрийте суть змін просторів елементарних подій як області визначення випадкового процесу у перетвореннях вимірювального сигналу.

17. Наведіть математичну модель стаціонарної послідовності, реалізацією якої є цифровий часовий ряд на кінцевому інтервалі часу.

18. Як визначається статистика у статистичному оцінюванні k -ї характеристики стаціонарної послідовності?

19. Наведіть постановку задачі статистичного оцінювання стаціонарності досліджуваної послідовності дискретних випадкових величин.

20. Розкрийте основний зміст застосування статистичних критеріїв t -Стьюдента і F -Фішера у завданні статистичного оцінювання стаціонарності досліджуваної послідовності випадкових величин.

21. Які часові характеристики стаціонарних послідовностей визначаються під час їх статистичного оцінювання?

22. Наведіть основні формули обчислення за даними вимірів статистических оцінок середнього (математичного сподівання) і дисперсії стаціонарної послідовності.

23. Розкрийте суть оцінювання якості обчислення статистичної оцінки математичного сподівання стаціонарної послідовності на основі використання дисперсії оцінки.

24. Наведіть послідовність операцій обчислення і оцінювання якості статистичної оцінки кореляційної функції стаціонарної послідовності.

25. Розкрийте суть статистичного оцінювання одновимірної функції розподілу стаціонарної послідовності.

26. Сформулюйте постановку задачі статистичного оцінювання двовимірної функції розподілу стаціонарної послідовності.

27. Дайте приклад статистичного оцінювання двовимірної функції розподілу стаціонарної послідовності за даними вимірювань з гаусовим розподілом.

28. Які спектральні характеристики стаціонарної послідовності

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

обчислюють для їх статистичного оцінювання за даними вимірів?

29. Назвіть основні прикладні методи статистичного оцінювання спектральної щільності потужності стаціонарної послідовності.

30. Які операції необхідно виконати для центрування і нормування даних вимірювань у вигляді цифрових часових рядів?

31. Наведіть основні формули застосування дискретного перетворення Фур'є в задачі статистичного оцінювання спектральної щільності потужності стаціонарної послідовності.

32. Розкрийте основну суть застосування вагових вікон для статистичного оцінювання спектральної щільності потужності стаціонарної послідовності

Conclusions

Statistical evaluation of random processes' characteristics – is a determination of statistical estimates of these characteristics based on measurement data of processes values.

Correlation theory is used for studying and statistical estimation of the first two moment functions of random processes. In theory and practice of statistical estimation most fully developed are: models, algorithms and software of calculating of statistical estimates of stationary random processes characteristics.

Characteristics of stationary processes are: mathematical expectation, dispersion, correlation function, spectral power density and one-dimensional distribution function.

Mathematical models, algorithms and software of calculating of statistical estimates of stationary processes characteristics referred to information support of stationary processes statistical estimation.

Statistical estimates of mathematical expectation, dispersion, correlation function, spectral power density of stationary processes are calculated based on measurement data of stationary processes values.

Application of Birkhoff-Khinchin ergodic theorem gives an opportunity, while statistically estimating stationary processes characteristics, to use time averaging implementations of investigated processes.

Sufficient and necessary conditions for ergodicity of a stationary process, with respect to mathematical expectation, are based on the use of known correlation function of the process.

Characteristic features of statistical estimation as a measurement process are: formation specifics of experimental data about studying processes and their consequent digital statistical processing.

Experimental data are conditionally divided into measurement data and other observational data that gives an opportunity to provide statistical estimation of stationary processes characteristics in narrow and broad sense.

Digital time series as an implementation of studied process were used during digital statistical estimation of stationary processes characteristics.

Among basic methods of statistical estimation two methods are highlighted, one of them gives an opportunity to make estimations by averaging over the set (ensemble) of implementations of the process studied; other – by averaging of process over time.

Application of method of implementation averaging over time allows for a considerable simplification of solving calculations problems of statistical estimations of stationary processes characteristics in practice.

Mathematical model, the implementation of which is a digital time series obtained as the measurement data, is presented as a random sampling of volume n of a stationary sequence of discrete random quantities.

Statistics, when statistically estimating k -th characteristics of the stationary sequence, is defined as the corresponding function of the random sampling of a stationary sequence of discrete random quantities.

Problem of statistical estimation of studied sequence stationary is solved by applying t-Student and F-Fisher statistical criteria. Here, t-Student statistical criterion is used on the basis of statistical estimates determination of mathematical expectations and F-Fisher statistical criterion - on the basis of statistical dispersion estimates determination of two implementations on disjoint time intervals.

When calculating, based on measurement data, statistical estimates of the average (mathematical expectation) and dispersion of stationary sequence is used ergodic hypothesis.

Statistical estimation of correlation function of stationary sequence is obtained based on data measurement in form of digital time series and calculated dispersion of the estimate obtained as a central moment of the fourth order.

Obtaining technique of statistical estimation of one-dimensional distribution function of stationary sequence based on measurements data is similar to obtaining technique of statistical estimation of random quantity distribution function, but has a completely different physical interpretation: as an estimate of stationary sequence of discrete random quantities.

To determine the two-dimensional distribution function of stationary sequence based on measurement data for Gaussian distribution law statistical estimates of mathematical expectation, dispersion and correlation function are defined.

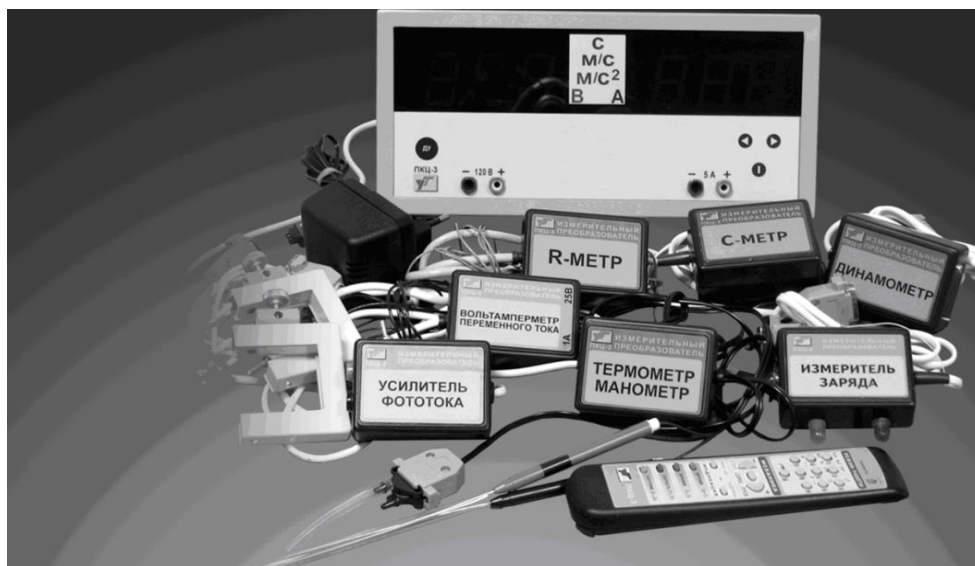
The main spectral characteristic of stationary sequence is its spectral power density. Basic applied methods of spectral density statistical estimation are: discrete Fourier transformation of studied stationary sequence implementations; discrete Fourier transformation of sequence correlation function; linear narrowband filtration.

In modern measuring information technologies is used statistical estimation of spectral density based on the use of discrete Fourier transformation of studied stationary sequence implementations. Information measurement technology of statistical estimation of stationary sequence spectral density is based on operations: centering, discrete Fourier transformation, the use of appropriate time windows and averaging over an ensemble of calculations results.

Ключові слова

Розділ 6. Статистичне оцінювання характеристик випадкових процесів

кореляційна теорія	correlation theory
стаціонарний процес	stationary process
ергодична теорема	ergodic theorem
методи усереднення	averaging methods
статистичні гіпотези	statistical hypotheses
статистичне оцінювання	statistical estimation
дисперсія	dispersion
кореляційна функція	correlation function
спектральна щільність	spectral density



ПЕРЕТВОРЮВАЧІ ІВС

7

- 7.1. Класифікація та основні характеристики вимірювальних перетворювачів
- 7.2. Різновидності сенсорів
- 7.3. Перетворювачі на операційних підсилювачах та їх похибки

7.1. Класифікація та основні характеристики вимірювальних перетворювачів

Вимірювальний перетворювач (ВП) – це пристрій з нормованими метрологічними характеристиками, що перетворює одну вимірювану величину в іншу величину або у вимірювальний сигнал, зручний для опрацювання, зберігання, перетворення, індикації або передавання.

За характером перетворення розрізняють:

аналоговий ВП, що перетворює одну аналогову величину (вимірюваний сигнал) в іншу аналогову величину (вимірюваний сигнал);

аналого-цифровий ВП призначений для перетворення аналогового вимірювального сигналу в цифровий код;

цифро-аналоговий ВП призначений для перетворення цифрового коду в аналогову величину.

За місцем у вимірювальному ланцюгу розрізняють:

сенсори, на який безпосередньо діє вимірювана фізична величина і який є першим перетворювачем у вимірювальному ланцюгу приладу;

ВП параметрів сигналів і ланцюгів – вимірювальний перетворювач, який у вимірювальному ланцюгу займає місце після сенсора (це вимірювальні підсилювачі, масштабні перетворювачі, детектори тощо).

За принципом дії сенсори поділяються на генераторні та параметричні. Процес перетворення включає трансформацію виду енергії і складається зі сприйняття вхідної вимірюваної величини за допомогою чутливого елемента і подальшого перетворення її в іншу форму за допомогою "перетворювального елемента". Таку комбінацію відображає узагальнена модель сенсора на рис. 7.1.

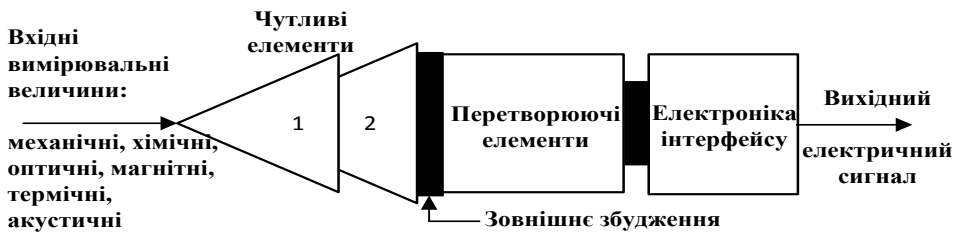


Рис. 7.1. Узагальнена модель сенсора

Деякі ВП допускають трансформацію виду енергії і в прямому і в зворотному напрямках. Така властивість характерна, наприклад, п'єзоелектричним перетворювачам (прямий і зворотний п'єзоєфекти).

Класифікація ВП приведена на рис. 7.2. Їх прийнято поділяти на дві групи: до першої відносяться перетворювачі (сенсори), в яких

відбувається зміна виду вхідної величини; до другої групи відносять ВП без зміни виду вхідної величини або ВП параметрів сигналів і ланцюгів. Найпоширенішими є сенсори з електричним вихідним сигналом і електричним сигналом активування (збудження).



Рис. 7.2. Класифікація ВП

Генераторні сенсори (датчики) створюють необхідну напругу або струм в процесі поглинання енергії вимірюваної фізичної величини. Параметричні сенсори живляться енергією зовнішнього джерела, але вони також можуть поглинати частину енергії вимірюваної величини.

Дія сенсорів базується на одному або декількох фізичних явищах або ефектах, до числа яких відносяться ємнісний, індуктивний і електромагнітний, резистивний і терморезистивний, п'єзоелектричний, ефекти Холла, Гаусса, Доплера, тунельний, магніторезистивний, термоелектричні (Зеєбека і Пельтьє), іонізаційні, фотогальванічні та ін.

На практиці найчастіше виникає необхідність у вимірювання наступних фізичних величин і характеристик об'єктів: лінійні переміщення, кутове положення об'єктів у просторі, швидкість, прискорення, сила, кутова швидкість обертання, параметри вібрації, температура, тиск, відносна вологість, шорсткість поверхні, концентрація газів і водневий показник, парціальний тиск O_2 і CO_2 у крові, обертальний момент, характеристики електричного, магнітного та акустичного полів. У табл. 7.1 наведені типи сенсорів і деякі вимірювані фізичні величини. Назва кожного типу визначає принцип перетворення.

Метрологічні характеристики ВП. ВП відрізняються від інших перетворювачів тим, що мають нормовані метрологічні характеристики (МХ). *Метрологічні характеристики* – це технічні характеристики ВП, що описують їх властивості, необхідні для виконання ними функціонального призначення і впливають на результати і похибки вимірювань.

Розділ 7. Перетворювачі ІВС

В документації на ВП вказуються наступні МХ:
номінальна функція перетворення (номінальна статична характеристика перетворення) вимірюваної величини у вигляді функції $f(x)$, графіка або таблиці; якщо номінальна функція перетворення лінійна і проходить через початок координат, вказують лише значення номінального коефіцієнта перетворення;

Таблиця 7.1

Вимірювані велечин, фізичні поля та випромінювання	Класифікація сенсорів																									
	Переміщення	Положення	Швидкість	Прискорення	Сила	Навантаження	Розтягнення	Кругильний момент	Лін. і круг. перетворення	Вібрація	Потік	Температура	Тиск	Вакуум	Відносна вологість	Атомний контур	Газова концентрація	Стан крові, рН	Оптичні поля	ІЧ випромінювання	Магнітні поля	Акустичні поля	Аудіополя і шуми	Рентгенівські промені	Кутова швидкість	Гамма-випромінювання
Ємнісні	●	●	●	●	●	●	●	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Індуктивні	●	●	●	●	●	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Електромагнітні	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Резистивні	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Тензодатчики	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
На явищі механічної деформації	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Терморезистори	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Термістори	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
На ефектах Гаусса і Холла	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
На хімічних польових транзисторах	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Хвилеводні	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
П'єзоелектричні	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
На ефекті Доплера	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Полімерні	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
На поверхневих акустичних хвилях	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Тунельні	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Термодинамічні	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Іонізуючі	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Фотонні	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Оптоелектронні (фізичні)	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
Оптоелектронні (хімічні)	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

чутливість – похідна від номінальної функції перетворення за інформативним параметром вхідного сигналу;

поріг чутливості – найменша зміна вимірюваної величини, яка викликає помітну зміну вихідного сигналу ВП;

похибка – різниця між вихідним сигналом ВП та його очікуваним значенням; обумовлена відмінністю між номінальною і реальною функціями перетворення ВП; їх класифікують за способом вираження (абсолютні, відносні), за характером прояву (систематичні, випадкові), за умовами застосування (основні, додаткові);

діапазон вимірювань – область значень величини, для якої

нормовані допустимі межі похибки;

динамічні характеристики (повні і часткові) – описують перетворення змінюваних в часі сигналів вимірюваної величини або сигналів вимірювальної інформації; до повних динамічних характеристик відносяться *імпульсна перехідна характеристика* і *комплексна частотна характеристика*, *амплітудно-частотна характеристика* у сукупності з *фазо-частотною характеристикою*; до часткових динамічних характеристик відносяться *час реакції* ВП і *граничні значення частот*, в межах яких амплітудно-частотна характеристика ВП відхиляється від номінального значення не більше ніж на задану величину.

Імпульсна перехідна характеристика – це вихідний сигнал лінійного аналогового ВП, що виникає як реакція на вхідний сигнал у вигляді δ -функції, тобто на дуже короткий імпульс, потужність якого достатня для отримання помітного вихідного сигналу.

Перехідна характеристика – це вихідний сигнал лінійного аналогового ВП як реакція на сигнал у вигляді одиничного стрибка.

Амплітудно-частотна характеристика – це функція від частоти відношення амплітуди синусоїдального вихідного сигналу до амплітуди синусоїдального сигналу на вході ВП.

Фазочастотная характеристика – це функція від частоти зсуву фази між вихідним і вхідним синусоїдальними сигналами ВП; ці характеристики для фізично реалізованих динамічних ланок негативні.

Комплексна частотна характеристика – це відношенням двох поліномів з аргументом $j\omega$, в якому ступінь полінома чисельника не перевищує ступеня полінома знаменника.

Час реакції (час відгуку) – інтервал часу між моментом стрибкоподібної зміни сигналу на вході ВП і моментом, починаючи з якого вихідний сигнал відрізняється від усталеного значення не більше ніж на задану величину (наприклад, не більше, ніж на межу допустимої основної похибки).

Важливою характеристикою сенсорів є імпеданс, який визначає вимоги щодо узгодження їх з наступними ВП. Імпедансні характеристики деяких сенсорів наведені в табл. 7.2.

Таблиця 7.2

Тип сенсорів	Модуль імпеданса
Термопары	Низький (< 20 Ом)
Терморезистори	Високий (> 1 кОм)
Резистивні сенсори температури	Низький (< 1 кОм)
Сенсори тиску	Високий (> 1 кОм)
Тензодатчики	Низький (< 1 кОм)

Розділ 7. Перетворювачі ІВС

Скляний рН електрод	Дуже високий (10^9 Ом)
Вимірювальний потенціометр	Високий (500 Ом...100 кОм)

ВП для сигналів низьких рівнів з виходу високоімпедансних сенсорів повинні бути розміщені в безпосередній близькості від нього.

7.2. Основні типи сенсорів

Ємнісні сенсори. Ці сенсори використовують для вимірювання швидкості, прискорення, сили, тиску, рівня рідини в ємності, вологості та інших фізичних величин. Ємність паралельних електропровідних пластин з діелектричним матеріалом між ними:

$$C = \epsilon_0 \epsilon_r S(n-1)/d, \quad (7.1)$$

де $\epsilon_0 = 8,854 \cdot 10^{-12}$ Ф/м – діелектрична стала; n – кількість пластин, ϵ_r – відносна діелектрична проникність матеріалу; S – площа, спільна для двох пластин конденсатора, м²; d – відстань між пластинами, м.

Будь-які зміни параметрів у (7.1) під дією вимірюваної величини забезпечують практичні механізми перетворення. На рис. 7.3 показані

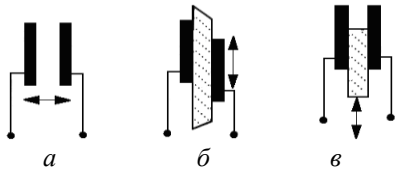


Рис. 7.3. Ємнісні сенсори для вимірювання зміщення

прикладі конфігурацій ємнісних сенсорів з плоскими електродами, які використовують для вимірювання різних фізичних величин, зокрема сенсори типу (а, б) – для визначення зміщення, сили, потоку, вакууму і тиску, а типу (в) – для оцінювання

варіації ϵ_r , викликані поглинанням води або хімічною реакцією з діелектричним матеріалом.

На рис. 7.4 зображений ємнісний сенсор тиску, корпус якого виготовлений з плавленого кварцу, що має незначний температурний коефіцієнт розширення.

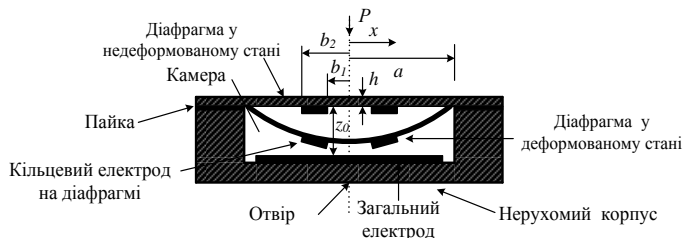


Рис. 7. 4. Ємнісний сенсор тиску

Перетворювач складається з кругової діафрагми, жорстко приєднаної до корпусу. Порожнина малого розміру з відстанню Z_0 між пластинами конденсатора поєднана з атмосферою невеликим отвором, що забезпечує вимірювання атмосферного тиску. Діафрагма має кільцеподібний електрод з металізацією хромом або золотом на внутрішній поверхні і загальний електрод, розміщений у нижній частині порожнини. Вимірювання ємності перетворювача теля вимірюється в пікофарад як функція тиску P [кг/см²]:

$$\Delta C(P) \cong 8,85 \cdot \pi (1 - \mu^2) a^4 (b_2^2 - b_1^2) P / 16 E Z_0^2 h^3, \quad (7.2)$$

де μ – коефіцієнт Пуассона (0,17 для плавленого кварца); E – модуль пружності Юнга ($745 \cdot 10^3$ кг·см⁻² для плавленого кварца); a – радіус порожнини м; b_1 , b_2 – внутрішній та зовнішній радіуси електрода, м; h – товщина діафрагми, м; Z_0 – глибина полості, м.

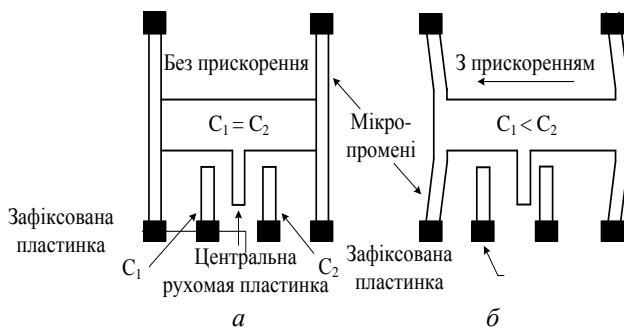


Рис. 7.5. Ємнісний акселерометр

Структуру монолітного ємнісного акселерометра у спокійному стані подано на рис. 7.5,а. Акселерометр складається з декількох диференціальних конденсаторів, а кожна конденсаторна секція –

з центральної рухомої і двох нерухомих зовнішніх пластин. На рис. 7.5,б показано відхилене положення центральної пластини, коли на сенсор діє сила прискорення. Зміна ємності пропорційна прискоренню.

У багатьох додатках ємнісні сенсори включають в схему моста змінного струму для перетворення ємності в електричний сигнал. За диференціального включення ємнісних сенсорів чутливість

Розділ 7. Перетворювачі ІВС

подвоюється, а температурна залежність зменшується. В інших додатках ємнісні сенсори включаються в коливальний контур генератора, частота генерації якого пропорційна значенням вимірюваної величини.

Індуктивні сенсори. Механізм перетворення полягає в тому, що власна індуктивність однієї котушки або взаємна індуктивність двох котушок змінюються під дією вимірюваної величини. Вимірюваними величинами можуть бути лінійне або обертальне переміщення, тиск, сила, обертальний момент, віброшвидкість, прискорення. Зміни індуктивності викликаються переміщенням феромагнітного осердя. Індуктивність котушки збільшується, якщо осердя входить в котушку, і досягає максимального значення, коли воно відцентроване по довжині котушки.

Один з варіантів індуктивного сенсора схематично зображений на рис. 7.6,а і відомий як перетворювач лінійного переміщення на базі диференціального трансформатора.

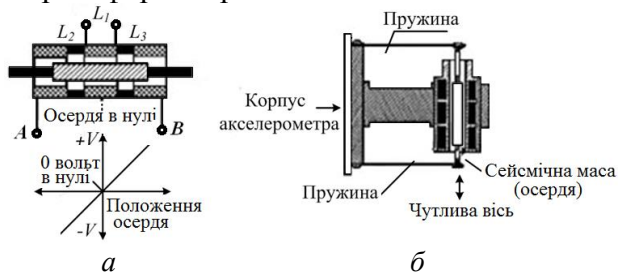


Рис. 7.6. Індукційні сенсори на диференціальних трансформаторах:
а – лінійного переміщення, б – прискорення

Сенсор складається з первинної L_1 та двох зв'язаних L_2 і L_3 котушок і спільного феромагнітного осердя. Котушки намотані на немагнітну трубку, всередині якої вздовж вісі ковзає осердя. Сенсор призначений для визначення сили, навантаження, тиску, обертального моменту. Схема акселерометра на його основі подана на рис. 7.6,б. Осердя служить сейсмічної масою, а пружини забезпечують компенсуючу силу.

Частота напруги збудження для котушки L_1 знаходиться в діапазоні від 1 до 10 кГц. Котушки L_2 і L_3 намотані протифазно таким чином, щоби напруги, індуковані в них котушкою L_1 , були зсунуті по фазі на 180° . Отже, напруга на клеммах А-В дорівнює нулю в разі, коли осердя розташоване симетрично відносно котушок L_2 і L_3 . Коли воно зміщується відносно початкового положення, напруга на клеммах А-В змінюється за амплітудою і фазою.

Такі сенсори здатні працювати в діапазоні переміщень від ± 1 до ± 300 мм з чутливістю від $\sim 1,7$ до 250 мВ/В/мм.

Той самий механізм перетворення на базі диференціального трансформатора використовується у сенсорах кутових переміщень і обертального моменту. Для досягнення необхідної лінійності кут обертання обмежений значеннями $\pm 40^\circ$. Крім того такі сенсори можуть використовуватися з трубками Бурдона, сильфонами і манометрами для вимірювання сили та тиску.

Електромагнітні сенсори. Якщо провідник довжиною l або котушка такої ж довжини рухається зі швидкістю ds/dt поперек і перпендикулярно лініям магнітного потоку з індукцією B , в провіднику (котушці) відповідно закону Фарадея генерується електрорушійна сила (е.р.с.) $e(t) = Bl ds/dt$, де $l ds$ – елемент площі, яку перетинають лінії потоку за час dt , $Bl ds$ – диференціал магнітного потоку $d\psi$ крізь цю площу. Якщо котушка має W витків, то

$$e(t) = W d\psi/dt. \quad (7.3)$$

Такий принцип перетворення використовується в сенсорах швидкості. На рис. 7.7, а показані дві котушки L_1 і L_2 з'єднані протифазно.

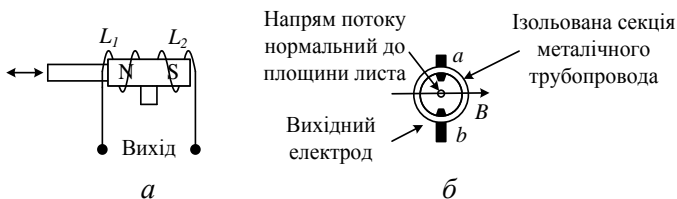


Рис. 7.7. Електромагнітні сенсори: а – швидкості; б – швидкості потоку рідини

Вимірювана величина пов'язана з рухомим магнітом, що вільно переміщується в котушках. Вихідна напруга пропорційна швидкості руху магніту. Такі сенсори відомі як лінійні перетворювачі швидкості.

Іншим застосуванням такого механізму перетворення є вимірювання швидкості потоку електропровідної рідини (рис. 7.7,б). Потік спрямований перпендикулярно площині рисунка. Магнітне поле перпендикулярне потоку. Е.р.с., що генерується вздовж діаметра перпендикулярно потоку і магнітного поля, пропорційна швидкості потоку.

Резистивні сенсори. Такі сенсори застосовують для перетворення переміщень, механічної напруги, тиску, сили, температури і швидкості потоку в електричні величини. Механізми перетворення засновані на зміні електричного опору під дією вимірюваних величин. Нижче розглянуті основні типи резистивних датчиків.

Розділ 7. Перетворювачі ІВС

Потенціометричний сенсор – це механічно керований резистор. Він складається з постійного дротового резистора і механічного ковзаючого контакту, який рухається по ньому і торкається різних сегментів резистора, як це схематично показано на рис. 7.8,а,б, де К позначає частину резистора між одним з його виводів і контактом.

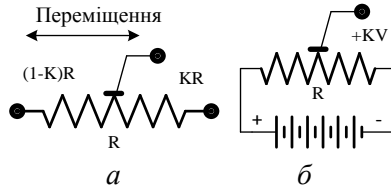


Рис. 7.8. Потенціометричні сенсори переміщення: а – опір пропорційний переміщенню; б – напруга пропорційна переміщенню.

Вимірюване переміщення визначається розташуванням механічного електричного контакту відносно виводів резистора і пропорційне частині опору або частини напруги. Розрізнявальна здатність такого сенсора залежить від діаметра дроту (крока намотування потенціометра). Для підвищення розрізнявальної здатності використовують металізовані керамічні або плівкові резистори. Якщо провід намотати на кільцеподібну трубку, сенсор перетворюватиме кутове переміщення механічного контакту в опір. Вихідна напруга, що відповідає переміщенню, силі або тиску є частиною зовнішньої значної за величиною напруги, що виключає необхідність її підсилення.

Тензорезистори. Вуглецеві гранули, упаковані в малому об'ємі, включені послідовно з джерелом напруги і резистором навантаження, раніше використовувались як мікрофони. Той же принцип використовують сенсори деформації на основі вуглецевих пластин. Пізніше з'явилися дротові, фольгові і напівпровідникові тензорезистори.

Дротові тензорезистори. Опір відрізка дроту є функцією питомого опору матеріалу ρ , його довжини L і площі поперечного перерізу S :

$$R = \rho L / S. \quad (7.4)$$

Якщо провід розтягувати (в межах пружної деформації), його довжина збільшується на ΔL , діаметр зменшується на Δd , а опір збільшується на ΔR (за незмінного ρ), тому з рівняння (7.4) маємо

$$\Delta R / R \approx (1 + 2\mu) \Delta L / L, \quad (7.5)$$

де μ – коефіцієнт Пуассона (відношення поперечного стиснення до осевого розтягування). Коефіцієнт тензочутливості визначається як

$$(\Delta R / R) / (\Delta L / L) = 1 + 2\mu . \quad (7.6)$$

У сенсорах струнного типу провід натягнутий між ізольованими стійками: одна стійка приєднана до нерухомої частини, а інша до переміщуваного елементу конструкції. Зміна відстані між стійками призводить до розтягування дроту і зміни його опору. Цей принцип застосовується в перетворювачах сили і прискорення.

У тензорезисторах, жорстко з'єднаних з поверхнею досліджуваного об'єкта, сітка високоякісного дротяного опору в формі меандру затиснута між двома тонкими шарами паперу, просоченого смолою для забезпечення необхідної міцності. Такі сенсори жорстко механічно з'єднуються з конструктивними елементами для визначення деформації в бажаному місці досліджуваного об'єкта і в необхідному напрямку.

Фольгові тензорезистори виготовляють методами фотолітографії з листового константана або хромонікелевого сплаву товщиною 2,5 ... 5 мкм. Такий резистор у формі меандра затиснутий між двома тонкими шарами паперу, просоченими для забезпечення необхідної міцності, жорстко з'єднується з поверхнею деформованого досліджуваного зразка. Коефіцієнт тензочутливості таких сенсорів сягає величини 2,7.

На рис. 7.9 зображений двоелементний тензорезистор, що дає змогу визначати деформації в двох взаємно перпендикулярних напрямках.

Ширина і довжина тензорезистору визначають площу поля вимірювання деформації. Сенсори з великим полем використовують для вимірювання середньої деформації в залізобетонних брусках.

Напівпровідникові сенсори деформації працюють відповідно до механізму перетворення, відомого як *n'єзорезистивний ефект*. Він полягає у зміні електричного питомого опору, викликаного пружним полем деформації. У деяких напівпровідниках цей ефект досить відчутний (в сто разів більший, ніж в металевих тензорезисторах).

У випадку коли напівпровідниковий сенсор деформований, розподіл кількості носіїв заряду і їх рухливість змінюються, отже, змінюється і питомий опір. Їх коефіцієнт тензочутливості дорівнює

$$G_F = (\Delta R / R) / (\Delta L / L) = 1 + 2\mu + \pi_L Y , \quad (7.7)$$

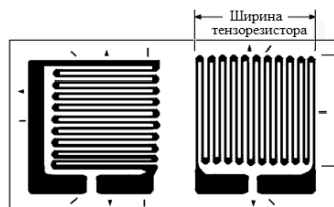


Рис. 7.9. Двоелементний фольговий тензорезистор

Розділ 7. Перетворювачі ІВС

де Y – модуль пружності Юнга; π_L – поздовжній п'єзорезистивний коефіцієнт. Перші два члена у (7.7) відповідають змінам розміру як у дротових і фольгових тензорезисторів, третій член обумовлений п'єзорезистивністю. Значення π_L більше за $(1+2\mu)$, приблизно в 100 разів. Величина G_F залежить і від напрямку прикладеного зусилля.

Основною схемою для вимірювання тензорезисторів є мостова схема постійного струму Уїтстона (рис 7.10,а), яку можна представити двома паралельно з'єднаними подільниками напруги (рис. 7.10,б).

Деформація тензорезисторів викликає зміну вихідної напруги

$$\Delta E_0 = Va(\Delta R_1/R_1 - \Delta R_2/R_2 + \Delta R_3/R_3 - \Delta R_4/R_4)/(1+a)^2, \quad (7.8)$$

де $a = R_1/R_2 = R_4/R_3$. Якщо R_3 і R_4 – постійні резистори, а R_1 та R_2 – металеві тензорезистори і $a=1$, то $\Delta R_3 = \Delta R_4 = 0$ і формула (7.8) спрощується:

$$\Delta E_0 = V(\Delta R_1/R_1 - \Delta R_2/R_2)/4. \quad (7.9)$$

Терморезистивні сенсори. Питомий електричний опір металів і напівпровідників збільшується з нагріванням. Цей механізм використано в термометрах опору (ТО), що виготовляються з нікеля, ніхрома, вольфрама, міді і платини. Платинові ТО мають практично лінійну та високостабільну залежність опору від температури $R_t(t^\circ)$, яка описується рівнянням Ван–Дюзена

$$R_t(t^\circ) = R_0 \left[1 + At^\circ + B(t^\circ)^2 + C(t^\circ - 100)t^{\circ 3} \right], \quad (7.10)$$

де $R_t(t^\circ)$ – опір ТО за температури t° ; R_0 – опір за температури таяння льоду ($0,01^\circ\text{C}$); A, B, C –

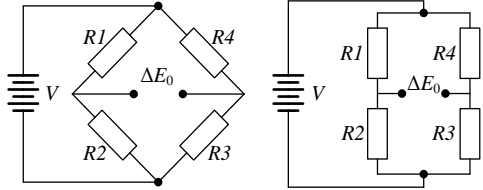


Рис. 7.10. Мост Уїтстона (а) і його представлення подільниками напруги (б)

Уїтстон, який винайшов багатоох вирахов. У 1837 році у співавторстві з Вільямом Куком отримав патент на електромагнітний телеграф і створив першу діючу телеграфну лінію в Англії. У 1843 році запропонував пристрій для вимірювання опору – вимірювальний міст.

константи Ван-Дюзена ($C = 0$ для $t^{\circ} > 0$).

Типові величини коефіцієнтів за стандартом DIN43760 дорівнюють: $A = 3,9080 \cdot 10^{-3} \text{ }^{\circ}\text{C}^{-1}$; $B = -5,8019 \cdot 10^{-7} \text{ }^{\circ}\text{C}^{-2}$; $C = 4,2735 \cdot 10^{-12} \text{ }^{\circ}\text{C}^{-4}$.

Як правило, ТО включають в плече мосту Уїтстона. В такому разі виникає необхідність компенсації опорів з'єднувальних проводів, що реалізується за трипровідною (рис. 7.11) або чотирипровідною схемою.

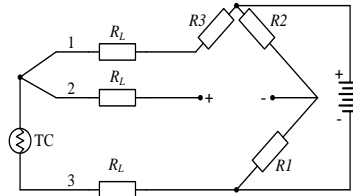


Рис. 7.11. Трипровідний інтерфейс для ТО

Вплив резисторів R_L на рис. 7.11 зкомпенсовано, оскільки вони включені в протилежні плечі мосту разом з резистором R_3 і ТО.

Термістори – теплочутливі опори, на відміну від ТО мають більш відчутну залежність опору від температури, крім того їх опір з ростом температури зменшується. У термісторах використовують сплави оксидів марганцю, нікелю, кобальту, міді, заліза і титану у формі невеликих кульок, пластинок або стрижнів. Залежність $R_t(t^{\circ})$ описується поліномом третього ступеня

$$\ln R_t(T^{\circ}) = a_0 + a_1/T^{\circ} + a_2/T^{\circ 2} + a_3/T^{\circ 3}, \quad (7.11)$$

де $R_t(T^{\circ})$ – значення опору за температури $T^{\circ} [K]$; a_0, a_1, a_2, a_3 – числові коефіцієнти, визначаються за чотирма точками градування.

Мостові схеми Уїтстона є типовими і для термісторів. Іноді вони включаються як елемент дільника напруги (рис. 7.12). У такій схемі вихідна напруга збільшується з ростом температури:

$$e_0(t) = V_s R / [R - R_t(t^{\circ})]. \quad (7.12)$$

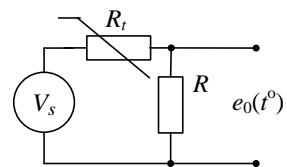


Рис. 7.12. Схема вмикання термістора з подільником напруги

Правильний вибір величини R забезпечує можливість роботи на квазілінійній ділянці функції $e_0(t^{\circ})$, де t° – температура в $^{\circ}\text{C}$.

Термоанемометр. У цьому сенсорі використано властивість потоку газу або рідини, оточуючої чутливий елемент з електричним підігрівом, охолоджувати його.

Розділ 7. Перетворювачі ІВС

За постійної потужності підігріву температура чутливого елемента залежить від швидкості потоку. За постійної температури зі швидкістю потоку пов'язана споживана потужність нагріву. Термоанемометр використовують для аналізу мікроструктури потоків газів і рідин, що необхідно під час проектування аеродинамічних схем літаків, систем вентиляції, вимірювань швидкості крові в медичних дослідженнях. Його чутливий елемент має малі розміри, виготовляється з тонких проводів і плівок для забезпечення високої просторової роздільної здатності та швидкодії.

Принцип роботи таких сенсорів полягає у зміні опору чутливого елемента, викликаному конвекцією тепла в потоці теплоносія. Опір терморезистора анемометра вимірюють мостом Уїтстона. Вихідна напруга моста в цьому випадку:

$$e_0^2(V) = A + BV^{0.5} + CV, \quad (7.13)$$

де V – швидкість теплоносія; A, B, C – коефіцієнти.

Термоанемометр складається з терморезистивного чутливого елемента, розташованого між двома стрижнями на опорі перетворювача. Чутливий елемент виготовляють з платиного або вольфрамового дроту малого діаметра (рис. 7.13).

Чутливим елементом плівкового термоанемометра є нікелеве наплення на кварцовому волокні діаметром 70 мкм. Типове значення опору чутливого елемента складає 3,5 Ом, а $V \in (0,2 \dots 500)$ м/с.

Сенсори на ефекті Холла застосовують для виміру магнітної індукції постійних, змінних і імпульсних магнітних полів. Як матеріал перетворювача використовують напівпровідники p - або n -типу (арсенід індію, фосфід арсеніду індію, легований кремній), в яких ефект Холла найбільш відчутний. Структуру сенсора Холла утворює кремнієвий p -легований напівпровідник, розташований в магнітному полі з індукцією B_z в



Едвін Герберт Холл (Edwin Herbert Hall, 1855 - 1938),

американський фізик відкрив і дослідив «ефект Холла».



Рис. 7.13. Конструкція термоанемометра

золотій пластинці, розміщеній на склі, через яку пропуская електричний струм. На бічних краях пластини в магнітному полі виникала різниця потенціалів – електрорушійна сила Холла.

напрямку z і електричному полі напруженістю E_x в напрямку x (рис. 7.14).

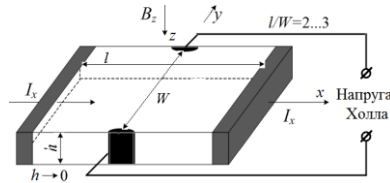


Рис. 7.14. Структура сенсору на ефекті Холла

Робочий струм I_x проходить в напрямку x . На рухомі носії зарядів в магнітному полі діє сила Лоренца, внаслідок чого концентрація зарядів в околі бічних країв сенсора змінюється і виникає е.р.с. Холла

$$V_H(B) = R_H(I_x B_z)/h, \quad (7.15)$$

де R_H – постійна Холла (від’ємна величина для напівпровідників n -типу), h – товщина чутливого елемента.

З ефектом Холла тісно пов’язаний магніторезистивний ефект (ефект Гаусса). Якщо якщо в сенсорі $l \ll W$, е.р.с. Холла може розглядатись як майже замкнута накоротко. Як наслідок, носії заряду рухаються під кутом Холла в напрямку x . Збільшення довжини пробігу носіїв заряду викликає збільшення опору сенсора, яке і складає сутність магніторезистивного ефекту.

Сенсори на основі хімічних польових транзисторів (ChemFET).

В цих сенсорах використано принцип дії метал-окисел-напівпровідникового польового транзистора (МОНПТ) з індукованим каналом. На затвори таких транзисторів наносяться один або декілька шарів спеціальних покриттів, здатних реагувати на певні хімічні речовини. Останні, впливаючи на затвор транзистора, змінюють його провідність між стоком і витокком. Величина цих змін визначається типом хімічної речовини. На рис. 7.15 приведена структура сенсору, що містить витік, стік і канал проходження струму з витокка до стоку.

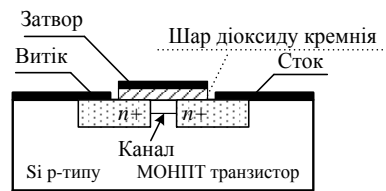


Рис. 7.15. Структура хімічного польового транзистора

Для створення активних поверхневих мембран використовують різні оксиди для виявлення різних іонів. Відомі сенсори водню, аміаку, сірководню, углеводнів, спиртів та ін.

Двома важливими прикладами ChemFET є МОНПТ із затвором з паладію (Pd), чутливі до газоподібного водню і іонів. Молекулярний водень у повітрі поглинається поверхнею затвора, де він піддається каталітичному

Розділ 7. Перетворювачі ІВС

розкладанню на атомарний водень (H). Останній, поширюючись на більшу частину електрода з паладію, формує дипольний шар на поверхні $Pd - SiO_2$. Дипольна поляризація модулює струм каналу прямо пропорційно концентрації іонів H . В іншому *ChemFET* платинова плівка завтовшки 10 нм, нанесена на поверхню Pd електрода затвора, дає змогу вимірювати концентрацію аміаку (NH_3). Якщо електрод затвора виконаний з перфорованої платинової плівки, то сенсор забезпечує вимірювання концентрації моноокси вуглецю. Для збільшення швидкодії) використовують зовнішній нагрівач (до 150 °C).

Такий принцип використовують для створення біосенсорів. Як селективну мембрану використовують органічні матеріали. Їх можна застосовувати безпосередньо на живому організмі.

П'єзоелектричні сенсори. П'єзоелектрику названо так від грецького слова "piezein" – пресувати. Коли п'єзоелектричний кристал деформується прикладеним тиском, в ньому виникає електрична поляризація, пропорційна за величиною і знаком механічному напруженню – це прямий п'єзоелектричний ефект. Зворотний ефект має місце, коли електричне поле викликає пружну деформацію кристала.

Типовими матеріалами, використовуваними в таких сенсорах є кристалічний кварц, нібат літію, деякі види здатної до поляризації фероелектричної кераміки, фероелектричні полімери, осаджені або напилені плівки сульфід кадмію і оксида цинку.

Фероелектрична кераміка набуває п'єзоелектричних властивостей тільки внаслідок поляризації. Поляризація досягається прикладанням високої напруги постійного струму до електродів, розміщених на бічних поверхнях пластини за високої температури.

Перетворювачі бувають різної форми – прямокутної, кругової з паралельними поверхнями, зі сферичною кривизною. Ефективність перетворення сенсорів визначається коефіцієнтом електромеханічного зв'язку:

$$K^2 = \frac{\text{механічна енергія, перетворена в електричну енергію}}{\text{вхідна механічна енергія}}$$

для прямого п'єзоелектричного ефекта, і

$$K^2 = \frac{\text{електрична енергія, перетворена в механічну енергію}}{\text{вхідна електрична енергія}}$$

для зворотного п'єзоелектричного ефекта.

Типовими величинами K є: для кварца – 0,1, для фероелектричної кераміки – 0,5...0,7, для фероелектричних полімерів – 0,2...0,3.

Важливим параметром для розрахунку сенсорів і їх узгодження з об'єктом контролю, є акустичний імпеданс $Z_0 = V\rho$, де V – швидкість поздовжньої акустичної хвилі в матеріалі сенсора; ρ – питома щільність матеріалу сенсора.

П'єзoeлектричні сенсори зазвичай використовують для приймання/генерування поздовжніх або поперечних хвиль, що поширюються в об'єктах контролю зі швидкостями V_L і V_s відповідно. У поздовжньої хвилі сила стиснення і рух частинок ідуть в напрямку поширення хвилі, в поперечній хвилі – перпендикулярно напрямку поширення хвилі.

У деяких сенсорах, використовуються поверхневі акустичні хвилі, що поширюються у приповерхневому шарі твердого середовища, де зосереджена велика частина їх енергії зі швидкістю $V_R < V_L, V_s$.

Максимум енергії переноситься від сенсора в інше середовище за однакових акустичних імпедансів матеріалу сенсора і середовища, тобто $Z_0 = Z_C$. Для оптимізації умов передавання використовують чвертьхвильовий узгоджувальний шар з акустичним імпедансом $Z_{ш} \approx \sqrt{Z_0 Z_C}$ завтовшки у чотири рази менше за довжину поздовжньої хвилі в матеріалі об'єкта контролю. Для поліпшення узгодження можливе використання кількох проміжних шарів.

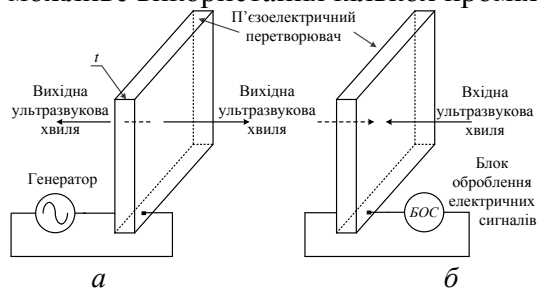


Рис. 7.16. Перетворювач хвилі поздовжнього типу в режимах:

- а – генерування (зворотний п'єзоefект);
- б – приймання (прямий п'єзоefект)

П'єзоперетворювачі поздовжньої хвилі. На рис. 7.16 показані п'єзоперетворювачі з двома нанесеними на їх поверхні електродами, орієнтованими перпендикулярно напрямку випромінювання/ приймання. Генератор електричних сигналів частотою f_0 з'єднується електрично з електродами випромінювача.

Утворені поздовжні хвилі зі швидкістю V_L поширюються в матеріалі. Максимальне збудження має місце, коли товщина h активного елемента п'єзоперетворювача дорівнює половині довжини λ в ньому:

$$h = V_L / 2f_0. \quad (7.16)$$

Найчастіше одноелементні п'єзоперетворювачі використовують в системах ультразвукового контролю і вимірювань. У неруйнівному

Розділ 7. Перетворювачі ІВС

контролі і медичних ультразвукових обстеженнях застосовують перетворювачі в діапазоні частот від 100 кГц до 50 МГц.

На рис. 7.17 показана осьова (пунктирна лінія) і поперечні області ультразвукового поля круглого поршневого перетворювача.

Осьове поле акустичного тиску складається з двох областей – ближньої (область Френеля) протяжністю $r_{\text{бл}}$ і дальньої (область Фраунгофера). В області Френеля поле має серію максимумів і мінімумів і закінчується останнім енергетичним максимумом, який вважається *ефективним фокусом перетворювача* (межею ближньої зони). Межа ближньої зони визначається як

$$r_{\text{бл}} = \left(D^2 / 4\lambda - \lambda / 4 \right), \quad (7.17)$$

где D – діаметр п'єзопластини; $\lambda = V_M / f_0$ – довжина хвилі в середовищі поширення; V_M – швидкість хвилі в середовищі.

Діаметр променя (на рівні - 6 dB) і ефективний фокус пов'язані

$$\text{Діаметр променя} = 1,028 F V_M / f_0 D, \quad (7.18)$$

де F – ефективна фокусна відстань у середовищі.

Фокусування випромінювання досягається застосуванням акустичних лінз або використанням сферичного п'єзоперетворювача.

Перетворювачі для медичних досліджень. Ультразвукові формувачі зображень внутрішніх органів використовуються для візуалізації і відображення неоднорідності акустичного імпедансу в біологічних середовищах шляхом поширення в ньому короткого ультразвукового імпульсу, як показано на рис 7.18. Цей метод відомий як технологія *імпульсної луналакації*.

Ультразвуковий приймач-передавач (ПП) збуджується електричним імпульсом і відповідний акустичний імпульс поширюється в середовищі зі швидкістю V і взаємодіє з неоднорідністю акустичного імпедансу на відстані d . На неоднорідності частина енергії імпульсу відбивається в зворотному напрямку. Перші відбиті імпульси повертаються до того ж перетворювача і ПП, який перетворює акустичний сигнал в електричний. Повний час

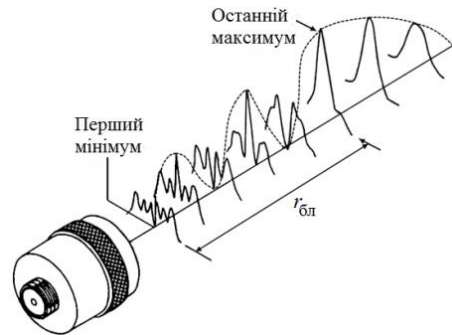


Рис. 7.17. Розподіл осьового поля акустичного тиску одноелементного перетворювача

поширення сигналу в середовищі на відстань $2d$ визначається як $t_V = 2d/V_M$, звідки $d = 0,5V_M t_V$.

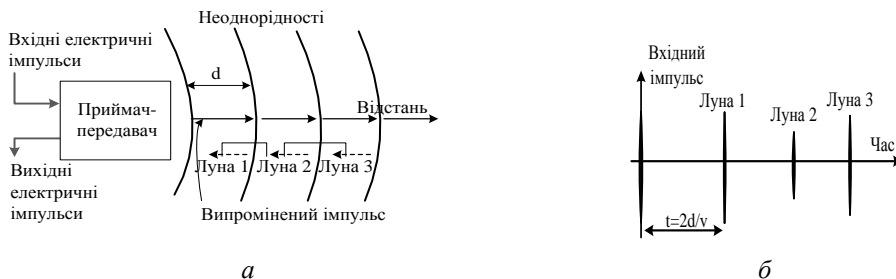


Рис. 7.18. Ілюстрація принципу імпульсної луналакації: *а* – формування ультразвукових сигналів; *б* – часове положення електричних лунаїмпульсів

На рис.7.19 показано систему візуалізації внутрішніх стінок кровносної судини. У центрі артерії знаходиться мініатюрний одноелементний перетворювач, який посилає ультразвуковий імпульс до її стінок і залишається в тому ж положенні до отримання відгуку від неоднорідності середовища. Після цього перетворювач повертається на деякий кут і отримує відгук з іншого напрямку. Процес сканування повторюється до повного обороту перетворювача. Всі відбиті імпульси запам'ятовуються і обробляються комп'ютером, що формує зображення перетину кровносної судини.

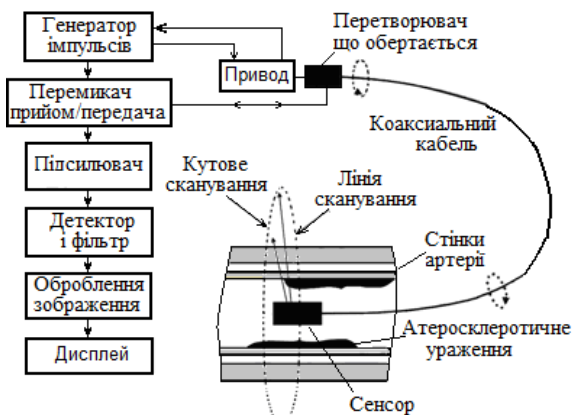


Рис. 7.19. Система візуального дослідження внутрішніх стінок артерії одноелементного перетворювача

Перетворювачі кута променя використовуються в неруйнівному контролі для трансформації поздовжніх ультразвукових хвиль в інші. Такі перетворювачі ґрунтуються на тому явищі, що поздовжня хвиля під час переходу межі двох твердих тіл 1 і 2 за кута падіння θ_{L1} (в межах першого і другого критичних кутів) трансформується в заломлені на поперечну і поздовжню хвилі, що поширюються в твердому тілі 2 (рис. 7.20).

Розділ 7. Перетворювачі ІВС

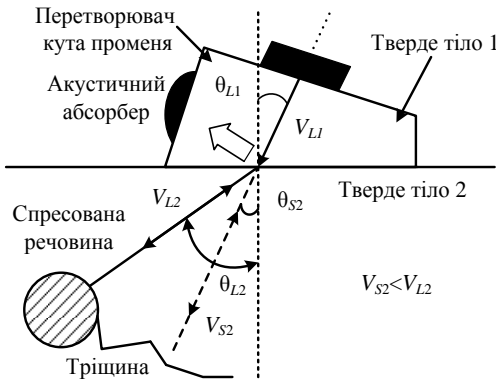


Рис. 7.20. Перетворювач кута променя

Напрямок заломлення хвиль підпорядкований закону Снелла: $\sin \theta_{L1}/V_{L1} = \sin \theta_{L2}/V_{L2} = \sin \theta_{S2}/V_{S2}$. Ці хвилі використовують для виявлення тріщин в зварних конструкцій, інших структурних дефектів.

Відповідно до закону Снелла, якщо збільшується кут θ_{L1} , то θ_{L2} і θ_{S2} також збільшуються. Якщо θ_{L1} сягає критичного значення, θ_{L2} стає рівним 90° і поздовжня хвиля зі швидкістю V_{L2} зникає – у твердому тілі 2 поширюється тільки зсувна хвиля зі швидкістю V_{S2} . Якщо θ_{L1} значно перевищує критичне значення, хвиля зсуву перетворюється в поверхневу хвилю, яка поширюється в поверхневому шарі твердого тіла 2. Цей тип хвиль використовується в неруйнівному контролі для виявлення поверхневих тріщин.

Сенсори на ефекті Доплера. Коли пружна хвиля частотою f_0 відбивається від рухомого із швидкістю V об'єкта, частота відбитого сигналу відмінна від f_0 . Такий зсув частоти називається *ефектом Доплера*.

Частота збільшується у випадку руху об'єкта до спостерігача, і зменшується під час віддалення:

$$f_D(V) = 2f_0 V \cos \theta / C, \quad (7.19)$$

де f_D – зсув частоти Доплера; C – швидкість хвилі в середовищі; θ – кут між векторами швидкості об'єкта і напрямком поширення хвилі.



Вілленброрд Снелл
(Willebrord Snel van Royen,
1580 - 1626),

голландський математик, фізик і астроном. Запропонував використовувати метод подібності трикутників в геодезичних вимірюваннях. У 1621 році відкрив закон заломлення світла. Однак результати численних експериментів з оптики опубліковані не були. Пізніше вони були виявлені в архівах Рене Декартом, і використані в його працях.



**Крістіан Доплер
(Christian Doppler,
1803 -1853),**

австрійський фізик. Наукові праці в області оптики, акустики. Основні праці виконані по аберації світла, теорії мікроскопа і оптичного дальноміра, теорії кольорів. Теоретично обґрунтував залежність частоти коливань, які сприймаються спостерігачем, від швидкості і напрямку руху джерела хвиль відносно спостерігача.

Для вимірювання швидкості потоку рідини на основі ефекту Доплера використовують два методи.

Метод безперервного випромінювання застосовують для визначення швидкості потоку суспензії в трубі (рис. 7.21,а). Передавальний сигнал від перетворювача 1 частково відбивається частинками або бульбашками газу в рідині. Сенсор 2 приймає розсіяний сигнал, який порівнюється за частотою з випроміненим сигналом для визначення f_D . Швидкість V визначається за виразом (7.19).

Другий метод – метод пульсуючої хвилі – використовується в медичній діагностиці для вимірювання швидкості течії крові в невеликому об'ємі або на окремих ділянках коронарної артерії (рис. 7.21,б).

В одній з реалізацій перетворювач мав 0,45 мм в діаметрі і складався з гнучкого з'єднувального кабелю, в наконечнику якого знаходився випромінювач сигналів частотою 12 МГц.

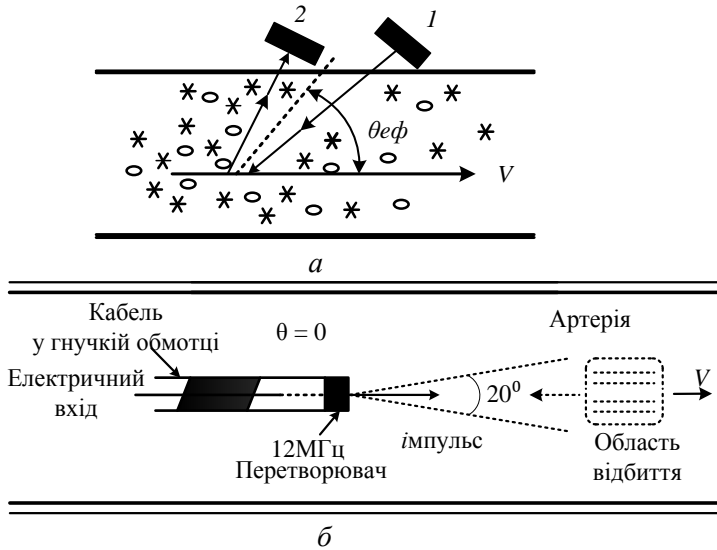


Рис. 7.21. Вимірювання швидкості потоку рідини за ефектом Доплера: а – метод неперервного випромінювання; б – метод пульсуючої хвилі

Перетворювачі переміщення на тунельному ефекті. У 1986 році вчені Герд Біннінг і Генріх Рорер отримали Нобелівську премію з фізики за розробку скануючого тунельного мікроскопа (СТМ). Фізичною основою перетворювача профілю поверхні (ППП) в СТМ є явище електронного тунелю, що представляє потік електронів між двома електропровідними поверхнями під впливом напруги зсуву. Результуючий струм, або *тунельний ефект*, є мірою відстані між цими провідними поверхнями. Рис. 7.22 ілюструє механізм ППП.

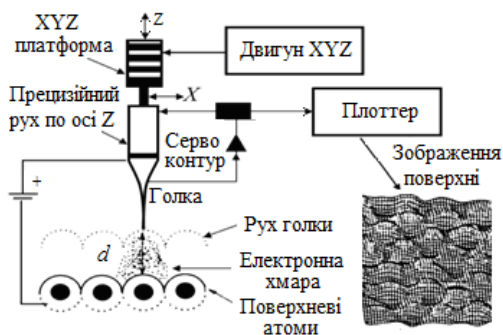


Рис. 7.22. Ілюстрація принципу дії ППП

На практиці один з провідників замінюється надзвичайно гострим наконечником голки, а інший – поверхнею з вимірюваним профілем. Коли хмари електронів, які оточують наконечник і вимірювану поверхню, між якими прикладено напругу зсуву, частково перекриваються, між ними виникає тунельний струм $I_t(d) \equiv e^{-2d/S_0}$, де d – відстань між наконечником і поверхнею.

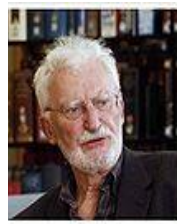
Тунельний струм – надзвичайно чутлива функція відстані d . Наприклад, зміна d на величину одного атомного діаметру, викликає зміну струму в 1000 разів. ППП використовує цю чутливість у СТМ для подання геометрії поверхні матеріалів на атомарному рівні.

Термодинамічні перетворювачі. У термодинамічних системах перенесення тепла від одного елемента до іншого виникає кожен раз, коли між ними є різниця температур. Перенесення тепла



Герд Карл Бінніг (Gerd Karl Binnig, 1947),

німецький фізик, в 1986 році разом зі швейцарським фізиком Генріхом Рорером отримав Нобелівську премію з фізики за винахід скануючого тунельного мікроскопа. Розробки Бінніга розширили можливості нанотехнологій, дозволивши не тільки візуалізувати окремі атоми, але й маніпулювати ними.



Генріх Рорер (Heinrich Rohrer, 1933- 2013)

здійснюється завдяки теплопровідності, конвекції і випромінювання.

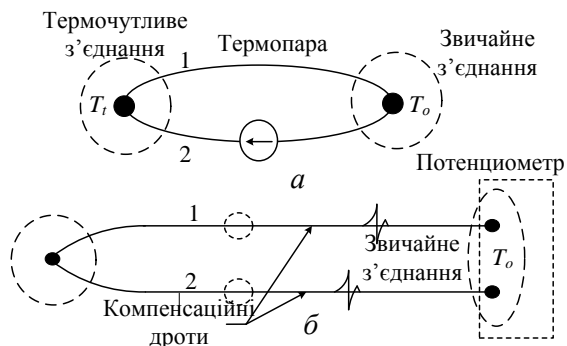
Калібрування платинових термометрів опору, термопар і вимірювання теплових потоків вимагають точного визначення температури. Міжнародний комітет мір і ваг прийняв Міжнародну температурну шкалу ITS-90. Ця шкала ґрунтується на шести реперних точках температурних фазових рівноваг – затвердіння і кипіння, а також на інтерполяційних рівняннях між ними.

Наприклад, в діапазоні температур від $-182,97\text{ }^{\circ}\text{C}$ до $630,5\text{ }^{\circ}\text{C}$ як еталонний інтерполяційний прилад використовують платинові термометри опору. Згідно ITS-90 вони в діапазоні від $0\text{ }^{\circ}\text{C}$ до температури затвердіння алюмінію, калібруються за температурами потрійної точки води ($0,01\text{ }^{\circ}\text{C}$) і точках затвердіння олова ($231,928\text{ }^{\circ}\text{C}$), цинку ($419,527\text{ }^{\circ}\text{C}$) і алюмінію ($660,323\text{ }^{\circ}\text{C}$).

Як одиницю вимірювання використовують кельвін – одиницю термодинамічної шкали. Залежно від того, яка температура приймається за нуль, використовують шкалу Кельвіна T_{90} (абсолютний нуль 0 K) або шкалу Цельсія t_{90} (потрійна точка води $0,01\text{ }^{\circ}\text{C}$). Перерахунок з однієї шкали в іншу виконується як: $t_{90} = T_{90} - 273,15$.

Термопари. Коли два металевих різнорідних провідники 1 і 2 з'єднуються кінцями, які утримують за різних температур T_i і T_o , в такому ланцюзі виникає струм в результаті *ефекту Зеєбека* (рис. 7.23, а). Якщо розірвати такий ланцюг, на вільних кінцях виникне напруга, прямо пропорційна $T_i - T_o$. Коефіцієнт пропорційності називається *коефіцієнтом Зеєбека*. Якщо $T_o = \text{const}$, вимірювана напруга прямо пропорційна вимірюваній температурі T_i .

Зворотний *ефект Пельтьє* полягає в тому, що температура спаю різнорідних металів може підвищуватися або знижуватися в залежності від напрямку струму зовнішнього джерела, включеного в цей контур. Прилади, що використовують такий принцип називаються *елементами Пельтьє*.



Розділ 7. Перетворювачі ІВС

Рис. 7.23. Виникнення термо-е.р.с. за ефектом Зеебека (а) та підключення термопар до потенціометру (б)

На рис. 7.23,б показано підключення термопар до віддаленого вимірювального обладнання. У цьому випадку використовуються з'єднувальні компенсаційні дроти, а еталонний спай формується безпосередньо біля вимірювального обладнання і повинен мати незмінну температуру. Найчастіше компенсація впливу зміни здійснюється автоматично внесенням поправок. Традиційно для вимірювання е.р.с. термопар використовують вольтметр компенсаційного типу (потенціометр). Властивості деяких термопар наведені в табл. 7.3.



Томас Йоганн Зеебек (Thomas Johann Seebeck, 1770 -1831),

німецький фізик. Зробив кілька відкриттів в оптиці, акустиці, вченні про теплоту, але особливо – в області електрики, у 1821 р. відкрив і досліджував термоелектричні струми.

Таблиця 7.3

Тип	Склад	Діапазон, °С	Середовище (оголений дріт)	Напруга, мВ для 500 °С
В	Платина 30 % + родій з платиною 6% + родій	0÷1700	Окислена або інертна; не встановлюється в металеві труби; використовується в склоіндустрії	1,241
Е	Нікель+хром (хромель) з мідь + нікель (константан)	-200÷900	Окислена або інертна; обмежено застосування у вакуумі	36,99
J	Залізо з мідь + нікель	0÷750	Вакуум, інертна; обмежено застосування в окисленому середовищі за високої t°	27,388

Термопары використовують для вимірювання температури об'єктів та середовищ та в системах управління і контролю.

Сенсори для вакууму. Термопары можуть також використовуватися для вимірювання тиску, меншого за атмосферний (вакуум) в діапазоні від 2 до 10^{-3} Торр (1 Торр = 1 мм рт. ст. – позасистемна одиниця, названа на честь Е. Торрічеллі). Сенсор складається з нитки розжарювання, по якій протікає постійний струм від джерела струму, і термопары, що знаходиться в тепловому контакті з ниткою для вимірювання її температури. Сенсор розміщують у

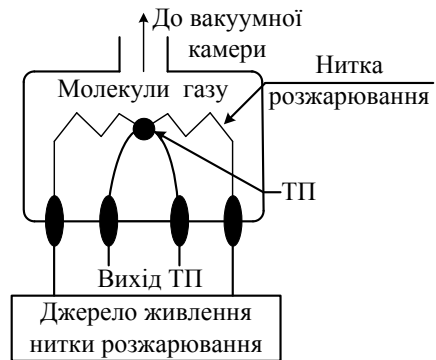


Рис. 7.24. Термоелектричний сенсор вакууму

вимірювальну камеру, з'єднану з ємністю в якій вимірюють тиск (рис.7.24). Принцип перетворення базується на залежності теплопровідності газу від тиску за малих тисків, коли вільний пробіг молекул газу є порівнюваним з розмірами нитки розжарювання. Вихідна напруга термопары пропорційна вимірюваному тиску. Для збільшення чутливості термопары з'єднують послідовно в термобатарей.

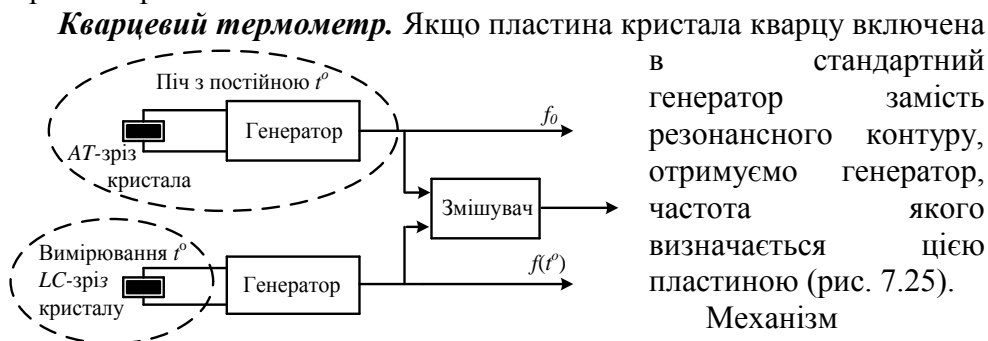


Рис. 7.25. Кварцевий термометр

Механізм перетворення таких термометрів базується на залежності пружних властивостей кварцової пластини від температури. Частота генератора визначається многочленом третього ступеня від температури t^0 з точкою калібрування $t_k^0 = 25^{\circ}C$

$$f(t^0) = f_{25} \left[1 + A(t^0 - 25) + B(t^0 - 25)^2 + C(t^0 - 25)^3 \right], \quad (7.20)$$

де A, B, C – коефіцієнти; f_{25} – частота генератора для $t_k = 25^{\circ}C$.

Хаммонд виявив зріз з найбільшим значенням A і $B = C = 0$ й позначив його як зріз з лінійним коефіцієнтом (LC). LC -пластини використовують як чутливі елементи в кварцових термометрах.

Вихідні сигнали двох генераторів змішуються. Різниця частоти сигналу змішувача лінійно залежить від t^0 . Кварцовий термометр вимірює температуру в діапазоні від -80 до $250^{\circ}C$ з похибкою $\pm 0,075^{\circ}C$. Чутливість складає $1000 \text{ Гц}/^{\circ}C$, що відповідає $0,0001^{\circ}C$.

Фотометричні сенсори використовуються для вимірювання інтенсивності світла шляхом визначення квантів світла або фотонів. У цих пристроях енергія фотонів перетворюється в пропорційний електричний сигнал за допомогою низки детекторів: *фотоємисійних детекторів*, в яких енергія фотона вибиває електрон з металеві поверхні (зовнішній фотоэффект), розміщеної в вакуумному або газонаповненому балоні; *фоторезистивних детекторів*, в яких вхідні фотонна енергія створює електронно-діркові пари (внутрішній фотоэффект), які змінюють провідність або опір напівпровідника,

Розділ 7. Перетворювачі ІВС

збільшуючи число вільних носіїв заряду; *фотодіодних детекторів*, в переходах яких вхідні фотони також утворюють електронно-діркові пари, що створюють додаткові умови для зміни електропровідності. Такі детектори можуть використовуватися в фоторезистивному (пасивному) або фотогальванічному (активному) режимах. Енергія фотона, виражена в електрон-вольтах, становить:

$$E_{ph} = (4,13 \cdot 10^{-15}) f . \quad (7.21)$$

Енергія E_{ph} пропорційна частоті світла f і обернено пропорційна довжині світлової хвилі λ . Існує межа для λ , вище якої детектор буде закритий, оскільки E_{ph} менше за енергію виходу електрона у фотоелектричному детекторі і менше за енергію подолання потенційного бар'єру в фотодіодному детекторі.

Фотоемісійні або фотоелектричні перетворювачі діють на основі зовнішнього фото ефекту, який полягає у тому, що електрони залишають металеву поверхню після поглинання енергії фотона. Вся енергія фотона передається одному електрону в металі, який отримавши енергію виходу, відділяється від металеві поверхні з кінетичною енергією, значення якої для фото ефекту визначається за рівнянням Ейнштейна:

$$mv^2/2 = E_{ph} - E_w , \quad (7.22)$$

де v – швидкість електрона що виходить; m – маса електрона; E_w – енергія виходу, необхідна електрону на рівні Фермі щоби полишити поверхню металу з найбільшою кінетичною енергією.

Фотоемісійні детектор – основний компонент вакуумних або газонаповнених фотоелементів або фотопомножувачів. Фотоелемент складається з напівкруглого фотоемісійного катода, концентричного з центральним стрижневим анодом, розміщеними в спеціальному скляному балоні. Падаюче на фотокатод світло (пучок фотонів) створює потік електронів, які залишають поверхню фотокатода і спрямовуються до анода під дією зовнішньої електричної напруги. Анодний струм пропорційний числу поглинених катодом фотонів.

Фотопомножувальна трубка діє аналогічно фотоелементу, але має можливість для реалізації великих коефіцієнтів підсилення струму. Спрощена схема фотопомножувача

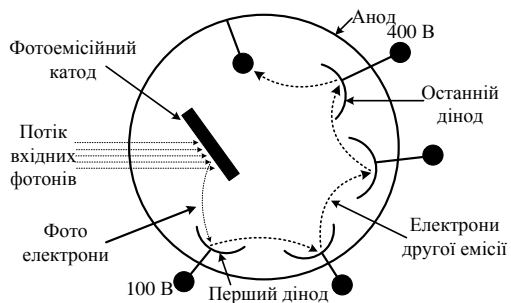


Рис. 7.26. Фотопомножувальна трубка

зображена на рис. 7.26. Фотоелектрони генеруються так само, як у фотоелементі. Вони фокусуються на першому елементі ланцюга фотопомножувача, що зветься *дінодом*. Фотоелектрони від першого дінотронного електрода направляються до наступного, який має більший позитивний потенціал. Електрони прискорюються між дінотронними електродами і генерують більше електронів. Це явище відоме як *вторинна емісія*.

Фоторезистивний сенсор використовує ефект фотопровідності, який полягає в збільшенні електропровідності напівпровідника під впливом електромагнітного випромінювання. Відкритий у 1873 р. У. Смітом (США). На рис. 7.27,а схематично показано конструкцію фоторезистивного сенсора, який складається із зигзагоподібної доріжки довжиною L і шириною W , що виконана з тонкої плівки напівпровідника на керамічній основі. По обидва боки доріжки знаходяться металізовані області під'єднувальних електричних контактів. Електричний опір доріжки за будь-якого освітлення:

$$R_0 = \rho_R W / L, \quad (7.23)$$

де ρ_R – поверхневий опір напівпровідникового матеріалу.

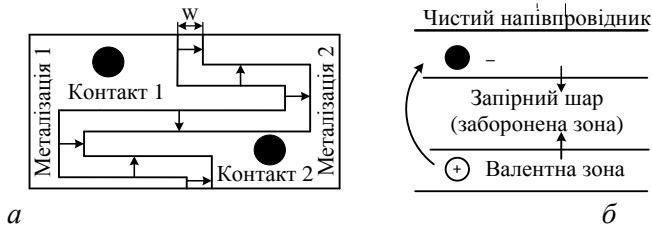


Рис. 7.27. Фотопровідний сенсор (а) і діаграма енергетичних рівнів для чистого напівпровідника (б)

Якщо створити різницю потенціалів V між електричними контактами сенсора, виникне струм $I_0 = VL / \rho_R W = \gamma_C VL / W$. Величина $1 / \rho_R = \gamma_C$ визначається як *поверхнева провідність*. Саме γ_C модулюється світловим потоком і викликає зміни опору та струму.

Механізм перетворення пояснює діаграма рівнів енергії для чистого (без домішок) напівпровідника (7.27,б). Останній поглинає фотон, коли його енергія ледь перевищує енергію забороненої зони, і одночасно створює пару «електрон – дірка». Електрон робить єдиний перехід через заборонену зону в зону провідності. Цей електрон і відповідна дірка в зоні валентності сприяють збільшенню повного струму. Як наслідок, зі збільшенням потоку фотонів електропровідність зростає, а опір фотодетектора падає. В напівпровідниках з домішками за рахунок останніх утворюються іони донори і акцептори, що додатково підвищують концентрацію носіїв заряду в зоні провідності і у валентній зоні. Завдяки цьому провідність зростає ще більше. Домішки збільшують квантову чутливість і граничну довжину хвилі

Розділ 7. Перетворювачі ІВС

фотодетектора. Такі сенсори виробляють з сульфиду свинцю (PbS) або селеніду свинцю (PbSe), і використовують в інфрачервоному діапазоні з довжиною хвиль від 1 до 3 мкм і від 1 до 6 мкм відповідно.

Акустооптичні сенсори (АОС) переміщення використовують для виявлення вібрацій шляхом перетворення їх у фазомодульований електричний сигнал. Механізм перетворення проявляється у взаємодії оптичної хвилі з акустичною та утворенні нової оптичної хвилі. На рис. 7.28 схематично показано пристрій АОС, відомий як комірка Брегга. Акустооптичне перетворення відбувається в пластині з оптично якісного скла.

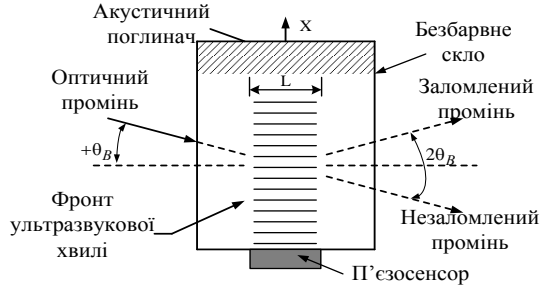


Рис. 7.28. Комірка Брегга

Сенсор має оптичні вікна на двох протилежних сторонах пластини для введення і виведення оптичного променя. Високочастотний ультразвуковий перетворювач прикріплений до однієї з двох інших сторін, збуджує в оптичній пластині поздовжні хвилі і створює області стиснення і розтягування, в яких коефіцієнт заломлення змінюється. Ці області еквівалентні тривимірній дифракційній решітці.

Генерований лазером оптичний промінь входить в комірку Брегга під кутом θ , проходячи відстань довжиною L крізь хвилі періодичних змін о коефіцієнта заломлення з частотою збудження ω_m . Оптичний промінь на виході комірки модулюється за фазою і складається з оптичної несучої частотою ω_c і пар бічних оптичних смуг з частотами $\omega_c \pm N\omega_m$,

$$(7.25)$$

де N – коефіцієнт, що відповідає порядку бічних смуг.

Є особливе значення кута $\theta = \theta_B$, що відповідає оптичній довжині хвилі λ і акустичній довжині хвилі Λ , коли тільки одна бічна смуга першого порядку зростає за рахунок підсилюючої дії інтерференції, в той час як інші бічні смуги вищих порядків зменшуються завдяки загасаючої дії інтерференції. Це явище відоме як *ефект Брегга*, а кут θ_B – як *кут Брегга*, який визначається з дифракційного рівняння Брегга $\sin \theta_B = \lambda/2\Lambda = \lambda\omega_m/4\pi v_m$,

$$(7.26)$$

де v_m – швидкість поширення ультразвукової хвилі. Кут між дифракційним променем і незаломленим променем дорівнює $2\theta_B$.

На рис. 7.29 показано систему віброметрії з коміркою Брегга. У наведеній системі як джерело оптичного променя використано He – Ne лазер. Комірка Брегга використовується для вимірювання поверхневих вібрацій з

частотою ω_v . У фотодіоді відбувається де-модуляція оптичних сигналів – вихідний струм фотодіода має складову з частотою $2\omega_m + \omega_v$. Приймач/передавач метрового діапазона вимірює амплітуду бічної смуги, пропорційну амплітуді вимірюваних вібрацій. Мінімальна амплітуда вимірюваного зсуву дорівнює $2 \cdot 10^{-12}$ м і відповідає сигналу на вході фотодіода ~ 300 мкВт.

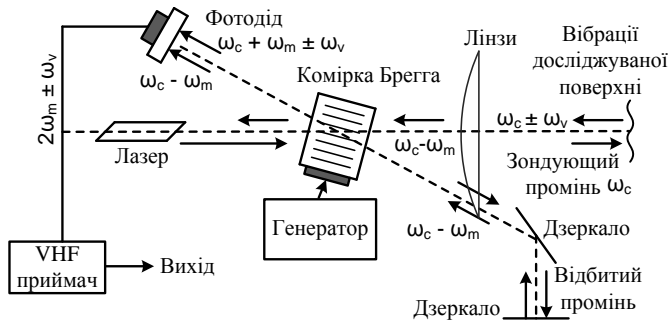


Рис. 7.29. Система бесконтактного вимірювання вібрацій з коміркою Брегга

Волоконнооптичні сенсори поділяються на зовнішні і внутрішні. У зовнішніх процес перетворення має місце зовні оптичного волокна, яке відіграє пасивну роль – використовується лише як провідник світла. У сенсорах внутрішнього типу перетворення має місце в оптичному волокні. У цьому випадку вимірювана величина модулює один з його параметрів, наприклад показник заломлення. Волокно, у свою чергу, модулює світло, що поширюється в його серцевині. Оптичне волокно в таких перетворювачах, складається з серцевини і оболонки, виконаної також зі скла але з іншим показником заломлення. Промінь світла у волокні поширюється переважно в серцевині завдяки повному внутрішньому віддзеркаленню.



Рис. 7.30. Волоконнооптичний сенсор нерівностей

Прикладом використання оптоволокна для вимірювання фізичних величин є перетворювач мікрозгинів (рис 7.30). Волокно деформується під впливом вимірюваної величини – тиску, сили, переміщення, натягу, прискорення, температури.

Механізм перетворення полягає у зменшенні інтенсивності променя світла на виході оптоволокна через його деформацію, що призводить до втрати енергії променя. На виході волокна оптичний сигнал перетворюється фотодіодом в електричний. На рис. 7.31 показаний сенсор тиску (а) і температури (б), відомий як *інтерферометр Фабрі – Перо*.

Розділ 7. Перетворювачі ІВС

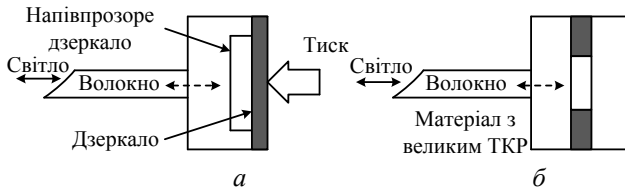


Рис. 7.31. Волоконнооптичний сенсор: *a* – тиску; *б* – температури

Будь-яка зміна інтервалу між дзеркалами, викликана зміною вимірюваної величини, призводить до зміни коефіцієнта перетворення. На виході волокна оптичний сигнал перетворюється фотодетектором в електричний. Такі сенсори мають невеликі розміри і виконуються методами інтегральної технології.

Оптичні волокна використовуються також для вимірювання параметрів магнітних і акустичних полів. У сенсорії магнітного поля, показаному на рис. 7.32, механізм трансформації складається з поздовжнього натягу в магнітострикційному матеріалі, викликаному магнітним полем, і відповідної деформації в серцевині оптоволокна. Ця деформація модулює показник заломлення серцевини і вносить залежний від поля фазовий зсув у вихідний оптичний промінь. Перетворювач акустичного поля також може бути виконаний уздовж оптоволоконних ліній, замінюючи магнітострикційну основу акустичночутливим покриттям навколо оболонки оптоволокна. В цьому випадку акустичне поле викликає відповідні фазові зміщення оптичного променя.

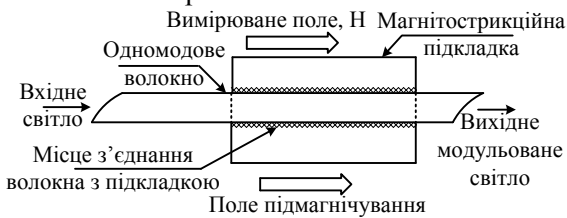


Рис. 7.32. Волоконнооптичний сенсор магнітного поля

Флуоресцентні сенсори хімічних величин.

Знання величин парціального тиску кисню (PO_2), рН і парціального тиску вуглекислоти (PCO_2) в артеріальній крові важливе в лікуванні критично хворих пацієнтів. Сенсор цих величин уявляє собою



Жан-Батіст Альфред Перо (Jean-Baptiste Alfred Perot, 1863 - 1925),

французький фізик. Дослідження в області електрики, тепла, спектроскопії, оптики. Вивчав приховану теплоту випаровування води, вимірював довжину світлової хвилі. Спільно з Ш. Фабрі розвинув новий метод оптичної інтерферометрії і побудував інтерферометр, названий їх іменами – інтерферометр Фабрі – Перо.

три оптичних волокна з доданими до їх наконечників чутливими елементами зі специфічними хімічними реактивами для кожної з вимірюваних величин. Волокна розміщуються в єдиному катетері, який вводиться в артерію через надріз для безперервного моніторингу кров'яних газів і рН.

Випромінювання лазерного діода певної довжини хвилі через оптоволокну потрапляє на чутливий елемент, розташований на периферійному кінці волокна всередині артерії (рис. 7.33).

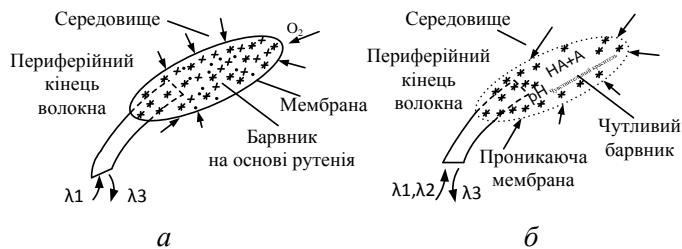


Рис.7.33. Волоконнооптичні флуоресцентні сенсори: *а* – кисню; *б* – рН

Чутливий елемент містить барвник з флуоресцентними молекулами, які абсорбують фотони і переходять у високоенергетичний стан (збуджуються). У цьому стані молекула знаходиться короткий час (близько 10 нс), після чого втрачає частину енергії, переходячи до основного стану. Фотони, що утворюються при зворотному переході, мають меншу енергію, ніж енергія збуджуючих фотонів, тому довжина хвилі флуоресцентного випромінювання більше довжини хвилі збуджуючого випромінювання (правило Стокса), оскільки енергія фотона обернено пропорційна довжині хвилі. Присутність вимірюваних газів змінює інтенсивність флуоресцентного випромінювання.

Інтенсивність може бути виміряна за допомогою спектральночутливих пристроїв і фотометричних детекторів, вихідний сигнал яких пропорційний концентрації вимірюваної речовини. Кисень, наприклад, зменшує інтенсивність флуоресцентного випромінювання. Внутрішньосудинний сенсор парціального тиску кисню в крові (рис. 7.33,*а*) заснований на ефекті поглинання флуоресцентного випромінювання. Співвідношення між PO_2 і флуорісцентними характеристиками молекул барвника представляється у вигляді рівняння Стерна – Волмера:

(7.27)

$$PO_2 = (I_0/I - 1)K,$$

де I , I_0 – інтенсивності флуоресцентного випромінювання відповідно за присутності та відсутності кисню; K – константа, що визначається нахилом функції $PO_2(I_0/I)$.

Аналогічно створюються та діють і інші сенсори цього типу.

7.3. Перетворювачі на операційних підсилювачах

та їх похибки

Схемотехніка вимірювальних перетворювачів передбачає широке використання універсальних і оригінальних засобів аналогового перетворення сигналів, виконаних на основі операційних підсилювачів.

Операційний підсилювач (ОП) – це диференціальний підсилювач напруги з безпосередніми зв'язками, коефіцієнтом підсилення понад 1000 і можливістю формування частотної характеристики введенням негативного зворотного зв'язку (НЗЗ). Для отримання розрахункових значень коефіцієнтів підсилення схем використовують модель ідеального ОП, еквівалентна схема якого, його підключення до джерел вхідних сигналів $U_{вх1}$, $U_{вх2}$ і напруг живлення $\pm E_{ж}$ показані на рис. 7.34.

ОП формує вихідну напругу за допомогою керованого джерела напруги $e_{вих}$, вихід якого пропорційний різниці напруг $e_{д} = U_{вх2} - U_{вх1}$ на входах ОП. Позитивне напруга $U_{вх1}$, що надходить на інвертуючий

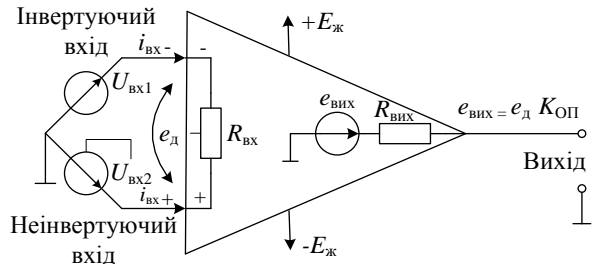


Рис. 7.34. Еквівалентна схема ідеального ОП

вхід ОП, дає негативний сигнал на виході ОП (за біполярного живлення), в той час як позитивний вхідний сигнал $U_{вх2}$ на неінвертуючому вході ОП створює на виході позитивний сигнал. Коефіцієнт підсилення ОП за напругою визначається як $K_{ОП} = e_{вих} / e_{д}$. Ідеальний ОП має такі параметри: вхідний опір $R_{вх} \rightarrow \infty$; вихідний опір $R_{вих} = 0$; коефіцієнт підсилення напруги $K_{ОП} \rightarrow \infty$; смуга робочих частот (за рівнем -3 дБ) $\Delta f_{-3дБ} \rightarrow \infty$; нульове початкове зміщення нуля ОП. Ці характеристики дають змогу сформулювати три аксіоми теорії ідеального ОП: диференціальна вхідна напруга ОП в лінійному режимі завжди дорівнює нулю, тобто $e_{д} = 0$; вхідні струми ОП дорівнюють нулю – $i_{вх}^+ = i_{вх}^- = 0$; за замкненого НЗЗ маємо $U_{вх1} = U_{вх2}$, оскільки початкова напругу зміщення нуля відсутня.

Все розмаїття схем на ОП зводиться до шести базових схем: інвертуючого та неінвертуючого підсилювачів; диференціального підсилювача; диференціатора; інтегратора; інвертуючого суматора.

Інвертуючий підсилювач (рис. 7.35,а). У цій схемі неінвертуючий вхід ОП з'єднаний із загальною шиною, сигнал надходить на інвертуючий вхід через резистор R_1 , а НЗЗ забезпечується резистором R_2 . З огляду на те, що $e_{\text{Д}}=0$ і $U_{\text{ВХ}2}=0$ маємо $U_{\text{ВХ}1}=0$. Оскільки $i_{\text{ВХ}}^+ = i_{\text{ВХ}}^- = 0$, маємо $i_1 = i_2$. З урахуванням цього струм в резисторі R_1 дорівнює $i_1 = U_{\text{ВХ}}/R_1$, а в резисторі R_2 $-i_2 = -U_{\text{ВИХ}}/R_2$ (знак мінус з'являється в результаті інвертування вхідного сигналу ОП). На рис. 7.35, б зображені діаграми вхідної і вихідної гармонічних напруг.

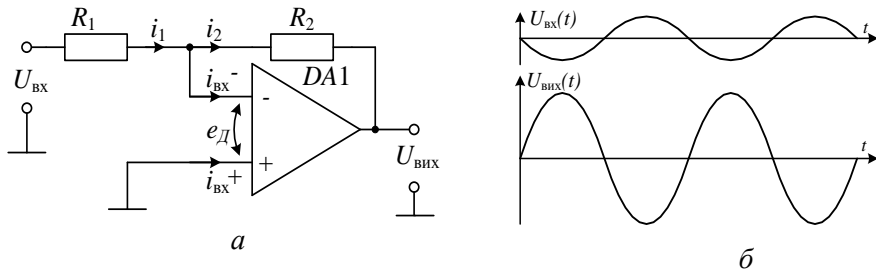


Рис. 7.35. Інвертуючий підсилювач (а) і часові діаграми сигналів (б)

Прирівнюючи праві частини виразів для струмів i_1 і i_2 , отримаємо вираз для вихідної напруги інвертуючого підсилювача

$$U_{\text{ВИХ}} = -R_2 U_{\text{ВХ}} / R_1. \quad (7.28)$$

Характеристики ідеального інвертуючого підсилювача на ОП: коефіцієнт підсилення за напругою $K_U = -R_2/R_1$; вхідний опір $R_{\text{ВХ}} = R_1$; вихідний опір $R_{\text{ВИХ}} = 0$.

Неінвертуючий підсилювач (рис.7.36,а).

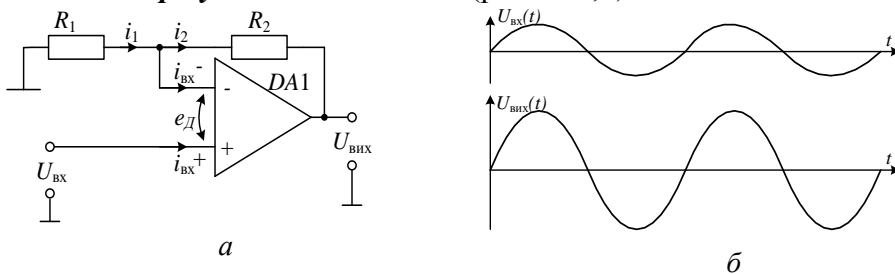


Рис. 7.36. Неінвертуючий підсилювач (а) і часові діаграми сигналів (б)

Вхідна напруга надходить на неінвертуючий вхід ОП, а частина вихідної через резистивний дільник напруги R_1, R_2 подається на інвертуючий вхід ОП. З огляду на те, що $i_{\text{ВХ}}^+ = i_{\text{ВХ}}^- = 0$, маємо $i_1 = i_2$. Відповідно до третьої

Розділ 7. Перетворювачі ІВС

аксіоми ідеального ОП напруга на інвертуючому вході ОП дорівнює $U_{\text{вх}}$, тому струм в резисторі R_1 дорівнює $i_1 = U_{\text{вх}}/R_1$, а вихідна напруга – $U_{\text{вих}} = i_2 R_2 + U_{\text{вх}}$. З огляду на, що $i_1 = i_2$, вихідна напруга неінвертуючого підсилювача:

$$(7.29)$$

$$U_{\text{вих}} = (R_1 + R_2)U_{\text{вх}}/R_1.$$

Характеристики ідеального неінвертуючого підсилювача: $K_U = 1 + R_2/R_1$;

$R_{\text{вх}} \rightarrow \infty$; $R_{\text{вих}} = 0$. На рис. 7.36,б зображені діаграми напруг неінвертуючий підсилювача в разі підключення на його вхід гармонічного сигналу.

Диференціальний підсилювач (рис. 7.37,а) є комбінацією двох попередніх схем. Його діаграми напруг подані на рис. 7.37,б.

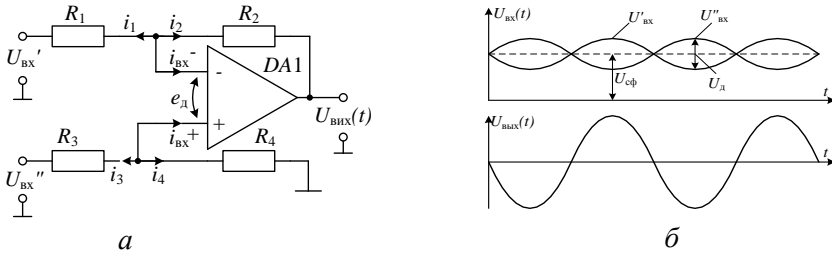


Рис. 7.37. Диференціальний підсилювач (а) і часові діаграми сигналів (б)

Вхідні гармонічні сигнали $U'_{\text{вх}}$ і $U''_{\text{вх}}$ можна уявити синфазною $U_{\text{сф}} = 0,5 \cdot (U'_{\text{вх}} + U''_{\text{вх}})$ і диференціальною $U_{\text{д}} = U'_{\text{вх}} - U''_{\text{вх}}$ складовими. Тоді, $U'_{\text{вх}} = U_{\text{сф}} + 0,5U_{\text{д}}$, а $U''_{\text{вх}} = U_{\text{сф}} - 0,5U_{\text{д}}$. Диференціальний підсилювач отримав назву завдяки тому, що він підсилює диференціальну складову вхідних напруг ОП і пригнічує синфазну складову. У загальному вигді його вихідна напруга:

$$(7.30)$$

$$U_{\text{вих}} = K_{\text{д}}U_{\text{д}} + K_{\text{сф}}U_{\text{сф}},$$

де $K_{\text{д}}$ – коефіцієнт підсилення диференціального сигналу; $K_{\text{сф}}$ – коефіцієнт передачі синфазного сигналу. Визначимо за яких умов матимемо $K_{\text{сф}} = 0$. З урахуванням того, що $e_{\text{д}} = 0$, $i_{\text{вх}}^+ = i_{\text{вх}}^- = 0$, маємо $i_1 = i_2$, $i_3 = i_4$ і, використавши метод накладання, подамо $U_{\text{вих}}$ через $U'_{\text{вх}}$ і $U''_{\text{вх}}$

$$U_{\text{вих}} = U''_{\text{вх}} R_4 / (R_3 + R_4) \cdot (R_1 + R_2) / R_1 - U'_{\text{вх}} R_2 / R_1. \quad (7.31)$$

Позначивши $R_4 / (R_3 + R_4) \cdot (R_1 + R_2) / R_1 = A_2$, $-R_2 / R_1 = A_1$ і взявши до уваги, що $U_{\text{сф}} = (U'_{\text{вх}} + U''_{\text{вх}}) / 2$, а $U_{\text{д}} = U'_{\text{вх}} - U''_{\text{вх}}$, отримаємо

$$U_{\text{вих}} = U''_{\text{вх}} A_2 + U'_{\text{вх}} A_1 = U_{\text{сф}} (A_1 + A_2) + 0,5 \cdot U_{\text{д}} (A_1 - A_2).$$

Співставляючи отриманий вираз з (7.30), побачимо, що $K_{\text{сф}} = A_1 + A_2$, а $K_{\text{д}} = (A_1 - A_2)/2$, отже остаточно отримаємо:

$$K_{\text{сф}} = (R_1 R_4 - R_2 R_3) / [R_1 (R_3 + R_4)]; \quad (7.32)$$

$$K_{\text{д}} = -(R_2 R_3 + 2R_4 R_2 + R_1 R_4) / [2R_1 (R_3 + R_4)]. \quad (7.33)$$

Якщо $R_2/R_1 = R_4/R_3$ маємо $K_{\text{сф}} = 0$ і $K_{\text{д}} = -R_2/R_1$.

Характеристики ідеального диференціального підсилювача: коефіцієнт підсилення диференціального сигналу $K_{\text{д}} = -R_2/R_1$; коефіцієнт підсилення синфазного сигналу $K_{\text{сф}} = 0$; вхідний опір інвертуючого входу $R_{\text{вх}}^- = R_1$; вхідний опір неінвертуючого входу $R_{\text{вх}}^+ = R_3 + R_4$; вхідні опори для диференціального і синфазного сигналів – $R_{\text{вх.д}} = R_1 + R_3$ і $R_{\text{вх.сф}} = (R_1 + R_2)(R_3 + R_4) / (R_1 + R_2 + R_3 + R_4)$; вихідний опір $R_{\text{вих}} = 0$.

Інвертуючий сумматор (рис. 7.38), використовує властивість інвертуючого входу ОП – рівність потенціалів інвертуючого входу

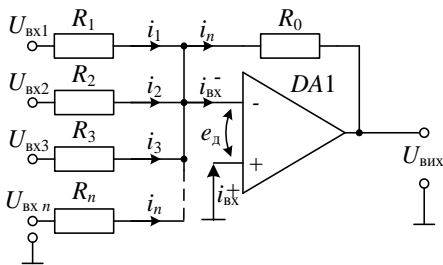


Рис. 7.38. Інвертуючий сумматор

потенціалу другого входу, з'єднаного з загальною шиною. Це дає змогу застосувати схему для підсумовування струмів. З огляду на те, що, $e_{\text{д}} = 0$, $i_{\text{вх}}^+ = i_{\text{вх}}^- = 0$, мають місце наступні представлення вхідних струмів:

$$i_i = U_{\text{вх } i} / R_i \Big|_{i=1,2,\dots,n}, \quad i_0 = \sum_1^n i_i.$$

Беручи до уваги, що $i_0 = -U_{\text{вих}}/R_0$, маємо:

$$U_{\text{вих}} = -(U_{\text{вх } 1} R_0 / R_1 + U_{\text{вх } 2} R_0 / R_2 + \dots + U_{\text{вх } n} R_0 / R_n) = -\sum_1^n k_i \cdot U_{\text{вх } i},$$

де $k_i = R_0/R_i \Big|_{i=1,\dots,n}$ – коефіцієнт передачі напруги за i -м входом.

Перевагою інвертуючого сумматора у порівнянні з пасивним резистивним сумматором є взаємна незалежність коефіцієнтів k_i .

Характеристики ідеального інвертуючого сумматора: кількість вхідних сигналів не обмежена; коефіцієнт підсилення i -ї вхідної напруги $k_i = R_0/R_i$; вхідний опір i -го входу $R_{\text{вх } i} = R_i$; вихідний опір $R_{\text{вих}} = 0$.

Інтегратор (рис. 7.39,а) є модифікацією інвертуючого підсилювача з конденсатором в контурі НЗЗ замість резистора. З огляду на те, що $e_{\text{д}} = 0$,

Розділ 7. Перетворювачі ІВС

$i_{\text{вх}}^+ = i_{\text{вх}}^- = 0$, маємо $i_1 = i_2$. Підставляючи в це рівняння значення струмів $i_1(t) = U_{\text{вх}}(t)/R$ і $U_C(t) = U_{\text{вих}}(t)$ отримаємо $U_{\text{вх}}(t)/R = -C dU_{\text{вих}}/dt$. (7.34)

Інтегруючи ліву і праву частини рівняння (7.34) за часом, маємо

$$U_{\text{вих}}(t) = -\frac{1}{RC} \int_0^t U_{\text{вх}}(\tau) d\tau. \quad (7.35)$$

На рис. 7.39,б показана реакція інтегратора на стрибок вхідної напруги. На лінійній ділянці вихідна напруга визначається як

$$U_{\text{вих}}(t) = -U_{\text{вх}}(t)t/RC. \quad (7.36)$$

Якщо на вхід інтегратора приходять синусоїдальна напруга $U_{\text{вх}}(t) = U_a \sin \omega t$, де U_a – амплітудне значення напруги, ω – кругова частота, то напруга на його виході: $U_{\text{вих}}(t) = U_a \cos \omega t / \omega RC$. (7.37)

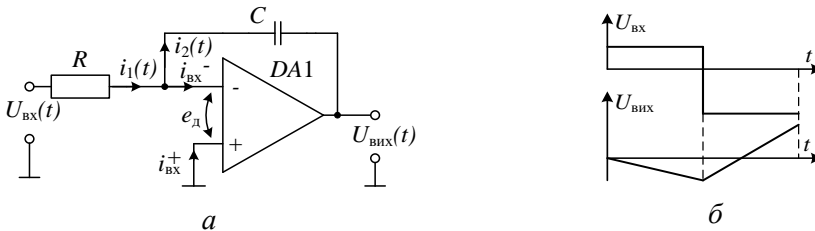


Рис. 7.39. Інтегратор (а) та часові діаграми сигналів (б)

З (7.37) випливає, що зі збільшенням частоти модуль коефіцієнта передачі ідеального інтегратора збігає до нуля, а фазовий зсув відносно вхідного сигналу на будь-якій частоті дорівнює 90° . Діаграми Бode для ідеального інтегратора зображені на рис. 7.40, де а, б – відповідно амплітудно-частотна фазочастотна характеристики (АЧХ і ФЧХ).

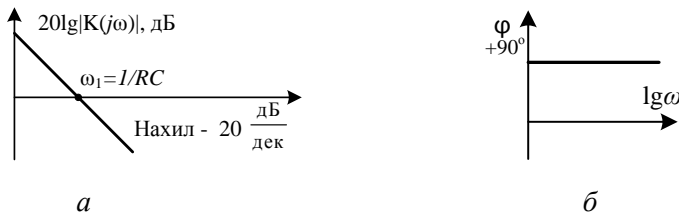


Рис. 7.40. Частотні характеристики інтегратора: а – АЧХ, б – ФЧХ

Характеристики ідеального інтегратора: залежність $U_{\text{вих}}(t) = F(U_{\text{вх}}(t))$ задається рівнянням (7.35); $R_{\text{вх}} = R$; $R_{\text{вих}} = 0$.

Диференціатор. Базова схема диференціатора на ОП зображена на рис. 7.41,а, а його часові діаграми, що пояснюють його роботу – на рис. 7.41,б. Він також є різновидом інвертуючого підсилювача, в якому

замість резистора R_1 включений конденсатор C .

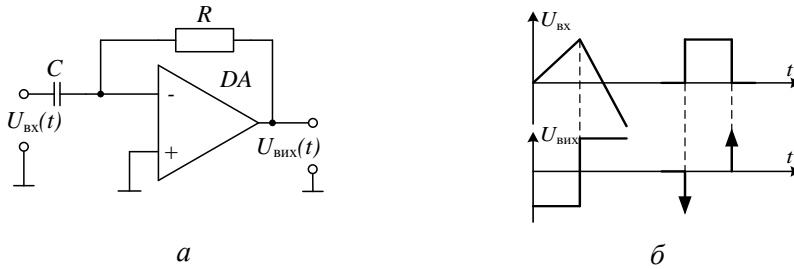


Рис. 7.41. Диференціатор (а) і часові діаграми його сигналів (б)

Застосувавши аксіоми ідеального ОП ($e_{\text{д}} = 0$, $i_{\text{вх}}^+ = i_{\text{вх}}^- = 0$), маємо $i_1 = i_2$. З огляду на те, що $i_1(t) = C dU_{\text{вх}}(t)/dt$ (оскільки напруга на конденсаторі дорівнює $U_{\text{вх}}$) і $i_2(t) = -U_{\text{вих}}(t)/R$, отримаємо рівняння для вихідної напруги

$$U_{\text{вих}}(t) = -RC dU_{\text{вх}}(t)/dt. \quad (7.38)$$

На рис. 7.41,б показані графіки вихідної напруги для лінійно імпульсно змінюваних входних сигналів. Для синусоїдної входної напруги $U_{\text{вх}}(t) = U_a \sin \omega t$ вихідна напруга: $U_{\text{вих}} = -RC\omega U_a \cos \omega t$. (7.39)

З рівняння (7.39) випливає, що модуль коефіцієнта передачі за напругою ідеального диференціатора збільшується разом зі збільшенням частоти сигналу. Зсув фази між входною і вихідною напругами дорівнює мінус 90° на будь-якій частоті. Діаграми Боде для ідеального диференціатора показані на рис. 7.42.

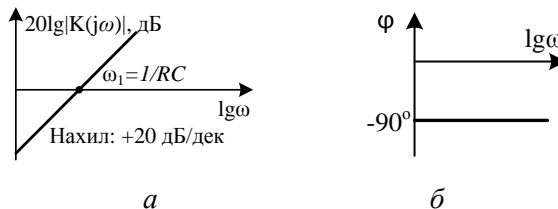


Рис. 7.42. Частотні характеристики диференціатора: а – АЧХ, б – ФЧХ

Характеристики ідеального диференціатора: $U_{\text{вих}}(t) = F(U_{\text{вх}}(t))$ задається рівнянням (7.38); входна ємність $C_{\text{вх}} = C$; $R_{\text{вих}} = 0$.

Реальні ОП відрізняються від ідеальних своїми параметрами та характеристиками, що є причиною виникнення статичних, динамічних і випадкових похибок перетворення сигналів.

Адитивні похибки. Розрахунок адитивних складових похибок схем на ОП проводять в масштабі вихідної величини. Це пояснюється

Розділ 7. Перетворювачі ІВС

тим, що коефіцієнт підсилення напруги зсуву не залежить від способів включення ОП (на відміну від коефіцієнта підсилення корисного сигналу). Коефіцієнт підсилення напруги зсуву називається *коефіцієнтом шуму* і визначається як величина обернена коефіцієнту β передачі ланцюга НЗЗ, тобто β^{-1} . Наприклад, для схеми інвертуючого підсилювача (рис. 7.35, а) маємо $\beta = R_1 / (R_2 + R_1)$.

$$\text{Абсолютна адитивна похибка: } \Delta U_{\text{вих}} = \beta^{-1} \sum_{i=1}^5 e_{\text{зм}i}, \quad (7.40)$$

де $e_{\text{зм}i}$ – складові напруги зміщення нуля реального ОП.

$$\text{Відносна адитивна похибка: } \delta = \Delta U_{\text{вих}} / U_{\text{вих}}, \quad (7.41)$$

$$\text{а зведена адитивна похибка – } \delta_{\text{зв}} = \Delta U_{\text{вих}} / U_{\text{вих ном}}, \quad (7.42)$$

де $U_{\text{вих ном}}$ – номінальна вихідна напруга (зазвичай $U_{\text{вих ном}} = 10$ В).

Сумарна напруга зсуву ОП має п'ять складових, що визначаються: початковою напругою зміщення і температурним дрейфом; струмами зміщення; неоптимальним послабленням синфазного сигналу; пульсаціями в ланцюгах живлення; шумами вхідних ланцюгів ОП.

Початкова напруга зсуву і температурний дрейф. Ця складова визначається як сума початкової напруги зміщення нуля ОП $e_{\text{зм}0}$ і температурного дрейфум напруги зсуву $e_{\text{др}}$:

$$e_{\text{зм}1} = e_{\text{зм}0} + e_{\text{др}}. \quad (7.43)$$

Складова $e_{\text{зм}0}$ враховує технологічне зміщення нуля і компенсується за допомогою або змінного опору, підключеного до коригувальних виходів ОП, або спеціальними схемами корекції. Ця компонента похибки трактується як основна. Складова $e_{\text{др}}$ враховується як основна за нормальних умов експлуатації і як додаткова похибка за робочих умова. Якщо за результатами аналізу виявиться, що ця складова дає істотний внесок в загальну похибка, слід використовувати спеціальні схеми автоматичної корекції або термостабілізацію вимірювального перетворювача.

Струми зміщення ОП і викликають зміщення напруги

$$e_{\text{зм}2} = R_1 R_2 I_{\text{зм}}^- / (R_1 + R_2) - R_3 I_{\text{зм}}^+, \quad (7.44)$$

де $I_{\text{зм}}^-$, $I_{\text{зм}}^+$ – середні вхідні струми зміщення для інвертуючого і неінвертуючого входів ОП.

Якщо резистори підібрані з умови $R_3 = R_1 R_2 / (R_1 + R_2)$, то для ОП на біполярних транзисторах (різниця вхідних струмів зміщення $\Delta I_{3M} \ll I_{3M}$) e_{3M2} помітно зменшується: $e_{3M2} = R_3 \Delta I_{3M}$.

Для оцінки впливу температури на ΔI_{3M} користуються залежністю

$$e_{3M2} = (\Delta I_{3M} + \Delta_{\Delta I_{3M}} \Delta T) R_3. \quad (7.45)$$

де $\Delta_{\Delta I_{3M}}$ – температурний дрейф вхідних струмів.

З формули (7.45) випливає, що за наявності резистора R_3 в схемі інвертуючого підсилювача зміщення e_{3M2} залежить від неідентичності вхідних некерованих струмів зміщення і їх зміни у разі варіації температури. Перший доданок суми (7.45) пов'язаний з основною похибкою перетворення і може бути скомпенсований, а другий – з додатковою і основною (аналогічно e_{3M1}).

Неідеальність послаблення синфазного сигналу обумовлена неповним заглушенням синфазного сигналу і обчислюється за формулою

$$e_{3M3} = U^+ / K_{НЗЗ}, \quad (7.46)$$

де U^+ – напруга на неінвертуючому вході.

В інвертуючих каскадах $e_{3M3} = 0$, а в неінвертуючих є мультиплікативною складовою зі значенням $\gamma_{3M3} = K_{НЗЗ}^{-1}$

В диференціальних каскадах необхідно виконувати додаткові розрахунки для визначення напруги U^+ .

Пульсації в ланцюгах живлення призводять до напруги зсуву

$$e_{3M4} = K_{ПЛЖ} \Delta U_{ж}, \text{ мкВ}, \quad (7.47)$$

де $\Delta U_{ж}$, В – варіація напруги живлення; $K_{ПЛЖ}$, мкВ/В – коефіцієнт впливу пульсацій по ланцюгам живлення ОП. Якщо в документації приведено коефіцієнт $K_{ПЛЖ}^*$ в дБ, то $e_{3M4} = 10^{\frac{K_{ПЛЖ}^*}{20}} \Delta U_{ж}$, В ..

Частотні залежності коефіцієнтів $K_{ОП}$, $K_{НЗЗ}$, $K_{ПЛЖ}^*$ мають однакові частоти зрізу і спаду АЧХ, але відрізняються значенням коефіцієнтів на постійному струмі (рис. 7.43, а). Для компенсації збільшення $K_{ПЛЖ}$ з ростом частоти, яке призводить до збільшення похибки e_{3M4} , в ланки живлення ОП включають RC-фільтри (рис. 7.43, б).

Розділ 7. Перетворювачі ІВС

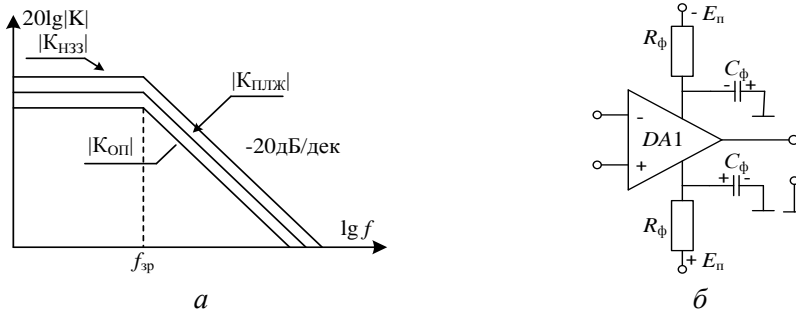


Рис. 7.43. Операційний підсилювач: *a* – АЧХ; *б* – вмикання RC- фільтрів в ланцюгах живлення ОП

Постійну часу фільтра τ_ϕ ланцюга живлення визначають з урахуванням частоти зрізу $f_{зр}$ ОП: $\tau_\phi = (2\pi f_{зр})^{-1}$. Ємність C_ϕ конденсатора фільтра розраховують із співвідношення $\tau_\phi = R_\phi C_\phi$. Значення R_ϕ обирають з діапазону 10 – 100 Ом таким чином, щоби падіння напруги на R_ϕ від струмів живлення ОП не перевищувало 200 – 500 мВ.

Для забезпечення необхідного імпедансу фільтра в смузі частот до 100 кГц рекомендується застосовувати танталові ємності. Якщо використовуються електролітичні конденсатори, для компенсації їх паразитної індуктивності необхідно їх шунтувати керамічними або плівковими конденсаторами невеликих номіналів.

Шумові властивості входних ланцюгів ОП. Шумова складова $e_{зм5}$ має особливість: вона є випадковою і завжди знакозмінною. Для підрахунку $e_{зм5}$ використовують середньоквадратичне значення шуму

$$e_{ш} = \sqrt{\left[G_{ш}^2 + 2(J_{ш} R_3)^2 + 8kTR_3 \right] \Delta f_{ш}}, \quad (7.48)$$

де $J_{ш} \left[A/\sqrt{\Gamma\text{ц}} \right]$ – спектральна щільність еквівалентного входного шумового струму (в документації на ОП наводяться графіки $J_{ш}(f)$ подібний до поданого на рис. 7.44 або проміжні точки цього графіка); $G_{ш} \left[B/\sqrt{\Gamma\text{ц}} \right]$ – спектральна щільність еквівалентної входної шумової напруги (в документації на ОП наводиться графік $G_{ш}(f)$ (рис. 7.45) або проміжні точки цього графіка); $\Delta f_{ш} = f_{ср} \pi/2$ – еквівалентна шумова смуга ОП; $k \approx 1,38 \cdot 10^{-23}$ Дж/К – постійна Больцмана; T –

температура, К.

Якщо розрахунок роблять на фіксованій частоті $f_p \in [f_1, f_2]$, значення $J_{ш}$ і $G_{ш}$ оцінюють за графіками $J_{ш}(f)$ і $G_{ш}(f)$ (рис. 7.44, 7.45).

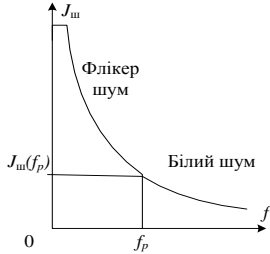


Рис. 7.44. Залежність $J_{ш}(f)$ і її використання на фіксованій частоті

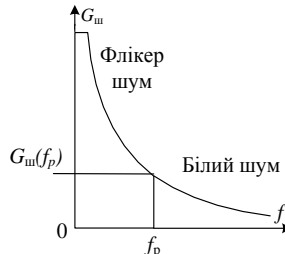


Рис. 7.58. Залежність $G_{ш}(f)$ і її використання на фіксованій частоті

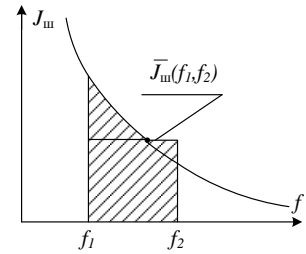


Рис. 7.59. Залежність $J_{ш}(f)$ її використання в діапазоні частот

Якщо розрахунок виконують для смуги частот, то обчислюють площу фігури під кривою $J_{ш}(f)$ (рис. 7.46); в межах f_1, f_2 будують прямокутник тієї ж площі; ордината прямокутника приймається за еквівалентне значення шуму $\bar{J}_{ш}(f_1, f_2)$.

Мультиплікативні похибки. Ці похибки мають п'ять складових, обумовлених такими чинниками: обмеженістю $K_{оп}$ і його зміною; шунтуючою дією на ланцюг НЗЗ вхідних імпедансів ОП; обмеженістю $R_{вих}$ ОП; нестабільністю ланцюга НЗЗ; співвідношенням вхідного і вихідного опорів (похибкою узгодження).

Обмеженість $K_{оп}$ і його зміна – складова, що визначається кінцевим значенням $K_{оп}$ і його залежністю від частоти, температури і зміни напруги живлення.

Коефіцієнт підсилення ОП з НЗЗ в загальному випадку має вигляд

$$K_{НЗЗ} = K_{оп} / (1 + K_{оп}\beta) = (1/\beta) / (1 + 1/K_{оп}\beta) = (1 - 1/K_{оп}\beta) / \beta. \quad (7.49)$$

Графіки АЧХ і ФЧХ для ОП ($|K_{оп}|$ і $\varphi_{оп}$ відповідно) та АЧХ ОП з ланцюгом НЗЗ ($|\dot{K}_{НЗЗ}|$) зображені на рис. 7.47,а.

Розділ 7. Перетворювачі ІВС

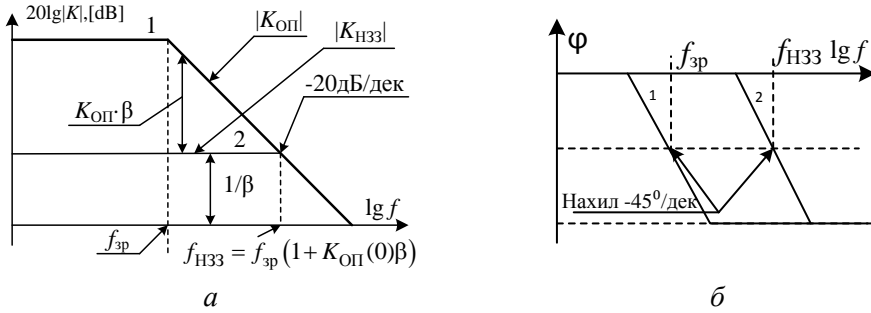


Рис. 7.47. Визначення петльового підсилення за частотними характеристиками: а) 1 – АЧХ для ОП, 2 – АЧХ ОП з ланцюгом НЗЗ ($|\dot{K}_{\text{НЗЗ}}|$); б) – ФЧХ

Елемент $1/\beta$ у формулі (7.49) дає розрахункове або ідеальне значення коефіцієнта підсилення ОП з НЗЗ послідовного типу (для НЗЗ паралельного типу ідеальне значення коефіцієнта підсилення ОП з НЗЗ дорівнює $(1-1/\beta)$), а інший елемент $(1-1/K_{\text{ОП}}\beta)$ – містить значення першої складової мультиплікативної похибки $-\gamma_1 = |1/K_{\text{ОП}}\beta|$. Цей вираз дійсний на частотах, менших за $f_{\text{зр}}$, коли $\varphi_{\text{ОП}} \approx 0$, а значення $K_{\text{ОП}}$ практично не змінюється. На високих частотах ($f \gg f_{\text{зр}}$) набуває 90° зсуву (рис. 7.47, б). У цьому випадку формула (7.49) набуває виду

$$|K_{\text{НЗЗ}}| = \left| \frac{jK_{\text{ОП}}(f)}{1 + jK_{\text{ОП}}(f)\beta} \right| = \left| \frac{1/\beta}{1 - j/K_{\text{ОП}}(f)\beta} \right| = \frac{1}{\beta} \left[1 - \frac{1}{2} \left(\frac{1}{|K_{\text{ОП}}(f)|\beta} \right)^2 \right]. \quad (7.50)$$

Значення першої складової мультиплікативної похибки $\tilde{\gamma}_1$ на високих частотах:

$$\tilde{\gamma}_1 = (1/|K_{\text{ОП}}(f)|\beta)^2 / 2. \quad (7.51)$$

З підвищенням частоти петльове підсилення $K_{\text{ОП}}(f)\beta$ зменшується, що призводить до збільшення похибки $\tilde{\gamma}_1$.

Для врахування температурного чинника і залежно $K_{\text{ОП}}(U_{\text{ж}})$ позначимо відносну зміну $K_{\text{ОП}}$, викликане цими факторами, як $\gamma_k = \Delta K_{\text{ОП}}/K_{\text{ОП}}$. Вираз для розрахунку похибки схеми з НЗЗ, зумовленої нестабільністю:

$$\gamma_1(K_{\text{ОП}}) = \Delta K_{\text{ОП}} \partial K_{\text{ОП}} / K_{\text{НЗЗ}} \partial K_{\text{ОП}} = \gamma_k / (1 + K_{\text{ОП}}\beta). \quad (7.52)$$

Враховуючи, що $K_{\text{ОП}}\beta \gg 1$, а $\gamma_k \leq 0,2$ в робочих умовах, ця похибка у більшості випадків не перевищує 0,1 %.

Шунтуюча дія вхідних імпедансів ОП на ланцюг НЗЗ. Значення вхідних опорів ОП не є нескінченними, а вхідних ємностей – не

нульові. Розглянемо їх вплив у двох основних варіантах включення ОП. Для інвертуючого підсилювача комплексні вхідні опору ОП диференціальному $\dot{Z}_{дф} = R_{дф} \parallel (j\omega C_{дф})^{-1}$ і синфазному $\dot{Z}_{сф} = R_{сф} \parallel (j\omega C_{сф})^{-1}$ сигналам враховуються в коефіцієнт підсилення схеми (рис. 7.48) наступним чином

$$K_{НЗЗ}(j\omega) = (\dot{Z}_2 / \dot{Z}_1) / \left(1 + (K_{ОП} \dot{\beta}')^{-1} \right), \quad (7.53)$$

де $\dot{\beta}'$ – коефіцієнт перетворення ланцюга НЗЗ,

$$\dot{\beta}' = 1 / \left(1 + \dot{Z}_2 / \dot{Z}_1 + \dot{Z}_2 / \dot{Z}_{дф} + \dot{Z}_2 / \dot{Z}_{сф} \right). \quad (7.54)$$

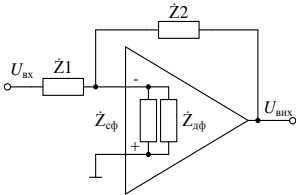


Рис. 7.48. Еквівалентна схема для врахування впливу вхідного імпедансу для інвертуючого включення ОП

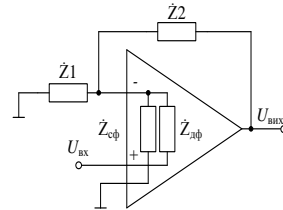


Рис. 7.49. Еквівалентна схема для врахування впливу вхідного імпедансу для неінвертуючого включення ОП

З (7.54) витікає, що за інвертуючого включення ОП вплив вхідного імпедансу є малим, оскільки $\dot{Z}_2 / \dot{Z}_{дф}$, $\dot{Z}_2 / \dot{Z}_{сф}$ діють лише на похибку у (7.53).

Для неінвертуючого включення ОП коефіцієнт підсилення ОП з НЗЗ (рис. 7.49) визначається виразом

$$K_{НЗЗ}(f) = \left(1 + \dot{Z}_2 / \dot{Z}_1 + \dot{Z}_2 / \dot{Z}_{дф} + \dot{Z}_2 / \dot{Z}_{сф} \right) / \left[1 + (K_{ОП} \dot{\beta}')^{-1} \right], \quad (7.55)$$

де $\dot{\beta}'$ задається аналогічно виразу (7.54).

Вплив обмеженості вихідного опору ОУ. Типове значення вихідного опору ОП становить 300 ÷ 500 Ом. Коефіцієнт підсилення ОП (рис. 7.50) з урахуванням опору $R_{вих}$

$$K'_{ОП} = K_{ОП} / \left(1 + R_{вих} / R_H \parallel R_2 \right), \quad (7.56)$$

де R_2 – опір ланцюга НЗЗ; R_H – еквівалентний опір навантаження.

Розділ 7. Перетворювачі ІВС

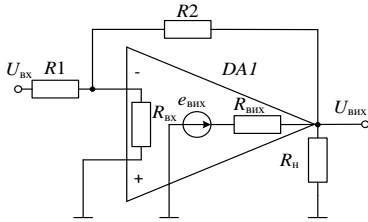


Рис. 7.50. Еквівалентна схема для врахування впливу опору ОП

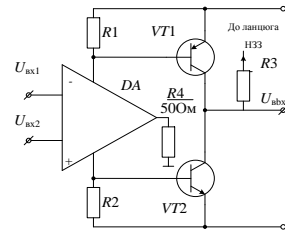


Рис. 7.51. Вихідний каскад ОП типу Rail-to-Rail

Якщо вхідний опір наступного каскаду перевищує 2 кОм, вихідним опором ОП можна знехтувати.

Розглянемо роботу ОП з зовнішнім вихідним ланцюгом, виконаним за схемою Rail-to-Rail (рис. 7.51), з опором навантаження ОП $R_{н} \approx 50$ Ом. Управління $p-n$ -переходами транзисторів $VT1$ і $VT2$ здійснюється завдяки падінню напруги на резисторах R_1 і R_2 в ланцюгах живлення ОП. Вихідний каскад на $VT1$ і $VT2$ дозволяє віддавати в навантаження сигнал значно більшої потужності і з більшою амплітудою, аніж типовий вихідний каскад на ОП.

Вплив нестабільності ланцюга НЗЗ є найбільш значущим в аналізі похибок схем на ОП. Проаналізуємо її для трьох схем включення ОП. Розглянемо спочатку схему інвертуючого підсилювача (рис. 7.35,а). Для цієї схеми коефіцієнт підсилення дорівнює $K_{НЗЗ} = -R_2/R_1$, а відносна похибка перетворення визначається

$$\gamma_4 = \sum_{i=1}^2 \frac{\partial K_{НЗЗ}}{\partial R_i} \frac{\Delta R_i}{K_{НЗЗ}} = \gamma_{R_2} - \gamma_{R_1}, \quad (7.57)$$

де $\gamma_{R_1} = \Delta R_1/R_1$, $\gamma_{R_2} = \Delta R_2/R_2$ – відносні відхилення опорів R_1 , R_2 від їх номінальних значень (клас точності резисторів). Якщо резистори R_1 , R_2 неідентичні, то ΔR_1 , ΔR_2 можуть мати ізні знаки, і максимальне значення цієї похибки дорівнює $\gamma_4 = |\gamma_{R_1}| + |\gamma_{R_2}|$, якщо вони ідентичні, то похибки будуть менші (див. (7.57)).

Розглянемо схему неінвертуючого підсилювача на ОП (рис. 7.36,а). Коефіцієнт підсилення ОП з НЗЗ дорівнює $K_{НЗЗ} = 1 + R_2/R_1$, а відносна похибка перетворення для цієї схеми

$$\gamma_4 = \sum_{i=1}^2 \frac{\partial K_{НЗЗ}}{\partial R_i} \frac{\Delta R_i}{K_{НЗЗ}} = (\gamma_{R_2} - \gamma_{R_1}) \frac{R_2}{R_1 + R_2}. \quad (7.58)$$

Для третьої схеми – диференціального підсилювача (рис. 7.37,а) коефіцієнти передачі синфазного і диференціального сигналів

визначаються відповідно як (7.32) і (7.33). Якщо в ці вирази підставитимо

$$R_2/R_1 = R_4(1+\delta)/R_3, \quad (7.59)$$

де $\delta = 4\gamma_R$; γ_R – клас точності резисторів $R_1 - R_4$, вони набудуть виду

$$K_{\text{сф}} = \delta R_4 / (R_3 + R_4), \quad (7.60)$$

$$K_{\text{л}} = -R_2(1 - \delta\beta/2) / R_1, \quad (7.61)$$

де $\beta = R_1 / (R_1 + R_2)$ – коефіцієнт передачі ланцюга НЗЗ.

Вплив δ на значення коефіцієнта $K_{\text{сф}}$ вельми значне, що є основним недоліком цієї схеми. На практиці його мінімізують, застосовуючи вимірювальні підсилювачі. Найбільш поширеним варіантом такого підсилювача є схема "на трьох ОП".

Вплив співвідношення вхідного і вихідного опорів (похибка узгодження) визначається співвідношенням вихідного опору $R_{\text{вих}}$ попереднього каскаду і вхідного опору наступного каскаду $R_{\text{вх}}$. Каскади узгоджуються за напругою ($R_{\text{вих}} \ll R_{\text{вх}}$) чи за струмом ($R_{\text{вих}} \gg R_{\text{вх}}$).

Спочатку розглянемо з'єднання з наступним каскадом за критерієм *узгодження за напругою*. На рис. 7.52 зображена еквівалентна схема вихідного каскаду ОП для низьких частот і вхідного каскаду наступного блоку з передачею сигналу між каскадами через конденсатор C_p .

$$\text{Відносна похибка узгодження: } \gamma_c = |(U_x - E_x)/E_x| = 1/2(\omega R_{\text{вих}} C_p)^2, \quad (7.62)$$

де C_p – ємність роздільного конденсатора; \bar{E} – напруга паразитної синфазної (постійної) складової напруги, обумовленої струмами зміщення нуля ОП; E_x – напруга корисного сигналу ОП; U_x – вхідна напруга наступного каскаду.

На рис. 7.53 показано еквівалентну схему, що аналогічна схемі на рис. 7.52, але для діапазона високих частот (за безпосереднього зв'язку між каскадами). Відносна похибка узгодження визначається як

$$\gamma_c = |(U_x - E_x)/E_x| = R_{\text{вих}} \sqrt{1 + (\omega\tau_{\text{вх}})^2} / R_{\text{вх}}, \quad (7.63)$$

де $\tau_{\text{вх}}$ – постійна часу вхідного ланцюга наступного каскаду. Вираз (7.63) у випадку $(\omega\tau_{\text{вх}})^2 \gg 1$ набуває вид: $\gamma_c = \omega R_{\text{вих}} C_{\text{вх}}$.

Розділ 7. Перетворювачі ІВС

Похибка $\bar{\gamma}_c$ узгодження за напругою на постійному струмі з виразу (7.63) визначається як $\bar{\gamma}_c = R_{\text{вих}} / R_{\text{вх}}$.

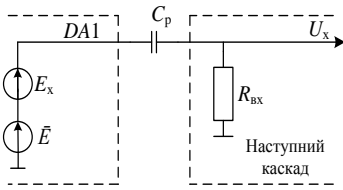


Рис. 7.52. Визначення похибки узгодження на низьких частотах

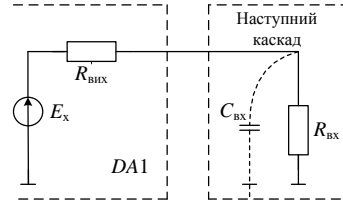


Рис. 7.53. Визначення похибки узгодження на високих частотах

Розглянемо другий варіант – узгодження за струмом (рис. 7.54). Для виконання умови $R_{\text{вих}} \gg R_{\text{вх}}$ вихідний каскад на ОП повинен працювати в режимі джерела струму, а для вхідного бажаною є схема перетворювача «струм - напруга» (рис. 7.55).

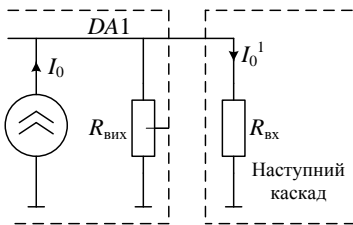


Рис. 7.54. Узгодження за струмом

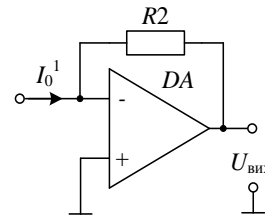


Рис. 7.55. Перетворювач струм – напруга

Похибка узгодження визначається як

$$\gamma_c = (I_0' - I_0) / I_0 = R_{\text{вх}} / R_{\text{вих}}, \quad (7.64)$$

Для перетворювача «струм – напруга» значення вихідної напруги дорівнює $U_{\text{вих}} = I_0' R_2$, а коефіцієнт перетворення ланцюга НЗЗ, враховуючи нерівність $R_{\text{вих}} \gg R_2$, дорівнює $\beta = R_{\text{вих}} / (R_{\text{вих}} + R_2) \approx 1$.

Вхідний опір цієї схеми визначається виразом

$$R_{\text{вх}} = R / (1 + K_{\text{ОП}} \beta). \quad (7.65)$$

Вхідні опори для інвертуючого (рис. 7.35, а) і неінвертуючого (рис. 7.36, а) підсилювачів визначаються відповідно як

$$R_{\text{вх}} \approx R_1 + R_2 / (1 + K_{\text{ОП}} \beta), \quad (7.66)$$

$$R_{\text{вх}} = R_{\text{д}} (1 + K_{\text{ОП}} \beta) \parallel R_{\text{сф}}. \quad (7.67)$$

Для цих схем вихідні опори однакові і дорівнюють

$$R'_{\text{вих}} = R'_{\text{вих ОП}} / (1 + K_{\text{ОП}} \beta). \quad (7.68)$$

Основні висновки

Вимірювальний перетворювач (ВП) – це пристрій з нормованими метрологічними характеристиками, призначений для перетворення вимірюваної величини в іншу величину або вимірювальний сигнал, зручний для опрацювання, зберігання, перетворення, передавання або індикації.

Метрологічні характеристики (МХ) ВП – це технічні характеристики ВП, що описують їх властивості, які важливі для виконання ними функціонального призначення і впливають на результат і похибку вимірювання. Найвживаніші МХ ВП: номінальна функція перетворення, чутливість, поріг чутливості, похибка, діапазон вимірювань, динамічні характеристики (імпульсна перехідна характеристика, комплексна частотна характеристика, амплітудно-частотна характеристика, фазо-частотна характеристика, час реакції ВП, граничні значення частот), імпеданс.

Найвживаніші типи сенсорів: ємнісні, індуктивні, електромагнітні, резистивні (потенціометричні, тензо-, терморезистивні, термісторні), сенсори Холла і Гаусса, сенсори на основі хімічних польових транзисторів, п'єзоелектричні, сенсори на ефекті Доплера, сенсори переміщення на тунельному ефекті, термодинамічні, фотометричні сенсори (фотоємісійні, фоторезистивні і фотодіодні детектори), акустооптичні, волоконнооптичні і флуоресцентні сенсори хімічних величин.

Схемотехніка вторинних ВП передбачає широке використання універсальних і оригінальних засобів аналогового перетворення сигналів, виконаних на основі операційних підсилювачів (ОП).

ОП – це диференційний підсилювач напруги з безпосередніми зв'язками, коефіцієнтом підсилення понад 1000 і можливістю формування частотної характеристики за допомогою введення негативного зворотного зв'язку.

Для отримання розрахункового значення коефіцієнта підсилення ВП користуються моделлю ідеального ОП, для якого мають місце аксіоми: диференціальна вхідна напруга ОП в лінійному режимі завжди дорівнює нулю; вхідні струми ОП дорівнюють нулю; за замкненого зворотного зв'язку напруга на інвертуючому вході ОП дорівнює напрузі на неінвертуючому вході.

Всі схеми на ОП зводяться до шести базовим схем: інвертуючого підсилювача; неінвертуючого підсилювача; диференціального підсилювача; диференціатора; інтегратора; інвертуючого суматора.

Характеристики реальних ОП поділяються на дві групи: межі допустимих параметрів і електричні характеристики. Перші визначають межі експлуатаційних параметрів, за яких ОУ зберігає здатність нормально функціонувати. Електричні характеристики дозволяють розрахувати похибки перетворення сигналів в схемах на ОП.

Аддитивна похибка ВП визначається напругою зміщення, яка має п'ять складових, обумовлених такими чинниками: початковою напругою зміщення

Розділ 7. Перетворювачі ІВС

і температурним дрейфом; струмами зміщення; кінцевим послабленням синфазного сигналу; пульсаціями в ланцюгах живлення; шумовими властивостями вхідних ланцюгів ОП.

Мультиплікативні похибки також мають п'ять складових, що визначаються такими факторами: обмеженістю $K_{\text{ОП}}$ і його зміною; шунтуючою дією вхідних імпедансів ОП на ланцюг зворотного зв'язку; обмеженістю вихідного опору ОП; нестабільністю ланцюга зворотного зв'язку; співвідношенням вхідного і вихідного опорів (похибка узгодження).

Запитання для самоконтролю

1. Дайте визначення і наведіть узагальнену модель сенсора.
2. Дайте загальну характеристику сенсорів генераторного типу.
3. Назвіть основні метрологічні характеристики ВП.
4. Наведіть приклади сенсорів фізичних величин параметричного типу.
5. Дайте приклади активних перетворювачів і їх загальну характеристику.
6. Наведіть приклади фізичних явищ, ефектів, які покладені в основу роботи сенсорів.
7. Розкрийте принцип дії, особливості роботи і основні характеристики ємнісних сенсорів; індуктивних сенсорів; електромагнітних сенсорів.
8. Поясніть принцип дії і особливості роботи, схеми включення та основні характеристики тензодатчиків; терморезистивних сенсорів.
9. Розкрийте принцип дії, особливості роботи і основні характеристики сенсорів на основі ефекту Холла.
10. Розкрийте принцип дії, особливості роботи і основні характеристики сенсорів на основі хімічних польових транзисторів.
11. Розкрийте принцип дії, особливості застосування та основні характеристики п'єзоелектричних сенсорів.
12. Поясніть фізичну сутність ефекту Доплера і приклади його застосування у вимірвальній техніці.
13. Розкрийте принцип дії, особливості застосування та основні характеристики термопар.
14. Наведіть принцип дії термоелектричних перетворювачів на основі кварцових генераторів. Дайте стисло загальну характеристику термодинамічних сенсорів і приклади їх застосування.
15. Розкрийте принцип дії, особливості застосування та основні характеристики фотометричних сенсорів.
16. Наведіть загальну характеристику волоконнооптичних сенсорів.
17. Поясніть принцип дії коїрки Брегга.
18. Назвіть параметри ідеального ОУ.
19. Дайте загальний опис основних базових схем на ОП.
20. Наведіть ідеальні характеристики інвертуючого та неінвертуючого підсилювачів на ОП. Дайте необхідні пояснення.

21. Наведіть ідеальні характеристики інвертуючого суматора на ОП. Дайте необхідні пояснення.
22. Назвіть характеристики ідеального інтегратора і диференціатора на ОП. Дайте необхідні пояснення.
23. Виконайте аналіз меж допустимих параметрів ОП.
24. Виконайте аналіз вхідних характеристик ОП.
25. Виконайте аналіз вихідних характеристик ОП.
26. Виконайте аналіз динамічних характеристик ОП.
27. Назвіть складові мультиплікативної похибки ОП?
28. Які складові включає адитивна похибка ОП?
29. Поясніть різницю в узгодженні по струму і напрузі схем на ОП.
30. Як необхідно узгоджувати частоту одичного підсилення ОП, частотний діапазон вхідних сигналів та коефіцієнт підсилення для підсилювачів напруги на ОП?

Conclusions

Measuring transducer (MT) is a device with normalized metrological characteristics serving to convert measuring quantity to other quantity or measurement signal, convenient for processing, storage, subsequent conversions, indication or transmission.

Metrological characteristics (MC) are technical characteristics of sensors that describe their properties needed to perform their functionality mission and that have an impact on measurement results and errors. MT documentation indicate following MCs: nominal conversion function, sensitivity, sensitivity threshold, accuracy, measurement range, dynamic characteristics (unit-impulse response, complex frequency response, amplitude-frequency response, measurement mean response time, frequency boundary values) impedance.

Most commonly used types of MT are: capacitive, inductive, electromagnetic, resistive (potentiometric, strain gauges, thermoresistive, thermistor), Hall and Gauss converters, sensors based on chemical field effect transistors, piezoelectric converters on Doppler effect, displacement transducer on tunnel effect, thermodynamic, photometric converters (photoemissive, photoresistive and photodiode detectors), acousto-optic, fiber-optic and fluorescent sensors of chemical quantities.

Circuit technique of secondary MT provides extensive use of universal and original means of analog signal conversion performed on the basis of operational amplifiers.

Operational amplifier (OA) is a differential voltage amplifier with direct coupling, amplifying coefficient over 1000 and possibility to generate frequency response by introducing a negative feedback (NFB).

To obtain the estimated amplifying coefficient value of MT the model of an ideal OA is used, where following axioms are true: differential input voltage of OA

Розділ 7. Перетворювачі ІВС

in linear regime is always equal to zero; OA input currents equal to zero; voltage at inverting input of OA equals to voltage at non-inverting input with closed NFB.

All of the schemes at OA are reduced to six basic schemes: inverting amplifier; non-inverting amplifier; differential amplifier; differentiator; integrator; inverting adder.

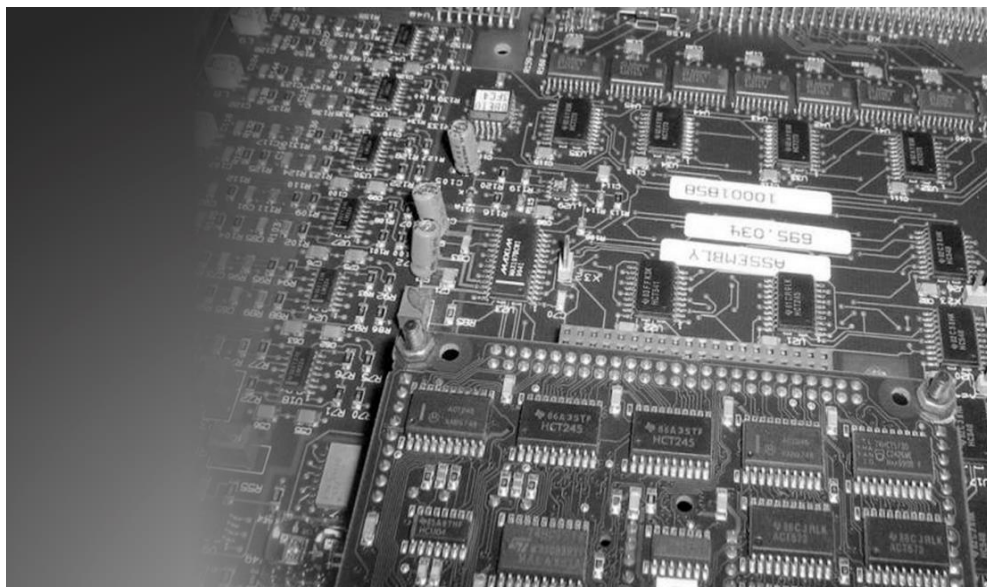
Characteristics of real OA are divided into two groups: limits of permissible parameters and electrical characteristics. At first, defined operating parameters limits at which OA retains the ability to function normally. Electrical characteristics allow calculating signal conversion error in circuits at OA.

On direct current MT' additive error is determined by voltage bias, which has five components caused by such factors: initial offset voltage and temperature drift; displacement currents; finite common mode rejection; pulsations in power circuits; noise properties of MT input circuits.

Multiplicative errors also have five components, determined by such factors: limitation of $K_{оп}$ and its change; shunting effect of OA input impedance on NFB circuit; OA output impedance; NFB circuit instability; input and output impedances ratio (matching error).

Ключові слова

вимірювальний перетворювач	measuring transducer
сенсор	sensor
метрологічні характеристики	metrological characteristics
міст Уїтстона	wheatstone bridge
коефіцієнт підсилення	amplifier gain
операційний підсилювач	operational amplifier
адитивна похибка	additive error
мультиплікативна похибка	multiplication error
коефіцієнт підсилення	amplifying coefficient



АНАЛОГОВІ ТА ЦИФРОВІ ІНТЕРФЕЙСИ

8

- 8.1. Аналогові інтерфейси та їх метрологічні характеристики
- 8.2. Обчислювальні ресурси комп'ютеризованих систем
- 8.3. Цифрові інтерфейси
 - 8.3.1. Паралельні інтерфейси
 - 8.3.2. Послідовні інтерфейси
 - 8.3.3. Інтерфейс USB

8.1. Аналогові інтерфейси та їх метрологічні характеристики

Аналоговий інтерфейс (АІ) – це сукупність засобів вимірювальної техніки, які є складовою частиною вимірювального каналу ІВС між сенсором і АЦП системи збору даних. АІ є однією з найбільш важливих складових ІВС з огляду на формування їх метрологічних характеристик. Узагальнена структура АІ та його місце у вимірювальному каналі наведено на рис. 8.1.



Рис. 8.1. Структура вимірювального каналу з аналоговим інтерфейсом

АІ в ІВС виконує такі функції.

Масштабне перетворення забезпечує підвищення відношення сигнал/шум на вході АЦП. Підсилення сигналу відбувається у передавальній частині АІ якомога ближче до сенсора.

Фільтрація обмежує смугу частот шуму, заглушує небажані частотні складові вимірювального сигналу (наприклад, мережеві наведення), зменшує ефект накладення спектрів під час дискретизації аналогових сигналів.

Ізолювання АІ забезпечує розрив паразитних контурів заземлення, можливість роботи за значних синфазних складових сигналу; досягається за допомогою трансформаторних, ємнісних, оптичних пристроїв гальванічної розв'язки.

Компенсація температури холодних спайв термопар зменшує методичну похибку термопар.

Лінеаризація характеристик сенсорів застосовується у випадках, коли функція перетворення сенсора є нелінійною; реалізується включенням в передавальну частину АІ програмно-керованих підсилювачів або функціональних перетворювачів. Необхідна для зменшення впливу адитивних похибок наступних ланок вимірювального каналу.

Ініціалізація пасивних сенсорів застосовується для підвищення

енергоєфективності і надійності вимірювального каналу, оскільки більшість сенсорів (наприклад, термометри опору, тензорезистивні, індуктивні, ємнісні сенсори) вимагають підведення додаткової енергії у вигляді напруги або струму для отримання активного вихідного електричного сигналу.

Мультиплексування каналів полягає у почерговому зчитуванні всіх сигналів в ІВС шляхом їх почергового підключення. Застосовується для підвищення ефективності використання цифрових процесорів і інших компонент ІВС, скорочення часу їх простоювання і зменшення апаратних витрат і вартості ІВС. Реалізується введенням в структуру АІ мультиплексорів або комутаторів, найчастіше на 10 ... 100 каналів. Мультиплексори будуються на основі електронних ключів. У найпростішому випадку мультиплексор містить k ключів, що включають один з k входів системи до одного виходу.

У разі комутації малих за рівнем сигналів у присутності завад важливо забезпечити захист від них. На рис. 8.2,*а* показано передавання сигналів значних рівнів. Навіть у випадку, коли джерело сигналу V_v знаходиться на деякій відстані від підсилювача A , для напруги між "землями" підсилювача і джерела V_g , яка виникає через протікання по земляній шині паразитних струмів, має місце нерівність $V_g \ll V_v$. За такої умови застосують мультиплексор з *асиметричною* (відносно землі) структурою. Якщо ця вимоги є нереалізовною, застосовують *симетричний* мультиплексор.

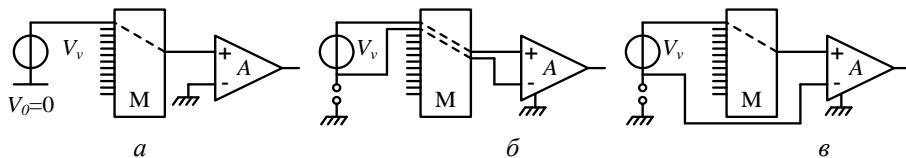


Рис. 8.2. Мультиплексори: асиметричний (*а*), симетричний (*б*) і псевдосиметричний (*в*)

В симетричному мультиплексорі (рис. 8.2,*б*) одночасно переключуються два входи кожного каналу. У цьому випадку важливо, щоби підсилювач A мав великий коефіцієнт ослаблення синфазного сигналу. Більш просте, але менш ефективне технічне рішення приведено на рис. 8.3,*в*. Застосовуючи *псевдосиметричний* мультиплексор, допускають, що для всіх джерел вхідних сигналів у першому наближенні напруги завад V_g однакові.

За необхідності прецизійного вимірювання малих за рівнем сигналів необхідно застосовувати екранування. Таке мультиплексування передбачає відключення і екранів, тому для одного каналу необхідно три ключа.

Цифровий процесор системи збирання даних здатний сприймати інформацію тільки в дискретні моменти часу. За необхідності одночасного

Розділ 8. Аналогові та цифрові інтерфейси

опрацювання декількох вхідних сигналів передбачається, що вибірки сигналів беруться в одні і ті самі моменти часу. На практиці вибірки різних сигналів робляться послідовно, що обумовлює їх зсув у часі. Затримка між вибірками декількох паралельно опрацьовуваних сигналів називається *рознесенням*. Її можна мінімізувати

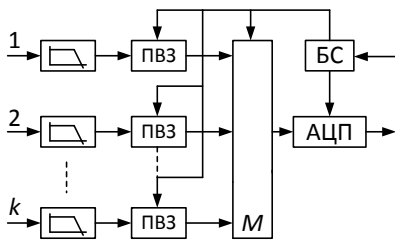


Рис. 8.3. Синхронний мультиплексор з ПВЗ

синхронним мультиплексуванням, що реалізується за допомогою ПВЗ (рис. 8.3). Кожен із вхідних сигналів мультиплексора (М) надходить через свій ФНЧ і ПВЗ. Блок синхронізації

(БС) одночасно перемикає всі ПВЗ в режим "зберігання" в кожному циклі сканування. Мультиплексор зчитує і передає "утримувані" ПВЗ напруги в АЦП. З метою забезпечення високого рівня синхронності вибірок необхідно забезпечити ідентичність характеристик всіх фільтрів.

Важливою характеристикою мультиплексора є *максимальна швидкість опитування* (сканування) каналів, яка визначається найбільшою кількістю каналів, які можна підключити до виходу мультиплексора за одну секунду. Обмеження цієї швидкості обумовлено кінцевою тривалістю перехідних процесів, що виникають внаслідок підключення нового каналу. Час до моменту, коли вихідний сигнал увійде в допустимі межі в околі кінцевого значення, називають *часом встановлення*.

Електронний мультиплексор не є ідеальним пристроєм: він не забезпечує нульовий опір між входом і виходом в режимі "включено", вносить невелику помилку зміщення нуля і не забезпечує повної ізоляції розімкнених каналів. Тому в стані "вимкнено" незначна частина вхідного сигналу проходить на вихід мультиплексора, що є причиною додаткових похибок за рахунок мультиплексування.

Додаткові функції АІ отримують у разі використання в їх складі засобів обчислювальної техніки (інтелектуалізації АІ), що розширює можливості сенсоров і покращує їх технічні характеристики. До кола таких функцій належать: автокалібрування; самодіагностика; інформаційні функції; управління параметрами вимірювального каналу тощо.

Функція корекції. Корекція вихідного сигналу виконується зазвичай за поточними параметрами навколишнього середовища (температурою, атмосферним тиском тощо), якщо останні впливають на сигнал.

Функція самодіагностування. У процесі функціонування інтелектуальні сенсори/інтерфейси аналізують їх роботу: за наявності збоїв, порушень і несправностей, локалізують місце їх виникнення,

оцінюють можливу причину, фіксують вихід похибки за допустимі межі, обліковують чинники, за якими корегують вихідні показання сенсорів тощо.

Інформаційні функції. Інтелектуальні сенсори/інтерфейси зберігають в пам'яті і за дистанційними запитами користувача видають дані про його властивості, характеристики і параметри (тип, заводський номер, діапазон вимірювання, шкалу, параметри налаштування, версію програмного забезпечення, архів метрологічних повірок тощо).

Управління параметрами вимірювального каналу дозволяє дистанційно регулювати (модифікувати) параметри АІ: установку нуля, діапазон вимірювання, параметри фільтрації, одиниці вимірювань тощо.

Класифікація АІ представлена на рис. 8.4.

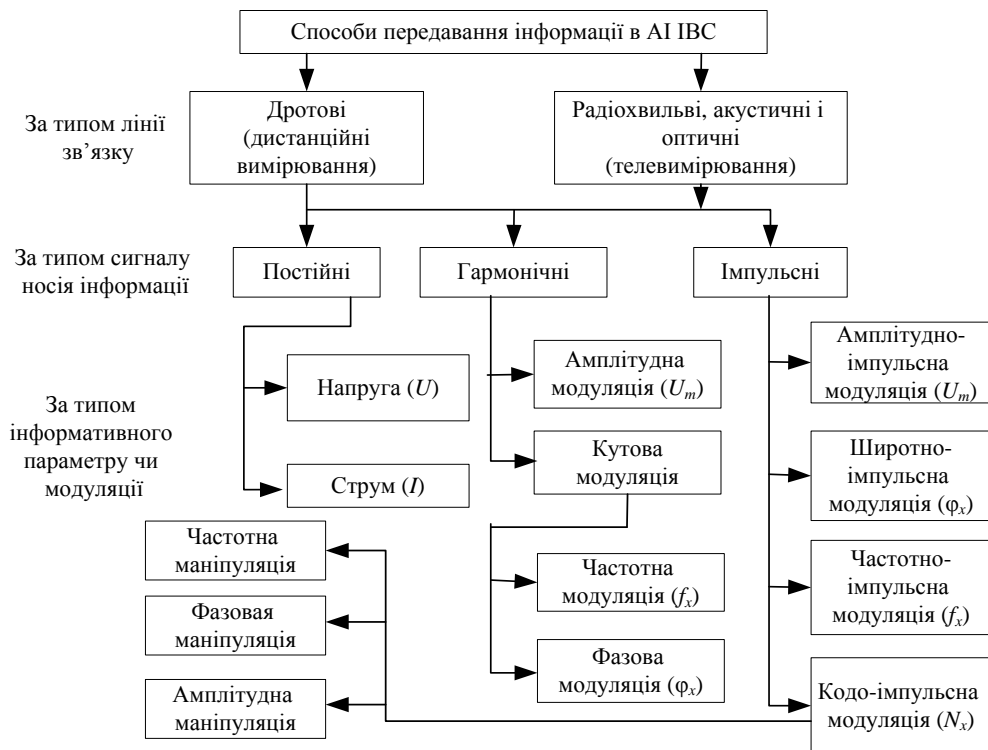


Рис. 8.4. Класифікація способів передавання інформації в АІ

Постійний струм, завдяки високій заводозахисності та практично відсутності впливу паразитних параметрів лінії зв'язку, використовується в розподілених у просторі системах, коли відстань між системою збору даних і об'єктом вимірювань перебуває в межах від декількох до сотень метрів. У

Розділ 8. Аналогові та цифрові інтерфейси

випадках безпосередньої близькості сенсора і системи збирання даних за сигнали вимірювальної інформації використовують напругу.

За типом засобу зв'язку розрізняють: дротові та бездротові АІ – оптичні, радіохвильові та акустичні АІ.

АІ з дротовими лініями зв'язку поділяють на:

двопровідні, в яких по лініях зв'язку передаються як вимірювальні, так і сигнали ініціалізації резидентної частини АІ; найбільш поширений струмовий уніфікований сигнал в діапазоні 4-20 мА;

трипровідні, в яких струм живлення резидентної частини і вихідний струм сенсора передаються в лініях зв'язку через окремі дроти;

чотирипровідні, в яких контури живлення і вихідного сигналу ізолювані один від одного.

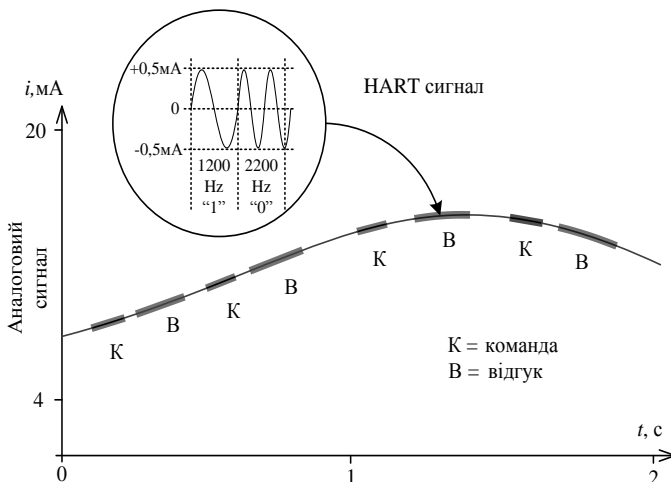
Оптичні АІ більш заводозахиснені, але дорожчі. Радіохвильовий спосіб передавання сигналів в АІ використовують, як правило, спільно з кодоімпульсною модуляцією. АІ на акустичних хвилях застосовуються в спеціальних умовах, наприклад на підводних човнах.

Реалізація дво- і трипровідних АІ істотно спрощується завдяки тому, що універсальні приймально-передавальні вимірювальні перетворювачі для уніфікованих вимірювальних сигналів постійного струму випускаються у вигляді інтегральних мікросхем. АІ з чотирипровідними лініями зв'язку використовуються, як правило, спільно з сенсорами загального застосування і вихідними сигналами (0-5 В, 0-10 В).

У зв'язку зі стійкою тенденцією інтелектуалізації віддалених сенсорів, необхідністю підтримки значної кількості службових функцій останні все частіше забезпечуються засобами передавання цифрової інформації. Цей процес розвивається у двох напрямках: оснащення сенсорів вбудованими цифровими контролерами для узгодження з універсальними мережевими інтерфейсами та поєднання технологій передавання аналогової і цифрової інформації через одну лінію зв'язку.

У першому випадку для взаємодії сенсора і контролера систем введення/виведення використовують протоколи CAN, PROFIBUS-PA, Foundation Fieldbus H1, Interbus, Devicenet, LON, As-i.

У другому випадку забезпечується можливість обміну цифровою інформацією на основі



використання стандартної мережі з *HART-протоколом* (*Highway Addressable Remote Transducer*). Згідно

Рис. 8.5. Обмін даними за HART-протоколом

семирівневої моделі взаємодії відкритих систем (модель OSI), HART-протокол реалізує наступні три рівні: фізичний, каналний і прикладний. HART побудований за принципом «ведучий/веде-ний» – віддалений пристрій відповідає на запит системи. Для передачі цифрової інформації в HART-протоколі (рис. 8.5) використаний принцип частотної маніпуляції: логічна одиниця відповідає одному повному періоду частоти 1200 Гц, а логічний нуль – двум періодам синусоїди 2200 Гц. Двійкові значення передаються зі швидкістю 1200 бод. Цифровий та аналоговий сигнали передаються по одній парі дротів шляхом накладання HART на струмову петлю. Оскільки середнє значення синусоїди за період дорівнює нулю, то HART сигнал не впливає на аналоговий сигнал 4-20 мА. Приймання інформаційного сигналу і настроювання датчика здійснюється за допомогою HART-модему або HART-комунікатора.

Недоліками таких систем є обмеження значенням ~ 10 Гц смуги аналогових сигналів внаслідок необхідності частотного розділення аналогових і цифрових сигналів, а також обмеження накладені топологією ІВС.

Особливості АІ в ІВС з різними топологіями. Різновиди ІВС за типом загальної структури (топології) наведено на рис. 8.6, на якому позначено: ПК – персональний комп'ютер; СЗД – система збирання даних; ЛАШ – локальна аналогова шина; ШВВ – шина введення-виведення; ШУ – шина управління; С – сенсор; μ С – мікроконтролер.

Розділ 8. Аналогові та цифрові інтерфейси

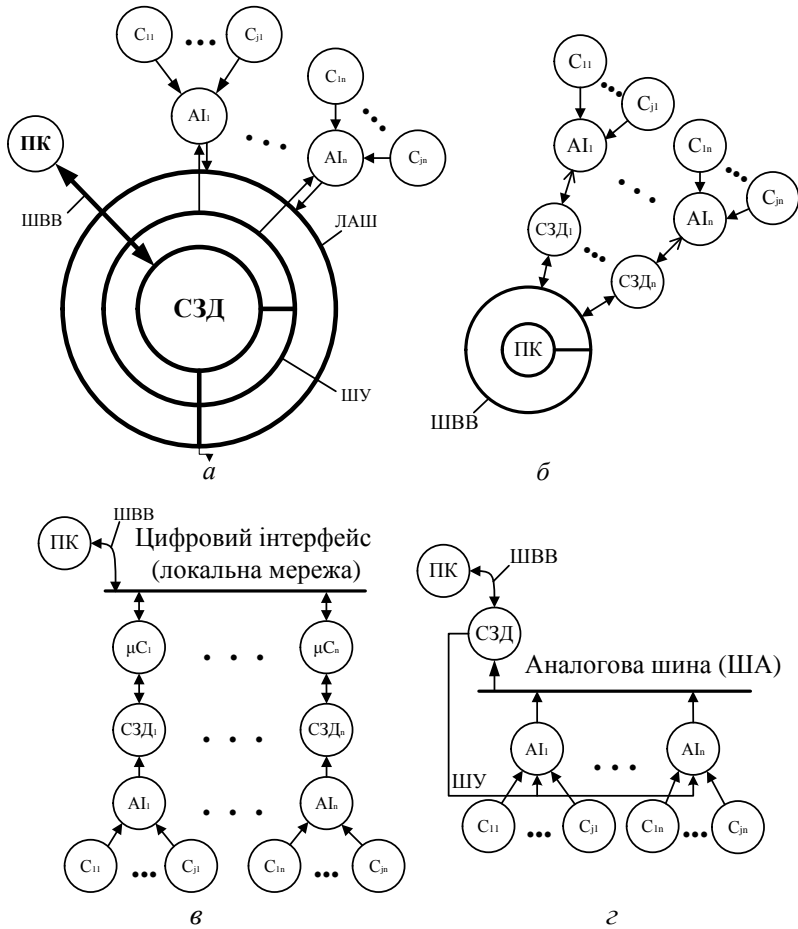


Рис. 8.6. Різновиди топологічних схем ІВС: централізованого (а) та децентралізованого типу (б), з цифровою (в) та аналоговою шиною (г)

За структурою ІВС поділяються на чотири типи. *ІВС централізованого типу* (рис. 8.6,а) є найбільш поширеними, для них характерно використання багатоступневих аналогових мультиплексорів та ЛАШ для збільшення кількості вимірювальних каналів. У випадку збільшення кількості каналів можливості стандартної СЗД, що має 16 диференціальних або 32 уніполярних входів, можуть бути розширені за допомогою зовнішнього мультиплексора. Зв'язок системи збирання даних з ПК реалізується за допомогою однієї з ШВВ: ISA (EISA), PCI, PCMCIA, FireWire, USB. У разі повномасштабних версій ІВС як ШВВ використовуються перші дві, а в портативних варіантах – одна з останніх трьох.

Перевагою систем централізованого типу є можливість реалізації багатofункціональних багатоканальних (до 3000 каналів) систем для лабораторних досліджень з високими метрологічними характеристиками (клас точності 0,05 ... 0,1). Певними недоліками таких систем є громіздкість

(нарощується ІВС шляхом підключення до СЗД аналогової шини і шини управління додатковими АІ), конфлікт між швидкістю і точністю, складність роботи з віддаленими об'єктами. В ІВС *децентралізованого типу* (рис. 8.6,б) для кожного датчика реалізується окремий вимірювальний канал зі своїм АЦП, підключеним до цифрового мультиплектора. Це дає можливість одночасного опитування багатьох каналів з великою швидкістю. Такі системи також називають *паралельними*. За такою схемою будують, як правило, багатofункціональні багатоканальні ІВС для лабораторних досліджень з високими метрологічними характеристиками. Проте в таких системах дещо зменшена щільність вимірювальних каналів і підвищена складність роботи з віддаленими об'єктами.

Багатоточкові ІВС з цифровою локальною мережею (рис. 8.6,в) є системами розподіленого типу. Передавання вимірної інформації в них реалізується за допомогою кодоімпульсної модуляції (що значно звуужує смугу частот корисного сигналу) в цифрових інтерфейсах мережевого типу, тобто використовуються окремі вимірювальні канали зі своїми АЦП, виходи яких об'єднані цифровим інтерфейсом, що утворює локальну мережу. У такій системі вимоги до метрологічних характеристик окремих вимірювальних каналів істотно зменшуються у зв'язку з відсутністю конфлікту між точністю і швидкістю.

Топологія ІВС з цифровою локальною мережею знаходить застосування для побудови багатоточкових ІВС, коли відстань до сенсорів значно перевищує 10 м. Головною перевагою таких систем над попередніми є їх здатність працювати на великих відстанях, а основним недоліком – низька швидкість обміну даними з ПК внаслідок проблем арбітражу і часовою надмірністю кодоімпульсної модуляції.

Багатоточкові ІВС з аналоговою шиною (рис. 8.6,г) використовують у випадку великої кількості аналогових сигналів – від декількох десятків до декількох тисяч каналів. Аналогова шина підключає один з аналогових каналів до єдиної СЗД. Це дозволяє використовувати від 2 до 5 ліній управління для перемикання необмеженої кількості сенсорів. Їх перевагою є можливість нарощування аналогових вимірювальних каналів без значних апаратних витрат. Недоліки таких систем – значний вплив опорів ізоляції роз'ємів, кабелів, комутаторів, оскільки вони підключаються паралельно, що обмежує точність у випадку використання постійних напруг і струмів як інформаційних сигналів; значні паразитні параметри лінії зв'язку обмежують використання модульованих сигналів частотно-модульованими.

Базові структури АІ. Відомо три основні структурні схеми АІ ІВС, які можна уявити комбінацією функціонального перетворювача (ФП) і формувача уніфікованого вихідного сигналу. Різниця між ними полягає в способі корекції похибки.

На рис. 8.7 показана найбільш поширена структура АІ з *ручною корекцією адитивної і мультиплікативної складових похибки*.

Розділ 8. Аналогові та цифрові інтерфейси

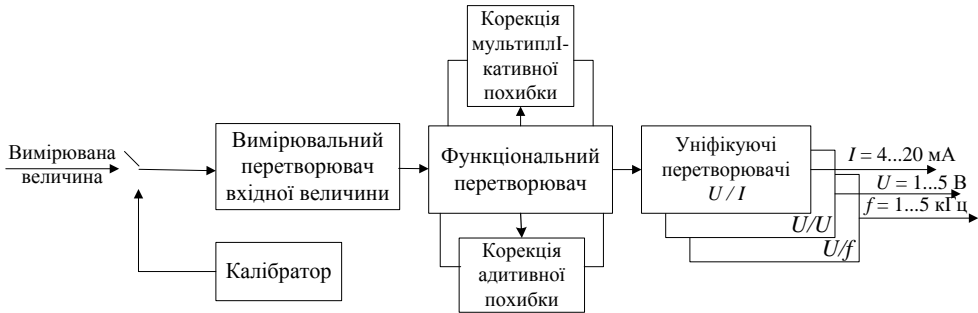


Рис. 8.7. Структурна схема АІ з ручною корекцією похибки

Під час корекції сенсор замінюється калібратором вимірюваної фізичної величини (зразковою мірою) з відповідними параметрами вихідного сигналу в точках калібрування. Оператор за допомогою ручних регулювань компенсує адитивну і мультиплікативну складові похибки АІ. Крім калібратора для такої процедури необхідний прецизійний вимірювач уніфікованого сигналу.

На рис. 8.8 представлена структура АІ з автоматичною корекцією похибок в аналого-цифровій формі, яка реалізується методом допоміжних вимірювань. Суть цього методу полягає в тому, що за допомогою додаткового вимірювального каналу визначається значення основної дестабілізуючої величини (найчастіше – температури). Двопровідні інтерфейси використовуються, як правило, в АІ без засобів автоматичної корекції класу точності 0,5 – 1,0. Внаслідок однополярного живлення їм властива велика похибка (до кількох відсотків) через значну синфазну заваду та обмежений коефіцієнт її заглушення (60 ... 80 дБ). Для зменшення цієї похибки, а також зменшення температурних похибок і похибок нелінійності сенсорів застосовують метод *допоміжних вимірювань* з попереднім індивідуальним градуюванням і аналого-цифровою корекцією. Суть такої корекції полягає в тому, що на етапі градуювання у програмований запам'ятовуючий пристрій з електричним зтиранням EEPROM записують керуючі коди у вигляді таблиць для ЦАП, що вносять поправки на адитивну, мультиплікативну і нелінійну складові похибки при певних температурах. Якщо функція перетворення сенсора лінійна, достатньо занести в EEPROM чотири величини: адитивну похибку сенсора (початковий зсув «нуля»); температурний коефіцієнт адитивної похибки; чутливість сенсора; температурний коефіцієнт чутливості. Недоліками такої корекції є наявність залишкової похибки від дрейфу в часі параметрів сенсора, похибки через зв'язок між каналами вимірювання і похибки ЦАП.

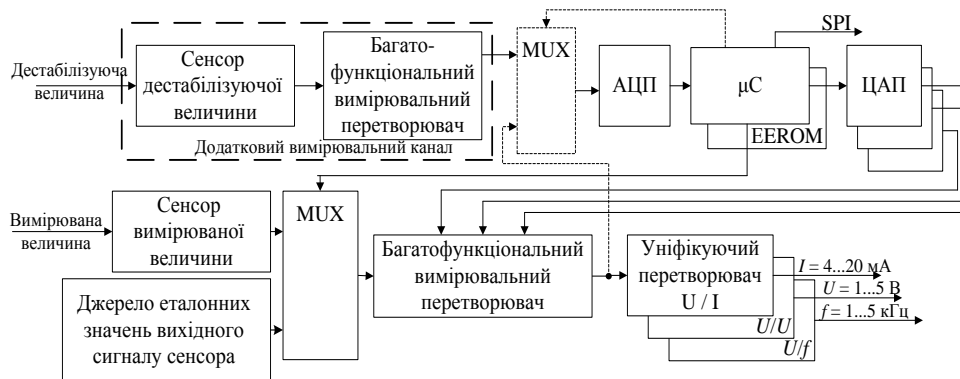


Рис. 8.8. Структурна схема АІ з аналого-цифровою корекцією

На рис. 8.9 представлена структурна схема АІ з автоматичною корекцією в цифровій формі. Через мультиплексор і АЦП значення основної і дестабілізуючої величин вводять в мікро-ЕОМ, де з використанням відомих функцій впливу розраховується скоригований результат, який за допомогою ЦАП і уніфікуючого вихідного перетворювача формує вихідний сигнал.

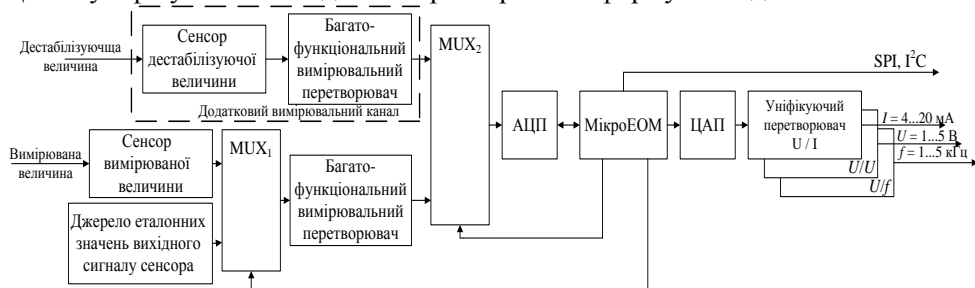


Рис. 8.9. Структурна схема АІ з цифровою корекцією похибки

Така схема застосовується у трипровідних АІ завдяки збільшеним можливостям живлення резидентної частини (струмами до 10 мА замість 4 мА). Завдяки тому, що мікро-ЕОМ має програмований таймер, АЦП може бути реалізований як АЦП інтегруючого типу з проміжним перетворенням «напруга – частота». Корекція вимірювального каналу може бути виконана за методом зразкових мір. АІ с цифровою корекцією використовується в інтелектуальних інтерфейсах с HART-протоколом.

У чотирьохпровідних інтерфейсах використовуються зовнішні мережеві ізольовані блоки живлення, які дозволяють, завдяки симетрично потужному біполярному харчування мостових схем, працювати з металевими тензорезисторами і уникнути в значній мірі вплив на точність вимірювання обмеження коефіцієнта пригнічення синфазних напруг в вимірювальних підсилювачах.

Узагальнена структурна схема АІ з уніфікованими сигналами

Розділ 8. Аналогові та цифрові інтерфейси

напруги представлена на рис. 8.10.

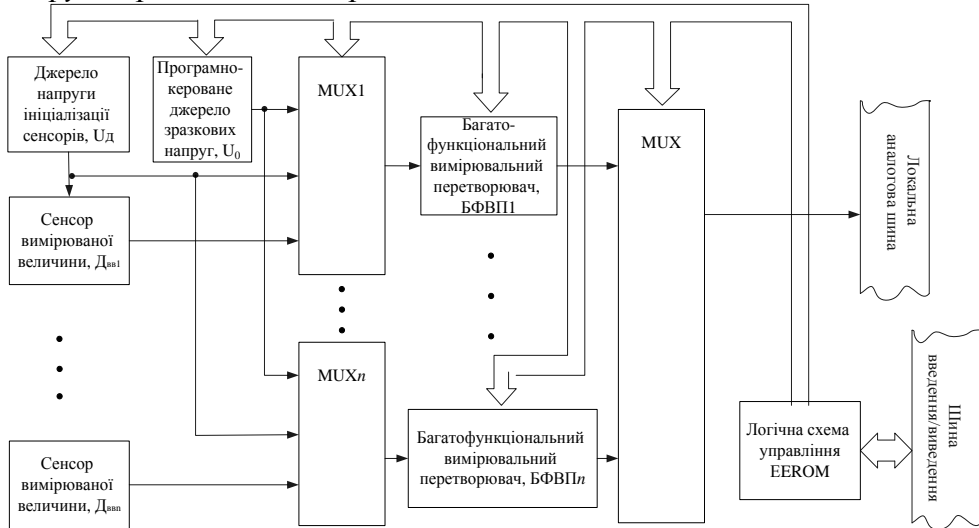


Рис. 8.10. Узагальнена структура АІ з уніфікованими сигналами напруги

Для кожного сенсора передбачений окремий функціональний перетворювач, що містить вимірювальний підсилювач напруги з програмно керованим коефіцієнтом підсилення і зміщенням нуля, програмно-керований ФНЧ, засоби первинного опрацювання сигналу. На вході кожного функціонального перетворювача включений мультиплексор для підключення програмно-керованого джерела зразкових напруг U_0 , а також програмно-керований генератор U_d для ініціалізації датчиків. Такі АІ призначені для лабораторних досліджень і роботи з вимірювальним обладнанням розташованим в безпосередній близькості від ПК (до 10 м) і вирізняються кращими метрологічними характеристиками, багатофункціональністю, високою щільністю каналів.

Якщо розглянуті засоби корекції реалізуються на одному чіпі з сенсором, тобто виконуються в одному технологічному процесі, такий перетворювач називається *інтелектуальним сенсором*. Якщо компоненти структурної схеми АІ є фізично окремими виробами, то такий перетворювач називається *інтелектуальним інтерфейсом*.

Метрологічні характеристики (МХ) аналогових інтерфейсів. Нормовані МХ АІ наведені в табл. 8.1.

Таблиця 8.1

Група	Підгрупа	Спосіб нормування
Характеристики функції		Номінальна функція перетворення
		Вихідний код
		Кількість розрядів кода

ТЕОРЕТИЧНІ ОСНОВИ ІНФОРМАЦІЙНО-ВИМІРЮВАЛЬНИХ СИСТЕМ

перетворення		Номінальна ціна одиниці найменшого розряду кода
Характеристики похибки	Характеристики похибки з поділом на складові	Характеристики систематичної складової похибки: - межа (довірча межа) допустимої систематичної складової похибки; - математичне сподівання і середньоквадратичне відхилення (СКВ) систематичної складової похибки
		Характеристики випадкової складової похибки: - межа допустимого СКВ випадкової складової похибки; - межа допустимої складової СКВ випадкової складової похибки, номінальна нормалізована (у долях дисперсії) автокореляційна функція випадкової складової похибки АІ та межа допустимого відхилення від неї; - межа допустимого СКВ випадкової складової похибки, номінальна спектральна щільність випадкової складової похибки АІ та межа допустимого відхилення від неї; - межа допустимої варіації вихідного аналогового чи цифрового сигналу АІ;
	Характеристики похибки без поділу на складові	Межі допустимої похибки
	Математичне сподівання і СКВ похибки	Межі (довірчі межі) допустимої похибки, межі (довірчі межі) припустимої зміни похибки
Характеристики чутливості до впливової величини		Номінальна функція впливу чи межі (довірчі межі) допустимих відхилень від неї
		Межі допустимих відхилень з МХ АІ від номінальних значень, викликаних впливових зовнішніх величин і міною неінформативних параметрів вхідного сигналу
		Динамічні характеристики впливу: - номінальна функція зв'язку між впливовою змінною в часі величиною та вхідним сигналом АІ і межа допустимого відхилення від неї; - час установлення показів чи вхідного сигналу АІ за заданого характеру змін в часі впливової величини
Динамічні характеристики	Повні динамічні характеристики	Перехідна характеристика
		Імпульсна перехідна характеристика
		Амплітудно-фазова характеристика
		Сукупність АЧХ та ФЧХ
		Передаточна характеристика
	Частинні динамічні характеристики	Час вимірювання (установлення сигналу) АІ
		Час затримки видавання результату
		Похибка часового прив'язування відліків
		Максимальна частота (швидкість) вимірювання
Характеристика взаємодії з об'єктом		Вхідний та/або вихідний імпеданс АІ
Характеристика		Параметри лінії зв'язку, за яких гарантовані МХ АІ

Розділ 8. Аналогові та цифрові інтерфейси

тики лінії зв'язку		Параметри лінії зв'язку, які дають змогу врахувати їх вплив на МХ АІ
--------------------	--	--

Шуми в АІ зазвичай займають широкую смугу частот. **Приклад** спектрального розподілу шуму і завад у вимірювальному каналі ІВС наведено на рис. 8.11. У центральній частині спектрограми шум є «білим», його спектр розподілений рівномірно за частотою. У лівій частині потужність, що припадає на одиничну смугу частот, падає зі збільшенням частоти – шум є «рожевим» (флікер-шум або шум $1/f$).

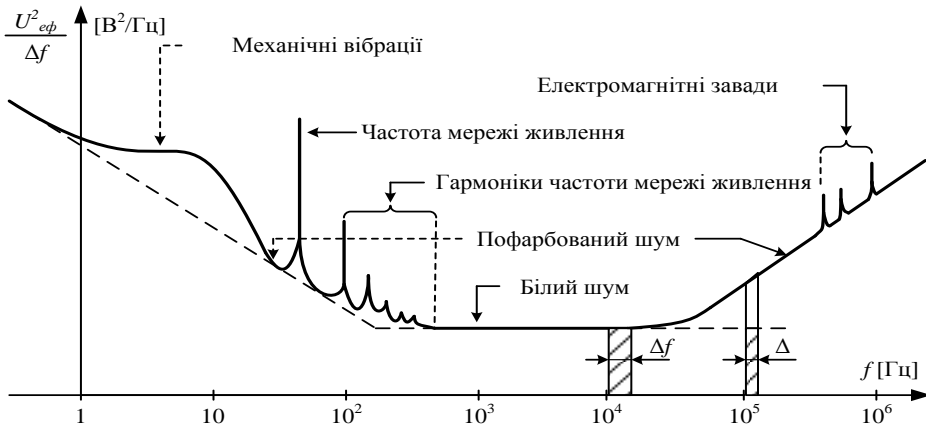


Рис. 8.11. Приклад спектрального розподілу шуму і завад у вимірювальному каналі ІВС

Класифікація шумів в АІ представлена на рис. 8.12. Під шумом розуміється будь-який електричний сигнал, що спотворює корисний сигнал. Рівень шуму визначає граничну можливість АІ правильно передавати інформацію, яка переноситься інформаційними сигналами. Джерела шумів доцільно згрупувати в три групи.



Рис. 8.12. Класифікація шумів АІ

Зовнішні шуми (завади). Ця група включає всі сигнали, які перешкоджають прийманню інформаційного сигналу. Прикладом завади цієї категорії є наведення на частоті 50 Гц і її гармоніках від між мережі живлення від потужних трансформаторів, електродвигунів тощо.

Дрейф. Навіть найвисокоякісніші засоби вимірювань мають обмежену

стабільність. Підсилювачі, джерела струму, тощо мають тенденцію до нестабільності в часі, яка може бути випадковою або мати характер повільних періодичних флуктуацій. Якщо під час вимірювання деякої величини за певних умов характеристики вимірювальної апаратури змінюються, в результат вимірювання вноситься деяка невизначеність за рахунок дрейфу і нестабільності.

Шум приладів. У цю категорію входять шуми, що притаманні будь-якому електронному компоненту: резистору, діоду, операційному підсилювачу і т. п. Вони обумовлені дискретністю заряду електрону. Напруги і струми в будь-якому сигналі є середніми величинами, які відображають в дійсності поведінку статистичного ансамблю електронів. Флуктуації сукупності електронів викликають появу шумів, тобто випадкових складових електричного струму і напруги.

Основні способи зменшення впливу шумів:

1. Для зменшення впливу електромагнітних завад використовується екранування (пасивне, активне).
2. Для зменшення уповільненого дрейфу застосовуються методи автоматичної корекції систематичних складових похибки.
3. Для зменшення шуму електронних приладів використовуються різні способи опрацювання сигналів.

Всі способи опрацювання сигналів ґрунтуються на зменшенні смуги частот сигналу для поліпшення відношення сигнал/шум. У загальному випадку вони можуть бути поділені на способи попереднього (pre-acquisition) і вторинного (post-acquisition) опрацювання. До попереднього опрацювання відносяться різні типи фільтрації – низькочастотна, високочастотна та смугова. Ширина смуги частот сигналу не повинна перевищувати частотного діапазону перетворювача. Способи вторинного опрацювання можуть бути представлені як способи цифрової фільтрації, найпростішим з яких є усереднення, що особливо ефективно для виключення певних частот, наприклад 50 Гц.

8.2. Обчислювальні ресурси комп'ютеризованих систем

Збільшення потоків вимірювальної інформації, ускладнення алгоритмів її отримання та опрацювання обумовили необхідність створення ІВС на базі комп'ютерних систем, що передбачає знання їх структури, принципів функціонування та введення цифрової інформації.

Комп'ютерна система (КС) – це єдиний комплекс, що складається з електронних модулів і пакету програм, призначений для вирішення універсальних або спеціалізованих завдань.

Історія створення КС, ідеологія їх побудови тісно пов'язані з історією математики та електроніки. Перший лічильний пристрій – арифметичну машину – винайшов у 1642 р. французький вчений Блез Паскаль. Майже

Розділ 8. Аналогові та цифрові інтерфейси

одночасно з Паскалем 1673 р. сконструював лічильну машину німецький математик Готфрід Лейбніц. У 1833 р. англійський вчений Чарльз Беббідж розробив проект «аналітичної машини», в якій були передбачені арифметичні і запам'ятовуючі пристрої. Його машина стала прообразом майбутніх комп'ютерів.

Поява електронних обчислювальних машин стала можливою завдяки відкриттям електрики, радіозв'язку та термоелектронної емісії, які дозволили створити електронну лампу. На базі цього електронного пристрою в 1927 р. у Массачусетському технологічному інституті був розроблений аналоговий комп'ютер. Можливість побудови цифрової обчислювальної машини була доведена англійським математиком Аланом Тьюрингом у 1936 р. Він показав, що будь-який алгоритм реалізується за допомогою дискретного автомата, який був названий машиною Тьюринга. У 1942 - 1943 рр. в Англії за його участі була створена обчислювальна машина «Колос», яка призначалася для розшифрування радіограм німецького вермахту. Першою електронною обчислювальною машиною прийнято вважати машину ENIAC (Electronic Numerical Integrator and Computer – електронний числовий інтегратор і обчислювач), розроблену в США у 1946 р.

Розвиток електронних обчислювальних машин (ЕОМ) в СРСР пов'язаний з ім'ям видатного вченого С.О. Лебедева. Незалежно від Джона фон Неймана він висунув, обґрунтував і реалізував принципи побудови ЕОМ з програмою, що зберігається в пам'яті машини. Мала електронна лічильна машина Лебедева містила більше 6000 електронних ламп. Перші ЕОМ мали значні габарити, енергоспоживання та низьку надійність. Винайдення напівпровідникового ефекту і перехід на нову елементну базу відкрило широкі перспективи розвитку ЕОМ.

Принципи побудови комп'ютерних систем. Перша структура КС запропонована в 1833 р. англійським математиком Чарльзом Беббіджом містила два основних пристрої: "млин" – для виконання обчислень і "склад" – для зберігання вихідних даних і результатів обчислень.

У 1944 р. американський математик Джон фон Нейман висловив нові продуктивні ідеї щодо організації структури і обчислень в КС. Він обґрунтував необхідність використання двійкової системи числення для подання чисел у ЕОМ та запропонував записувати програму обчислень у вигляді двійкових кодів в ту ж пам'ять, де зберігалися дані. Архітектура ЕОМ, в якій реалізовані принципи Джона фон Неймана зображена на рис. 8.13.

Архітектура комп'ютера – це сукупність основних пристроїв, блоків і їх взаємодія з опрацювання інформації. Основними блоками КС за Нейманом є ПУ і АЛП (зазвичай об'єднані в центральний процесор), пам'ять, зовнішня пам'ять, пристрої введення та виведення. Безперервні лінії зі стрілками на рис. 8.13 вказують напрямок потоків

інформації, пунктирні – керуючих сигналів.

Архітектура сучасних КС визначаються наступними принципами.

Принцип програмного управління, що забезпечує автоматизацію

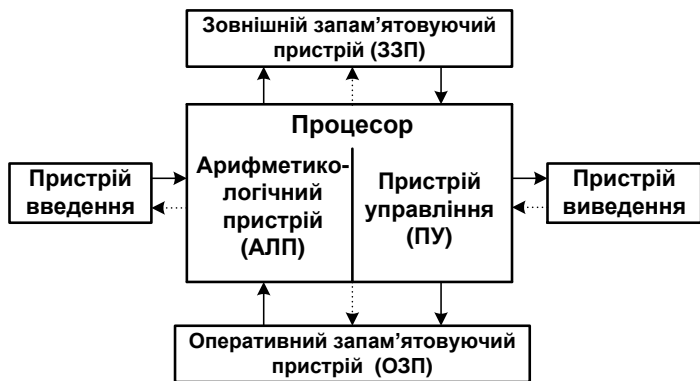


Рис. 8.13. Архітектура ЕОМ на принципах фон Неймана

процесу обчислень на ЕОМ.

Згідно з

цим принципом для вирішення кожного завдання складається програма, що визначає

послідовність дій комп'ютера.

Принцип програми, що зберігається в пам'яті, згідно з яким команди програми подаються і обробляються як дані, у вигляді чисел, а сама програма перед виконанням завантажується в ОЗП, що прискорює процес її виконання.

Принцип довільного доступу до пам'яті, згідно з яким елементи програм та даних можуть записуватися в довільне місце ОЗП, що дозволяє звертатися до будь-якої з них за заданою адресою без перегляду попередніх адрес.

Сучасний комп'ютер – це технічний пристрій (HardWare), який після введення в пам'ять початкових даних і програми їх оброблення (SoftWare) у вигляді цифрових кодів здатний автоматично здійснити обчислювальний процес і видати придатні для сприйняття людиною готові результати.

Архітектура персональних комп'ютерів (ПК) заснована на *магістрально-модульному принципі* (рис. 8.14), згідно якого інформаційний

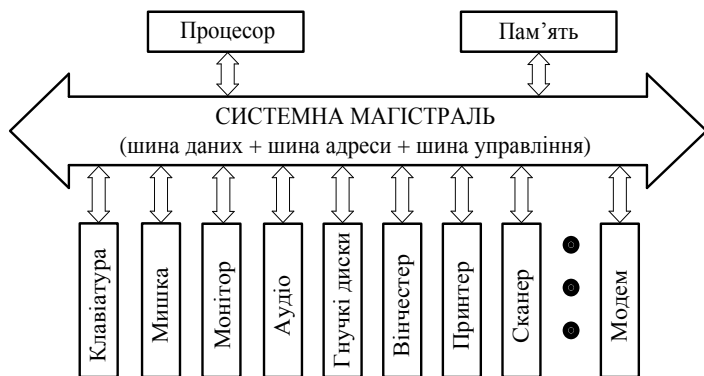


Рис. 8.14. Структура КС, що реалізує магістрально-модульний принцип

зв'язок між пристроями ПК здійснюється через системну шину (системну магістраль).

Обмін даними між окремими структурними елементами КС здійснюється

Розділ 8. Аналогові та цифрові інтерфейси

через системну магіст-раль, що складається з декількох

шин, по яких передаються адреси, дані і сигнали управління.

Шиною називається набір проводів, по яких передається інформація. Передавання інформації в системній магістралі здійснюється за паралельно-послідовним принципом, тобто послідовно в часі у вигляді блоків, кожний з яких складається з бітів, що надходять одночасно і паралельно по всім розрядам шин. Швидкість обміну інформацією через системну магістраль залежить як від розміру блоків, так і від швидкості їх передачі. Перша складова визначається розрядністю шин, а друга – тактовою частотою шини. Важливою характеристикою КС є *пропускна здатність* системної магістралі, яка визначається як добуток розрядності шини на тактову частоту.

Оскільки до системної шини підключаються пристрої з різними швидкостями опрацювання інформації, в сучасних КС застосовують принцип подвійної незалежної системної шини. Він полягає в тому, що всі структурні елементи КС поділяються на "швидкі" і "повільні", а обмін інформацією всередині кожної групи виконується за допомогою власної системної шини. До першої групи належить процесор і ОЗП, до другої – інші пристрої. Виняток становить відео підсистема, для якої використовують власну шину – AGP (Acceleration Graphic Port).

Шину для з'єднання процесора і ОЗП назвали внутрішньою системною шиною, її робоча частота пов'язана з частотою роботи процесорного ядра (дорівнює їй або менше в певну кількість разів).

Для другої групи пристроїв використовується зовнішня системна шина, її частота є фіксованою і не залежить від робочої частоти процесора. Завдяки цій шині КС може інтегруватись у склад ІВС.

КС характеризується різними показниками. Це набір команд, які ПК здатний розуміти, швидкодія центрального процесора, кількість периферійних пристроїв введення/виведення, які можна приєднати одночасно до ПК, тощо. На практиці користувача більше цікавить продуктивність ПК.

Під *продуктивністю* ПК розуміють кількість



**Блез Паскаль
(Blaise Pascal,
1623 - 1662),**

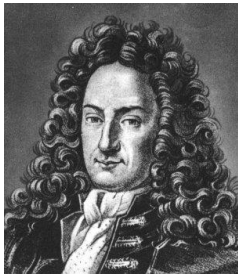
видатний французький фізик, математик, механік, літератор, філософ. Класик французької літератури, один із засновників математичного аналізу, теорії ймовірностей, проєктивної геометрії, гідростатики. Роботи з арифметики, теорії чисел, алгебри. Творець перших зразків лічильної техніки, автор основного закону гідростатики.



Рахункова машина
Паскаля

Розділ 8. Аналогові та цифрові інтерфейси

операцій, які він виконує в одиницю часу.



**Готфрід
Вільгельм
Лейбніц
(Gottfried
Wilhelm Leibniz,
1646 - 1716),**

видатний німецький філософ, логік, математик, механік, фізик, юрист, історик, дипломат, винахідник і мовознавець.

Засновник і перший президент Берлінської Академії наук, іноземний член Французької академії наук. Дав опис двійкової системи числення з цифрами 0 і 1, на якій заснована сучасна комп'ютерна техніка.



Механічний калькулятор Лейбніца

Розрізняють пікову і реальну продуктивність.

Під *піковою* розуміють продуктивність за умови, що всі пристрої ПК працюють в максимально продуктивному режимі. Пікова продуктивність є величина теоретична, практично недосяжна.

Реальна продуктивність залежить від взаємодії програмної моделі, в якій реалізовано додаток. На неї впливають апаратно-програмні особливості ПК – структура процесора, система команд, склад функціональних пристроїв, реалізація введення/виведення тощо.

Існує два способи оцінювання пікової продуктивності комп'ютера. Перший враховує число команд, виконуваних ПК за одиницю часу – MIPS (Million Instructions Per Second). Продуктивність в MIPS характеризує швидкість виконання ПК своїх інструкцій. Оскільки заздалегідь невідомо, якою кількістю інструкцій буде реалізована конкретна програма, дана характеристика дає лише загальне уявлення про продуктивності ПК. Другий спосіб оцінювання продуктивності полягає у визначенні числа операцій з плаваючою комою, виконуваних ПК за одну секунду – Flops (Floating point operations per second). Цей спосіб більш прийнятний для користувача, якому відома обчислювальна складність програми. Користуючись цією характеристикою можна отримати нижню межу часу її виконання.

Шляхи підвищення продуктивності КС полягають у використанні більш швидкодіючих мікроелектронних модулів, вдосконаленні архітектури КС і оптимізації взаємодії складових КС. Перший полягає в принциповому вдосконаленні елементної бази комп'ютерів, створенні нових, більш швидких і надійних мікропроцесорів (МП), запам'ятовуючих пристроїв, тощо. Однак швидкодію елементів неможливо збільшувати безмежно внаслідок технологічних обмежень.

Другий шлях пов'язаний зі збільшенням розрядності шин і оптимізацією обчислень. Перспективнішим є шлях оптимізації обчислень на мікропрограмному рівні. Комп'ютерні програми перетворюються в МП на послідовність мікрокоманд. За типом мікрокоманд МП поділяються на CISC і RISC архітектури.

Розділ 8. Аналогові та цифрові інтерфейси

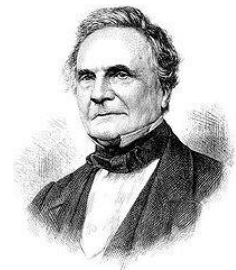
Архітектура CISC (Complex Instruction Set Computer) характеризується великим набором мікрокоманд, спрощує написання програм, однак ускладнює реалізацію кожної мікрокоманди.

Для CISC-процесорів характерне невелике число регістрів загального призначення (РЗП), велика кількість машинних команд, які виконуються за багато тактів, велика кількість методів адресації і форматів команд різної розрядності.

Основою архітектури сучасних процесорів є архітектура ПК зі скороченим набором команд – RISC (Reduced Instruction Set Computer). RISC-процесор характеризується невеликим (50 – 100) набором мікрокоманд, фіксованою довжиною машинних команд, спеціалізованими командами читання/запису в пам'ять, великою кількістю РЗП. Вимоги до наявності доступних обсягів пам'яті підвищились, створення програм ускладнилось, але структура RISC-процесорів стала гнучкішою, що істотно вплинуло на їх продуктивність. В обох архітектурах використовується паралельне виконання мікроінструкцій, яке забезпечується на апаратному рівні в МП і відомо як *конвеєрне опрацювання*. Його принцип полягає в тому, що в кожен момент часу процесор працює одночасно з різними мікроінструкціями декількох команд.

Іншим способом паралельного виконання мікроінструкцій є векторне опрацювання, яке передбачає застосування однієї операції до елементів вектора. Головний принцип обчислень на векторній машині полягає у виконанні певної елементарної операції або комбінації елементарних операцій повторно до деякого блоку даних.

Управління пам'яттю КС. *Обсяг пам'яті* ПК і її *швидкодія* є найважливішими характеристиками КС. *Запам'ятовуючі пристрої* (ЗП) класифікують за низкою ознак. За типом звернення ЗП вирізняють пристрої, що припускають читання інформації і запис – RAM (random access memory), та постійні запам'ятовуючі пристрої (ПЗП), призначені тільки для читання записаних в них даних – ROM (read only memory). ЗП першого типу використовують для зберігання виконуваних програм, вихідних даних, проміжних і кінцевих результатів. У ПЗП зберігають системні програми, необхідні для запуску комп'ютера, зокрема базова система введення/виведення BIOS (Basic Input/Output System).



Чарльз Баббідж
(Charles Babbage,
1797 - 1871),

видатний англійський математик, винахідник першої аналітичної обчислювальної машини. Іноземний член-кореспондент Імператорської академії наук у Санкт-Петербурзі. Праці з теорії функцій, механізації розрахунків в економіці.



Модель
обчислювальної
частини машини
Чарльза
Баббіджа

Розділ 8. Аналогові та цифрові інтерфейси



Джон фон Нейман
(John von Neumann,
1903 - 1957),

угорсько-американський математик єврейського походження, який зробив важливий внесок у квантову фізику, функціональний аналіз, теорію множин, інформатику, економіку і інші галузі науки. Найбільш відомий як людина, з іменем якого пов'язують архітектуру сучасних комп'ютерів (архітектура фон Неймана), застосування теорії операторів до квантової механіки (алгебра фон Неймана), а також як учасник Манхеттенського проекту і як творець теорії ігор та концепції клітинних автоматів

Основні характеристики ЗП – *ємність і швидкодія*. Ці параметри знаходяться у протиріччі: для певного типу ЗП поліпшення одного з них веде до погіршення іншого. Крім того, використання ЗП з покращеними характеристиками обмежене економічною доцільністю. Тому пам'ять ПК будується за ієрархічним принципом (рис. 8.15).

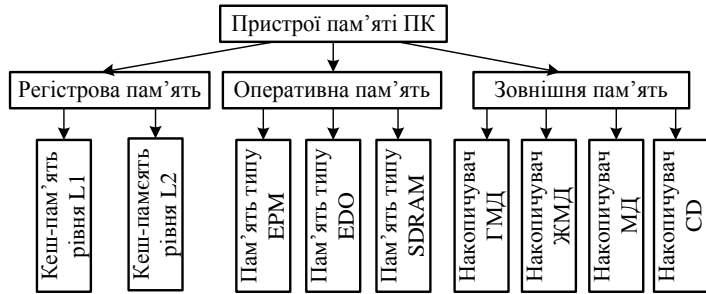


Рис.8.15. Ієрархія типів пам'яті комп'ютера

На нижньому рівні ієрархії знаходиться регістрова пам'ять – набір регістрів, в яких основним елементом зберігання інформації є тригер, елемент цифрової схемотехніки з двома стійкими станами. Один тригер зберігає один біт інформації. Для зберігання одного байта потрібно 8 тригерів, одного кілобайта - 8192 тригерів і т.д. Оскільки управління записом і читанням кожного біта здійснюється окремо для кожного тригера, обсяг регістрової пам'яті не буває великим і обмежений кількома десятками кілобайт.

Час доступу до такої пам'яті мінімальний і порівнюваний з періодом тактової частоти процесора, тому регістрова пам'ять застосовується для зберігання проміжних результатів обчислень і називається *надоперативним запам'ятовуючим пристроєм* (НОЗП) або кеш-пам'яттю.

У сучасних ПК кеш-пам'ять має два рівні: перший рівень (L1) розміщується безпосередньо в мікрочіпі процесора (внутрішня кеш-пам'ять), а другий (L2) – на материнській платі (зовнішня кеш-пам'ять). Оскільки кеш-пам'ять невелика за обсягом, вона грає роль буфера між оперативною пам'яттю (ОП) і процесором. У ній зберігаються копії даних, до яких процесор звертається найчастіше. ОП виконує основну роль зі збереження даних у ПК. Як елементи зберігання інформації

використовуються мікроконденсатори, які здатні деякий час

зберігати електричний заряд. Згодом за рахунок струмів витоку конденсатори розряджаються, і без додаткової підзарядки інформація може бути втрачена. Тому необхідна періодична підзарядка конденсаторів – регенерація пам'яті. Ця пам'ять працює тільки в динамічному режимі, тобто за умови безперервної підзарядки конденсаторів, тому називається *динамічною* (DRAM). Динамічна пам'ять досить проста, дешевше регістрової, але програє їй у швидкодії. Для прискорення процесу запису/зчитування комірки динамічної пам'яті розташовують у вигляді матриці – звернення відбувається відразу до одного рядка матриці. Такий режим називається режимом *швидкого посторінкового обміну* – FPM (fast page mode). Наступною модифікацією динамічної пам'яті, спрямованої на підвищення її продуктивності є EDO-DRAM (enhanced data out). Ця пам'ять містить регістр-засувку, що дозволяє виконувати одночасно зчитування з комірок і запис в них нової інформації.

Подальше підвищення швидкодії пам'яті можливе за рахунок синхронізації читання і запису з тактовою частотою. Такий вид пам'яті називається SDRAM - synchronous DRAM.

Для будь-якого виду пам'яті продуктивність розраховується як добуток розрядності шини на частоту, з якою інформація по цій шині зчитується/записується в пам'ять. Динамічна пам'ять найчастіше застосовується у вигляді модулів з розрядністю 1, 2, 4 або 8 байт, стандартизованих відповідним чином.

Зовнішня пам'ять організовується, як правило, на магнітних та оптичних дисках, магнітних стрічках. Остання внаслідок малої продуктивності і великого обсягу використовується тільки як пристрій резервного копіювання даних.

Основним видом зовнішньої пам'яті є накопичувачі на жорстких магнітних дисках – "вінчестери". Обсяг дискової пам'яті досягає десятків терабайт з часом звернення менше 1 мкс.

Оптичні CD-диски дозволяють зберігати значні обсяги інформації поза ПК.

Все більшої популярності набувають пристрої зовнішньої пам'яті на основі ПЗС-структур (Flash-



**Алан Метисон
Тьюрінг
(Alan Mathison
Turing,
1912 - 1954),**

англійський математик, логік, криптограф, що істотно вплинув на розвиток інформатики. Кавалер Ордена Британської імперії (1945), член Лондонського королівського товариства (1951). Запропонував абстрактну обчислювальну «Машину Тьюрінга» (1936), яка вважається моделлю комп'ютера загального призначення і використовується в безлічі теоретичних і практичних досліджень. Вважається батьком інформатики та теорії штучного інтелекту.

пам'ять, ПЗС-прилади із зарядовим зв'язком) і зовнішні накопичувачі на жорстких магнітних дисках.

8.3. Цифрові інтерфейси



**Сергій
Олексійович
Лебедев
(1902 - 1974),**

видатний радянський вчений, академік АН СРСР, творець першого в Україні та в континентальній Європі комп'ютера. Наукова школа Лебедева за своїми результатами успішно конкурувала з американською фірмою ІВМ. Під його керівництвом створені 15 типів високопродуктивних, найскладніших ЕОМ, починаючи з лампових (БЕСМ-1, БЕСМ-2, М-20) і закінчуючи суперкомп'ютерами на інтегральних схемах.

Управління обміном даними із зовнішніми пристроями. До складу КС входять різні периферійні пристрої. Їх з'єднання між собою здійснюється за допомогою цифрових інтерфейсів.

Цифровий інтерфейс (ЦІ) – це сукупність програмних і апаратних засобів, призначених для передачі інформації між компонентами КС, в тому числі електронні схеми, лінії, шини і сигнали адреси, даних і управління, алгоритми їх передавання і правила інтерпретації пристроями. ЦІ характеризуються такими параметрами: пропускну здатністю – кількістю інформації, яке може бути передана через ЦІ в одиницю часу; максимальною частотою передавання інформації через інтерфейс; максимальною відстанню між зв'язуваними пристроями; загальним числом проводів (ліній) в інтерфейсі; інформаційною шириною – числом бітів або байтів даних, що передаються одночасно через інтерфейс.

ЦІ пристроїв введення/виводу в КС повинен відповідати таким вимогам: модульність і можливість підключення до КС пристроїв різного типу; можливість виконання периферійними пристроями процедур введення/виведення без завантаження центрального процесора КС, універсальність програмування користувачем операцій введення/виведення незалежно від особливостей периферійного пристрою, а також можливість передавання в КС службових сигналів від периферійних пристроїв для реалізації апаратних переривань. Основним шляхом реалізації цих вимог є магістрально-модульний спосіб організації КС, коли всі пристрої комп'ютера представляються у вигляді модулів, з'єднаних між собою загальною магістраллю (рис.8.16).

Обмін інформацією по магістралі задовольняє вимогам загального інтерфейсу, встановленого для магістралі даного типу. Кожний модуль (МП, ОЗП, пристрої введення – ПВ) підключається до магістралі за допомогою спеціальних інтерфейсних схем – Імп, Іозп, Іпв₁, ..., Іпв_л. Інтерфейсні схеми вирішують завдання:

забезпечення функціональної та електричної сумісності сигналів і протоколів обміну модуля і системної магістралі, перетворення внутрішнього формату даних модуля в формат даних системної магістралі і навпаки, а також забезпечення сприйняття єдиних команд обміну інформацією і перетворення їх в послідовність керуючих внутрішніх сигналів. Інтерфейсні схеми можуть бути досить складними і за своїми можливостями відповідати універсальним мікропроцесорам. Такі схеми називають *контролерами*.

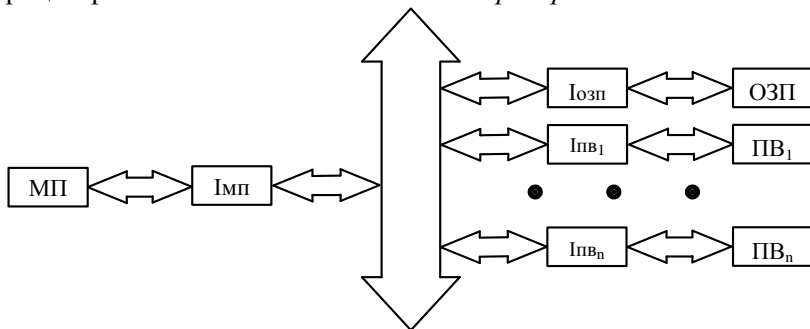


Рис. 8.16. Ілюстрація магістрально-модульного принципу побудови КС

Недоліком магістрально-модульного способу організації КС є неможливість одночасної взаємодії більше двох модулів, що обмежує продуктивність комп'ютера. Цей спосіб використовується в КС, до характеристик яких не висуваються високі вимоги щодо швидкодії.

Організація обміну інформацією між ОЗП і периферійними пристроями в КС реалізується двома способами: програмно-керованою передачею і прямим доступом до пам'яті. Перший здійснюється за безпосередньої участі і під управлінням процесора, що суттєво знижує його можливості одночасно виконувати обчислювальні дії.

Прямий доступ до пам'яті передбачає організацію спеціального каналу обміну інформацією між пам'яттю і периферійним пристроєм без участі процесора. У цьому разі функцію управління обміном приймає контролер прямого доступу до пам'яті. Однак і за його використання процесор повністю не звільняється від управління операціями введення/виводу. Це знижує ефективність роботи ПК, особливо у випадках, коли в КС використовується багато високошвидкісних зовнішніх пристроїв. Для вирішення цієї проблеми в склад високопродуктивних комп'ютерів включають спеціалізовані процесори введення/виводу, здатні повністю розвантажити основний процесор від управління операціями обміну з зовнішніми пристроями.

У сучасних КС всі програми поділяються на системні і прикладні (програми користувачів). Перші входять до складу операційних систем і утворюють операційне середовище, в якому виконуються програми користувачів.

Класифікація цифрових інтерфейсів (рис. 8.17). Основне призначення ЦІ полягає в забезпеченні можливості підключення до КС зовнішніх пристроїв, для чого інтерфейси повинні задовольняти

Розділ 8. Аналогові та цифрові інтерфейси

певним

стандартам.

За місцем інтерфейсів в структурі КС вони поділяються на внутрішні і зовнішні. Перший тип ЦІ забезпечує функціонування самої КС, з'єднання її окремих модулів. В першу чергу це інтерфейси внутрішньої системної шини (ISA, EISA, VLB, PCI), магістральні інтерфейси графічних адаптерів AGP, пристроїв зовнішньої пам'яті (IDE, EIDE, ATA, SCSI), розширення портативних комп'ютерів PCMCIA.

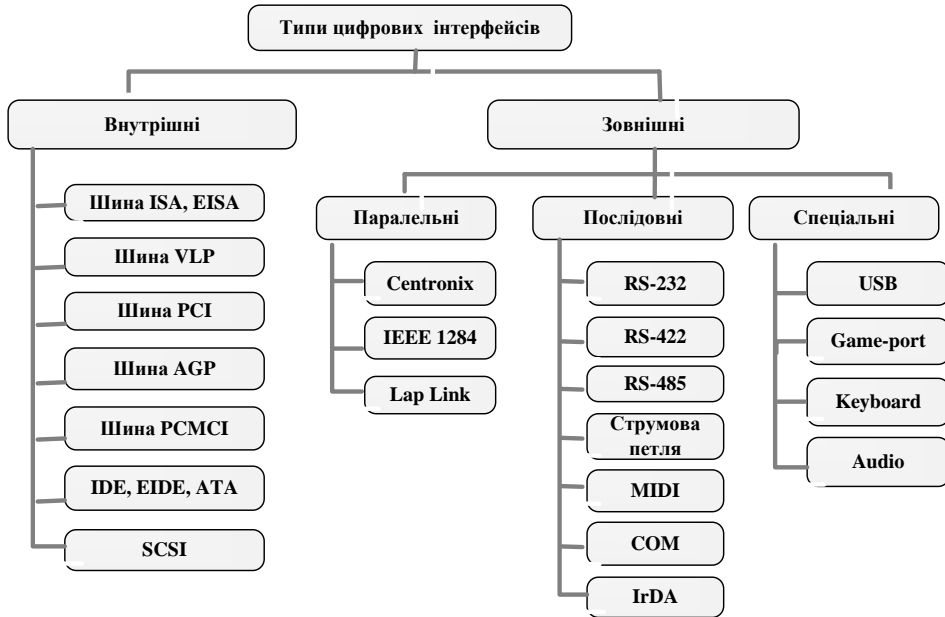


Рис.8.17. Класифікація цифрових інтерфейсів КС

Зовнішні інтерфейси призначені для підключення додаткових пристроїв, що розширюють можливості КС. За способом передавання інформації відомі послідовні, паралельні і спеціальні зовнішні інтерфейси. У послідовних біти інформації передаються послідовно один за одним. У паралельних інтерфейсах кілька біт передаються одночасно по паралельних лініях, що значно збільшує їхню пропускну здатність. До спеціальних належать інтерфейси клавіатури, маніпулятора "миша", аудіопристроїв, відеомоніторів, USB.

В ЦІ застосовують три режиму обміну інформацією. *Дуплексний режим* припускає передавання інформації по одному каналу зв'язку одночасно в обох напрямках. *Напівдуплексний режим* дозволяє передавати і приймати інформацію виключно по черзі, а інтерфейс має засоби перемикання напрямку каналу. *Симплексний* (односторонній) *режим* передбачає тільки один напрямок передавання інформації (у зустрічному напрямку передаються тільки допоміжні сигнали).

Іншим важливим параметром ЦІ є гранична відстань між з'єднаними пристроями. Вона обмежується як частотними властивостями кабелів, так і завадозахищеністю інтерфейсів. Частина завад викликана сусідніми лініями ЦІ – це наведені завади, для захисту від яких застосовують кручені пари проводів для кожної лінії. Інші завади викликають спотворення рівнів і форми сигналів внаслідок зміни хвильового опору ліній зв'язку.

Для шин USB і Fire Wire як характеристика інтерфейсу застосовують топологію з'єднання. Інтерфейсні шини USB і Fire Wire реалізують деревоподібну топологію, в якій зовнішні пристрої можуть бути як кінцевими, так і проміжними (розгалужувачами), що дозволяє підключати безліч пристроїв до одного порту USB або Fire Wire.

Важливою властивістю інтерфейсу є наявність/відсутність гальванічної розв'язки. З наведених інтерфейсів гальванічну розв'язку пристроїв забезпечують MIDI, «струмова петля», шина Fire Wire, мережеві інтерфейси Ethernet. Іншою важливою властивістю інтерфейсів є можливість «гарячого» підключення/відключення або заміни пристроїв (Hot Swap). Така властивість підтримують шини USB і Fire Wire. Ряд інтерфейсів підтримує автоматичне конфігурування – PnP (Plug and Play – вклячай і грай). Така функція закладена у інтерфейсах PCI, USB, Fire Wire, Bluetooth і легко реалізується.

8.3.1. Паралельні інтерфейси

В паралельних інтерфейсах для передачі кожного біта слова використовується окрема сигнальна лінія, тому біти передаються одночасно. Ці інтерфейси використовують логічні рівні TTL, що обмежує довжину кабелю через невисоку завадозахищеність TTL-інтерфейсу.

Спочатку паралельний інтерфейс призначався для підключення до ПК принтера по інтерфейсу Centronics, тому він отримав назву LPT-порт (Line PrinTer - порядковий принтер). Стандартний LPT-порт орієнтований на виведення даних, хоча з деякими обмеженнями може працювати і на введення. Іноді паралельний інтерфейс використовують для зв'язку між двома комп'ютерами – реалізується найпростіша мережу (LapLink), тобто «зроблена на коліні»

Існують різні модифікації LPT-порту. Оскільки спочатку вони були нестандартними технічними рішеннями окремих виробників, пізніше для них був прийнятий стандарт IEEE 1284.

Інтерфейс CENTRONICS. Поняття Centronics відноситься як до набору сигналів і протоколів взаємодії, так і до гнізда на принтерах. Сигнали порту виводяться на роз'єм DB-5S (розетка) (рис. 8.18).

Розділ 8. Аналогові та цифрові інтерфейси

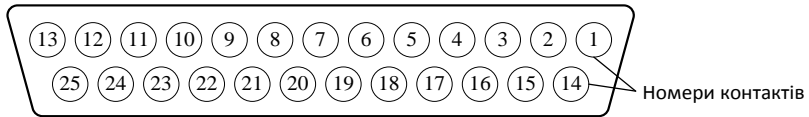


Рис. 8.18. Розміщення і нумерація контактів LPT-порту

LPT-порт має 8-бітну шину даних, 5-бітну шину сигналів стану і 4-бітну шину сигналів управління. Гальванічна розв'язка відсутня, тому порт є вразливим місцем комп'ютера. За неправильного підключення і заземлення пристроїв може вийти з ладу не тільки сам порт, а й електронні блоки з його найближчого оточення.

Логічна структура паралельного порту. Структура LPT-порту як логічного пристрою наведена на рис. 8.19 і має три групи сигналів – дані (8 біт), сигнали стану (5 біт), сигнали управління (4 біти).

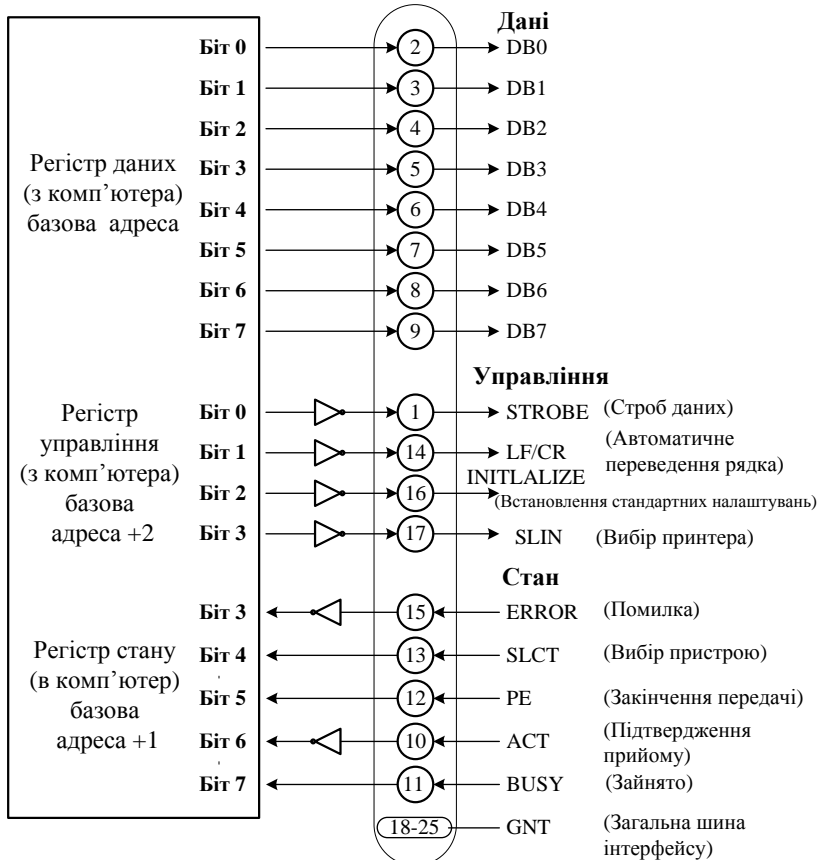


Рис. 8.19. Логічна структура LPT-порту

У КС паралельний порт розглядається як пристрій введення/виведення і

асоціюється з певним простором. ПК підтримує до чотирьох портів LPT1 – LPT4. Після включення або перезавантаження комп'ютера BIOS перевіряє наявність паралельних портів і для кожного порту формує адресний простір, що визначається базовою адресою, яка залежить від поточної конфігурації устаткування комп'ютера. Існує два способи отримання базової адреси: шляхом аналізу конфігурації обладнання комп'ютера або визначення адреси програмним шляхом за допомогою BIOS. Якщо LPT-порти виявляються, їх базві адреси (двобайтні слова) містяться в комірках ОЗП:

LPT1: 0000:0408 h - 0000:0409h	(стандартне значення – 378h)
LPT2: 0000:040Ah - 0000:040Bh	(стандартне значення – 278h)
LPT3: 0000:040Ch - 0000:040Dh	(стандартне значення – 3BCh)
LPT4: 0000:040Eh - 0000:040Fh	(стандартне значення – відсутнє)

Перша комірка містить молодший, друга – старший байт базової адреси кожного порту. Ще одна комірка пам'яті – 0000: 4011 h, в бітах 6 і 7 містить відомості про загальну кількість встановлених на ПК паралельних портів:

біт 7 = 0, біт 6 = 0	паралельні порти не встановлені,
біт 7 = 0, біт 6 = 1	встановлений один паралельний порт,
біт 7 = 1, біт 6 = 0	встановлено два паралельних порти,
біт 7 = 1, біт 6 = 1	встановлено три паралельних порти.

Базова адреса порту дозволяє програмним шляхом забезпечити введення/виведення інформації. Щоби проглянути значення того чи іншого біта або змінити його значення, потрібно звернутися до певних розрядів регістра LPT-порту. Наприклад, щоб перевірити готовність зовнішнього пристрою (принтера) до роботи, потрібно зчитати стан біту 7 (BUSY) регістра стану.

Стандарт IEEE 1284 описує порти SPP, EPP і ECP. Він визначає 5 режимів обміну даними, метод узгодження режиму, фізичний та електричний інтерфейси. Можливі наступні режими обміну даними.

Режим сумісності (Compatibility Mode) – однонаправлене виведення згідно протоколу Centronics; режим відповідає SPP-порту.

Напівбайтний режим (Nibble Mode) – введення байта в два цикли по 4 біта; використовує для прийому лінії стану; режим обміну прийнятний для будь-яких адаптерів.

Байтний режим (Byte Mode) – введення байта цілком, використовуючи для прийому лінії даних; режим працює на портах, що допускають читання вихідних даних (Bi-Directional або PS/2 Type 1).

Режим EPP (EPP Mode) – двонаправлений обмін даними (EPP – Enhanced Parallel Port – модернізований паралельний порт). Режим ефективний при роботі з пристроями зовнішньої пам'яті і адаптерами локальних мереж.

Режим ECP (ECP Mode, Extended Capability Port – порт з розширеними можливостями) – двонаправлений обмін даними з можливістю апаратного стиснення даних методом RLE (Run Length Encoding – кодування тривалості послідовності), використання FIFO-буферів і DMA.

Розділ 8. Аналогові та цифрові інтерфейси

Стандарт IEEE 1284 визначає фізичні характеристики приймачів і передавачів сигналів, сумісних за рівнями з ТТЛ. IEEE 1284 визначає два рівні інтерфейсної сумісності.

Перший рівень (Level I) визначено для пристроїв повільних, але таких, що використовують зміну напрямку передачі даних. *Другий рівень (Level II)* визначено для пристроїв, які працюють в розширених режимах, з високими швидкостями передачі даних і довгими кабелями.

Рекомендовані схеми прикінцевих ланцюгів наведені на рис. 8.20. Інтерфейсні кабелі мають від 18 до 25 проводів в залежності від числа провідників GND. Такі кабелі працюють ненадійно на швидкості передачі 2 Мбайт/с і довжині понад 2 м.

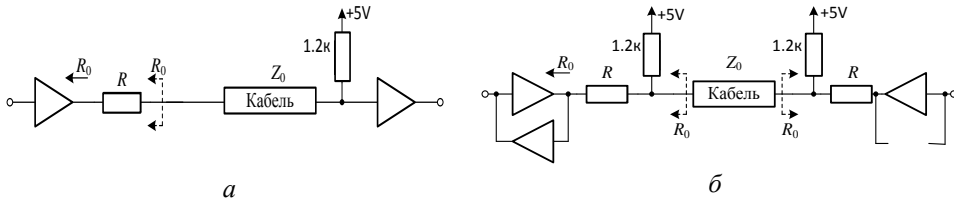


Рис. 8.20. Прикінцеві ланцюга ліній інтерфейсу IEEE 1284:
а – однонапрявлена лінія; б – двонапрявлена лінія

Стандарт IEEE 1284 регламентує такі властивості кабелів: всі сигнальні лінії повинні бути зв'язані з окремими зворотними (загальними) проводами; всі пари повинні мати імпеданс 62 ± 6 Ом в діапазоні частот 4 ... 16 МГц; рівень перехресних завад між парами не повинен перевищувати 10%; кабель повинен мати екран (фольгу), що покриває не менше 85% зовнішньої поверхні, на кінцях кабелю екран повинен бути окільцьований і з'єднаний з контактом роз'єму.

8.3.2. Послідовні інтерфейси

Послідовні інтерфейси мають одну сигнальну лінію, по якій інформаційні біти передаються послідовно в часі. Звідси англійська назва інтерфейсу і порту – *Serial Interface* і *Serial Port*. Послідовна передача дозволяє скоротити кількість сигнальних ліній і збільшити дальність зв'язку. Її особливістю є застосування сигналів з відмінними від ТТЛ рівнями. У ряді послідовних інтерфейсів застосовано гальванічну розв'язку зовнішніх сигналів від схемної землі, що дозволяє з'єднувати пристрої, що знаходяться під різними потенціалами.

Послідовна передача даних може здійснюватися в асинхронному або синхронному режимах. В *асинхронному* режимі кожному байту передують старт-біт, що сигналізує приймачу про початок посліжки, за

яким ідуть біти даних і, можливо, контрольний біт парності. Завершує послітку стоп-біт, що гарантує паузу між послітками (рис. 8.21). Старт-біт наступного байта посилається в будь-який момент після стоп-біта, тобто між передачами можливі інтервали довільної тривалості.



Рис. 8.21. Формат асинхронного передавання даних

У середині послітки між бітами існує жорстка синхронізація, підтримувана за рахунок використання приймачем і передавачем синхрогенераторів однієї частоти. Старт-біт задає початок відліку, відносно якого визначається часове положення бітів даних. Синхрогенератор приймача використовує лічильник-подільник опорної частоти, що устанавлюється в нуль в момент приймання старт-біта і генерує внутрішні строби, за якими приймач фіксує наступні одержувані біти.

Строби розташовуються в середині бітових інтервалів, що дозволяє приймати дані за незначної неузгодженості синхрочастот приймача і передавача. У разі передавання послітки з 10 бітів (8 біт даних, одного контрольного і одного стоп-біта) гранично допустима неузгодженість синхрочастот приймача і передавача не повинна перевищувати 5%, інакше інформація не буде правильно розпізнана.

Формат асинхронної передачі дозволяє виявляти такі помилки:

якщо прийнятий перепад, що сигналізує про початок послітки, а по стробу старт-біта зафіксований рівень логічної одиниці (Лог 1), старт-біт вважається помилковим і приймач переходить в стан очікування;

якщо у відведений під стоп-біт інтервал часу виявлений рівень логічного нуля (Лог 0), фіксується помилка стоп-біта;

якщо застосовується контроль парності, то після послітки бітів даних передається контрольний біт, що доповнює кількість бітів даних з одиницями до парного або непарного в залежності від прийнятого алгоритму контролю; порушення контролю фіксується як помилка.

Контроль формату дозволяє виявляти обрив лінії. У цьому випадку приймається Лог 0, який спочатку трактується як старт-біт, і нульові біти даних, потім спрацьовує контроль стоп-біта.

Для асинхронного режиму прийнятий ряд стандартних швидкостей обміну: 50, 75, 110, 150, 300, 600, 1200, 2400, 4800, 9600, 19200, 38400, 57600 і 115200 біт/с. Іноді використовують іншу одиницю виміру швидкості - «бод» (baud) – число змін стану лінії (від 0 до 1 і навпаки) за одну секунду. У бодах прийнято вимірювати частоту зміни стану лінії.

Розрядність даних в одній послітці може становити 5, 6, 7 або 8 біт, а кількість стоп-біт – 1, 1,5 або 2. Асинхронний обмін в ПК реалізується за

Розділ 8. Аналогові та цифрові інтерфейси

допомогою COM-портів з використанням протоколу RS-232C.

Синхронний режим передачі передбачає постійну активність каналу зв'язку. Посилка починається з синхробайта, за яким йде потік інформаційних бітів. Якщо у передавача немає даних для передачі, він заповнює паузу послією байтів безперервної синхронізації. Очевидно, що у випадку передачі великих масивів даних в синхронному режимі співвідношення інформаційної і службової складових послілки значно вище, ніж в асинхронному. Однак в синхронному режимі необхідна зовнішня синхронізація приймача з передавачем, оскільки навіть незначна неузгодженість їх робочих частот призводить до спотворення отриманих даних. Зовнішня синхронізація можлива за окремою лінією передачі синхросигналу, або за допомогою самосинхронізуючого кодування даних, для якого на стороні приймача з прийнятого сигналу можуть бути виділені імпульси синхронізації. Синхронний режим вимагає дорогих ліній зв'язку або кінцевого обладнання.

На фізичному рівні послідовний інтерфейс має різні реалізації, що різняться способом передачі сигналів і регламентовані рядом споріднених стандартів: RS-232C, RS-423A, RS-422A і RS-485. На рис.8.22 приведені схеми з'єднання приймачів і передавачів, обмеження на довжину лінії (L) і максимальну швидкість передачі даних (V).

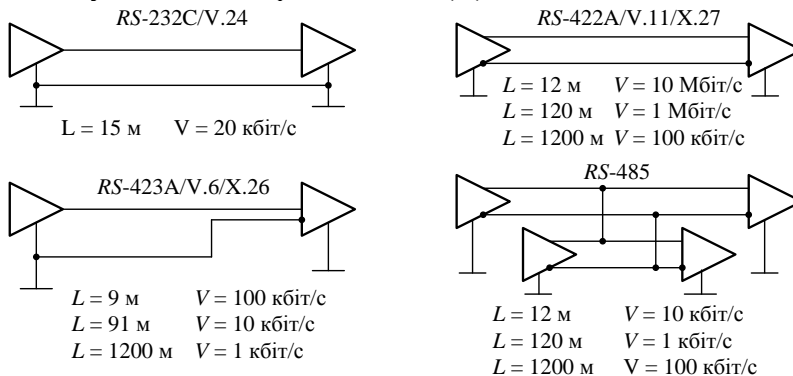


Рис. 8.22. Стандарти послідовних інтерфейсів

Протокол RS-232. Послідовний інтерфейс RS-232 – це промисловий стандарт послідовного двонапрявленого асинхронного передавання даних. Він використовується в ПК для підключення принтерів, модемів, миші, з'єднання двох комп'ютерів (нуль-модемного з'єднання). Відстань, на якому можна організувати зв'язок, не перевищує 15 м. Офіційна назва стандарту – Interface Between Data Terminal Equipment (DTE) and Data Circuit-Termination Equipment (DCE) Employing Serial Binary Data Interchange. У ролі DTE можуть бути комп'ютер, принтер, плоттер і інше периферійне устаткування. Як DCE зазвичай виступає модем. Стандарт описує сигнали управління інтерфейсу, пересилання даних, електричний інтерфейс і типи роз'ємів.

Згідно RS-232 послідовність бітів передається напругою. Передавач і приймач є несиметричними: сигнали передаються відносно загального проводу (на відміну від симетричної передачі протоколу RS-485 або RS-422). Рівню Лог 0 на вході приймача відповідає напруга +3...+12В, а Лог 1 – -12...-3В. Діапазон -3...+3В – це зона нечутливості, що забезпечує гістерезис приймача (передавача). Інтерфейс не забезпечує гальванічної розв'язки пристроїв. Підключення та відключення інтерфейсних кабелів пристроїв з незалежним живленням повинно проводитися при відключеному живленні. Недоліком стандарту RS-232 є низька завадозахищеність і обмежена довжина лінії передачі.

Протокол RS-422A. Стандарт RS-422A (інша назва ITU-T V.11) визначає електричні характеристики симетричного цифрового інтерфейсу. Він передбачає роботу на більш високих швидкостях (до 10 Мбіт/с) і великих відстанях (до 1000 м) в інтерфейсі DTE-DCE. Для його практичної реалізації, на відміну від RS-232, потрібні по два фізичні дроти на кожен сигнал. Реалізація симетричних ланцюгів забезпечує найкращі вихідні характеристики. Даний стандарт є простим описом електричних характеристик інтерфейсу і не визначає параметри сигналів, типи роз'ємів і протоколи управління передачею даних. Стандарт RS-422A, який розроблено спільно з RS-423A, дає змогу розміщувати лінії цих інтерфейсів в одному кабелі. Він несумісний з RS-232, а взаємодія між RS-422A і RS-232 забезпечується за допомогою спеціального інтерфейсного конвертера.

Протокол RS-423A. Стандарт RS-423A (інша назва V.6) визначає електричні характеристики несиметричного цифрового інтерфейсу. "Несиметричність" означає, що даний стандарт, подібно RS-232, для кожної лінії інтерфейсу використовує тільки один провід і для всіх ліній використовується єдиний загальний провід. Стандарт RS-423A передбачає максимальну швидкість передачі 100 кбіт/с. Цей стандарт не визначає сигнали, конфігурацію виведень або роз'єми. Він містить тільки опис електричних характеристик інтерфейсу.

Протокол RS-485. Повна назва цього стандарту TIA / EIA-485A. "Додаток А" цього стандарту описує принципи роботи пристроїв на основі RS-485. На додаток до опису стандарту був опублікований документ TSB89, який називається "Application Guidelines for TIA / EIA-485A" і роз'яснює застосування пристроїв, що визначаються стандартом TIA / EIA-485A для фізичних мереж. Принципова відмінність передавачів RS-485 – можливість їх переведення в третій стан. Передавачі RS-422/485 сумісні з приймачами RS-423.

Протокол RS-499. Стандарт RS-449, на відміну від RS-422A і RS-423A, містить інформацію про параметри сигналів, типах роз'ємів, розташування контактів і т. п. RS-449 є доповненням до стандартів RS-422A і RS-423A. Стандарту RS-449 відповідає міжнародний стандарт V.36. Стандарт RS-449 визнає 30 сигналів інтерфейсу, більшість з яких мають еквівалент-стрічки в RS-232. Десять сигналів RS-449 визначені як лінії першої категорії. Це такі

Розділ 8. Аналогові та цифрові інтерфейси

сигнали як "Передані дані", "Отримані дані", "Синхронізація терміналу". Швидкість їх передачі залежить від довжини кабелю. Для ліній цієї категорії на швидкостях до 20 кбіт/с можуть використовуватися стандарти RS-422A або RS-423A, на швидкостях від 20 кбіт/с до 2 Мбіт/с – тільки RS-422A. Залишені 20 ліній класифікуються як лінії другої категорії і використовуються стандартом RS-423A. До них відносяться такі керуючі лінії, як "Якість сигналу", "Вибір швидкості передачі" і ін. Стандарт RS-449 визначає тип роз'єму і, на відміну від RS-232, розподілення контактів роз'єму. Роз'єми мають 37 контактів для прямого і 9 контактів для зворотного каналу.

Співставний аналіз RS-протоколів. У табл. 8.2 наведені результати порівняння чотирьох найбільш поширених RS-протоколів. Несиметричні інтерфейси RS-232C і RS-423A мають найнижчу захищеність від синфазної завади. Кращі параметри має двоточковий інтерфейс RS-422A і його магістральний аналог RS-485, що працюють на симетричних лініях зв'язку. Вони передбачають використання для передачі диференціальних сигналів окремої (звитої) пари проводів.

Таблиця 8.2

Характеристика	RS-232CV.24	RS-422AV.11	RS-423AV.6	RS-485
Максимальна довжина кабеля, м	15	1200	1200	1200
Число приймачів	1	10	10	32
Число передавачів	1	1	1	32
Максимальна швидкість передачі, кбіт/с	20	100	10	10

У перерахованих стандартах сигнал є потенціалом. Існують послідовні інтерфейси, в яких інформативним параметром є електричний струм, який протікає в ланцюжку передавач/приймач – «струмова петля» і MIDI. Для зв'язку на короткі відстані прийняті стандарти бездротового інфрачервоного зв'язку. Найбільшого поширення в ПК отримав реалізований COM-портами стандарт RS-232C. У вимірювальній техніці широко застосовують RS-485 і RS-422A

Логічна структура послідовного порту (рис. 8.23). Для звернення до окремих регістрів послідовного порту використовуються адресні зміщення відносно базових адрес подібно до паралельних портів.

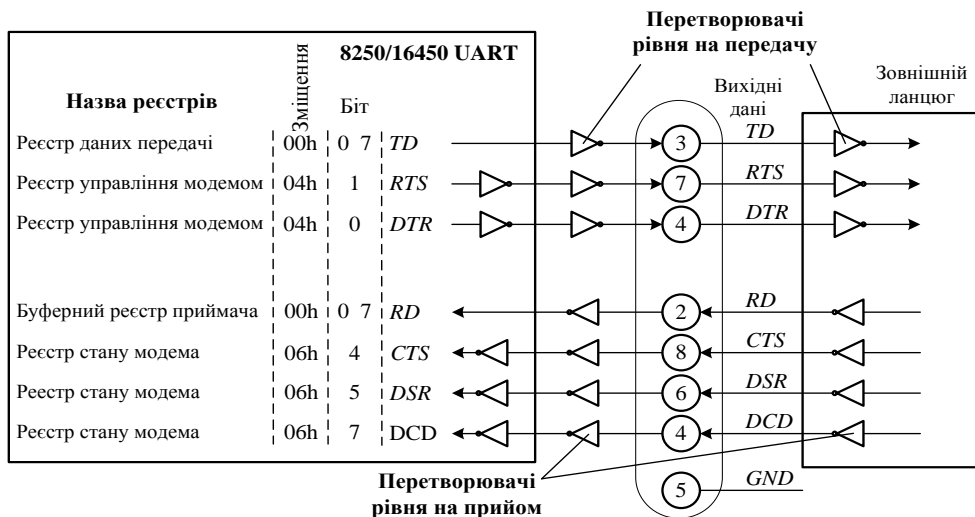


Рис. 8.23. Логічна структура послідовного порту

Призначення сигналів інтерфейса RS-232:

TD (*Transmit Data*) – вихід передавача;

RTS (*Request To Send*) – запит передавання даних;

DTR (*Data Terminal Ready*) – вихід готовності терміналу до обміну даними;

RD (*Receive Data*) – вхід приймача;

CTS (*Clear To Send*) – вхід дозволу терміналу передавати дані;

DSR (*Data Set Ready*) – вхід готовності від апаратури передавання даних (модему);

DCD (*Data Carrier Detected*) – вхід визначення несучої віддаленому модему.

ПК може мати до чотирьох послідовних портів COM1 – COM4, базові адреси яких зазвичай мають значення 3F8h, 2F8h, 3E8h і 2E8h відповідно. Під час завантаженні ПК наявність послідовних портів і їх тестування виконує POST BIOS, в результаті чого базові адреси виявлених послідовних портів поміщаються в комірки BIOS Data Area 0: 0400h, 0402h, 0404h, 0406h. Нульове значення адреси означає відсутність порту з даним номером. В комірках 047C, 047D, 047E, 047F містяться константи, що задають тайм-аути для портів, а молодші біти комірки 0: 0411h – інформацію про кількість встановлених COM-портів. Для портів COM3 і COM4 можливі альтернативні адреси - 3E0h, 338h і 2E0h, 238h відповідно. Для PS/2 стандартними для COM3 – COM8 є адреси 3220h, 3228h, 4220h, 4228h, 5220h і 5228h.

8.3.3. Інтерфейс USB

USB (Universal Serial Bus) – універсальна послідовна шина, є промисловим стандартом розширення архітектури ПК. Інтерфейс USB регламентується стандартами USB 1.1 і USB 2.0. Обмін даними по шині USB версії 1.1 виконується зі швидкістю до 12 Мбіт/с, а по шині USB версії 2.0 – зі швидкістю до 480 Мбіт/с. Привабливість інтерфейсу USB визначається

такими його особливостями: простотою кабельної системи і підключень, можливістю автоматичного підключення та налаштування пристроїв, підключенням/відключенням пристроїв без перезавантаження операційної системи.

Складові USB. Для реалізації інтерфейсу USB в ПК застосовується шинна технологія. Пристрої до ПК можуть бути підключені за кількома шинами USB, кожній з яких управляє окремий *хост-контролер* або *хост*, встановлений в системному блоці ПК (host – центральний або керуючий контролер в мережевому середовищі або середовищі з розподіленим опрацюванням даних). На шині USB допускається наявність тільки одного хоста. Пристрій може бути *хабом*, *функцією* або їх комбінацією і підключається через *порт*.

Хаб (hub – концентратор) – пристрій, що забезпечує додаткові порти на шині USB шляхом розгалуження. Архітектура шини USB допускає каскадне з'єднання до п'яти хабів. Хаб розпізнає підключення і відключення пристроїв до портів і може керувати подачею живлення на порти і обмежувати споживаний портом струм. Кожен порт може бути дозволений або заборонений і налаштований на повну або обмежену швидкість обміну. Хаб ізолює низькошвидкісні сегменти від високошвидкісних. *Кореневий хаб* – це хаб у складі хоста.

Функція – це периферійний пристрій (ПП) або окремий блок периферійних пристроїв, здатний передавати і приймати інформацію по шині USB. Кожна функція надає конфігураційну інформацію, що описує можливості ПП і вимоги до ресурсів. Перед використанням функція повинна бути налаштована хостом – для неї має бути виділена смуга частот в каналі і вибрані опції конфігурації.

Властивості хабів. Хаб виконує комутацію сигналів і видачу напруги живлення, а також відстежує стан підключених до нього пристроїв, повідомляючи хост про зміни. Хаб складається з двох частин – контролера (Hub Controller) і повторювача (Hub Repeater).

Контролер містить регістри для взаємодії з хостом. Доступ до регістрів здійснюється за специфічними командам звернення до хабу, які дозволяють конфігурувати хаб, управляти спадними портами і опитувати їх стан. Повторювач є керованим ключем, що з'єднує вихідний порт зі входним і має засоби скидання і припинення передачі сигналів.

Порти хабів можуть перебувати в наступних станах:

живлення вимкнено (Powered off) – порт не підключений до джерела живлення; вихідні буфери переводяться в високоімпедансний стан;

від'єднаний (Disconnected) – порт не передає сигнали ні в одному напрямку, але здатний виявити підключення пристрою;

заборонено (Disabled) – порт передає тільки сигнал скидання (за командою контроллера), сигнали від порту не сприймаються;

дозволено (Enabled) – порт передає сигнали в обох напрямках; по команді контроллера або після виявлення помилки порт переходить в стан заборонено, а після виявлення відключення – в стан відключений;

припинено (Suspended) – порт передає сигнал переведення в стан зупинки ("сплячий режим"), якщо хаб знаходиться в активному стані сигнали через порт не пропускаються ні в одному напрямку.

До хабу USB можуть підключатися інші хаби USB або пристрої з інтерфейсом USB. Використовуючи хаби USB можна створити мережу пристроїв підключених до хабів USB, один з яких повинен бути кореневим, підключеним до шини USB материнської плати ПК.

Для підключення пристроїв до хабів USB і зв'язків хабів між собою використовуються кабелі USB. Вони мають чотири дроти, два з яких призначені для живлення, а два інших – для передачі даних. На кінцях кабелів USB знаходяться роз'єми одного з двох типів – А і В. Роз'єми А застосовуються для підключення хабів (Upstream Connector). Вилки встановлюються на кабелях, що не від'єднуються від пристроїв (наприклад, клавіатура, миша і т.п.). Гнізда встановлюють на низхідних портах (Downstream Port) хабів. Роз'єми В (Downstream Connector) встановлюють на пристроях, від яких з'єднувальний кабель може від'єднуватися (принтери). Роз'єми А і В відрізняються механічно (рис. 8.24), що виключає неприпустимі петльові з'єднання портів хабів. Конструкція роз'ємів забезпечує більш пізнє з'єднання і раннє від'єднання сигнальних ланцюгів порівняно з ланцюгами живлення.

Для зв'язку ПК один з одним через інтерфейс USB потрібен спеціальний кабель USB, відмінний від кабелю, що використовується для підключення до ПК периферійних пристроїв. Справа в тому, що стандартний кабель USB містить провідники для подачі живлення на пристрої, які підключаються, і при з'єднанні кабелем ПК відбувається замикання двох джерел живлення, наслідком чого може бути вихід з ладу ПК.

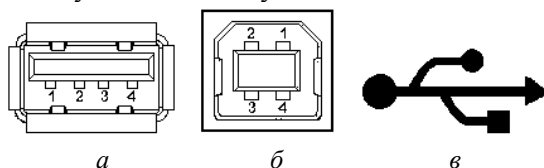


Рис. 8.24. Гнізда USB: *а* – роз'єми А, *б* – роз'єми В, *в* – символічне позначення гнізда

При плануванні архітектури власної мережі використання інтерфейсу USB виправдано лише для підключення окремих ПК один до одного, наприклад, ноутбука до настільного ПК, оскільки довжина кабелю USB

обмежена п'ятьма метрами.

Режими передавання даних. Смуга пропускання ділиться між усіма підключеними до шини пристроями. Шина USB має три режими передавання даних: низькошвидкісний (LS, Low-speed); повношвидкісний (FS, Full-speed) і високошвидкісний (HS, High-speed, тільки для USB 2.0).

Швидкість передавання даних USB-каналом дозволяє використовувати USB-шини для підключення вимірювальних приладів таких як генератори сигналів, цифрові осцилографи, логічні аналізатори тощо. У таких пристроях USB використовується як для передавання даних в ПК з метою подальшого опрацювання і відображення, так і для завдання параметрів приладів і режимів вимірювання.

Фізичний інтерфейс. Стандарт USB визначає електричні і механічні специфікації шини. Інформаційні сигнали і напруга живлення 5 В передаються по чотирипровідному кабелю. Використовується диференційний спосіб передачі сигналів D_+ і D_- по двох проводах. Рівні логічних сигналів передавачів в статичному режимі повинні бути нижче 0,3 В або вище 2,8 В. Передавачі повинні мати можливість переходити в високоімпедансний стан для двостороннього напівдуплексного передавання даних.

Крім диференціального приймача кожен пристрій має лінійні приймачі сигналів D_+ і D_- . Це дозволяє розрізняти більше двох станів лінії, що використовуються для організації апаратного інтерфейсу. Стани $Diff0$ і $Diff1$ визначаються за різницею потенціалів на лініях D_+ і D_- більше 200 мВ за умови, що на одній з них потенціал вище порога спрацьовування. Стан, за якого на обох входах присутній низький рівень, називається *лінійним нулем* (SE0 – Single-Ended Zero).

Шина USB має два режими передавання. Повна швидкість передавання становить 12 Мбіт/с, низька – 1,5 Мбіт/с. Для повної швидкості використовується екранована кручена пара з імпедансом 90 Ом і довжиною сегмента до 5м (рис. 8.25), для низької – нескручений неекранований кабель до 3м. Швидкість, яка використовується підключеним до порту пристроєм, визначається хабом за рівнями сигналів на лініях D_+ і D_- , зміщених резисторами навантажень $R2$ приймачів (у низькошвидкісному ресивері $R2$ підключається до лінії D_-). Кабель також має лінії V_{bus} і GND для передавання до пристроїв напруги живлення 5В.

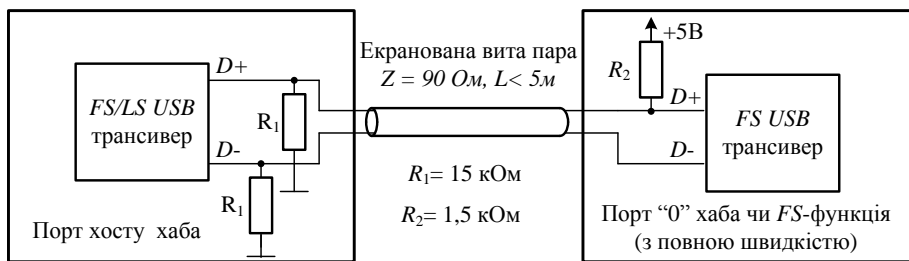


Рис. 8.25. Підключення повношвидкісного пристрою

Кодування даних. Всі дані кодуються за допомогою методу *NRZI with bit stuffing* (NRZI – Non Return to Zero Invert, метод без повернення до нуля з інвертуванням) (рис. 8.26 а). За такого кодування логічна 1 передається незмінним протягом бітового інтервалу рівнем, а логічний 0 – зміною рівня на протилежний. Якщо поточний біт має значення 0, то поточний потенціал є інверсією потенціалу попереднього біта незалежно від його значення; якщо поточний біт дорівнює 1, поточний потенціал повторює попередній.

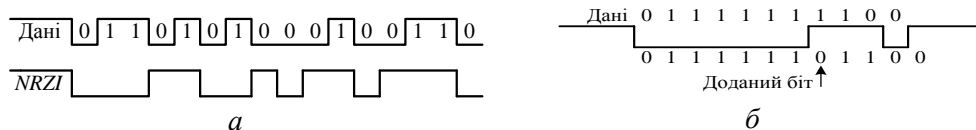


Рис.8.26. Кодування за методом NRZI: а – без стафінга, б – зі стафінгом

Кожному пакету передуює поле синхронізації SYNC, що дозволяє приймачу налаштуватися на частоту передавача.

Якщо дані містять нулі, то в системі досить легко підтримувати синхронізацію – рівень сигналу буде постійно змінюватися. Коли ж дані містять довгу послідовність одиниць, рівень сигналу не буде змінюватися і можлива розсинхронізація. Для надійної передачі даних з кодів виключають наддовгі послідовності одиниць. Ця дія називається *стафінгом* (Bit stuffing): після кожних шести одиниць автоматично додається 0 (рис.8.26 б). Стафінг може збільшити число переданих біт до 17%.

Принципи передавання даних. Механізм передавання даних є асинхронним і блоковим. Блок переданих даних називається *USB-фреймом* або *USB-кадром* і передається за фіксований часовий інтервал. Оперування командами і блоками даних реалізується за допомогою логічної абстракції, званої *каналом*. Зовнішній пристрій також поділяється на логічні абстракції, що називаються *кінцевими точками*.

Каналом (Pipe) в USB називають модель передачі даних між хост-контроллером і кінцевою точкою пристрою (рис.8.27) (логічну зв'язку між хост-контроллером і кінцевою точкою зовнішнього пристрою).

Розділ 8. Аналогові та цифрові інтерфейси

Існує два типи каналів: потоки (Stream) і повідомлення (Message). Потік доставляє дані від одного кінця каналу до іншого. Потік завжди однонаправлений. Потік може реалізовувати такі типи обміну: суцільний, ізохронний і переривання. Доставка завжди йде за типом FIFO.

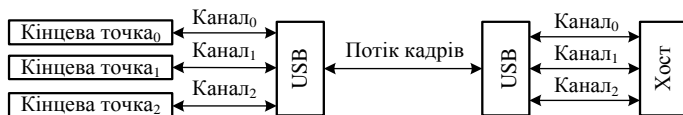


Рис. 8.27. Канали USB

Повідомлення мають визначений специфікацією LJSB формат. Хост посилає запит до кінцевої точки, після чого передається (приймається) пакет повідомлення, за яким слідує пакет з інформацією про стан кінцевої точки. Наступне повідомлення не може бути нормально направлено до закінчення опрацювання попереднього, але при відпрацюванні помилок можливе скидання необслуговуваних повідомлень.

Комунікаційні канали USB поділяються на два типи:

Потоковий канал (streaming pipe) доставляє дані від одного кінця каналу до іншого, він завжди однонаправлений. Передача даних в різних поточкових каналах не синхронізована. Це означає, що запити клієнтських драйверів для різних каналів, що поставлені в певному порядку один до одного, можуть виконуватися в іншому порядку. Якщо під час виконання запиту відбувається помилка, потік зупиняється.

Канал повідомлень (message pipe) є двонаправленим. Передавання повідомлень у зустрічних напрямках синхронізована між собою і строго впорядкована. На кожне повідомлення протилежна сторона зобов'язана відповісти підтвердженням його прийому і опрацювання.

Основними характеристиками каналів є смуга пропускання каналу, використовуваний каналом тип передавання даних, характеристики кінцевої точки – напрямок передавання даних і максимальний розмір пакету.

Смуга пропускання шини розділяється між усіма встановленими каналами. Виділена смуга закріплюється за каналом, і якщо установа нового каналу вимагає такої смуги, яка не вписується в уже існуючий розподіл, запит на виділення каналу відхиляється. Архітектура USB передбачає внутрішню буферизацію всіх пристроїв, причому, чим більшої смуги пропускання вимагає пристрій, тим більше повинен бути його буфер. Шина USB повинна забезпечувати обмін з такою швидкістю, щоб викликана буферизацією затримка даних у пристрої не перевищувала декількох мілісекунд.

Основні висновки

Аналоговий інтерфейс (AI) – це сукупність засобів вимірювальної

техніки, які є складовою частиною вимірювального каналу між датчиком і АЦП. АІ в ІВС виконують функції масштабного перетворення та фільтрації вимірювальних сигналів, їх комутації, гальванічної розв'язки передавальної і приймальної частин, лінеаризації характеристик сенсорів, ініціалізації пасивних сенсорів, передавання вимірювальної та службової інформації.

У багатоканальних ІВС зчитування всіх сигналів здійснюється їх почерговим підключенням за допомогою мультиплексорів, важливою характеристикою яких є максимальна швидкість опитування каналів.

За типом лінії зв'язку АІ поділяються на провідні і безпровідні (радіохвильові, акустичні, оптичні); за типом інформаційних сигналів – на АІ з передачею на постійному струмі, гармонічними і імпульсними сигналами; за інформативними параметрами – на АІ з передаванням напруги, струму, параметрами і характеристиками модульованих сигналів. Постійний струм використовується в АІ для передавання сигналів на відстань до сотень метрів.

Інтелектуальні сенсори оснащені вбудованими контролерами для узгодження з універсальними мережевими інтерфейсами або для об'єднання технологій передавання аналогової і цифрової інформації однією лінією зв'язку. В останньому випадку обмін цифровою інформацією здійснюється на основі використання стандартної мережі з HART-протоколом.

Для АІ нормують наступні метрологічні характеристики: функцію перетворення, похибку, чутливість, динамічні характеристики, імпеданс, характеристики лінії зв'язку. Рівень шуму визначає граничні можливості АІ правильно передавати вимірювальну інформацію. Вплив завод зменшується екрануванням, корекцією систематичних похибок, різними способами первинного і вторинного опрацювання.

Збільшення потоків вимірювальної інформації передбачає використання в ІВС обчислювальних ресурсів. Комп'ютерна система – це єдиний комплекс, що складається з електронних модулів і пакету програм. В основу будови комп'ютерних систем покладені ідеї Джона фон Неймана: в цифрових системах числа подаються двійковій системі числення; програми обчислень і дані у вигляді двійкових кодів зберігаються в одній пам'яті. За Нейманом основні блоки комп'ютерних систем – це пристрій управління, арифметико-логічний пристрій, пам'ять, зовнішня пам'ять, пристрої введення/виводу даних.

Архітектура ПК ґрунтується на магістрально-модульному принципі, згідно якого інформаційний зв'язок між пристроями відбувається через системну магістраль, що складається з шин для передачі адреси, даних і сигналів управління. Інформація по шинам передається за паралельно-послідовним принципом: блоками послідовно в часі, а в межах блоку – бітами, що надходять одночасно, паралельно в усіх розрядах шини. Продуктивність ПК – це кількість виконуваних в одиницю часу операцій, оцінюється числом команд, виконуваних ПК за одиницю часу (MIPS), або числом операцій з плаваючою комою за одну секунду (Flops).

Цифровий інтерфейс – це сукупність програмних і апаратних засобів,

Розділ 8. Аналогові та цифрові інтерфейси

призначених для передавання інформації між компонентами ПК та зовнішніми пристроями. За місцем інтерфейсів в структурі ПК вони поділяються на внутрішні (для з'єднання модулів ПК) і зовнішні (для підключення зовнішніх пристроїв). За способом передавання даних інтерфейси поділяються на паралельні, послідовні і спеціальні. У паралельних передавання кожного біта виконується по окремій сигнальній лінії, всі біти передаються одночасно. Ці інтерфейси використовують логічні рівні ТТЛ, що обмежує довжину кабелю через невисоку завадозахищеність. Гальванічна розв'язка відсутня.

Послідовні інтерфейси використовують одну сигнальну лінію, по якій інформаційні біти передаються послідовно в часі. У цих інтерфейсах застосовуються сигнали з рівнями, відмінними від рівнів ТТЛ, що дозволяє збільшити дальність зв'язку. У ряді послідовних інтерфейсів застосовується гальванічна розв'язка. Послідовне передавання даних може здійснюватися в асинхронному або синхронному режимах.

Інтерфейс USB (універсальна послідовна шина) відрізняється простотою кабельної системи, можливістю автоматичного підключення та налаштування пристроїв, підключення і відключення пристроїв без перезавантаження операційної системи. Зовнішні пристрої можуть бути підключені до ПК по декільком шинам USB, кожним з яких управляє окремий хост-контролер системного блоку ПК. Хост – це центральний (керуючий) контролер в мережевому середовищі або середовищі з розподіленим опрацюванням даних.

Пристрій може бути хабом, функцією або їх комбінацією і підключається через порт. Хаб або концентратор – пристрій, що забезпечує додаткові порти на шині USB шляхом розгалуження. Функція – це периферійний пристрій або окремий блок периферійних пристроїв, здатний передавати і приймати інформацію по шині USB.

Шина USB має три режими передавання даних: низкошвидкісний, повношвидкісний і високошвидкісний. Висока швидкість передавання даних дозволяє використовувати USB-шини для підключення вимірювальних пристроїв. Дані для передавання по шині USB кодуються методом NRZI – «без повернення до нуля з інвертуванням», в якому логічний 0 визначається як зміна напруги, а логічна 1 – як незміна напруги.

Каналом в USB називається модель передачі даних між хост-контролером і кінцевою точкою пристрою. Кінцева точка – це частина USB-пристрою, яка має унікальний ідентифікатор і є одержувачем або відправником інформації. У USB використовуються два типи каналів: потоки і повідомлення. Доставка даних йде в порядку «першим надійшов – першим вийшов». Механізм передачі даних є асинхронним і блоковим. Блок надісланих даних називається USB-фреймом або USB-кадром і передається за фіксований часовий інтервал.

Питання для самоконтролю

1. *Вкажіть місце і функції AI в структурі ІВС.*
2. *У чому полягає особливість комутації сигналів низького рівня?*
3. *У чому полягає сенс синхронного мультиплексування?*
4. *Чим визначається максимальна швидкість сканування каналів?*
5. *Перелічіть основні функції інтелектуальних інтерфейсів.*
6. *Як класифікуються AI за типом засобів зв'язку?*
7. *Дайте загальну характеристику різних топологій інтерфейсів.*
8. *Обґрунтуйте необхідність гальванічної розв'язки для AI та ЦІ.*
9. *Охарактеризуйте різні типи завод, що виникають в AI.*
10. *У чому полягає суть ідей, сформульованих Нейманом щодо архітектури обчислювальних машин?*
11. *У чому суть магістрально-модульного принципу організації ПК?*
12. *Які показники швидкодії ПК вам відомі?*
13. *У чому полягає різниця CISC і RISC архітектури комп'ютера?*
14. *Яка роль конвеєрної обробки мікрокоманд?*
15. *Які різновиди багатопроцесорних систем вам відомі?*
16. *Які відомі принципи організації та види комп'ютерної пам'яті?*
17. *Перелічіть основні характеристики комп'ютерної пам'яті.*
18. *Наведіть відмінність між внутрішніми і зовнішніми комп'ютерними інтерфейсами?*
19. *Перерахуйте основні характеристики ЦІ.*
20. *Які режими обміну забезпечують інтерфейси введення/виведення?*
21. *У чому переваги і недоліки паралельних інтерфейсів?*
22. *Базова адреса інтерфейсу і способи її визначення.*
23. *Назвіть основні принципи програмного управління введенням/виведенням інформації через паралельні порти ПК.*
24. *Які ви знаєте режими синхронізації послідовної передачі даних?*
25. *Охарактеризуйте переваги і недоліки інтерфейсу USB.*
26. *Назвіть основні складові інтерфейсу USB.*

Conclusions

Analogue interface (AI) is a set of measuring instruments that are an integral part of measuring channel between primary measuring transducer and ADC. AI in IMS perform the following functions: scale transformation of measurement signals; filtering of measurement signals; temperature compensation of thermocouples cold junctions; galvanic separation of transferring and receiving parts; linearization of PMT characteristics; passive sensors initialization; measurement signals multiplexing; measurement and service information transmission.

In multichannel IMS reading of all signals is carried by connecting them alternately using multiplexers. In sequential multiplexer signals switching occurs as a result of order from synchronization block. In random access multiplexers signals connection is performed to CPU' specified address. An important multiplexer

characteristic is channel polling maximum speed defined by the highest number of channels that can be connected to multiplexer output in one second.

By type of communication line AI are divided into wired and wireless (radio wave, acoustic, optical); by information signals type: AI with DC transmission, harmonic and pulse signals; by informative parameters - AI with transmission of voltage, current, parameters and characteristics of modulated signals. Direct current is used in AI when the distance between data collection system and object of measurement is in range from few meters to hundreds of meters.

Intelligent sensors are supplied with digital information transmission means by equipping them with built-in digital controllers to match the universal network interfaces or by combining technology of analog and digital information transmission in a single communication line. In the last case digital information exchange is carried out by using a standard network with HART-protocol.

Following metrological characteristics are rationed for AI: conversion function, accuracy, sensitivity, dynamic characteristics, impedance and characteristics of communication line. Noise level determines the limited possibility of AI to correctly transmit measurement information. Basic ways to reduce noise influence are: shielding, systematic error components correction (to reduce slow drift), different ways of pre-processing and secondary processing of signals.

Increasing flows of measuring information, complication of algorithms for its obtaining and processing suggests the use of computer systems computational resources in IMS. Computer system is a single complex, consisting of electronic modules and a software package designed to solve universal or specialized tasks. Construction basis of computer systems is based on John von Neumann ideas: for representation of numbers in digital systems it is necessary to use binary number system; calculation programs and data in binary form should be stored in a single memory. Basic blocks of computer systems according to Neumann are: control device, arithmetic logic unit, memory, external memory, input and output devices.

PC architecture is based on trunk-modular principle, according to which information communication between devices is carried out via system trunk, consisting of several buses that carry address, data and control signals. Information is transmitted through buses according to parallel-serial principle, i.e. sequentially over time in form of blocks; each block consists of bits, coming simultaneously, parallel over all bus bits. PC performance is a number of operations that it carries out in one time unit. There are two methods to evaluate computer performance. First uses a number of computer-executable instructions per time unit (MIPS), second - number of floating-point operations per second (Flops).

Digital interface is a set of hardware and software for information transmission between components. By interfaces place in PC structure they are divided into internal (connecting separate PC modules) and external (connecting external devices). By data transmission methods interfaces are divided into parallel, serial and special. In parallel interfaces every bit is transmitted via separate signal line, all bits are transmitted simultaneously. Galvanic separation is absent.

Serial interfaces use one signal line, through which information bits are transmitted sequentially over time. In these interfaces signals with levels different from TTL levels are used, which allows to increase communication range. In a number of serial interfaces galvanic separation is used. Serial data transfer can be performed in asynchronous or synchronous modes.

USB interface differs by simplicity of cable system and connectivity, an ability to automatically connect and configure devices, connection and disconnection of devices without rebooting the operating system. Devices could be connected to PC via several USB buses, each of them are operated by host controller, mounted in PC system block. Host is a central (operational) controller in network environment or distributed data processing environment.

Device may be a hub, function or a combination of them and is connected via port. Hub is a device that provides additional ports on USB bus by branching. Function is a peripheral device or a separate block of peripheral devices capable of transmitting and receiving information via USB bus.

USB bus has three modes of data transmission: low-speed, full-speed and high-speed. High speed data transmission of USB channel allows the use of USB bus for connection of measuring devices. Data for transmission via USB bus are encoded by NRZI – “Non Return to Zero Inverted”, where logical 0 is defined as voltage changed, and logical 1 – as voltage unchanged.

Channel in USB is a model of data transmission between host-controller and endpoint of device. Endpoint is a part of USB-device that has unique identifier and acts as information recipient or sender. USB uses two types of channels: streams and messages. Data delivery is always going in order "first in - first out» (FIFO). Message channel is a channel for data transmission, which structure is defined by USB specification.

Ключові слова

Українською	Англійською
аналоговий інтерфейс	analogue interface
мультиплексор	multiplexer
апаратне забезпечення	hardware
программне забезпечення	software
системна шина	system bus
цифровий інтерфейс	digital interface
послідовний інтерфейс	serial interface
паралельний інтерфейс	parallel interface
драйвер шини	bus driver



ПЕРЕДАВАННЯ ІНФОРМАЦІЇ В КАНАЛАХ ІВС

9

- 9.1. Системи передавання даних
- 9.2. Види інформаційних каналів, їх математичні моделі та характеристики
- 9.3. Швидкість передачі інформації в каналах зв'язку
- 9.4. Ситез елементів систем. Оптимальний приймач
- 9.5. Багатоканальні мережі передавання даних
- 9.6. Завадостійкість систем передавання інформації

9.1. Системи передавання даних

Важливо забезпечити вірність, цілісність, доступність інформації на всіх етапах опрацювання, зберігання, та передавання даних по інформаційним каналам зв'язку ІВС.

Сигнали, які передаються по каналам зв'язку, під впливом завад спотворюються. Внаслідок цього прийняте повідомлення лише в якійсь мірі точності відповідає переданому. Ступінь відповідності прийнятого повідомлення переданому називається *вірністю* інформації. Вірність інформації є одним із основних показників якості систем передавання даних. Вірність інформації тісно пов'язана з поняттям *завадостійкості*.

Канал зв'язку є сукупністю технічних засобів між джерелом повідомлення і споживачем. Технічні пристрої, що входять до складу каналу зв'язку, призначені для того, щоб повідомлення дійшли до споживача найкращим чином – для цього сигнали опрацюють та перетворюють згідно визначених алгоритмів. Такими перетвореннями сигналу є, наприклад, модуляція і перетворення неперервних сигналів в дискретні. Відповідно до *типу сигналів*, канали поділяють на *безперервні* та *дискретні*.

Перетворення дискретного повідомлення в інформаційний сигнал каналу зв'язку, здійснюється у вигляді операцій – *форматування, кодування і модуляції*. *Кодування* є перетворенням повідомлень в послідовність кодових символів, а *модуляція* – перетворенням цих символів в сигнали, придатні для передавання по інформаційному каналу.

При кодуванні відбувається процес перетворення елементів повідомлень у відповідні їм числа (кодові символи). Кожному елементу повідомлення присвоюється визначена сукупність кодових символів, яка називається *ковою комбінацією*. Сукупність кодових комбінацій утворює код. Основні показники якості дискретних систем зв'язку наведені на рис. 9.1.

В результаті перетворень неперервного сигналу, які називають *дискретизацією* і *квантуванням* отримують відліки, що розглядаються, як числа в тій або іншій системі числення, що є дискретними сигналами. Ці числа перетворюють в кодові комбінації електричних сигналів, які і передають по лініям зв'язку як неперервні. При використанні носієм інформації постійного струму отримують послідовність відеоімпульсів. При необхідності цю послідовність модулюють гармонічним коливанням і отримують послідовність радіоімпульсів.



Рис. 9.1. Основні показники якості дискретних систем зв'язку

У загальному вигляді модель системи передавання даних представлена на рис. 9.2. Хоча ця модель і містить основні елементи, властиві будь-якій системі передавання інформації, вона може служити лише простою ілюстрацією до опису *інформаційної системи передавання даних*, оскільки практично не відображає тих дій, які повинні (або можуть) виконуватися над інформацією в процесі її передавання від джерела до споживача.

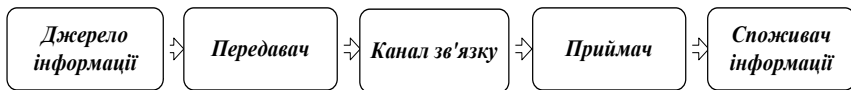


Рис. 9.2. Узагальнена схема передавання даних

Значно повнішою в цьому сенсі є модель системи передавання (і зберігання) інформації приведена на рис. 9.3.

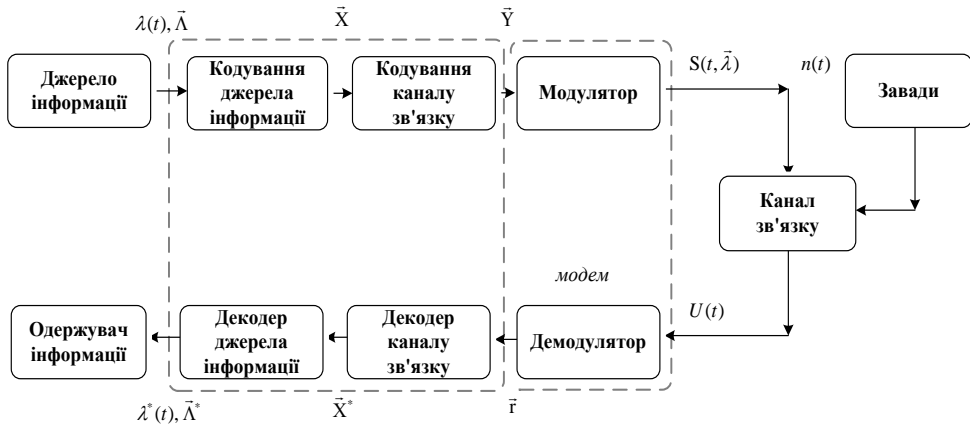


Рис. 9.3. Структурна схема системи передавання інформації

Слід відзначити, що насправді проблеми, що виникають при передаванні (причому не тільки з використанням радіохвиль) і зберіганні інформації (на оптичних дисках, магнітних носіях і в пам'яті комп'ютерів) схожі, тому методи їх рішення і структура технічних пристроїв також багато в чому ідентичні.

Охарактеризуємо призначення і функції елементів цієї моделі.

1. Джерело інформації або повідомлення - це фізичний об'єкт, система або явище, що формують повідомлення для його передачі від джерела повідомлення до споживача інформації.

Повідомлення – це значення або зміна деякої фізичної величини, що відображає стан об'єкту (системи або явища). Як правило, первинні повідомлення – мова, музика, зображення, вимірювання параметрів навколишнього середовища і т.д. – є функціями часу неелектричної природи. З метою передачі по каналу зв'язку ці повідомлення перетворюються в електричний сигнал, зміни якого в часі $\lambda(t)$ відображає повідомлення, що передається. Значна частина таких повідомлень, за своєю природою не є сигналами – це масиви чисел, текстові або інші файли тощо. Повідомлення такого типу можна подати у вигляді деяких векторів $\bar{\Lambda}$.

2. Кодування джерела інформації. Переважна частина інформаційних повідомлень містять в собі надлишкову інформативність та мають повільно-змінну частоту, що в загальному випадку погано пристосована для її ефективного передавання по каналах зв'язку. Тому повідомлення ($\lambda(t)$ або $\bar{\Lambda}$), як правило, піддаються кодуванню. Процедура кодування включає попередньо дискретизацію безперервного повідомлення $\lambda(t)$, тобто його перетворення в послідовність елементарних дискретних повідомлень $\{\lambda_i\}$.

Кодування джерела інформації може переслідувати різні цілі: скорочення обсягу переданих даних (стиснення даних), збільшення кількості переданої за одиницю часу інформації, підвищення достовірності передавання, забезпечення секретності при передаванні і т.д.

Таким чином, на виході *кодера джерела* за переданим повідомленням $\lambda(t)$ або $\bar{\Lambda}$ формується послідовність кодових символів \bar{X} , яка називається *інформаційною послідовністю*, що допускає точне або наближене відновлення початкового інформаційного повідомлення.

3. Кодування в каналах зв'язку. При передаванні інформації по каналу зв'язку з завадами в прийнятих даних можуть виникати помилки. Якщо такі помилки невеликі або виникають достатньо рідко, інформація може бути використана споживачем. При великому числі помилок, отриманою інформацією користуватися не можливо.

Кодування в каналі, або завадостійке кодування, це спосіб оброблення переданих даних, що забезпечує зменшення кількості помилок, що виникають в процесі передачі по каналу з завадами. Існує багато різних методів завадостійкого кодування інформації, але всі вони базуються на наступному: при завадостійкому кодуванні до переданих повідомлень вноситься спеціальним чином організована надлишковість (у кодові послідовності, що передаються, вводяться надлишкові символи), що дозволяє на приймальній стороні виявляти і виправляти виникаючі помилки. Таким чином, якщо при кодуванні джерела проводиться усунення природної надлишковості, що має місце в повідомленні, то при кодуванні в каналі надлишковість до переданого повідомлення свідомо вноситься. На виході кодера каналу в результаті формується послідовність кодових символів $Y(X)$, яка називається *ковою послідовністю каналу зв'язку*.

Слід відзначити, що як завадостійке кодування, так і стиснення даних не є обов'язковими операціями при передаванні інформації. Ці процедури (і відповідні блоки в структурній схемі системи передавання інформації) можуть бути відсутніми. Проте це може привести до істотних втрат в завадостійкості системи, зменшенню швидкості передавання і зниженню якості передачі інформації. Тому всі сучасні системи включають і ефективно і завадостійке кодування даних.

4. Модулятор. *Функції модулятора в інформаційних системах передавання даних - це узгодження повідомлення джерела або кодових послідовностей, що виробляються кодером, з властивостями каналу зв'язку і забезпечення можливості одночасного передавання великого числа повідомлень по загальному інформаційному каналу зв'язку.*

Дійсно, більшість безперервних $\lambda(t)$ і дискретних $\vec{\Lambda}$ повідомлень, що підлягають передаванню, а також результати їх кодування - послідовності кодових символів \vec{X} і \vec{Y} - є порівняно низькочастотними сигналами з широкою смугою ($\Delta F \leq 1\text{МГц}$, $\Delta F \sim f_0$). В той же час ефективна передача з використанням електромагнітних коливань (радіохвиль) можлива лише для достатньо високочастотних сигналів ($f_0 \geq 1 \dots 1000 \text{ МГц}$) з відносно вузькосмуговими спектрами ($\Delta F \ll f_0$). Тому модулятор повинен перетворити повідомлення джерела $\lambda(t)$ або \vec{X} відповідні їм кодові послідовності \vec{X} і \vec{Y} в сигнали $S(t, \lambda(t))$, $S(t, Y(\lambda(t)))$ (накласти повідомлення на сигнали), властивості яких забезпечували б її можливість ефективного передавання по радіоканалу (або іншим каналам зв'язку - телефонним, оптичним і т.д.).

Існує значна кількість методів модуляції сигналів, що забезпечують передавання інформації з різною ефективністю та якістю. Найпростішими з них є амплітудна, частотна і фазова модуляції неперервних сигналів.

5. Канал зв'язку – це система передавання інформації, що використовує як носій від джерела до споживача електромагнітні хвилі або радіохвилі, а як середовище розповсюдження - навколишній простір або радіоканал.

Розглянемо інформаційний канал зв'язку у вигляді ланки радіотехнічної системи передавання інформації, на вхід якого поступає сигнал передавача $S(t, \lambda(t))$, а на виході виходить сигнал, який зазвичай називають прийнятим коливанням..

Існує багато моделей інформаційних каналів більшої або меншої складності, проте в загальному випадку сигнал $S(t, Y(\lambda(t)))$, проходячи по каналу зв'язку, піддається ослабленню, набуває деякої тимчасової затримки (або фазове зрушення) і зашумлюється. Коливання $U(t)$, що приймається, в цьому випадку матиме вигляд

$$U(t) = \varepsilon S(t - \tau, Y(\lambda(t))) + n(t) \quad (9.1)$$

де ε - згасання, τ - тимчасове запізнювання, $n(t)$ - завади в каналі зв'язку.

6. Одержувач інформації. Функції одержувача системи передавання інформації – з максимально можливою точністю по прийнятому коливанню $U(t)$ відтворити на своєму виході передане повідомлення $\lambda(t)$ або $\vec{\Lambda}$.

Прийняте (відтворене) повідомлення через завади в загальному випадку відрізняється від переданого. Прийняте повідомлення називатимемо оцінкою

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

(мається на увазі оцінкою повідомлення) і позначатимемо тим же символом, що і відправлене повідомлення, але із знаком *: $\lambda^*(t)$ або $\bar{\lambda}^*$.

7. Демодулятор. Для відтворення оцінки повідомлення $\lambda^*(t)$ або $\bar{\lambda}^*$ приймач системи повинен по прийнятому коливанню $U(t)$ і з урахуванням відомостей про використані при передачі види сигналу і способи модуляції отримати оцінку кодової послідовності $Y(\lambda^*(t))$, яка називається прийнятою послідовністю \vec{r} . Ця процедура називається демодуляцією, детектуванням або прийомом сигналу. При цьому демодуляція повинна виконуватися так, щоб прийнята послідовність \vec{r} в мінімальному ступені відрізнялася від переданої кодової послідовності \vec{Y} . Завдання демодуляції прийнятого коливання $U(t)$ співпадає з різними варіантами завдання оптимального прийому сигналу на фоні завад (оптимальне виявлення, оптимальне розрізнення двох або декількох сигналів і т.д.).

8. Декодер каналу зв'язку. Прийняті послідовності \vec{r} в загальному випадку можуть відрізнятися від переданих кодових слів \vec{Y} , тобто містити помилки. Кількість таких помилок залежить від рівня завад в каналі зв'язку, швидкості, вибраної для передавання сигналу і способу модуляції, а також від способу прийому (демодуляції) коливання $U(t)$.

Функція декодера каналу зв'язку – виявити і, по можливості, виправити помилки що виникли при впливі завад та відтворити з максимальним наближенням переданий сигнал від джерела інформації до споживача. Процедура відтворення повідомлення, виявлення і виправлення помилок в прийнятій послідовності \vec{r} називається декодуванням каналу.

Результатом декодування \vec{r} є оцінка інформаційної послідовності \vec{X}^* . Вибір завадостійкого коду, способу кодування, а також методу декодування повинні проводитися так, щоб на виході декодера каналу залишилося якомога менше невиправлених помилок.

Питанням завадостійкого кодування/декодування в системах передавання (і зберігання) інформації приділяється велика увага, оскільки цей прийом дозволяє істотно підвищити якість її передавання. У багатьох випадках, коли вимоги до достовірності інформації, що приймається, високі (у комп'ютерних мережах передавання даних, в дистанційних системах управління і т.п.), передавання без завадостійкого кодування взагалі неможлива.

9. Декодер джерела інформації. Оскільки інформація джерела ($\lambda(t)$, $\bar{\lambda}$) в процесі передавання піддавалася кодуванню з метою її компактнішого

(або зручнішого) представлення (стиснення даних, економне кодування, кодування джерела), необхідно відновити її до результуючого (або початкового вигляду) за прийнятою послідовністю \vec{X}^* . Процедура відновлення $\vec{\Lambda}^*$ по \vec{X}^* називається *декодуванням джерела* і може бути зворотною операції кодування (неруйнівне кодування/декодування), або відновлювати наближене значення $\vec{\Lambda}^*$, що більшою чи меншою мірою відрізняється від $\vec{\Lambda}$ (руйнуюче кодування/декодування). До операції відновлення $\vec{\Lambda}^*$ по \vec{X}^* відноситимемо також відновлення, якщо в цьому є необхідність, безперервної функції $\lambda^*(t)$ по набору дискретних значень оцінок $\vec{\Lambda}^*$.

Таким чином, коротко розкривши загальну структуру інформаційної системи передавання даних, перейдемо до детальнішого вивчення її основних елементів.

9.2. Види інформаційних каналів, їх математичні моделі та характеристики

Класифікація каналів за частотним розподілом. У методах та засобах передавання даних знаходять застосування механічні, акустичні, оптичні, електричні й радіоканали, що розрізняються за технічними характеристикам та фізичною природою сигналів. Основною ознакою кожного з таких видів каналів звичайно служить діапазон робочих частот. Класифікація каналів за частотним діапазоном приведена на рис. 9.4, а в табл. 9.1 подані також відповідні дані по частотним діапазнам.



Рис. 9.4. Класифікація інформаційних каналів

Таблиця 9.1

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

Вид каналу	Класифікаційні ознаки	Різновиди каналів	Частота сигналів f	Довжина хвилі λ
Механічний	Середовище передачі	Жорсткі Гідравлічні Пневматичні		
Акустичний	Діапазон частот	Звукові Ультразвукові	$< 20 \text{ кГц}$ $> 20 \text{ кГц}$	$> 15 \text{ км}$ $< 15 \text{ км}$
Оптичний	Діапазон частот	Видимого спектра Інфрачервоне випромінювання	$400\text{-}1000 \text{ ГГц}$ $0,3\text{-}400 \text{ ГГц}$ $1000\text{-}3000 \text{ ГГц}$	$0,3\text{-}0,75 \text{ мкм}$ $0,75\text{-}1000 \text{ мкм}$ $0,01\text{-}0,3 \text{ мкм}$
Електричний	Діапазон частот	Ультрафіолетове випромінювання Підтональних частот Тональних частот	200 ГГц $300\text{-}3400 \text{ Гц}$ $4000\text{-}8500 \text{ Гц}$ $> 10 \text{ кГц}$	$> 1500 \text{ км}$ $90\text{-}1000 \text{ км}$ $32\text{-}75 \text{ км}$ $< 30 \text{ км}$
Радіо	Діапазон частот	Надтональних частот Високочастотні Довгохвильового діапазону Середньохвильового діапазону Проміжний діапазон Короткохвильового діапазону Ультракороткохвильового діапазону	$< 300 \text{ кГц}$ $300\text{-}1500 \text{ кГц}$ $1,5\text{-}6 \text{ МГц}$ $6\text{-}30 \text{ МГц}$ $30\text{-}30 \cdot 10^5 \text{ МГц}$	$> 1000 \text{ м}$ $200\text{-}1000 \text{ м}$ $50\text{-}200 \text{ м}$ $10\text{-}50 \text{ м}$ $0,0001\text{-}10 \text{ м}$

Механічний канал передавання даних. Механічні канали застосовують для передачі на короткі відстані (до 500 м) сигналів у вигляді механічних зусиль або тисків. Застосовуються наступні різновиди механічних каналів:

Жорсткі або власне механічні канали. Найпростішим прикладом служить трос керування дросельною заслінкою карбюратора. Довжина таких каналів може доходити до декількох десятків метрів.

Гідравлічні канали, у яких передавальним середовищем служить рідина. Їхнім недоліком є погані динамічні властивості, у зв'язку із чим довжина цих каналів обмежується декількома метрами.

Пневматичні канали. По пневматичних каналах передається сигнал у вигляді тиску. Середовищем для передачі звичайно служить повітря. Довжина пневматичних каналів досягає декількох сотень метрів.

Акустичні канали передавання даних. Акустичні канали призначаються для передавання коливань. Середовищем для передавання можуть служити будь-які звукопровідні матеріали й середовища.

Акустичні сигнали й канали знайшли різноманітне застосування в техніці автоматичного контролю, виявлення й зв'язку: акустичний контроль стану працюючих механічних об'єктів, ультразвукова дефектоскопія, акустичне виявлення об'єктів (підводних човнів, літаків), гідролокація,

акустичний зв'язок і ін. (в табл. 9.2 наведені довідкові дані щодо швидкості поширення звукових хвиль у різних середовищах).

Таблиця 9.2

Середовище	Швидкість поширення звукових хвиль, м/с	Середовище	Швидкість поширення звукових хвиль, м/с
Повітря	331,45	Дерево	3350
Вода прісна	1430	Скло	5400
Вода морська	1500	Сталь	6100

При пасивному передаванні сигналів у процесі контролю або виявлення джерелами звуку є контрольовані або виявлені об'єкти. При активному передаванні (ультразвукова дефектоскопія, локація, зв'язок) акустичні сигнали створюються спеціальними генераторами.

Оптичний канал передавання даних. За діапазоном використовуваних частот (або довжиною хвилі) оптичні канали розділяють (див. табл. 9.1) на наступні групи:

канали видимої частини спектра оптичних сигналів (з довжиною хвилі $0,3 < \lambda < 75 \text{ мкм}$);

канали інфрачервоної частини спектра ($0,75 < \lambda < 1000 \text{ мкм}$);

канали ультрафіолетової частини спектра ($\lambda < 0,3 \text{ мкм}$).

Пристрої, що працюють із інфрачервоним випромінюванням, знайшли широке застосування через ряд переваг перед видимим і ультрафіолетовим випромінюванням:

менше ослаблення інфрачервоного випромінювання атмосферою в порівнянні з випромінюванням видимої й ультрафіолетової частини спектра;

поширення інфрачервоного випромінювання в темряві, скритність передавання.

Перспективи в техніці передачі інформації має застосування квантових генераторів світла – лазерів. Розроблені численні конструкції кристалічних і газових лазерів, що працюють у різних частинах оптичного діапазону. Як приймачі використовуються фоторезистори, фотодіоди й фотомножники.

Електричний канал. Електричні канали – це канали із застосуванням комутативних (провідних) мереж зв'язку. Для передавання інформації використовуються, як спеціально виділені мережі, так і мережі, споруджені для інших цілей, наприклад, широко застосовуються мережі енергопостачання, високовольтні мережі електропередачі, телекомунікаційні мережі.

Шкала частот яку, займають сигналами в електричному каналі зв'язку, умовно ділиться на чотири діапазони:

підтональні частоти 0 - 200 Гц;

тональні частоти 300 - 3400 Гц;

надтональні частоти 4000 - 8500 Гц;

високі частоти понад 10 кГц.

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

Радіоканали передавання даних. За діапазоном частот радіосигналів розрізняють наступні канали:

довгохвильового діапазону (300-30 кГц.)

середньохвильового діапазону (3-0,3 мГц.);

короткохвильового діапазону (3-30 мГц.);

ультракороткохвильового діапазону (< 20 мГц.).

Класифікація і способи поширення радіохвиль наведені в табл. 9.3 і 9.4.

Таблиця 9.3

Вид радіохвиль	Тип радіохвиль	Діапазон радіохвиль (довжина хвилі)	Номер діапазону	Діапазон частот	Вид радіочастот
Міріаметрові	Наддовгі	10..100 км	4	3...30 кГц	Дуже низькі (ДНЧ)
Кілометрові	Довгі	1..10 км	5	30...300 кГц	Низькі (НЧ)
Гектометрові	Середні	100..1000 м	6	300..3000 кГц	Середні (СЧ)
Декаметрові	Короткі	10..100 м	7	3..30 МГц	Високі (ВЧ)
Метрові		1..10 м	8	30..300 МГц	Дуже високі (ДВЧ)
Дециметрові	Ультракороткі	10..100 дм	9	300...3000 МГц	Ультрависокі (УВЧ)
Сантиметрові		1..10 дм	10	3...30 ГГц	Надвисокі (НВЧ)
Міліметрові		1..10 мм	11	30...300 ГГц	Вкрайвисокі (КВЧ)
Дециміліметрові		0.1..1 мм	12	300...3000 ГГц	Гіпервисокі (ГВЧ)

Таблиця 9.4

Вид радіохвиль	Основні способи поширення радіохвиль	Дальність зв'язку
Міріаметрові і кілометрові (наддовгі і довгі)	Дифракція Відбиття від Землі й іоносфери	До тисячі км Тисячі км
Гектометрові (середні)	Дифракція Переломлення в іоносфері	Сотні км Тисячі км
Декаметрові (короткі)	Преломлення в іоносфері і відображення від Землі	Тисячі км
Метрові і більш короткі	Вільне поширення і відображення від Землі Розсіювання в тропосфері	Десятки км Сотні км

Напруженість поля в точці прийому залежить також від довжини хвилі, освітленості земної атмосфери сонцем і ряду інших факторів.

Розподіл радіохвиль на діапазони встановлено Міжнародним регламентом радіозв'язку МСЕ-Р. Радіохвилі, що випромінюються передавальною антеною, перш ніж потрапити на приймальну антену,

проходять у загальному випадку складний шлях. На величину напруженості поля в точці прийому впливає ряд факторів:

- відбиття електромагнітних хвиль від поверхні Землі;
- переломлення (заломлення) в іонізованих шарах атмосфери (іоносфері);
- розсіювання на діелектричних неоднорідностях нижніх шарів атмосфери (тропосфері);
- дифракція на сферичній поверхні Землі.

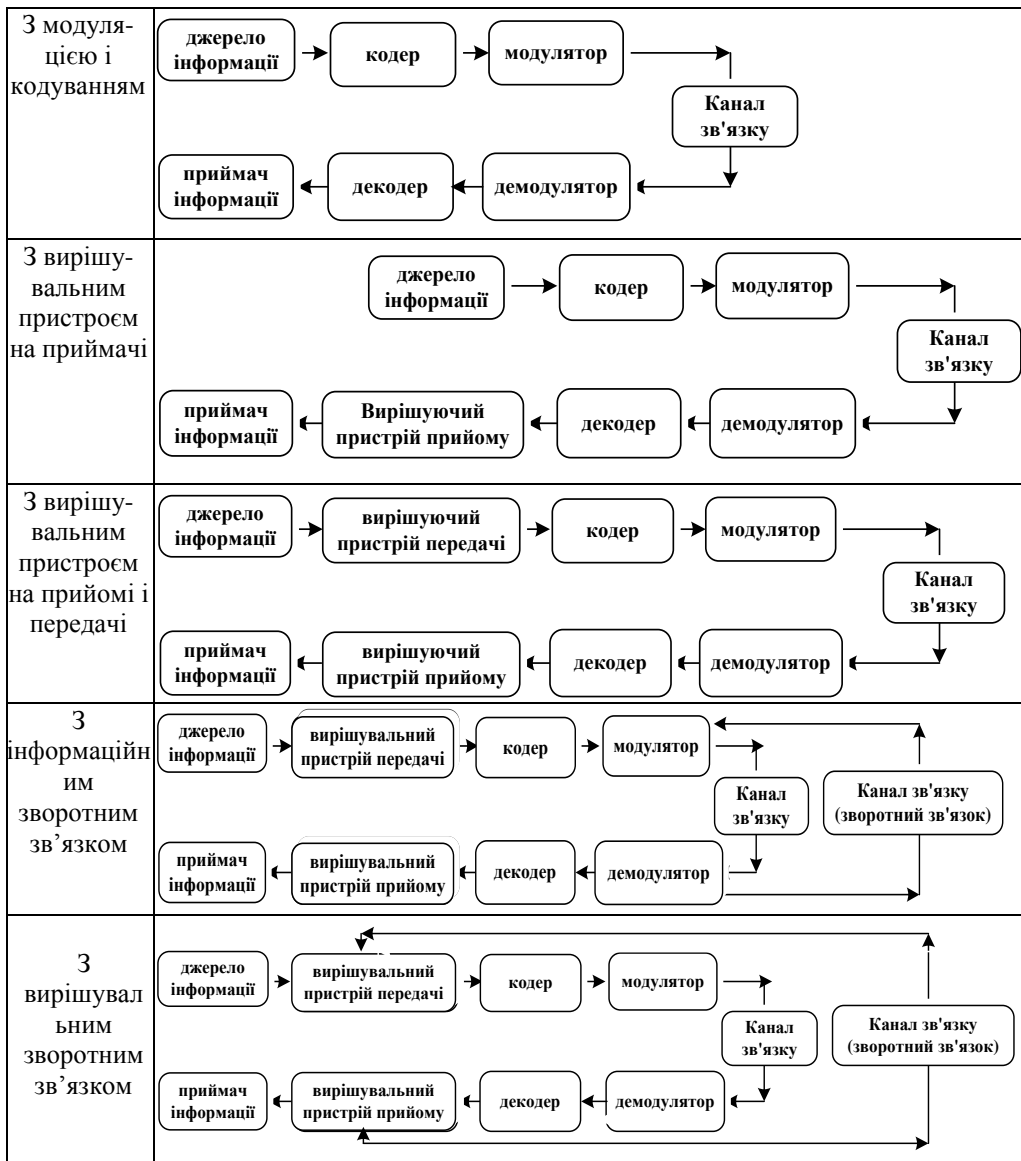
Класифікація каналів за структурою. Канал передавання інформації складається з мережі зв'язку, модулятора й демодулятора (крім випадку, коли для передавання використовується проста модуляція, при якій сигнал у мережі збігається із сигналом сенсору, що кодує й декодує, а також вирішувальних пристроїв, що дозволяють із високим ступенем достовірності прийняти й передати повідомлення, для збільшення надійності передавання застосовуються також канали зворотного зв'язку. Варіанти структур каналів наведені в табл. 9.5.

Вирішувальні пристрої служать для класифікації сумнівних сигналів, ототожнюючи їх з досить високим ступенем ймовірності зі станом джерела інформації або з певним кодом. Кількість інформації, що міститься в окремому сигналі, що має ймовірність p становить $\log_2(1/p)$ біт і може бути значно більшим. Однак у середньому в досить довгому повідомленні двійковий сигнал переносить не більше одного біта інформації. У загальному випадку в повідомленні, що складається із символів алфавіту ємністю h , середня кількість інформації не перевищує однієї одиниці інформації цього алфавіту на символ. Проте не всі сигнали несуть корисну інформацію. Імпульс завади, що випадково виникає у мережі, інформації не несе й заважає передаванню. Іноді

Таблиця 9.5

Вид структур	Схема
Елементарні	
З модуляцією	

Розділ 9. Передавання вимірювальної інформації в каналах ІВС



додаткові (надлишкові) сигнали вводяться спеціально для підвищення завадостійкості.

У цих випадках середня кількість інформації, що переноситься поодиноким сигналом, зменшується. Випадок, коли повідомлення складається із символів, що несуть максимальну кількість інформації, зустрічається рідко, виникає він при відсутності завад у мережі, при рівномірному розподілі ймовірностей між символами та при відсутності статистичної залежності між послідовно переданими символами. При невиконанні однієї із цих умов кількість інформації виявляється меншою.

При заданому характері й рівні завод завдання підвищення ефективності передавання вирішується насамперед шляхом певного перетворення, що приводить до збільшення середньої кількості інформації.

Безпосереднім носієм інформації є сигнал, що випадково змінюється. Кожен сигнал має певну тривалість. Тривалість сигналу характеризує час передавання повідомлень, тривалість зайнятості інформаційного каналу, тобто *час передачі сигналу* T_c .

Кожен сигнал характеризується певним частотним спектром. Теоретично ширина спектра сигналу кінцевої тривалості необмежена. Однак вивчення спектрів реальних сигналів показує, що їхня спектральна щільність спадає з ростом частоти. Це дозволяє за певних умов розглядати сигнали як процеси з обмеженим спектром. Існують різні критерії обмеження спектра сигналу. Одним з таких критеріїв є *припустимі перекручування сигналу*. Наприклад, при передачі мовного сигналу розбірливість і якість мови практично повністю зберігаються при ширині спектра від 300 до 3400 Гц.

Таким чином, другою узагальненою характеристикою сигналу повинна бути *ширина частотного спектру* F_c .

Третьою важливою характеристикою сигналу є його енергетична характеристика - *середня потужність* P_c .

Однак оскільки при передаванні на сигнали завжди впливають завади, то як енергетичну характеристику сигналу доцільно брати відношення середньої потужності сигналу P_c до середньої потужності завади P_z .

Динамічним діапазоном інформаційного каналу – називають логарифмічну міру виражену у відношенні середньої потужності сигналу P_c до середньої потужності завади P_z : $D = \log_2 P_c / P_z$. (9.1)

При оцінці інформаційної змістовності зручно виражати динамічний діапазон через логарифм із основою 2.

Ємністю сигналу називається – добуток часу передачі сигналу T_c , ширини частотного спектра F_c та середньої потужності P_c визначеного сигналу: $V_c = T_c F_c P_c$. (9.2)

У геометричному представленні об'єм сигналу має вигляд паралелепіпеда з ребрами T_c , F_c й P_c (рис. 9.4).

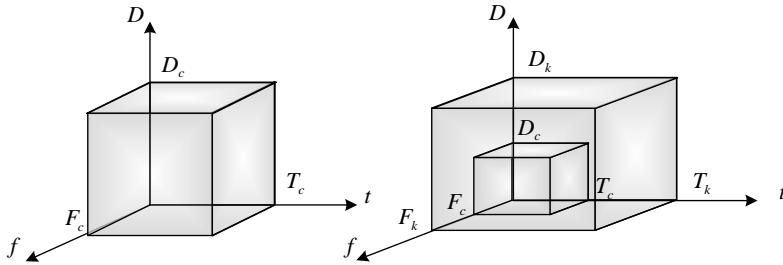


Рис. 9.4 . Геометрична модель ємності сигналу та каналу зв'язку
 Інформаційний канал характеризують також трьома параметрами: часом використання каналу T_k , шириною смуги частот, пропусканням каналу F_k , і динамічним діапазоном каналу D_k , що показує його здатність передавати різні рівні сигналу.

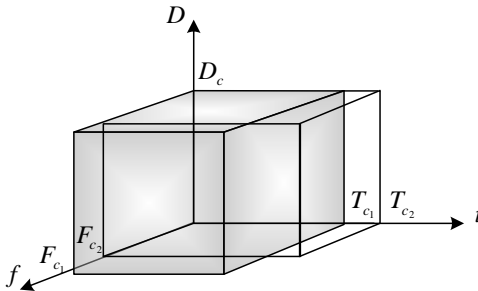


Рис. 9.5. Геометричне подання трансформації сигналу за параметрами F_c і T_c

Ємністю каналу – називається добуток часу зайнятості каналу T_k , ширини частотного спектра каналу зв'язку F_k та динамічного діапазону каналу D_k :

$$V_k = T_k F_k D_k. \quad (9.3)$$

Неспотворене передавання сигналів можливе лише за умови, що сигнал за своїм обсягом

еквівалентний чи менший за ємність каналу зв'язку (вміщується в ємність каналу (рис. 9.5)).

Загальна умова узгодження сигналу з каналом передавання інформації визначається співвідношенням:

$$V_c \leq V_k. \quad (9.4)$$

Однак співвідношення (9.4) відтворює необхідну, але недостатню умову узгодження сигналу з каналом.

Достатня умова узгодження за всіма параметрами:

$$T_c \leq T_k, F_c \leq F_k, D_c \leq D_k. \quad (9.5)$$

Якщо при виконанні умови (9.4) не забезпечується частина умов (9.5), то узгодження можна домогтися трансформацією сигналу при збереженні його об'єму.

Наприклад, якщо відсутнє узгодження мовного сигналу з каналом тільки по частоті, тобто мають місце співвідношення: $T_c < T_k$, $F_c > F_k$, $D_0 < D_k$,

то узгодження по частоті при передачі можна досягти, записуючи сигнал на стаціонарний накопичувач інформації з однією швидкістю й відтворюючи його з меншою швидкістю.

Припустимо, що при виконанні умов $V_c \leq V_k$ і $D_c \leq D_k$ частотний спектр сигналу в n раз ширше смуги пропускання каналу $F_c = nF_k$.

Для узгодження сигналу з каналом можна записати сигнал на стаціонарний накопичувач інформації зі швидкістю u_1 , а передавати зі швидкістю u_2 , у n раз меншого від швидкості u_1 . При цьому тривалість сигналу збільшується в n раз і в стільки ж раз зменшується ширина його спектра. Ємність сигналу при цьому залишається незмінною.

Швидкістю передавання інформації – називається середня кількість інформації, передана по каналу зв'язку за одиницю часу.

У загальному випадку швидкість передавання інформації залежить від тривалості передавання повідомлень T . При досить довгих повідомленнях швидкість передавання залишається постійною. У зв'язку із цим аналітично швидкість передавання інформації визначають як:

$$C = \bar{I}(Z, Y) = \lim_{T \rightarrow \infty} I(Z, Y) / T, \quad (9.6)$$

де $I(Z, Y)$ – кількість інформації, передана сигналом тривалістю T .

Пропускною здатністю каналу – називається максимально теоретично досяжна для даного каналу швидкість передавання інформації.

Пропускна здатність інформаційного каналу: $C = \max \{ \bar{I}(Z, Y) \}$.

Швидкість передавання інформації в загальному випадку залежить від статистичних властивостей повідомлення, методу кодування й властивостей каналу. Пропускна здатність - це характеристика каналу. Вона не залежить від фактичної швидкості передавання інформації.

З метою найбільш ефективного використання інформаційного каналу необхідно приймати міри до того, щоб швидкість передавання інформації була як найближчою до пропускну здатності каналу.

Разом з тим швидкість введення інформації в канал не повинна перевищувати пропускну здатність каналу, інакше не вся інформація буде передана по каналу.

Швидкість введення інформації – аналітичне співвідношення середньої кількості інформації, що закладена в інформаційному потоці на вході каналу до тривалості повідомлення:

$$\bar{I}(X) = \lim_{T \rightarrow \infty} I(X)/T, \quad (9.7)$$

де $I(X)$ — середня кількість інформації, що закладена в повідомленні на вході каналу; T — тривалість повідомлення. Таким чином, повинна бути здійснена *основна умова динамічного узгодження* інформаційного потоку джерела повідомлень й інформаційного каналу:

$$\bar{I}(X) \leq C. \quad (9.8)$$

Одним з основних питань у теорії передавання інформації є визначення залежності швидкості передавання інформації й пропускної здатності від параметрів каналу й характеристик сигналів і завад. Ці питання були вперше глибоко досліджені Клодом Шенноном

Розглянемо три види каналів: дискретний канал без завад, дискретний канал з завадами й безперервний канал з завадами.

Завади в безперервному каналі. *Завадою* називається будь-який небажаний процес або дія що впливає на сигнал та ускладнює його достовірний прийом.

Для розгляду завад сигнал на виході каналу подають у вигляді:

$$y(t) = x(t)n_m(t) + n_a(t),$$

де $x(t)$ — чистий або корисний інформаційний сигнал на вході передавача, $y(t)$ — сигнал на вході приймача, $n_m(t)$ — *мультиплікативна* завада, обумовлена випадковими змінами параметрів каналу, $n_a(t)$ — *адитивна* завада, що має розмірність сигналу і складається з сигналом. Зазвичай вважають, що адитивна завада виникає в мережі зв'язку.

Адитивні завади обумовлені численними чинниками: флуктуаційні шуми, викликані тепловими процесами в резисторах і інших елементах схем, індустриальні завади, зокрема, від наведень ліній електроживлення, контактної мережі, радіо станцій, інших мереж зв'язку.

Адитивні завади діляться на *зосереджені і флуктуаційні*. Зосереджені бувають вузькосмуговими (зосереджені у вузькій смузі частот) і імпульсними (зосереджені в часі). Вузькосмугові характерні для радіозв'язку (завади від сусідніх станцій), боротьба з ними ведеться методами підвищення селективності систем.

Імпульсні завади — випадкові послідовності імпульсів, що створюються промисловими установками, наприклад, в колі контактний дріт-пантограф. Завада вважається імпульсною, якщо її тривалість набагато менше тривалості сигналу $\tau_s \ll T_c$. Боротьба з

імпульсною завадою – застосування систем з широкосмуговими підсилювачами, обмежувачами, вузькосмуговими підсилювачами, з використанням низькочастотних і надвисокочастотних діапазонів, де має місце спад спектральної щільності потужності завад.

Флуктуаційна завада розподілена в широкому спектрі частот. Потужність теплового шуму на 1 Ом навантаження в смузі частот Δf визначається як: $P_3 = \sigma^2 = \Delta f 4KT$, де $K = 1,37 \cdot 10^{-23}$ Дж/Кельвін – постійна Больцмана, T° - абсолютна температура.

Загальна характеристика дискретного каналу. Дискретний канал має в своєму складі безперервний канал, через який дискретні сигнали (послідовність символів) проходять як безперервні, такі, що відрізняються один від одного кодovими ознаками. На ці сигнали діють завади, в результаті на виході дискретного каналу з'являється послідовність символів, частина яких виявилася спотвореною, тобто в каналі з'явилися помилки.

Дискретний канал в загальному вигляді є сукупністю дискретного модулятора на вході, безперервного каналу і дискретного демодулятора на виході (рис. 9.6):



Рис. 9.6. Дискретний канал в загальному вигляді

Дискретний канал характеризується:

алфавітом входу $B_k = (b_1; b_2 \dots b_i \dots b_{m-1})$, наприклад, 0,1,2...;

алфавітом виходу $B'_k = (b'_1; b'_2 \dots b'_i \dots b'_{m-1})$, причому алфавіт входу і виходу не обов'язково співпадають, наприклад, вихідний алфавіт може мати зайві символи (m - основа коду);

швидкістю передачі V_k симв/с. Ця швидкість визначається в основному властивостями безперервного каналу – його пам'яттю;

матрицею або графом переходів, тобто сукупністю умовної ймовірності $P(b'_i / b_j)$ того, що при вхідному символі b_j на виході каналу буде b'_i .

Окрім перерахованих характеристик, дискретний канал має такі характеристики, як *пропускна спроможність і кількість переданої через канал інформації*.

Моделі дискретних каналів передавання даних. Завади і спотворення в безперервному каналі викликають появу в дискретному каналі потоку помилок.

Залежно від властивостей потоку помилок дискретні канали можуть описуватися наступними моделями.

1. *Двійковий симетричний канал без пам'яті* характеризується тим, що кожен переданий символ може бути прийнятий або з помилкою P_0 , або правильно $1 - P_0$, причому у разі помилки переходи $1 \rightarrow 0$ і $0 \rightarrow 1$ рівноймовірні.

Ймовірність того, що при передачі b_j буде отриманий b'_i :

$$P(b'_i / b_j) = P_0 \text{ при } i \neq j, P(b'_i / b_j) = 1 - P_0 \text{ при } i = j.$$

Відсутність пам'яті виявляється в тому, що умовна ймовірність $P(b'_i / b_j)$ не залежить від попередніх подій, тобто від того, які символи передавалися раніше і як вони були прийняті. Ймовірність $P(b'_i / b_j)$ називається *ап'юріорною ймовірністю*, вона для даного каналу визначається на основі тривалих експериментів.

Фізичний сенс симетрії каналу полягає в тому, що в дискретному демодуляторі, який називають також *першою розрахунковою схемою*, пороговий рівень вибраний точно посередині між середніми значеннями сигналів, що відповідають нулю і одиниці. Якби цей рівень був вибраний біля значення 0, то за рахунок завади ймовірність переходу $0 \rightarrow 1$ була б більше, ніж $1 \rightarrow 0$.

Ймовірність переходів в двійковому симетричному каналі показана у вигляді графа на рис. 9.7,а.

2. *Двійковий несиметричний канал без пам'яті* відрізняється від попередньої моделі нерівною ймовірністю переходів від $0 \rightarrow 1$ і $1 \rightarrow 0$. Графи переходів для даної моделі показані на рис. 9.7,б.

Повна ймовірність прийому b'_i символу для двійкового каналу на базі ймовірності вхідних символів і умовної ймовірності переходів визначається як:

$$P(0') = P(0)P(0' / 0) + P(1)P(0' / 1); P(1') = P(1)P(1' / 1) + P(0)P(1' / 0).$$

Середня ймовірність помилки в каналі:

$$P_0 = P(0) P(1' / 0) + P(1) P(0' / 1). \quad (9.9)$$

3. *Симетричний канал без пам'яті із стиранням* відрізняється від симетричного каналу тим, що на виході алфавіт каналу містить додатковий символ, який з'являється тоді, коли демодулятор не може надійно впізнати переданий символ.

Природно, що в такому каналі ймовірність помилкових переходів зменшується за рахунок ймовірності q появи символу стирання θ . Граф переходів для цього каналу представлений на рис. 9.7,в. Фізичний сенс каналу із стиранням полягає в тому, що в демодуляторі створюється два порогові рівні, і фіксується 0, якщо прийнятий безперервний сигнал менше

нижнього рівня, і 1, якщо сигнал більше верхнього рівня. Символ стирання θ фіксується (його можна і не фіксувати) при попаданні сигналу в проміжок між рівнями.

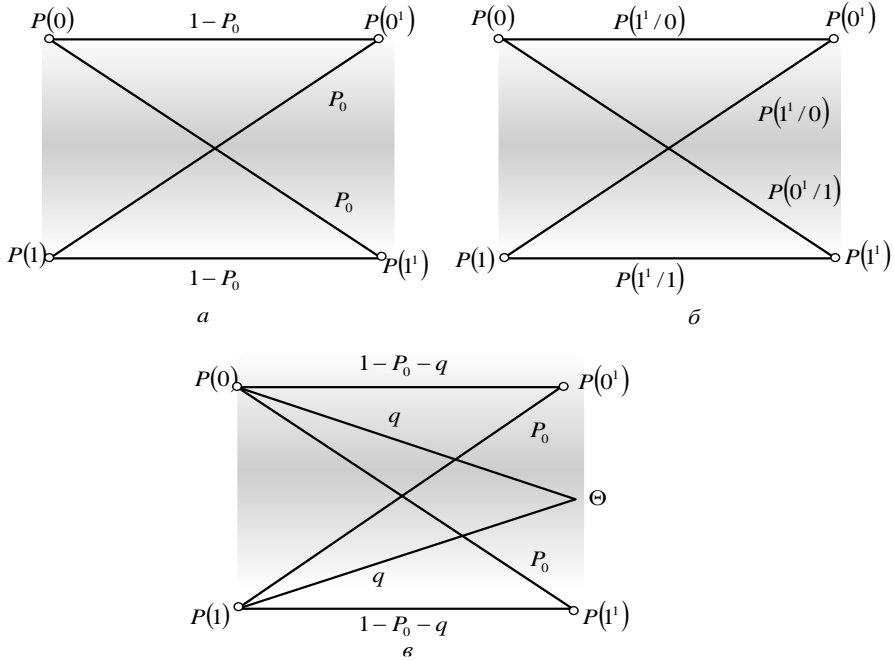


Рис. 9.7. Графи ймовірностей переходів в двійкових симетричних каналах без пам'яті: *a* - з рівними переходами, *б* - з не рівними переходами, *в* - зі стиранням пам'яті

Основне завдання прийому дискретних сигналів. Заданим називається дискретний канал, якщо відомі його алфавіти входу (0,1) і виходу ($0',1'$), а апіорна ймовірність переходу символів вхідного алфавіту в символи вихідного визначена умовними ймовірностями: $P(0'/0)$, $P(1'/0)$, $P(1'/1)$, $P(0'/1)$.

Зворотнє завдання прийому - визначення апостеріорної ймовірності того, що при прийомі було визначено підтвердження зазначеної гіпотези, наприклад, при прийомі нуля був переданий нуль. Оскільки нуль в приймачі міг бути отриманий також і за рахунок помилкового переходу $1 \rightarrow 0$, у разі двійкового каналу необхідно розглядати дві гіпотези появи нуля на виході приймача і визначати ймовірність цих гіпотез за формулою Байеса (рис. 9.8):

$$\begin{aligned} P(0/0') &= P(0)P(0'/0)/[P(0)P(0'/0) + P(1)P(0'/1)]; \\ P(1/0') &= P(1)P(0'/1)/[P(1)P(0'/1) + P(0)P(0'/0)]. \end{aligned} \tag{9.10}$$

Ймовірність кожної гіпотези визначається відношенням ймовірності прийому даного символу за рахунок відповідного переходу до повної ймовірності прийому символу. Порівнюючи ймовірність гіпотез, особливо при прийомі малоїмовірного символу, коли гіпотеза правильного переходу приблизно рівна ймовірності помилкового переходу, роблять висновок про погану роботу каналу.

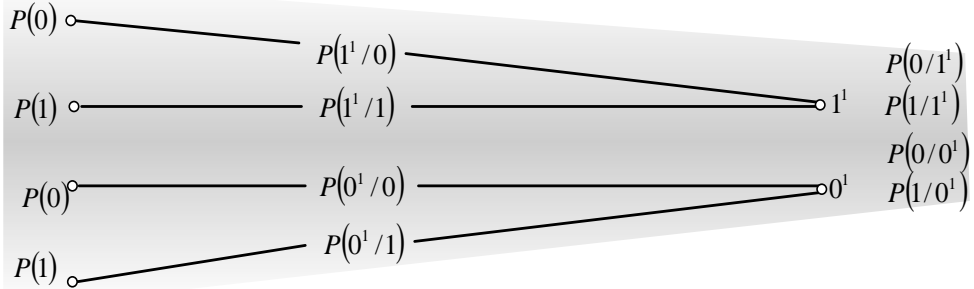


Рис. 9.8. Граф переходів в двійковому каналі

Поліпшення можна досягти зміною порогового рівня так, щоб апіорна ймовірність, наприклад прийому нуля при передачі одиниці $P(0^1/1)$, була б набагато менше ймовірності $P(1)$, тобто $P(0^1/1) \ll P(1)$.

9.3. Швидкість передавання інформації в каналах зв'язку

Швидкість передавання інформації та пропускна здатність дискретного каналу без завад. Дискретний канал передавання інформації – сукупність методів та засобів, призначених для передавання дискретних сигналів від джерела повідомлення до споживача.

На вхід такого каналу подаються дискретні повідомлення X , що утворюють первинний алфавіт x_1, x_2, \dots, x_n . Останні кодуються за допомогою кодерів (див. рис. 9.3) і перетворюються у кодовані повідомлення Y . Для кодування використовується деякий алфавіт символів y_1, y_2, \dots, y_m . Суть кодування зводиться до представлення окремих повідомлень або послідовностей повідомлень певними комбінаціями символів використовуваного алфавіту.

Швидкість введення інформації дискретного каналу:

$$\bar{I}(X) = H(X) / \bar{\tau}_X = \bar{V}_X H(X), \quad (9.11)$$

де $H(X)$ — середня ентропія одного повідомлення; $\bar{\tau}_X$ — середня тривалість повідомлення.

Швидкість видачі символів повідомлення джерелом: $\bar{V}_X = 1 / \bar{\tau}_X$.

Під тривалістю повідомлення розуміється інтервал часу, у який повідомлення на виході джерела інформації формується або існує.

Середня тривалість $\bar{\tau}_x$ інформаційного повідомлення на фоні джерела, при відсутності статистичних залежностей між повідомленнями визначається виразом:

$$\bar{\tau}_x = \sum_{i=1}^n p(x_i) \tau_{x_i}, \quad (9.12)$$

де $p(x_i)$ і τ_{x_i} – апіорна ймовірність і тривалість i -го повідомлення; n – кількість повідомлень.

У каналі без завад кожному певному вхідному сигналу завжди буде відповідати той самий сигнал на виході каналу, іншими словами, вхідні й вихідні сигнали зв'язані однозначною функціональною залежністю (рис. 9.9, а).

Середня кількість інформації у цьому випадку дорівнює ентропії символу на вході каналу: $I(Y) = H(Y)$ [дв.од./символ].

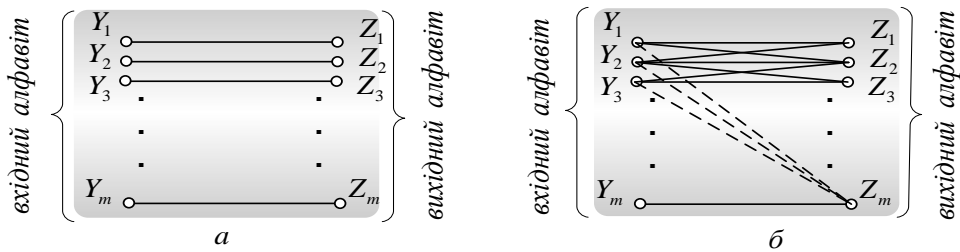


Рис. 9.9. Графи функціональних залежностей між входом і виходом дискретного каналу без завад (а) і з завадами (б)

Швидкість передачі інформації в дискретному каналі без завад:

$$I(Y) = \bar{U}_Y \cdot H(Y) \text{ [дв.од./символ]}, \quad (9.13)$$

де $\bar{U}_Y = 1 / \bar{\tau}_y$ – швидкість передачі елементарних символів сигналу; τ_y – середня тривалість елементарних сигналів.

Пропускна здатність дискретного каналу без завад: $C = \max\{\bar{U}_Y H(Y)\}$

Максимальна швидкість передачі інформації \bar{U}_Y буде забезпечена при максимальному значенні ентропії кодованого сигналу

$$C = \bar{U}_Y \max\{H(Y)\} = \bar{U}_Y \log_2 n, \quad (9.14)$$

тобто при рівномірному розподілі ймовірностей і статистичній незалежності символів алфавіту сигналів.

Таким чином, швидкість передавання інформації може бути максимальною за умови, якщо статистичні характеристики джерела повідомлень певним чином погоджені із властивостями інформаційного

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

каналу. Для кожного джерела повідомлень це узгодження може бути досягнуто спеціальним вибором способу кодування сигналів.

На питання про те, чи може швидкість передавання інформації бути наближеною до пропускної здатності інформаційного каналу, відповідає теорема К.Шеннона для дискретного каналу без завад.

Теорема К.Шеннона для дискретного каналу зв'язку без завад: якщо потік інформації від джерела, досить близький до пропускної здатності каналу, тобто: $\bar{I}(X) = C - \delta$, (9.15)

де δ – значно мала величина, то завжди можна знайти такий спосіб кодування, що забезпечить передавання всіх повідомлень від джерела до споживача, причому швидкість передавання інформації буде досить близька до пропускної здатності каналу $\bar{I}(Z, Y) = C - \delta$.

Зворотнє твердження теореми полягає в тому, що неможливо забезпечити тривале передавання всіх повідомлень, якщо потік інформації від джерела, перевищує пропускну здатність каналу $\bar{I}(X) > C$.

Таким чином, теорема Шеннона стверджує, що при виконанні умови (5.15) швидкість передавання інформації може бути в принципі значно наближена до пропускної здатності каналу. Це може бути забезпечено відповідним кодуванням сигналів. Однак розглянута теорема не відповідає на запитання, яким чином потрібно здійснювати кодування.

Швидкість передавання інформації й пропускна здатність дискретного каналу з завадами. При наявності завад у каналі передавання інформації порушується однозначна відповідність між вхідним і вихідним алфавітами каналу. Одному вхідному сигналу можуть відповідати різні вихідні сигнали (рис. 9.9,б). Внаслідок випадкового характеру завад неможливо заздалегідь точно встановити, який сигнал може бути прийнятий на виході каналу при надсиланні певного вхідного сигналу. Мова може йти тільки про ймовірності одержання на виході каналу елементарного сигналу z_j - за умови, що був відправлений відповідний елементарний сигнал.

Ймовірнісний характер зв'язку між вхідним і вихідним алфавітами каналу передачі інформації повністю визначається матрицею

перехідних ймовірностей:

$$\begin{vmatrix} p_{11} & p_{12} & p_{13} \cdots p_{1n} \\ p_{21} & p_{22} & p_{23} \cdots p_{2n} \\ \cdots & \cdots & \cdots \cdots \cdots \\ p_{n1} & p_{n2} & p_{n3} \cdots p_{nn} \end{vmatrix}$$

де p_{ij} – умовна ймовірність переходу i -го символу вхідного алфавіту в j -й символ вихідного алфавіту.

Очевидна справедливість наступної рівності: $\sum_{j=1}^n p_{ij} = 1$

Бінарним каналом називається дискретний канал, по якому передаються тільки два типа елементарних сигналів.

Матриця перехідних ймовірностей для такого каналу: $\begin{vmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{vmatrix}$.

Симетричним називається канал при умові, якщо всі ймовірності правильної передачі сигналів однакові й однакові всі ймовірності перекрученої передачі.

Для симетричного бінарного каналу матриця перехідних ймовірностей $\begin{vmatrix} p & q \\ q & p \end{vmatrix}$, де $p = p_{11} = p_{22}$ — ймовірність правильної передачі; $q = p_{12} = p_{21}$ — ймовірність перекрученої передачі.

Тому що в симетричному каналі ймовірності перекручування всіх символів сигналу однакові, то можна стверджувати, що в такому каналі завади не залежать від переданих сигналів. На рис. 9.10 наведений графік перехідних ймовірностей двійкового симетричного каналу.

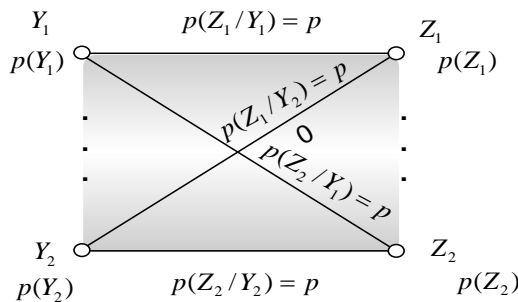


Рис. 9.10. Граф ймовірностей переходів в двійковому симетричному каналі без пам'яті з завадами

Швидкість передавання інформації дискретного каналу з завадами $\bar{I}(Z, Y) = \bar{U}_Y [H(Y) - H(Y/Z)],$ (9.16)

де $H(Y/Z)$ – залишкова ентропія сигналу, обумовлена дією завад.

Вираз для швидкості передавання інформації може бути представлений також у вигляді: $\bar{I}(Z, Y) = \bar{U}_Y [H(Z) - H(Z/Y)],$ (9.17)

де $H(Z)$ – ентропія вихідного сигналу; $H(Z/Y)$ – умовна ентропія вихідного сигналу при відомій ентропії вхідного сигналу.

Каналом без пам'яті називається дискретний канал, у якому на кожний переданий символ інформаційного сигналу завади впливають незалежно від того, які сигнали передавалися раніше.

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

У таких каналах завади не викликають додаткових кореляційних зв'язків між символами.

У випадку незалежності окремих символів сигналу вирази (9.16) і (9.17) для каналу без пам'яті приймуть вид

$$\bar{I}(Z, Y) = \bar{U}_Y \left[-\sum_{i=1}^n p(y_i) \log_2 p(y_i) + \sum_{i=1}^n \sum_{j=1}^n p(z_i) p(y_i/z_i) \log_2 p(y_i/z_i) \right]; \quad (9.18)$$

$$\bar{I}(Z, Y) = \bar{U}_Y \left[-\sum_{i=1}^n p(z_i) \log_2 p(z_i) + \sum_{i=1}^n \sum_{j=1}^n p(y_i) p(z_i/y_i) \log_2 p(z_i/y_i) \right]. \quad (9.19)$$

Як **приклад** розглянемо бінарний канал. Для такого каналу алфавіти вхідного Y і вихідного Z сигналів складаються із двох символів $Y = \{y_1, y_2\}$; $Z = \{z_1, z_2\}$.

Для скорочення запису позначимо ймовірності перекручування сигналів: $p(z_1/y_1) = 1$, $p(z_2/y_2) = q_2$. Очевидно, що ймовірності правильного передавання будуть: $p(z_1/y_1) = 1 - q_2$; $p(z_2/y_2) = 1 - q_1$.

При цьому вираз (9.19) прийме вигляд

$$\begin{aligned} \bar{I}(Z, Y) = \bar{U}_Y \{ & -p(z_1) \log_2 p(z_1) - p(z_2) \log_2 p(z_2) + \\ & + p(y_1)[(1 - q_2) \log_2(1 - q_2) + q_2 \log_2 q_2] + p(y_2)[q_1 \log_2 q_1 + (1 - q_1) \log_2(1 - q_1)] \}. \end{aligned}$$

Максимізуючи праву частину виразу (9.19), визначають пропускну здатність каналу. Очевидно, що цього можна досягти за рахунок оптимізації значень апіорних ймовірностей $p(y_1)$ і $p(y_2)$ передавання сигналів y_1 й y_2 , тому що ніякі інші параметри каналу міняти ми не можемо.

Розглянемо окремі випадки:

1. Ймовірності перекручувань сигналів $q_1 = q_2 = q$. Цей випадок відповідає симетричному каналу. Для симетричного каналу умовна ентропія

$$\begin{aligned} H(Z/Y) = -[p(y_1) + p(y_2)][q \log_2 q + (1 - q) \log_2(1 - q)] = \\ = q \log_2 q + (1 - q) \log_2(1 - q), \end{aligned} \quad (9.20)$$

тому що $p(y_1) + p(y_2) = 1$. З виразу (9.20) видно, що умовна ентропія не залежить від апіорних ймовірностей $p(y_1)$ та $p(y_2)$. Отже, максимальна швидкість передачі інформації виходить у цьому випадку при такому розподілі ймовірностей $p(y_1)$ та $p(y_2)$, при якому ентропія $H(Z)$ виявляється максимальною. Це буде мати місце при рівності апіорних ймовірностей $p(z_1) = p(z_2)$. При цьому максимальне значення ентропії

$$\begin{aligned} H(Z_{\max}) = 1[\text{дв.од}] \quad \text{і} \quad \text{пропускна} \quad \text{здатність} \quad \text{каналу} \\ C = \bar{U}_Y [\log_2 2 + q \log_2 q + (1 - q) \log_2(1 - q)]. \end{aligned} \quad (9.21)$$

Використовуючи відоме правило теорії ймовірностей

$$p(z_i) = \sum_{i=1}^n p(y_i) p(z_i/y_i), \quad (9.22)$$

можна показати, що рівність апіорних ймовірностей вихідних сигналів z_1 й z_2 для симетричного каналу буде мати місце при рівності апіорних ймовірностей вхідних сигналів.

Таким чином, у симетричному бінарному каналі з завадами максимальна швидкість передавання інформації виходить за такої ж умові, як й у каналі без завад. Однак з порівняння (9.21) і (9.14) видно, що наявність завад у каналі приводить до зменшення пропускної здатності каналу.

2. Ймовірності перекручувань сигналів $q_1 = 0, q_2 \neq 0$, тобто мають місце перекручування лише при передачі сигналів y_2 .

Із (9.19) з використанням (9.22) для цього випадку одержуємо $\bar{I}(Z, Y) = \bar{U}_Y \{-p(y_1) \log_2 p(y_1) q_2 \log_2 q_2 - [p(y_2) + p(y_1) q_2] \log_2 [p(y_2) + p(y_1) q_2]\}$. Максимальне значення швидкості передачі інформації буде за умови

$$p(y_1)_{\text{опт}} = 1 / \left(1 + q_2^{q_2 / (1 - q_2)} - q_2 \right). \quad (9.23)$$

Пропускна здатність такого каналу визначається виразом

$$C = \bar{U}_Y \{-p(y_1)_{\text{опт}} + \log_2 p(y_1)_{\text{опт}} + p(y_1)_{\text{опт}} q_2 \log_2 q_2 - [1 - p(y_1)_{\text{опт}} + p(y_1)_{\text{опт}} q_2] \log_2 [1 - p(y_1)_{\text{опт}} + p(y_1)_{\text{опт}} q_2]\}. \quad (5.24)$$

На рис. 9.11 наведені графіки: $C/\bar{U}_Y = C\bar{\tau}_Y$, $p(y_1)_{\text{опт}}$ і $p(y_2)_{\text{опт}} = 1 - p(y_1)_{\text{опт}}$ як функції від ймовірності q_2 . Як видно із графіків, зі збільшенням q_2 від 0 до 1 пропускна здатність падає від $U_Y = 1/\tau_Y$ до нуля.

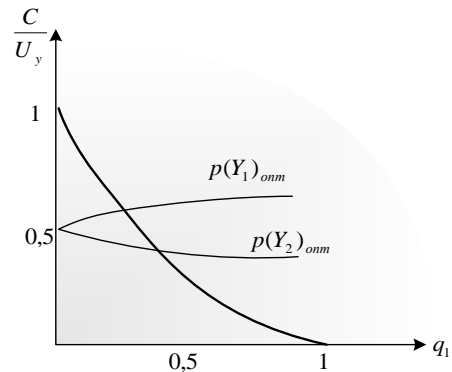


Рис. 9.11. Пропускна здатність каналу залежно від ймовірності спотворень

Теорема К. Шеннона для дискретного каналу з завадами: якщо потік інформації від джерела повідомлень, досить близький до пропускної здатності каналу, тобто якщо справедлива рівність: $\bar{I}(X) = C - \delta$, де δ - дуже мала величина, то завжди можна знайти такий спосіб кодування, що забезпечить передачу всіх повідомлень від джерела, а ймовірність помилкового впізнання будь-якого переданого повідомлення буде дуже малою, тобто $p_{\text{н.в}} \leq \eta$, де $p_{\text{н.в}}$ - ймовірність неправильного впізнання переданого повідомлення; η - дуже мала величина.

Зворотне твердження теореми – якщо потік інформації джерела перевищує пропускну здатність каналу, то не існує способу

кодування, що забезпечує передачу будь-якого повідомлення з малою ймовірністю помилки.

Таким чином, розглянута теорема визначає співвідношення між швидкістю створення повідомлень джерелом, пропускною здатністю каналу при наявності завад і ймовірністю передавання. Якщо для каналу без завад цікавим є ефективність передавання, то для каналу з завадами - ефективність і ймовірність передавання.

Дана теорема, як і теорема для каналу без завад, не відповідає на запитання, яким чином потрібно здійснювати кодування, щоб наблизити швидкість передавання інформації до пропускної здатності каналу. Але для наближення швидкості передавання до межі загальним методом як для каналу з завадами, так і для каналу без завад є кодування довгих повідомлень.

Швидкість передавання інформації й пропускна здатність безперервного каналу з завадами. *Безперервний канал передавання інформації* - сукупність методів та засобів, призначених для передачі безперервних сигналів від джерела повідомлення до споживача.

На відміну від дискретних каналів у неперервних каналах замість кодувальних і декодувальних пристроїв може використовуватися більш широкий клас різних перетворювачів. Для передавання інформації по каналу може застосовуватися модуляція одного або декількох параметрів сигналу. Незалежно від конкретного характеру перетворення сигналів вхідні й вихідні сигнали безперервного каналу задаються у вигляді ансамблів безперервних функцій з відповідними функціями щільності розподілу ймовірностей.

Нехай на вхід каналу надходить безперервний сигнал $Y(t)$ тривалістю T . Внаслідок впливу завад $\xi(t)$ вихідний сигнал $Z(t)$ буде відрізнятися від вхідного. Кількість інформації у випадковому сигналі $Z(t)$ щодо випадкового сигналу $Y(t)$, визначається як: $I_T(Z, Y) = H_T(Z) - H_T(Z/Y)$. (9.25)

Відповідно теорему Котельникова, безперервні сигнали $Y(t)$ і $Z(t)$ можуть бути представлені сукупностями відліків y_i та z_i у дискретні моменти часу (рис. 9.12), які є випадковими величинами.

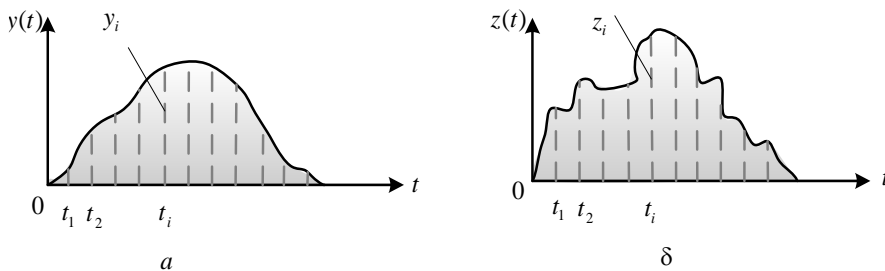


Рис. 9.12. Випадковий інформаційний сигнал (а), суміш інформаційного сигналу та завад (б)

Розподіл сукупності випадкових величин описується багатомірними щільностями розподілу ймовірності $w(y_1, y_2, \dots, y_m)$ та $w(z_1, z_2, \dots, z_m)$.

Тоді диференціальна ентропія сигналу на виході каналу

$$h_t(Z) = - \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} m \dots w(z_1, z_2, \dots, z_m) \log_2 w(z_1, z_2, \dots, z_m) dz_1, dz_2, \dots, dz_m.$$

Відповідно критерію М.А. Железнова при квантуванні випадкових сигналів за часом інтервал квантування необхідно встановлювати рівним інтервалу кореляції функції τ_0 , тоді випадкові величини $z_1, z_2, \dots, z_i, \dots, z_m$ можна вважати незалежними: $w(z_1, z_2, \dots, z_m) = w(z_1)w(z_2) \dots w(z_m)$.

Оскільки ентропія сукупності випадкових незалежних величин дорівнює сумі ентропії випадкових величин, одержимо наступний вираз для

диференціальної ентропії сигналу: $h_T(Z) = \sum_{i=1}^m h(z_i)$, де

$$h(z_i) = - \int_{-\infty}^{\infty} w(z_i) \log_2 w(z_i) dz_i$$
 - диференціальна ентропія i -го відліку

сигналу Z ; $m = T/\Delta t$ - загальна кількість відліків сигналу Z тривалістю T ;

Δt — інтервал тимчасового квантування. Обмежуючись розглядом стаціонарних процесів, одержимо:

$$w(z_1) = w(z_2) = \dots = w(z_m); h(z_1) = h(z_2) = \dots = h(z_m) = h(Z).$$

Тоді $h_T(Z) = m h(Z)$, де $h(Z)$ – диференціальна ентропія одного відліку.

Умовна диференціальна ентропія $h_T(Z/Y) = m h(Z/Y)$, де $h(Z/Y)$ – умовна диференціальна ентропія одного відліку. Таким чином, вираз для кількості інформації приймає вид $I_T(Z, Y) = m[h(z) - h(Z/Y)]$.

Швидкість передачі інформації в безперервному каналі з завадами:

$$\bar{I}_T(Z, Y) = m[h(z) - n(Z/Y)]/T = F_0[h(Z) - h(Z/Y)], \quad (9.26)$$

де $F_0 = m/T = 1/\Delta t$ — частота тимчасового квантування (відліку).

Пропускна здатність каналу в безперервному каналі з завадами:

$$C = \max[\bar{I}_T(Z, Y)] = F_0 \max[h(Z) - h(Z/Y)]. \quad (9.27)$$

Розглянемо деякі окремі випадки:

1. Сигнал обмеженої потужності передається по каналу, у якому діє адитивна завада обмеженої потужності типу білого гаусового шуму.

При адитивній заваді сигнал $Z(t)$ на виході каналу буде дорівнювати $Z(t) = Y(t) + \xi(t)$, де $\xi(t)$ – завада, що діє в каналі передачі інформації.

Середні потужності сигналу й завади: P_Y й $P_\xi = \sigma_\xi^2$. Смуга пропускання каналу обмежена межами 0 і F_k . Ширина спектра сигналу й завади обмежуються смугою пропускання каналу. Частота квантування обмеженого

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

по спектру сигналу відповідно теореми Котельникова $F_0 = 2F_k$. Тоді вираз (5.27) для пропускну́ї здатності каналу прийме вид:

$$C = 2F_k \max[h(Z) - h(Z/Y)]. \quad (9.28)$$

При взаємно незалежних сигналах Y і заваді ξ ймовірність того, що при передачі сигналу Y вихідний сигнал буде дорівнювати $Z = Y + \xi$, повинна визначатися ймовірністю того, що завада прийме дане значення $\xi = Z - Y$, тобто $w(Z/Y)dY = w(\xi)d\xi$. а

$$w(Z/Y) = w(Y + \xi/Y) = w(Y/Y + \xi/Y) = w(Y/Y) w(\xi/Y) = w(\xi). \quad (9.29)$$

З огляду на (9.29), вираз для умовної диференціальної ентропії $h(Z/Y)$ можна перетворити в такий спосіб:

$$\begin{aligned} h(Z/Y) &= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(Y) w(Z/Y) \log_2 w(Z/Y) dY dZ = \\ &= - \int_{-\infty}^{\infty} w(Y) \left[\int_{-\infty}^{\infty} w(\xi) \log_2 w(\xi) d\xi \right] dY = h(\xi) \int_{-\infty}^{\infty} w(Y) dY = h(\xi), \end{aligned} \quad (9.30)$$

де $h(\xi)$ — диференціальна ентропія завади.

Таким чином, у випадку адитивної завади умовна диференціальна ентропія $h(Z/Y)$ повністю визначається властивостями завади.

Раніше було встановлено, що диференціальна ентропія сигналу, розподіленого за нормальним законом: $h(\xi) = \log_2(\sqrt{2\pi e} \sigma_\xi)$,

$$(9.31)$$

де e - показникова функція $e = 2,7..$

Підставивши (9.31) в (9.28), одержимо вираз для пропускну́ї здатності каналу:

$$C = 2F_k \max\{h(Z) - \log_2 \sqrt{2\pi e} \sigma_\xi\}. \quad (9.32)$$

Оскільки σ_ξ задане, то максимальне значення виразу (9.32) буде забезпечено за умови максимізації диференціальної ентропії вихідного сигналу $h(Z)$. Середні потужності вхідного сигналу $Y(t)$ і завади $\xi(t)$ обмежені, тому середня потужність вихідного сигналу $Z(t)$ також обмежена. Величина диференціальної ентропії $h(Z)$ буде максимальною у випадку, якщо $Z(t)$ характеризується нормальним законом розподілу. Якщо ж сумарний сигнал $Z(t)$ і одна з його складових $\xi(t)$ розподілені за нормальним законом, то й друга складова, тобто вхідний сигнал $Y(t)$, також повинна мати нормальний закон розподілу. Отже, диференціальна ентропія вихідного сигналу

$$h(Z) = \log_2(\sqrt{2\pi e} \sigma_Z) = \log_2 \sqrt{(\sigma_Y^2 + \sigma_\xi^2) 2\pi e}. \quad (9.33)$$

Підставляючи (9.31) в (9.30), остаточно одержимо

$$C = 2F_k \left[\log_2 \sqrt{(\sigma_y^2 + \sigma_\xi^2) 2\pi e} - \log_2 \sqrt{\sigma_\xi^2 2\pi e} \right] = F_k \log_2 (\sigma_y^2 + \sigma_\xi^2) / \sigma_\xi^2 = F_k \log_2 (1 + P_y / P_\xi), \quad (9.34)$$

де $P_y = \sigma_y^2$ - середня потужність корисного сигналу; $P_\xi = \sigma_\xi^2$ - середня потужність завади.

Отже, можна зробити висновки: швидкість передавання інформації сигналами з обмеженою середньою потужністю через канал, у якому діє білий гаусів шум, виявляється максимальною при повній подібності між сигналом і завадою; максимальна швидкість передавання інформації буде забезпечена, якщо як фізичний носій інформації застосовувати стаціонарний випадковий процес у виді білого гаусового шуму.

Як видно (9.34), пропускну здатність каналу можна регулювати шляхом зміни F_k або P_y . Причому практично залежність пропускну здатності каналу від F_k при постійній потужності сигналу нелінійна. Це обумовлено тим, що потужність завади P_ξ також залежить від ширини частотного спектра. Дійсно, енергетичний спектр білого шуму рівномірний, тому потужність такої завади можна представити у вигляді:

$$(9.35)$$

де P_0 — потужність завади, що приходить на смугу в 1 Гц (спектральна щільність потужності завади). Підставивши (9.35) в (9.34), одержимо вираз, що визначає дійсний характер залежності пропускну здатності каналу від ширини його смуги пропускання:

$$C = F_k \log_2 (1 + P_y / P_0 F_k). \quad (9.36)$$

Характер залежності $C = f(F_k)$ представлений графічно на рис. 9.13.

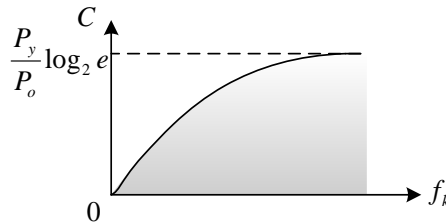
Визначимо межу, до якої прямує пропускну здатність каналу при необмеженому збільшенні його смуги пропускання:

$$\lim_{F_k \rightarrow \infty} C = \lim_{F_k \rightarrow \infty} \left[F_k \log_2 (1 + P_y / P_0 F_k) \right].$$

Розкриваючи невизначеність, одержимо граничне значення пропускну здатності каналу:

$$\lim_{F_k \rightarrow \infty} C = P_y \log_2 e / P_0. \quad (9.37)$$

З формули (9.37) видно, що максимальне значення, до якого прямує пропускну здатність каналу з ростом його ширини смуги пропускання, пропорційна відношенню середньої потужності сигналу до спектральної щільності потужності завади.



З приведенного аналізу можна зробити висновок:

немає рації сильно збільшувати смугу пропускання каналу, тому що в міру розширення смуги пропускання ріст пропускну здатності каналу

сповільнюється й на межі при $F_k \rightarrow 0$ пропускна здатність наближається до постійної величини. Має місце збільшення смуги пропускання приблизно до значення, рівного відношенню P_Y/P_0 .

Гранично можливе значення пропускної здатності може бути збільшене за рахунок збільшення відношення P_Y/P_0 . Сигнал обмеженої потужності передається по каналу, у якому діє адитивна завада у вигляді довільного шуму. При певному середньоквадратичному значенні (потужності) завади найбільшою ентропією володіє завада з нормальним законом розподілу ймовірностей. При будь-якому іншому законі розподілу ймовірностей завади її ентропія (при тій же середньоквадратичному значенні) виявляється меншою. Залежність ентропії від виду закону розподілу спонукала Шеннона характеризувати дію завади не її дійсною потужністю, а так названою ентропійною потужністю завади. Під ентропійною потужністю Шеннон розумів потужність еквівалентного білого шуму, що, має ту ж тривалість і ширину спектра, володіє такою ж, як дана завада, ентропією. Якщо довільна завада $\xi(t)$ характеризується ентропією на один відлік $h(t)$, то тоді потужність еквівалентного білого шуму, тобто ентропійна потужність $\sigma_{\xi e}^2$, може бути визначена з умови

$$h(\xi) = \log 2\sqrt{2\pi e\sigma_{\xi e}^2}, \sigma_{\xi e}^2 = 2^{p(\xi)}/2\pi e = 2^{h(\xi)-1}/\pi e. \quad (9.38)$$

Величина відношення середньої потужності завади σ_{ξ}^2 й її ентропійної потужності $\sigma_{\xi e}^2$ визначається законом розподілу завади.

$$\text{Позначивши це відношення через } k_e, \text{ одержимо: } \sigma_{\xi}^2 = k_e \sigma_{\xi e}^2. \quad (9.39)$$

Зокрема, для закону рівної ймовірності $k_e = 1,3$.

Використовуючи поняття ентропійної потужності, можна отримати вираз для пропускної здатності каналу, у якому діє довільна завада:

$$\begin{aligned} C &= 2F_k [\log_2 \sqrt{(\sigma_Y^2 + \sigma_{\xi}^2)2\pi e} - \log_2 \sqrt{\sigma_{\xi e}^2 2\pi e}] = \\ &= F_k \log_2 \left[(\sigma_Y^2 + \sigma_{\xi}^2) / \sigma_{\xi e}^2 \right] = F_k \log_2 \left[k_e (\sigma_Y^2 + \sigma_{\xi}^2) / \sigma_{\xi}^2 \right]. \end{aligned} \quad (9.40)$$

Для всіх законів розподілу (крім нормального) коефіцієнт $k_e > 1$, тому пропускна здатність каналу, у якому діє довільна завада, буде завжди більше пропускної здатності каналу, у якому діє білий шум з такою ж середньою потужністю, що й довільна завада.

Теорема Шеннона для безперервного каналу з завадами: якщо ентропія $\bar{N}_{\xi}(X)$ джерела безперервних повідомлень, що визначає кількість інформації за одиницю часу при заданій оцінці g вірності відтворення, значно близька до пропускної здатності каналу, тобто справедливе співвідношення: $\bar{N}_{\xi}(X) = C - \alpha$, де α — значно мала, то

існує метод передачі, при якому всі повідомлення від джерела, можуть бути передані, а вірність відтворення при цьому як завжди близька до g . Зворотне твердження цієї теореми говорить про те, що така передача неможлива, якщо $\overline{H}_e(X) > C$.

Теорема дозволяє знаходити гранично досягну ефективність безперервних каналів. Кількість інформації, що може бути передана по каналу за час його роботи T_k при впливі завад типу білого шуму:

$$I_T(Z, Y) = T_k F_k \log_2(1 + P_Y/P_\xi). \quad (9.41)$$

Практично в більшості випадків потужність корисного сигналу значно перевищує потужність завад. У цих випадках вираз (9.41) можна з достатнім наближенням представити у вигляді: $I_T(Z, Y) = T_k F_k \log_2 P_Y/P$. (9.42)

У формулах (5.41) і (5.42) члени $\log_2(1 + P_Y/P_\xi)$, $\log_2(P_Y/P_\xi)$ виражають максимально можливу кількість інформації на один відлік. Тоді, порівнюючи (5.41) і (5.42) з (5.3), можна стверджувати, що ємність каналу визначає максимально можливу кількість інформації, що може бути передана по каналу за час його роботи. Пропускна здатність виражає максимально можливу кількість інформації, що може бути передана по каналу за одиницю часу, тоді зв'язок між ємністю й пропускною здатністю каналу визначається залежністю:

$$V_k = T_k C. \quad (9.43)$$

9.4. Синтез елементів систем. Оптимальний приймач

Більшість інформаційних процесів пов'язана з рішенням проблеми вибору. Така ж проблема розв'язується при пошуку кращих в певному значенні алгоритмів обробки інформаційних повідомлень і сигналів, технічної реалізацій цих методів. Основні алгоритми перетворюючих елементів інформаційних систем виносять двійкові рішення «Так» або «Ні»: знайдено повідомлення чи ні, розрізняються сигнали на фоні завад чи ні, повідомлення спотворене чи ні, доступна інформація для споживача чи ні, тощо. Ці задачі розв'язуються при обміні даними, при формуванні та прийнятті рішень, при відображенні інформації тощо.

Завданням синтезу є визначення алгоритму функціонування інформаційних систем за заданим критерієм якості і інтерпретація чи впровадження цього алгоритму технічними засобами.

Завданням аналізу є розрахунок робочих характеристик та узагальнених структур інформаційних систем.

Повідомлення, що поступають в інформаційні системи, є випадковими процесами. Сигнали, як носії повідомлень, і діючі на них завади також є випадковими процесами. Тому пошук алгоритмів опрацювання сигналів з завадами здійснюється при використанні ймовірнісних моделей. Оцінки

(рішення), одержувані на підставі вибірок кінцевого розміру, називають статистичними характеристиками. При пошуку рішення завжди виникає ситуація невизначеності щодо розподілу ймовірності сигналів і завад, їх параметрів, додаткових обмежень. Якщо всі ці дані не відомі, то говорять про задачі статистичного синтезу в умовах апіорної невизначеності.

При повній відсутності апіорних даних вирішувати задачі оптимального синтезу неможливо. На практиці завжди знаходяться якісь апіорні відомості.

Розрізняють наступні основні типи задач статистичного синтезу:

1. *Виявлення сигналу на фоні завад.*
2. *Розрізнення сигналів на фоні завад.*
3. *Виявлення (розрізнення) сигналів і оцінка їх параметрів на фоні завад.*
4. *Виділення сигналів на фоні завад.*

Підсумком рішення перерахованих задач є відповідні алгоритми опрацювання сигналу з завадою, що приводять, як правило, до поліпшення співвідношення між корисним сигналом і завадою. Всі ці задачі мають багато загального, але мають і свої особливості, що відображаються на структурі алгоритмів.

Статистичні критерії виявлення сигналів на тлі завад. При прийомі інформаційних повідомлень, залежно від призначення системи і виду сигналів мають місце два типи завдань: виявлення сигналів та розпізнавання сигналів.

Виявлення сигналів. *Завдання виявлення інформаційного сигналу полягає в тому, щоб за наслідками процесу опрацювання прийнятого повідомлення, яке може бути тільки завадою або сумою переданого сигналу і завади, було прийняте рішення щодо наявності в прийнятому сигналі корисної інформації або її відсутності.*

При виявленні можуть бути дві помилки:

1) *помилка першого роду або «помилкова тривога»* – за відсутності корисного сигналу ухвалюється помилкове рішення про наявність сигналу;

2) *помилка другого роду або «пропуск цілі»* – за наявності корисного сигналу в повідомленні ухвалюється помилкове рішення про його відсутність.

Ці помилки кількісно оцінюються умовною ймовірністю α і β помилкових рішень про наявність сигналу, коли його немає, і про відсутність його, коли він є. Повна ймовірність помилкового рішення визначається виразом: $P_0 = q\alpha + p\beta$ де q і p відповідно апіорна ймовірність відсутності і присутності корисного сигналу.

Очевидно, що втрати, які може нести споживач інформації при помилках першого і другого роду, можуть бути далеко не однаковими.

Розпізнавання сигналів. *Передаванням з пасивною паузою* називається такий спосіб передавання інформації, при якому виявлення сигналів на наявність корисного повідомлення (посилка) відповідає прийому символу 1 (гіпотеза H_1), відсутність сигналу (пауза) - символу 0 (гіпотеза H_0).

Передаванням з активною паузою називається такий спосіб передавання інформації, при якому в процесі розрізнення двох сигналів, символу 1 відповідає сигнал x_1 , а символу 0 – x_0 ; причому сигнали x_0 і x_1 мають однакову енергію.

При розпізнаванні сигналів можуть також з'являтися *помилки першого і другого роду*. Величини α і β є умовною ймовірністю помилкових рішень про наявність сигналу x_1 , коли насправді прийнятий x_2 або навпаки. Завдання розпізнавання або розрізнення сигналів зустрічається в системах автоматизованого керування, радіолокації та радіонавігації, телекомунікаційних системах та мережах тощо.

Розглянемо графіки розподілів щільності умовної ймовірності того, що прийнятий сигнал x характеризує стан a_1 і a_2 (рис. 9.14).

Графіки перетинаються, і для симетричного каналу межа ab розрізнення ознак сигналів x_1 про стан a_1 і x_2 про стан a_2 лежить посередині інтервалу, що розділяє \bar{x}_1 і \bar{x}_2 . Нехай α – ймовірність рішення про прийом, коли насправді приймається x_2 , а β – рішення про прийом x_2 , коли приймається x_1 . У симетричному каналі $\alpha = \beta$. Якщо джерело інформації виробляє тільки два сигнали x_1 і x_2 , вони утворюють повну групу подій $P(x_1) + P(x_2) = 1$, то на виході приймача можуть утворюватися чотири події, також створюючи повну групу: $P(x'_1) + \alpha + P(x'_2) + \beta = 1$. Ці переходи показані на рис. 9.15.

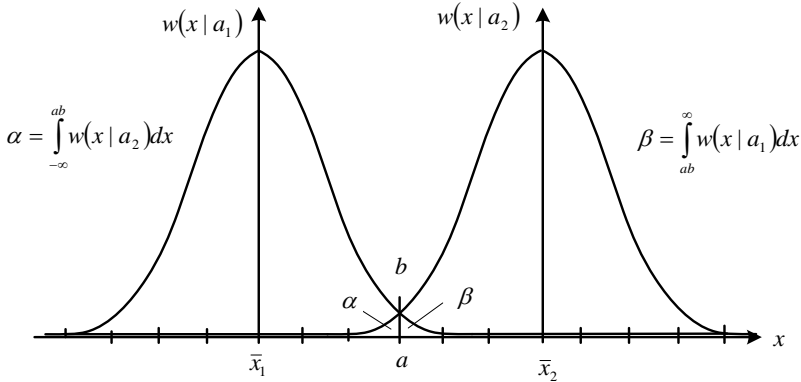


Рис. 9.14. Графіки розподілів щільності умовної ймовірності в симетричному каналі

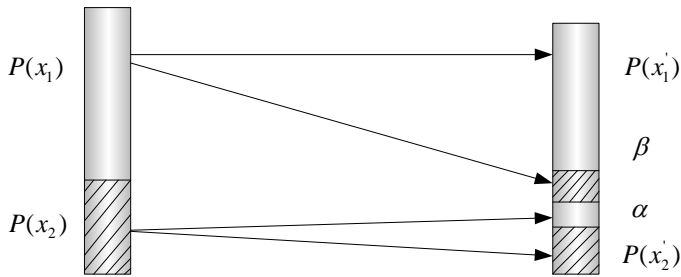


Рис. 9.15. Граф переходів ймовірностей при двох повідомленнях x_1 та x_2

Даний тип задач є базовим та широко розповсюдженим у техніці зв'язку, в автоматичі, в радіолокації, у вимірюваннях де широко застосовуються статистичні критерії виявлення і розпізнавання. Розглянемо деякі з них.

Загальні критерії розпізнавання. На графіку (рис. 9.14) межа розрізнення ознак ab проведена через точку перетину графіків. При $w(x|a_1) = w(x|a_2)$ реалізується критерій *максимальної правдоподібності* (не віддається перевага жодному з сигналів) $\alpha = \beta$. Ймовірність помилкових рішень залежить не тільки від вигляду, але і від апіорних значень ймовірності сигналів, відповідних a_i , станам об'єкту. З урахуванням «ваги» графіки $P(x_i) \cdot w(x|a_i)$ матимуть вигляд, показаний на рис. 9.16. Як і у попередньому випадку порогове значення між ознаками сигналів x_1 і x_2 проводиться через точку перетину графіків. В даному випадку буде реалізований *критерій мінімуму середньої помилки* або *критерій ідеального спостерігача Котельникова*. В даному випадку реалізується і критерій *максимальної апостеріорної ймовірності*.

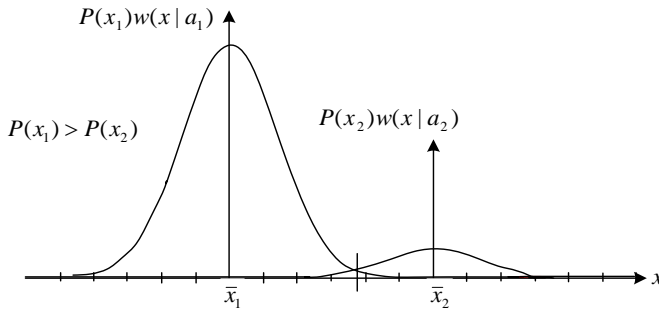


Рис. 9.16. Графіки розподілів щільності умовної ймовірності з урахуванням «ваги»

У тих випадках, коли крім відомої ймовірності сигналів x_i і x_2 відомі і *втрати*, які несе споживач від помилкових рішень, проведення межі розрізнення ознак проводиться з урахуванням відносних втрат r_{12} і r_{21} (рис. 9.17).

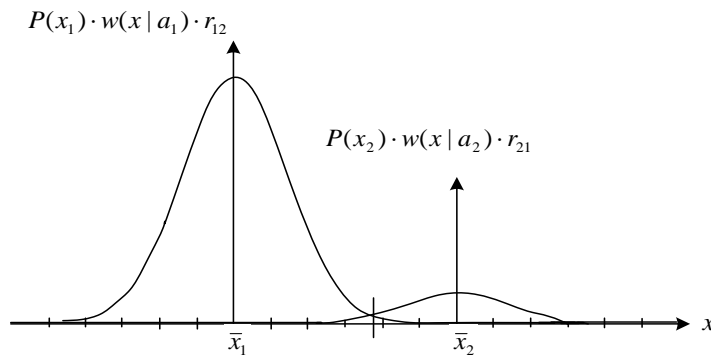


Рис. 9.17. Графіки розподілів щільності умовної ймовірності з урахуванням «ваги» та відносних втрат

В даному випадку реалізується *критерій мінімального ризику*. Усереднена величина ризику: $r = P(x_i) \cdot \alpha \cdot r_{12} + P(x_2) \cdot \beta \cdot r_{21}$. Цей критерій доцільно використовувати при розробленні таких ІВС, в яких помилки першого і другого роду приводять до *різних втрат*, але обидві помилки не створюють небезпечних ситуацій.

Виявлення сигналу методом одноразового відліку. Всі розглянуті критерії якості приводять, по суті, до одного правила ухвалення рішення. Воно полягає у визначенні відношення правдоподібності Λ і порівняння його з пороговим значенням Λ_n , залежним від вживаного критерію.

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

Завдання опрацювання реалізації $y(t)$ для отримання відношення правдоподібності Λ можна вирішити, якщо апіорно відомі хоча б деякі дані про корисний сигнал $s(t)$, характеристики ймовірності завади $n(t)$ і характер взаємозв'язку між корисним сигналом і завадою.

Розглянемо простий випадок виявлення при методі *одноразового відліку*.

Метод одноразового відліку - це процес обробки єдиного відліку $y(t_i) = y_i$ реалізації вхідного сигналу, взятого в деякий момент часу t_i .

На основі відліку $y(t_i) = y_i$ ухвалюється рішення про наявність або відсутність корисної складової в прийнятому інформаційному повідомленні.

Миттєве значення реалізації вхідного сигналу, як суми корисного сигналу та завади:

$$y_i = s_i + n_i \quad (9.44)$$

При відсутності корисного сигналу $s_i = 0$ і $n_i \neq 0$. Тоді

$$P(y/0) = P(y_i/0) = P(n_i) = w(n_i)dx_i = w(y_i)dy_i, \quad (9.45)$$

де $w(n_i)$ - одновимірна щільність ймовірності завади.

Ймовірність $P(y/s)$ отримання реалізації сигналу з завадою співпадає з ймовірністю отримання випадкової величини $(y_i - s_i)$, рівній величині n_i .

Тому $P(y/s) = P(y_i/s_i) = P[(y_i - s_i)] = w_i(y_i, s_i)dy_i$.

$$\Lambda = P(y/s) = P(y/0) = w_i(y_i, s_i)/w(y_i). \quad (9.47)$$

Заваду можна вважати стаціонарним нормальним випадковим процесом з нульовим середнім і дисперсією σ^2 . Тоді

$$w_1 = w(y_i, s_i) = \exp\{-(y_i - s_i)^2 / (2\sigma^2)\};$$

$$w = w(y_i) = \exp\{-(y_i)^2 / (2\sigma^2)\} / \sqrt{2\pi\sigma^2};$$

$$\Lambda = \exp\{(s_i / \sigma^2) / (y_i - 0,5s_i)\}. \quad (9.48)$$

З (9.48) витікає, що при відомих s_i і σ^2 відношення правдоподібності Λ і відлік y_i реалізації зв'язані між собою однозначно. Кожному відліку y_i відповідає цілком певне значення Λ . Тому виявляється достатнім порівнювати відліки y_i з деяким порогом, одержаним з (9.48) при $\Lambda = \Lambda_n$

$$\Lambda = \Lambda_n:$$

$$y_n = (\sigma^2 \ln \Lambda_n + 0,5s_i^2) / s_i. \quad (9.49)$$

При $y_i > y_n$ видається рішення "Так", при $y_i < y_n$ - рішення "Ні".

Основні недоліки методу, які вносять невизначеність в рішення задачі: невизначеність частоти послідовності відліків. При дуже рідкісних відліках сигнал може бути пропущений; визначення значення s_i у момент відліку. Безпосереднє вимірювання миттєвого значення s_i по набутому миттєвому значенню y_i неможливе через наявність випадкової величини n_i .

Відліки слід виробляти безперервно. В цьому випадку пропускання сигналу через дискретність відліків не можливе. Таким чином, сигнал $y(t)$ повинен поступати на вирішувальний пристрій безперервно і порівнюватися з порогом y_n (рис. 9.18). Якщо $s_i = a$ і $A_n = 1$, то $y_n = 0,5a$. При цьому:

$$\text{Ймовірність помилкової тривоги: } P_{\text{пт}} = P[(y_i > y_n) / 0] = \int_{y_n}^{\infty} w(y) dy, \quad (9.50)$$

де $y(t)$ не містить корисного сигналу.

$$\text{Ймовірність пропуску сигналу: } P_{\text{прп}} = P[(y_i < y_n) / s] = \int_{-\infty}^{y_n} w_i(y) dy, \quad (9.51)$$

де $y(t)$ – містить корисну складову.

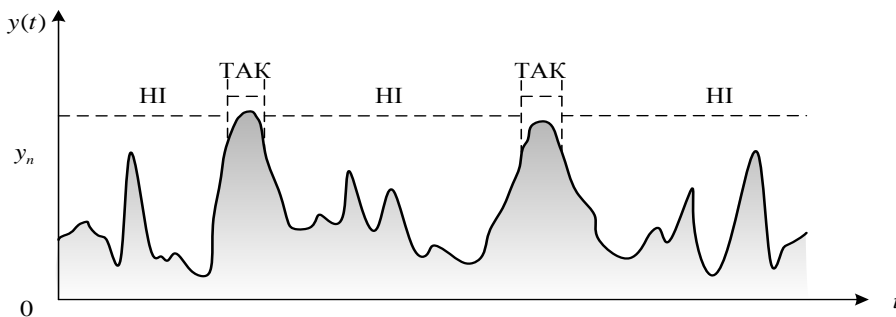


Рис. 9.18. Схема опрацювання сигналів на вирішувальному пристрої

Кореляційний метод виявлення сигналів на фоні завад. Ухвалення рішення не за одним значенням визначеної величини, а по великому числу N значень, дозволяє одержати більш позитивний ефект, якщо різні відліки взаємно незалежні. Для виконання цієї умови відліки повинні відрізнятися один від одного не менше, ніж на $\Delta\tau$ – інтервал кореляції завади. Для завади типу

“білого шуму” $\Delta\tau \rightarrow 0$. В цьому випадку:
$$\Lambda = \prod_{i=1}^N w_i(y_i, s_i) / \prod_{i=1}^N y_i$$

(9.52)

Для нормальних випадкових процесів з нульовим середнім і дисперсією σ^2 відношення правдоподібності запишеться у вигляді:

$$\Lambda = \exp \left\{ \sum_{i=1}^N [-(y_i - s_i)^2 / 2\sigma^2] \right\} / \exp \left\{ \sum_{i=1}^N (-y_i^2 / 2\sigma^2) \right\}. \quad (9.53)$$

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

Перетворення (9.53) зводяться до
$$\ln \Lambda = \left(\sum_{i=1}^N s_i y_i - 0,5 \sum_{i=1}^N s_i^2 \right) / \sigma^2. \quad (9.54)$$

Правило опрацювання зручніше представити у вигляді

$$(a_y)_N = \sum_{i=1}^N s_i y_i = \sigma^2 \ln \Lambda + 0,5 \sum_{i=1}^N s_i^2, \quad (9.55)$$

з якого виходить, що при опрацюванні необхідно визначити величину $(\sum_{i=1}^N s_i y_i)$ і порівняти її з порогом: $(a_y)_N^n = \sigma^2 \ln \Lambda_n + 0,5 \sum_{i=1}^N s_i^2,$ (9.56)

який знаходять з (9.54) при $\Lambda = \Lambda_n.$

При $(a_y)_N > (a_y)_N^{npz}$ видається рішення “Так”. Якщо вибірка береться на інтервалі $[0, t],$, а відліки в ній беруться через $\Delta\tau \rightarrow 0,$ то суми в (9.56) перетворюються на інтеграли, а величина $\sigma^2 \Delta\tau$ – в спектральну щільність потужності. Тому маємо: $\int_0^{t_0} s(t)y(t)dt = \ln \Lambda + 0,5 \int_0^{t_0} s^2(t)dt.$ (9.57)

Процедура ухвалення рішення згідно (9.57), полягає в перемножуванні реалізації $y(t)$ на очікуваний сигнал $s(t),$ інтеграції одержаного добутку в межах від нуля до t_0 і порівнянні результату з порогом.

Інтеграл $\int_0^{t_0} s(t)y(t)dt$ є мірою взаємної кореляції між реалізацією $y(t)$ і корисним сигналом $s(t).$ Тому його називають *кореляційним інтегралом,* а описану вище процедуру виявлення – *кореляційним методом.*

Функціональна схема пристрою для виявлення сигналів, яка працює на основі кореляційного методу, показана на рис. 9.19.

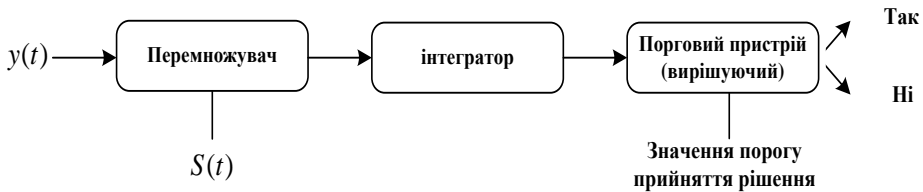


Рис. 9.19. Функціональна схема пристрою для виявлення сигналів кореляційним методом

Умовна ймовірність $P_{п.т}$ і $P_{прп}$ визначається за тією ж методикою, що і у разі одноразового відліку з попереднім уточненням закону розподілу.

Синтез приймачів безперервних сигналів. Узгоджений прийом. Ця задача відноситься до класу виділення сигналів на фоні завад (або

відтворення повідомлень, задачі фільтрації). *Вперше вона сформульована і вирішена Н. Вінером в 1941 р.*

Нехай на вхід приймача (рис. 9.20) поступає вхідний сигнал з завадою $y(t) = s(t) + n(t)$. Приймач характеризується передавальною функцією $K(i\omega)$, яку необхідно визначити. Етапи пошуку зводяться до наступного.



Рис. 9.20. Опрацювання сигналів в приймачі

$$\text{Сигнал на виході приймача: } \gamma(t) = \int_0^{\infty} y(t - \tau)g(\tau)d\tau, \quad (9.58)$$

де $g(\tau) = 1/2\pi \int_{-\infty}^{\infty} K(i\omega)e^{i\omega\tau}d\omega$ – імпульсна перехідна функція.

Помилка у відтворенні сигналу: $v(\tau) = \gamma(\tau) - s(t)$.

Задача оптимального прийому - знайти передавальну функцію $K(i\omega)$ приймача, яка забезпечує мінімальне значення дисперсії (середнього квадрата) помилки:

$$\bar{v}^2 = [\gamma(t) - s(t)]^2. \quad (9.59)$$

Після підстановки в (6.59) відомих і шуканих величин задача зводиться до інтегрального рівняння, вирішити яке можна методами варіаційного числення. Рішення, одержане Н. Вінером, має вигляд

$$K(i\omega) = S(\omega) / (S(\omega) + N(\omega)), \quad (9.60)$$

де $S(\omega)$, $N(\omega)$ – енергетичні спектри сигналу і завади.

Дана постановка задачі потребує уточнення критерію оптимальності щодо передавальної функції приймача.

Передавальна характеристика приймача повинна забезпечити максимум відношення сигналу до шуму на виході системи. Умову оптимальності в такому сенсі забезпечує фільтр (приймач) з передавальною функцією:

$$K(i\omega) = aS^*(i\omega)e^{-i\omega t_0}, \quad (9.61)$$

де a , t_0 – постійні значення, $S^*(i\omega)$ – спектр, зв'язаний зі спектром сигналу.

Узгодженою фільтрацією називається процес оптимізації передавальної характеристики фільтра та її узгодження із спектром вхідного сигналу з метою забезпечення максимуму відношення корисного сигналу до шуму на виході системи.

На виході приймача відношення сигналу до шуму: $h_{\max} = \sqrt{2E/N_0}$, де $E = 1/2\pi \int_{-\infty}^{\infty} S^2(\omega)d\omega$ – енергія сигналу на вході приймача, N_0 – енергетичний

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

спектр завад. Сигнал на виході узгодженого фільтру співпадає формою з автокореляційною функцією вхідного сигналу, оскільки

$$s_{\text{вих}} = \int_{-\infty}^{\infty} s_{\text{вх}}(t-x)g(x)dx, \text{ а для узгодженого фільтру } g(t) = A_0s_{\text{вх}}(t-t_0). \text{ Тому}$$

$$s_{\text{вих}} = A_0 \int_{-\infty}^{\infty} s_{\text{вх}}(y)s_{\text{вх}}(y-\tau)dy = A_0K_s(\tau), \text{ де } \tau = t-t_0.$$

Це означає, що приймач можна будувати за правилами обчислення взаємно кореляційної функції вхідного сигналу з очікуваним. На рис. 9.21 показана кореляційна функція K відрізка гармонічного коливання $K = f(t)$.

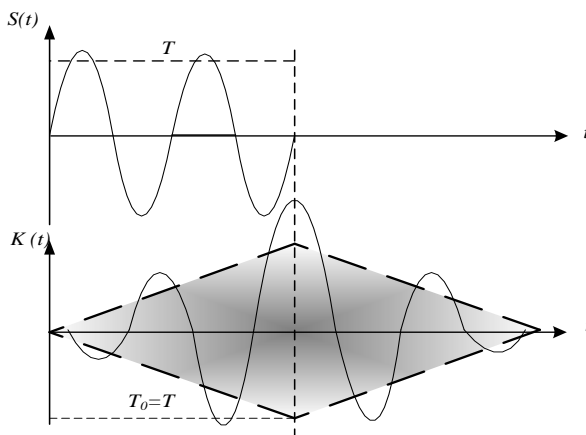


Рис. 9.21. Кореляційна функція відрізка гармонічного коливання

Ідеальний приймач Котельникова. Передавання даних може здійснюватися з використанням амплітудної, частотної або фазової модуляції будь-яких повідомлень дискретного або безперервного характеру (в останньому випадку після відповідної дискретизації відповідно теореми В.А. Котельникова).

У залежності від використовуваних при прийомі параметрів сигналів розрізняють *когерентний* і *некогерентний* прийоми. У першому випадку використовуються частота і фаза сигналу, у другому — частота, а фаза до уваги не приймається.

Порівняння методів передавання і приймання здійснюється за їх завадостійкістю. Для її оцінки використовується розроблена в 1946 р. В.А. Котельниковим теорія потенційної завадостійкості. В.А. Котельников обґрунтував, що для зменшення впливу флуктуаційних завад існує найкращий (ідеальний) приймач, що володіє найбільшою (потенційною) завадостійкістю для даного методу передавання.

Розглянемо підходи, покладені в основу побудови ідеальних приймачів при когерентному прийомі. Якщо сигналам 1 і 0 на вході

когерентного приймача відповідають сигнали $A(t)$ і $B(t)$ рівної тривалості t_c , а сума сигналу і флуктуаційної завади – $x(t)$, то, відповідно теореми В.А. Котельникова, приймач забезпечить найменшу ймовірність викривлення (викривлення символу), якщо буде видавати сигнал A при виконанні умови: $I(A) < I(B)$, (6.62)

інакше - сигнал B . У даному виразі: t_c – тривалість двійкового сигналу 0 або 1; $I(B) = \int_0^{t_c} [x(t) - B(t)]^2 dt$; $I(A) = \int_0^{t_c} [x(t) - A(t)]^2 dt$.

Оптимальний приймач за критерієм ідеального спостерігача (ідеальний приймач Котельникова) є технічний засіб, що реалізує алгоритм опрацювання вхідного інформаційного повідомлення на основі порівняння прийнятого сигналу з завадами із не викривленими (ідеальними) зразками сигналу, обчислює енергію різниці і відносить прийнятий сигнал до того зразка, для якого енергія різниці мінімальна.

Використовуючи вираз (9.62), можна побудувати ідеальний приймач Котельникова. Структурна схема такого приймача представлена на рис. 9.22.

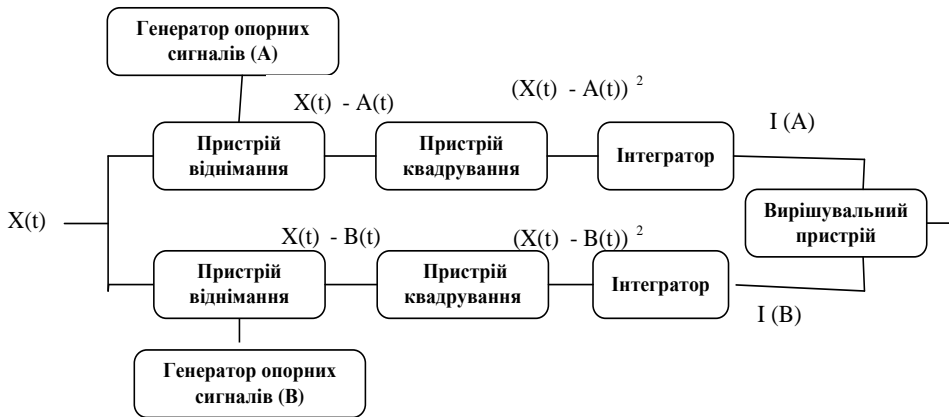


Рис. 9.22. Структурна схема ідеального приймача Котельникова

Вона включає: генератори опорних сигналів, що точно повторюють передані сигнали A і B ; пристрій віднімання, пристрій квадратування; інтегратор; вирішувальний пристрій. Для генерації опорних сигналів у приймачі повинні бути відомі всі параметри переданих сигналів A і B . Приймач має дві гілки, у кожній з яких обчислюється середня квадратична величина відхилення прийнятого колювання $x(t)$ від відомого сигналу: у першій гілці – від сигналу $A(t)$, у другій – від сигналу $B(t)$. Рішення приймається по тій гілці, де ця відмінність менша.

Оптимальні приймачі з активною і пасивною паузами.

Позначимо енергії сигналів A і B відповідно через E_A й E_B :

$$E_A = P_A, \quad t_c = \int_0^{t_c} [A(t)]^2 dt; \quad E_B = P_B, \quad t_c = \int_0^{t_c} [B(t)]^2 dt, \quad (9.63)$$

де P_A і P_B – відповідно питома середня потужність сигналів A і B (потужність, що розвивається на одиничному резисторі).

Сигналом з пасивною паузою називається сигнал, що випромінюється тільки при передачі одного із символів (наприклад, 1), передачі іншого символу (наприклад, 0) відповідають паузи між сигналами. У цьому випадку: $A(t) \neq 0, B(t) = 0, E_A = E, E_B = 0.$

$$(9.64)$$

Якщо підставити в (9.62) вирази (9.63) і (9.64), то одержимо вираз, що описує алгоритм роботи оптимального приймача сигналів з пасивною паузою:

$$\int_0^{t_c} x(t)A(t)dt \geq E/2 \text{ – прийнято сигнал } A;$$

$$(9.65)$$

$$\int_0^{t_c} x(t)A(t)dt \leq E/2 \text{ – прийнято сигнал } B.$$

Ліва частина цих виразів характеризує взаємну кореляцію між $x(t)$ і $A(t)$ (кореляційний інтеграл). Відповідно до правила (9.65), приймач повинен обчислити значення кореляційного інтеграла і порівняти його з деякою фіксованою величиною – порогом. Таким чином, оптимальний приймач сигналів з пасивною паузою – це кореляційний приймач (рис. 9.23), що складається з корелятора і вирішувального пристрою.

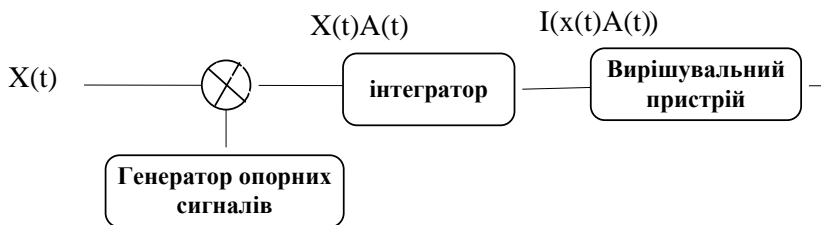


Рис. 9.23. Структурна схема оптимального кореляційного приймача з пасивною паузою

Корелятор містить у собі перемножувач, блок прийняття рішення, інтегратор і пристрій знімання відліків (на схемі не показано) у моменти часу, кратні тривалості сигналів: $(t_k = kt_c, k = 1, 2, \dots)$. Розглянутий приймач називають

ще когерентним приймачем з пасивною паузою. Якщо кореляційний інтеграл обчислювати за допомогою лінійного фільтра, погодженого із сигналом $A(t)$, то прийдемо до приймача (рис. 9.24) з оптимальним фільтром.

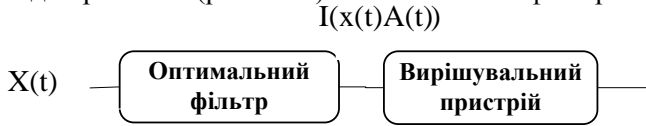


Рис. 9.24. Структурна схема приймача з оптимальним фільтром та пасивною паузою

Сигналом з активною паузою називають такий сигнал, що випромінюється для передавання будь-якого символу (0 або 1). Даний сигнал повинен мати двійкову структуру для прийняття двох різних значень. Найчастіше використовуються сигнали з рівними енергіями: $E_A = E_B = E$. Перетворимо (9.62) до виду:

$$2 \int_0^{t_c} x(t)A(t)dt - \int_0^{t_c} [A(t)]^2 dt > 2 \int_0^{t_c} x(t)B(t)dt - \int_0^{t_c} [B(t)]^2 dt,$$

і з урахуванням (9.63) і (9.64), а також того, що $\int_0^{t_c} [A(t)]^2 dt = \int_0^{t_c} [B(t)]^2 dt$, одержимо два варіанти запису. Перший з них:

$$\int_0^{t_c} x(t)A(t)dt - \int_0^{t_c} x(t)B(t)dt > 0 \quad - \text{прийнято сигнал } A,$$

$$\int_0^{t_c} x(t)A(t)dt - \int_0^{t_c} x(t)B(t)dt < 0 \quad - \text{прийнято сигнал } B,$$

Другий варіант: $\int_0^{t_c} x(t)\Delta A(t)dt > 0$ – прийнято сигнал A , (9.66)

$$\int_0^{t_c} x(t)\Delta A(t)dt < 0 \quad - \text{прийнято сигнал } B, \quad \Delta A(t) = A(t) - B(t).$$

Вирази (9.66) і (9.65) визначають різні алгоритми роботи оптимального приймача з активною паузою.

9.5. Багатоканальні мережі передавання даних

Поділом каналів називається процес розділення (розмежування) інформаційних трактів (мереж) на тлі інформаційних параметрів переданих сигналів з метою з'єднання кожного визначеного джерела повідомлень зі своїм приймачем.

Введемо позначення: $U_k(t)$ – сигнали сенсорів, що відображають інформаційні функції $x_k(t)$ й породжують модуляцію параметрів a_k носія;

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

$u_{xk}(t)$ – сигнали на виходах передавальних пристроїв окремих каналів; $u(t)$ – сумарний сигнал у лінії зв'язку.

Сигнал в одному каналі описується виразом:

$$u_x(t) = g[a_1, \dots, a_i(t), \dots, a_j, \dots, a_n],$$

де параметр $a_i(t)$ передає інформацію, а параметр a_j або кілька параметрів можуть бути використані для характеристики індивідуальних каналів. Такими параметрами можуть бути приналежність до певного електричного кола, частота або фаза носія, положення на часовій осі, форма тощо. Кожному каналу відповідає певне значення a_{jk} або область значень Δa_{jk} параметру a_j . При використанні цього позначення вираз для сигналу k -го каналу приймає вигляд $u_{xk}(t) = g[a_1, \dots, a_i(t), \dots, a_{jk}(t), \dots, a_n]$, або скорочено: $u_{xk}(t) = g_k[a_k(t)]$. У лінію зв'язку надходить складений сигнал, що звичайно є сумою сигналів окремих каналів: $u(t) = \sum_k g_k[a_k(t)]$.

Процес поділу можна розглядати як фільтрацію, що здійснює виділення u_{xk} або сигналу сенсора $U_k : \Phi_k \left\{ \sum_k g_k[a_k(t)] \right\} = g_k[a_k(t)]$.

Залежно від виду оператора фільтрації Φ , обумовленого у свою чергу обраним параметром a , розрізняють наступні методи розподілу каналів: просторовий (схемний), диференціальний, частотний, часовий, фазовий, кодовий, за рівнем, за формою, кореляційний (рис. 9.25).



Рис. 9.25. Класифікація методів розподілу каналів передавання даних

Багатоканальні системи зв'язку з просторовим поділом. Це найпростіший вид розподілу, при якому кожному каналу виділяється індивідуальна мережа зв'язку (рис. 9.26).

Прикладом служить просторовий розподіл сигналів за допомогою матричних комутаторів з n входами й n виходами. Інші форми розподілу

каналів припускають передавання повідомлень по одній мережі. У зв'язку із цим багатоканальне передавання називають також ущільненням каналів.

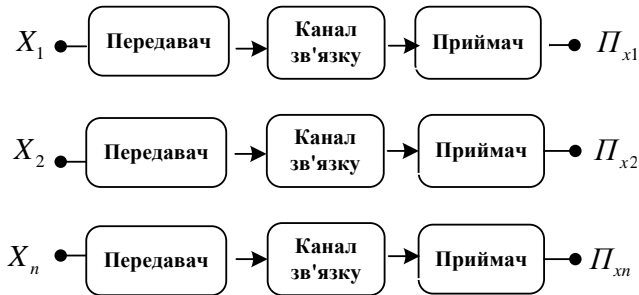


Рис. 9.26. Багатоканальна система з просторовим розподілом
 X_n - датчик n -го каналу; P_{xn} - приймач інформації n -го каналу

Багатоканальні системи зв'язку з диференціальним розподілом. На рис. 9.27 наведена широко розповсюджена схема використання комутативного (провідного) каналу для передачі інформаційних сигналів (від сенсора X до приймача). У лінію на передавальній і приймальній стороні включаються диференціальні трансформатори (ДТ). Середні точки їх з'єднують із X і P_x . Таким чином, телекомунікаційні сигнали не створюють завод у первинних колах диференціальних трансформаторів, пов'язаних з телефонними апаратами A . Завдяки диференціальному включенню телефонні сигнали також не створюють завод.

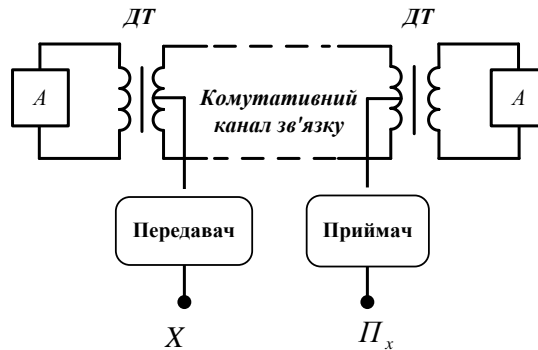


Рис. 9.27. Багатоканальна система з диференціальним розподілом:
 A – телефонний апарат; ДТ – диференціальний трансформатор;
 X – сенсор сигналів; P_x – приймач

Багатоканальні системи зв'язку з частотним розподілом. При частотному розподілі для різних каналів відводяться ділянки (робоча ширина спектру визначеного каналу Δf_k), що не перетинаються

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

$\Delta f_1, \Delta f_2, \dots, \Delta f_n$, на частотній шкалі f (рис. 9.28,а). Спектри сигналів u_{kx} відповідних каналів повинні вкладатися в межі Δf_k . Смуга пропускання мережі зв'язку $\Delta F_{\text{Д}}$ визначається крайніми частотами (мінімальною частотою інтервалу Δf_1 і максимальною частотою інтервалу Δf_n). На рис. 9.28,б наведена схема багатоканальної системи із частотним розподілом. Низькочастотні сигнали U_k сенсорів X модулюють по амплітуді або частоті високочастотні сигнали з несучими f_1, f_2, \dots, f_n , які виробляються спеціальними генераторами. Сигнали на виході модулаторів мають спектри Δf_k , положення яких на шкалі частот визначається несучими частотами f_k , а ширина залежить від ширини спектра сигналів датчиків. Смугові фільтри передавальної частини служать для обмеження смуги частот своїх каналів. На приймальній стороні фільтри розділяють сигнали, які, пройшовши через демодулятори, можуть бути сприйняті приймальними пристроями.

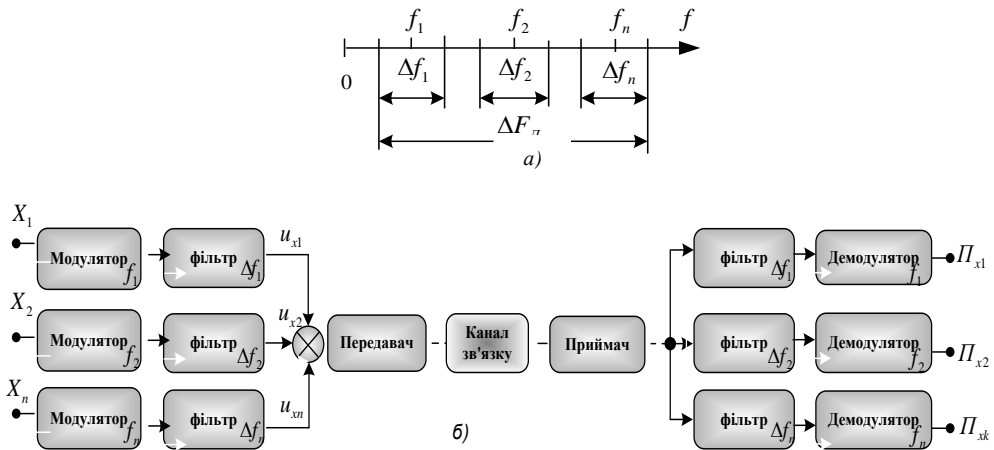


Рис. 9.28. Багатоканальна система з частотним розподілом:

а - розподіл каналів за шкалою частот; б - схема частотного розподілу

Перевагою систем із частотним розподілом є можливість одночасного передавання сигналів, що відносяться до різних каналів. Друга перевага полягає в можливості передавання сигналів від розосереджених об'єктів. Їхній недолік – порівняно великий взаємний вплив каналів через перекриття спектрів сигналів, неідеальності смугових фільтрів і поява паразитних частотних складових внаслідок нелінійності електричних кіл (так звана перехресна модуляція).

Багатоканальні системи зв'язку з часовим розподілом. При часовому розподілі сигнали u_{xk} сенсорів передаються тільки у

відведені для них відрізки часу Δt_k (рис. 9.29,а), які не пересікаються. Розподіл здійснюється розподільниками P (рис. 9.29,б), які повинні бути строго синхронізовані (тобто працювати з однаковою швидкістю) і синфазні (працювати без зрушення).

Взаємний вплив каналів при часовому розподілі зазвичай незначний, що дозволяє будувати системи з більшою кількістю каналів. Завдяки цій обставині, а також простоті технічних засобів цей метод використовується досить широко.

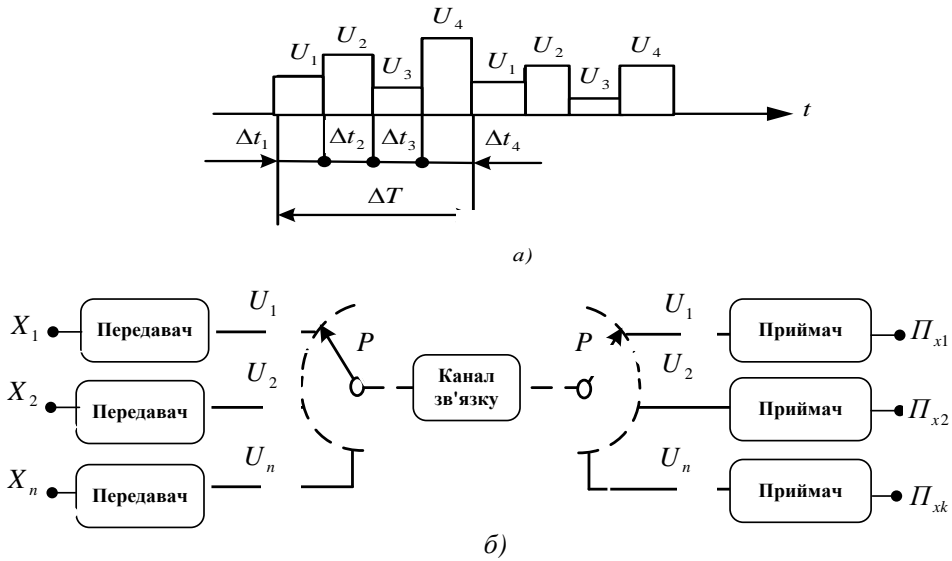


Рис. 9.29. Багатоканальна система з часовим розподілом:
 а – розподіл каналів за часом; б – схема часового розподілу

Багатоканальні системи зв'язку з фазовим розподілом. Фазовий розподіл застосовують у двоканальній системі (рис. 9.30) із синусоїдальними сигналами, фази яких відрізняються на 90° .

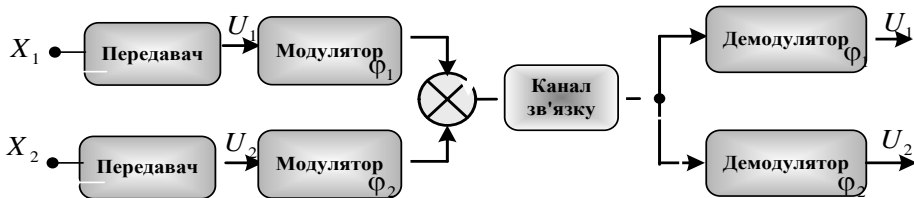


Рис. 9.30. Багатоканальна система з фазовим розподілом

Сигнали сенсорів X_n модулюють амплітуду синусоїдальних носіїв, що розрізняються по фазі. Таким чином, сигнали U_{xn} на виході модуляторів

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

мають амплітуди, обумовлені модульовальними функціями сенсорів, і фази відповідно φ_1 та $\varphi_2 = \varphi_1 + \pi/2$: $u_{x1} = U_1 \sin \omega_0 t$; $u_{x2} = U_2 \sin(\omega_0 t + \pi/2) = U_2 \cos \omega_0 t$. Фазові детектори виділяють відповідні модульовальні функції U_1 та U_2 .

Багатоканальні системи зв'язку з кодовим розподілом. При кодовому розподілі адреса каналу вказується кодованим сигналом, що посиляється в мережу зв'язку. Розподіл на приймальній стороні здійснюється декодувальним пристроєм, що направляє повідомлення по обраному каналу. Код адреси може бути як послідовним, так і паралельним. В останньому випадку використовується окрема мережа зв'язку або індивідуальний частотний канал на кожний розряд коду. Кодовий розподіл каналів дозволяє робити опитування каналів у довільному порядку, що робить зручним його використання в системах передачі даних.

Багатоканальні системи зв'язку з розподілом за формою. Для розподілу сигналів, що відрізняються формою, використовуються операції, найбільш чутливі до зміни форми, - звичайне диференціювання, інтегрування й віднімання. Розглянемо процедуру розподілу, коли функції носія отримують шляхом послідовного диференціювання.

Нехай, наприклад: $u_{x1}(t) = U_1$; $u_{x2}(t) = U_2 t$. У лінію надходить сума сигналів $u(t) = U_1 + U_2 t$. Процес розподілу має на меті виділення інформаційних параметрів U_1 та U_2 . Виділення U_2 здійснюється шляхом диференціювання функції $u(t)$. Інтегрування U_2 відновлює переданий сигнал другого каналу $u_{x2}(t)$. U_1 отримуємо шляхом віднімання $u_{x2}(t)$ з $u(t)$.

Багатоканальні системи зв'язку з кореляційним розподілом. У ряді випадків сигнали окремих каналів можуть бути представлені у вигляді: $u_{xk}(t) = g_k[a_k(t)] = a_k(t)g_k(t) = U_k(t)g_k(t)$, де функція $g_k(t)$ описує носій з деякою заданою величиною роздільного параметра a_{jk} , а інформаційний параметр $a_k(t)$, що модулює функцію $g_k(t)$ по амплітуді, дорівнює сигналу $U_k(t)$ відповідного сенсора. Цей параметр є функцією часу, що повільно змінюється в порівнянні з $g_k(t)$, і його можна вважати постійним. Сигнал у мережі є лінійною комбінацією функцій: $u(t) = \sum_k U_k g_k(t)$.

Якщо $g(t)$ – лінійно незалежні, вони можуть бути розділені лінійними фільтрами. Такі багатоканальні системи передачі зветься *лінійними*. До лінійних відносяться, зокрема, системи із частотним, часовим, фазовим поділом і поділом за формою. Важливим різновидом лінійно незалежних сигналів є ортогональні сигнали, для яких існує загальний метод розподілу із застосуванням оператора кореляційної фільтрації до сигналу, що надходить із мережі зв'язку. Для геометричної інтерпретації поняття ортогональності розглянемо дві функції $g_i(t)$ та $g_j(t)$ визначені на інтервалі $T_1 < t < T_2$. Представимо їх у вигляді сукупності дискретних відліків:

$$g_i^* = (g_i(t_1), g_i(t_2), \dots, g_i(t_m)); \quad g_j^* = (g_j(t_1), g_j(t_2), \dots, g_j(t_m)).$$

Ці сукупності можна розглядати як два вектори в m -мірному евклідовому просторі з координатами $g_i(t_k)$ та $g_j(t_k)$, $k = 1, 2, \dots, m$. Якщо вектори ортогональні, то скалярний добуток дорівнює нулю: $(g_i^*, g_j^*) = \sum_k g_i(t_k)g_j(t_k) = 0$.

При нескінченному збільшенні числа відліків m дискретні функції перетворюються в безперервні. За аналогією з попереднім їх також можна розглядати як вектори, але вже в нескінченно мірному просторі. Умова ортогональності приймає інтегральний вид і визначається на заданому інтервалі (T_1, T_2) :

$$(g_i(t), g_j(t)) = \int_{T_1}^{T_2} g_i(t)g_j(t)dt = 0.$$

Ортогональну систему утворюють наступні функції:

1. *Нескінченна множина функцій $\cos k\omega t$, $\sin k\omega t$ ортогональних на інтервалі $0 < t < 2\pi / \omega$, де k – не негативне число (рис. 9.31,а). Носій з такими функціями використовується в пристроях із частотним і фазовим поділом.*

2. *Множина довільних функцій, визначених на непересічних інтервалах часу, а в інший час тотожно рівних нулю (рис. 9.31,б), також є ортогональною системою, тому що при $i \neq j$: $(g_i(t), g_j(t)) = 0$.*

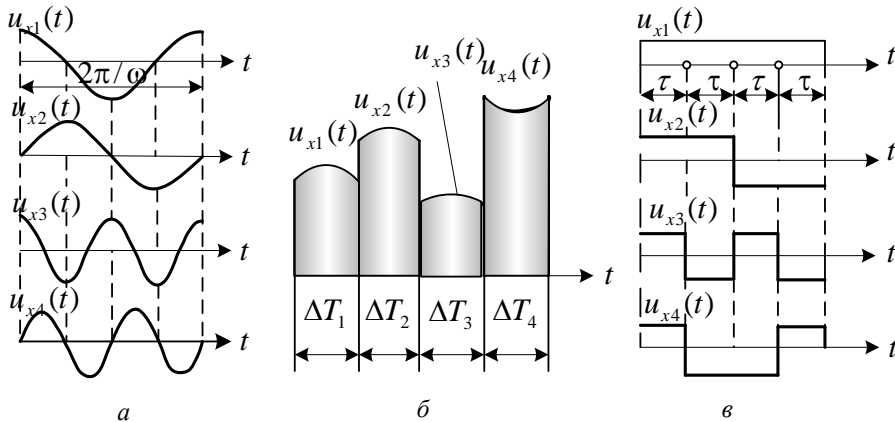


Рис. 9.31. Приклади ортогональних сигналів: *a* – з гармонічним носієм; *б* – із будь-яким носієм з розподілом за часом; *в* – з імпульсним носієм

3. Множина, що складається з дискретних знакозмінних функцій, які можна одержати за допомогою n - розрядного лічильника в режимі віднімання, що перебуває спочатку в заповненому стані (рис. 9.31,в), ортогональна на заданому інтервалі $T = n\tau$ (τ - тривалість імпульсу молодшого розряду). Коди, що відповідають цим послідовностям, відносяться до групових і зуться кодами Ріда-Мюллера 1-го порядку.

4. Ортогональними на певних інтервалах є багато спеціальних функцій: поліноми й функції Лежандра, Чебишева, Якобі, Ерміта, Лагерра, Хаара, Уолша, Радемахера та інші.

Ортогональну систему зручно використовувати в нормованому виді, при

якому виконується умова: $\int_{T_1}^{T_2} g_k^2(t) dt = 1$. Якщо $\phi_k(t)$ – ненормовані

ортогональні функції, то операція нормування виконується шляхом

множення на коефіцієнт $\lambda_k = 1 / \sqrt{\int_{T_1}^{T_2} \phi_k^2(t) dt}$. У цьому випадку в мережу

надходить сигнал виду: $u_{xk}(t) = U_k \lambda_k \phi_k(t) = U_k g_k(t)$, де $g_k(t)$ вже є

нормованими функціями, що утворюють ортонормовану систему. Для

виділення інформаційного параметра U_k потрібно помножити прийнятий

сигнал $u(t)$ на функцію $g_k(t)$ та інтегрувати отриманий добуток у межах

$T_1 \leq t \leq T_2$:

$$\int_{T_1}^{T_2} \left[\sum_i U_i g_i(t) \right] g_k(t) dt = U_k \int_{T_1}^{T_2} g_k^2(t) dt = U_k .$$

Множення сигналу лінії на всі функції $g_k(t)$ забезпечує повний поділ будь-яких ортогональних сигналів. Оператор поділу Φ , який виконує це перетворення, визначає по суті ступінь взаємної кореляції сигналів $u(t)$ і $g_k(t)$. Таким чином, багатоканальна система (рис. 9.32) на передавальній стороні містить генератори G_n ортогональних функцій і модулятори M_n з нормалізаторами, а на приймальній – такі ж генератори G_n і корелятори K_n .

Ефективність кореляційного методу поділу полягає в тому, що він дозволяє значно послабити вплив перехресних завад, а це істотно у випадку перекриваючих спектрів сигналів.

Розглянемо як **приклад** систему, що використовує поліноми Лежандра першого роду. Ці поліноми ортогональні на інтервалі $-1 < t < +1$ і описуються співвідношеннями:

$$\phi_0(t) = 1; \phi_1(t) = t; \phi_k(t) = \left[(2k-1) \cdot \int_{-1}^t \phi_{k-1}(t) dt + \phi_{k-2}(t) \right], k \geq 2.$$

Скалярний добуток двох функцій:

$$\int_{-1}^1 \phi_i(t) \phi_j(t) dt = \begin{cases} 0 & \text{при } i \neq j, \\ 2/(2i+1) & \text{при } i = j. \end{cases}$$

Отже, нормуючий множник: $\lambda_k = \sqrt{(2k+1)/2}$.

Сигнали утворюються на основі носія у вигляді постійної події, промодульованої за формою відповідно до наведених співвідношень. Таким чином, форма служить параметром поділу, а для перенесення інформації використовується амплітуда.

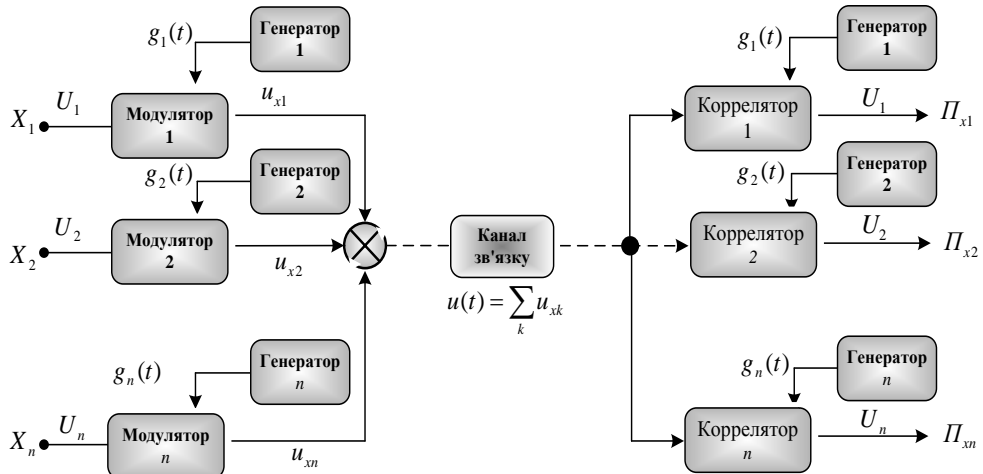


Рис. 9.32. Багатоканальна система з кореляційним розподілом

Функціональна схема генератора ортогональних функцій $\varphi_k(t)$ наведена на рис. 9.33.

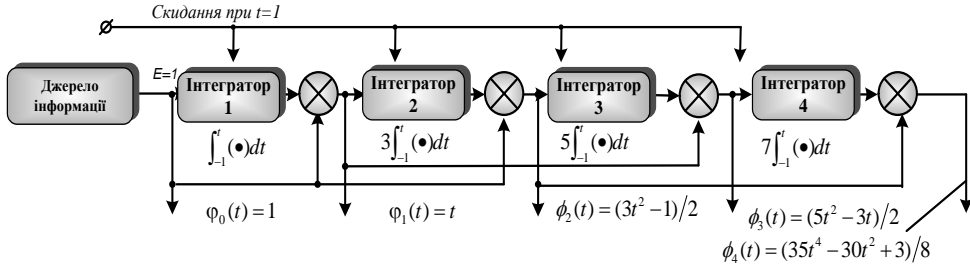


Рис. 9.33. Схема розподілу каналів на тлі ортогональних функцій

Джерело E створює постійний сигнал $E=1$, що відіграє роль функції $\varphi_0(t)$. Після інтегрування інтегратором I_1 і наступного віднімання $\varphi_0(t)$ з результату формується функція $\varphi_1(t)$. Формування наступних функцій здійснюється відповідно до вищенаведених рекурентних співвідношень шляхом послідовного інтегрування функцій $\varphi_{k-1}(t)$ із множенням результату на відповідний коефіцієнт $2k-1$ і додаванням функції $\varphi_{k-2}(t)$. Після нормування й модуляції утвориться сигнал: $u_{xk}(t) = U_k \lambda_k \varphi_k(t) = U_k g_k(t)$. Повний сигнал у мережі зв'язку: $u(t) = \sum_k U_k g_k(t)$. Поділ каналів виконується за загальною схемою рис. 9.32. Із цією метою на прийомній стороні використовується аналогічний генератор поліномів, синхронізований з першим, і корелятор, який виділяє інформаційні параметри U_k .

Багатоканальні системи зв'язку з комбінованими методами поділу. Одночасне використання декількох методів поділу дозволяє збільшити число каналів і зменшити їхній взаємний вплив. Так, поділ за формою в сполученні із частотним або часовим поділом подвоює загальне число каналів. Перехресні перекручування обмежують ємність системи із частотним поділом каналів. Крім того, апаратура із частотним поділом при великій кількості каналів виявляється складною через наявність роздільних фільтрів. Комбінування частотного методу з часовим усуває ці недоліки. При цьому часовий поділ здійснюється комутацією одного або декількох частотних каналів. Застосовуються також комбінації кодових і частотних, кодових і часових методів поділу й інші.

9.6. Завадостійкість систем передавання інформації

Критерії оцінки завадостійкості систем передавання. В основі методів підвищення завадостійкості лежать різні способи введення надлишковості. Підвищення функціональної надійності або завадостійкості систем передавання інформації досягають збільшенням тимчасової, частотної, енергетичної надлишковості тощо.

Завадостійкість системи – здатність системи протистояти впливу випадкових завад (або навмисних) на корисний сигнал та зменшенню (чи унеможливленню) достовірного передавання даних від джерела повідомлення до споживача.

Оскільки в результаті дії завади прийнятий сигнал завжди відрізняється від переданого, то завадостійкість характеризується ступенем відповідності прийнятого повідомлення переданому. При порівнянні різних систем, та з них буде більш завадостійкою, яка при даному рівні завад забезпечить меншу відмінність між прийнятим і переданим повідомленнями. Кількісна міра цієї відповідності вибирається по-різному залежно від характеру повідомлення.

При оцінці завадостійкості систем передачі безперервних повідомлень в каналах з флуктуаційною завадою типу «білого шуму» використовується відношення середніх потужностей корисного сигналу і завади h^2 . При порівнянні різних способів прийому використовують поняття «переваги» системи, де $h_{\text{вх}}^2$ і $h_{\text{вих}}^2$ – відношення середніх потужностей сигналу і завади на вході і виході пристрою. У разі передавання дискретних повідомлень як критерій завадостійкості (надійності або «вірності») використовують ймовірність правильного прийому символу, де P_0 – ймовірність помилки при прийомі символу. Величина помилки P_0 залежить від h^2 в безперервному каналі, через який проходять сигнали, відповідні переданим символам.

При оцінці способів передавання проводять оцінку завадостійкості різних видів модуляції порівняно з амплітудною. Введену надлишковість найлегше реалізувати в передавачі: підвищити потужність сигналу, застосувати, наприклад, більш високочастотний носій, нарешті, використовувати кодування, що коректує, тобто у всіх випадках збільшити об'єм переданого сигналу. Етап передавання в принципі змінити складніше. На вході приймача з'являються завади, що виникли в мережі зв'язку, які не можна виключити, але можна тим або іншим способом зменшити.

Будь-який дискретний приймач можна представити у вигляді структурної схеми (рис. 9.34).

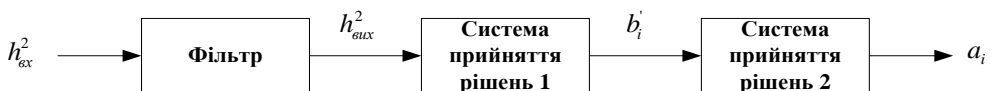


Рис. 9.34. Узагальнена структурна схема дискретного приймача

При цьому виконуються наступні операції:

Розділ 9. Передавання вимірювальної інформації в каналах ІВС

фільтрація або «очищення» сигналу від завади до максимально можливого значення $B = h_{\text{вих}}^2 / h_{\text{вх}}^2$;

забезпечення правильного вибору символу повідомлення – перша вирішальна система;

рішення задачі про ототожнення сигналу, що є комбінацією символів, з переданим корисним повідомленням – друга вирішальна система.

Поелементним прийомом (послідовним) називається такий прийом, при якому остаточний висновок про ухвалення повідомлення робиться шляхом аналізу кожного елемента сигналу.

Прийомом в цілому називається такий прийом, при якому прийнятий сигнал можна не розбивати на елементи, а висновок про прийняте повідомлення робити, аналізуючи всю послідовність символів.

У ряді випадків, коли передане повідомлення має надлишковість, прийом в цілому забезпечує велику надійність виводу, ніж поелементний прийом. В деякому розумінні такі пристрої можуть бути навіть простіше традиційних, оскільки в них немає першої системи на прийняття рішення щодо символів повідомлення.

Потенційна завадостійкість при прийомі дискретних сигналів.

При розрізненні двох сигналів за рахунок дії завади, крім правильних рішень можливі і помилкові.

Знайдемо ймовірність викривлення $P_{\text{зав}}$, що забезпечується ідеальним приймачем Котельникова (рис. 9.22). Ймовірність викривлення – ймовірність помилкового рішення приймача та дорівнює ймовірності невиконання умови реалізації оптимального приймача (9.62).

Викривлення виникають через те, що в прийнятий сигнал входить випадкова складова – флуктуаційна завада. У більшості випадків, заваду вважають стаціонарним ергодичним процесом з нормальним законом розподілу миттєвих значень напруг:

$$w(U_n) = \exp\left[-(U_n - \bar{U}_n)^2 / (2\sigma_n^2)\right] / \left[\sigma_n^2 \sqrt{2\pi}\right],$$

де U_n – миттєве значення напруги завади (перешкоди); \bar{U}_n – середнє значення напруги завади (постійна складова); σ_n^2 – дисперсія (середній квадрат відхилення від постійної складової) завади; $w(U_n) = \exp\left[-(U_n - \bar{U}_n)^2 / (2\sigma_n^2)\right] / \left[\sigma_n^2 \sqrt{2\pi}\right]$ – функція щільності розподілу ймовірностей згідно нормального закону.

Оскільки постійна складова завади дорівнює нулеві, то помилка відбувається тоді, коли миттєве значення напруги завади більше деякого

значення X . Ймовірність цієї події дорівнює площі під кривою, розподілу від X до ∞ (рис. 9.35): $P_{\text{зав}} = P(|U_n| > x) = \int_x^{\infty} w(U_n) dU_n = -\Phi(x)$, (9.67)

де $\Phi(\alpha) = \left(\int_{-\infty}^{\alpha} \exp[-z^2/2] dz \right) / [\sigma_n^2 \sqrt{2\pi}]$ – функція Лапласа; α – узагальнена

функція, обумовлена відношенням енергій сигналу і завади, а також ступенем розходження сигналів A і B :

$$\alpha = \sqrt{\int_0^{\alpha} [A(t) - B(t)]^2 dt / (2N_0)}, \quad (9.68)$$

де: $N_0 = U_{\text{п.еф}}^2 / \Delta F$ – спектральна щільність потужності завади; $U_{\text{п.еф}}^2$ – інтегральна потужність завади; ΔF – смуга пропускання приймального фільтра.

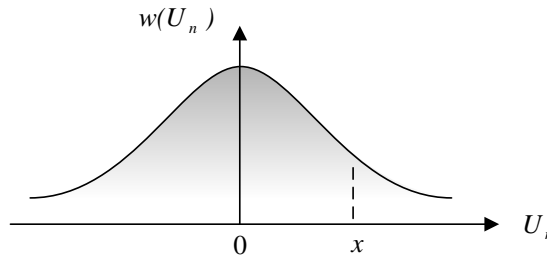


Рис. 9.35. До визначення ймовірності викривлення при оптимальному прийманні

Перетворимо вираз (9.68), для чого використовуємо (9.63), а також поняття питомої різниці енергій різницевого сигналу: $\Delta E = \Delta P_c \cdot t_c = \int_0^{t_c} \Delta A(t)^2 dt$, (9.69)

де ΔP_c – питома середня потужність різницевого сигналу $\Delta A(t)$. Тоді

$$\alpha = \sqrt{\Delta E / (2N_0)}. \quad (9.70)$$

Підставивши це значення в (5.67), одержимо

$$P_{\text{зав}} = 1 - \Phi \sqrt{\Delta E / (2N_0)}, \quad (9.71)$$

Ймовірність викривлення (9.71) визначає потенційну заводостійкість ідеального і будь-якого іншого оптимального приймача. Зменшити її можна за рахунок зменшення спектральної щільності потужності завод N_0 .

Потенційна заводостійкість прийому інформаційних сигналів за Котельниковим. Заводостійкість до дії завади типу «білого шуму» зветься потенційною (за Котельниковим), якщо вона забезпечує мінімальну середню ймовірність помилки розрізнення сигналів. Це досягається використанням критерію *ідеального спостерігача*, що

робить висновок про переданий код x_2 шляхом порівняння відношення правдоподібності з певним порогом

$$\omega(u_y | x_2) / \omega(u_y | x_1) > p(x_1) / p(x_2). \quad (9.72)$$

Припустимо, передані сигнали $u_{x_1}(t)$ і $u_{x_2}(t)$, що відповідають кодам x_1 і x_2 , мають довільну форму, а прийнятий сигнал $u_y(t) = u_x(t) + u_\xi(t)$. Представимо сигнал $u_y(t)$ у вигляді сукупності відліків u_{y1}, \dots, u_{yn} , розділених інтервалами $\Delta t = 1/2f_m$. Розподілу $\omega(u_y | x_i)$ у просторі сигналів u_y відповідає розподіл $w(u_\xi) = w(u_y - u_{xi})$ - у просторі сигналів u_ξ . Для гаусової завади при відсутності кореляції між значеннями шуму в моменти відліку

різницева щільність:

$$w(u_y - u_{xi}) = \prod_{j=1}^n e^{-\frac{(u_{yj} - u_{xij})^2}{2\sigma_\xi^2}} / \sqrt{2\pi\sigma_\xi} = e^{-\frac{1}{2\sigma_\xi^2} \sum_{j=1}^n (u_{yj} - u_{xij})^2} / (\sqrt{2\pi\sigma_\xi})^n,$$

де $\sigma_\xi = \sqrt{D[U_\xi]}$ - середнє квадратичне відхилення завади. Правило для

прийняття рішення: $e^{\left[\sum_{j=1}^n (u_y - u_{x1j})^2 - \sum_{j=1}^n (u_y - u_{x2j})^2 \right] / 2\sigma_\xi^2} > p(x_1) / p(x_2),$; або після логарифмування:

$$\left[\sum_{j=1}^n (u_y - u_{x1j})^2 - \sum_{j=1}^n (u_y - u_{x2j})^2 \right] / 2\sigma_\xi^2 > \ln p(x_1) - \ln p(x_2). \quad (9.73)$$

Помножимо обидві частини нерівності на Δt і перейдемо до неперервних у часі функцій, вважаючи $u_x(t)$ і $u_\xi(t)$ на малих інтервалах постійними. Суми можна замінити на інтеграли відповідно до наближених рівностей:

$$\sum_{j=1}^n (u_y - u_{x1j})^2 \Delta t \approx \int_0^x [u_y(t) - u_{x1}(t)]^2 dt; \quad \sum_{j=1}^n (u_y - u_{x2j})^2 \Delta t \approx \int_0^x [u_y(t) - u_{x2}(t)]^2 dt.$$

Оскільки $\Delta t = 1/2f_m$, то (9.20) приймає вид:

$$\int_0^{T_x} [u_y(t) - u_{x2}(t)]^2 dt - C_2 < \int_0^{T_x} [u_y(t) - u_{x1}(t)]^2 dt - C_1; \quad C_k = \sigma_\xi^2 \ln p(x_k) / f_m, \quad k=1,2.$$

Схема, що реалізує це правило, одержала назву *оптимального приймача* (рис. 9.36).

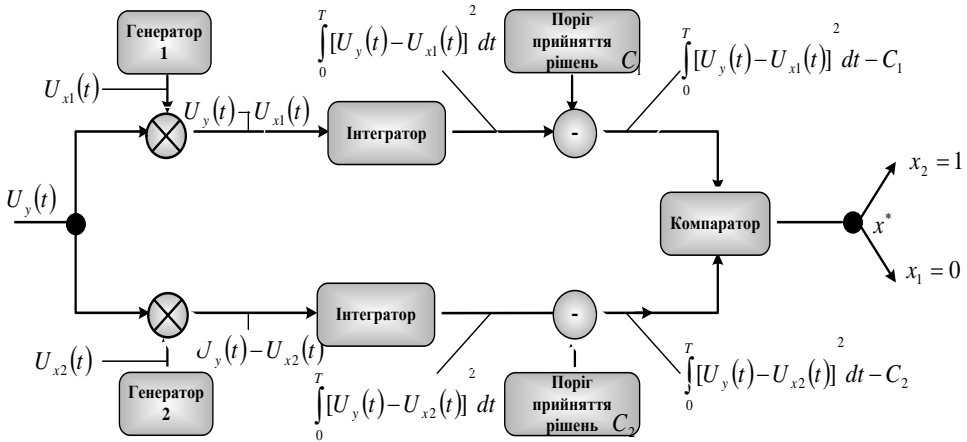


Рис. 9.36. Оптимальний приймач Котельникова з системою прийняття рішень на базі визначених порогів на тлі завад

Розглянемо важливий окремий випадок, коли ймовірності $p(x_1)$ і $p(x_2)$ рівні між собою, і функції $u_{x1}(t)$ і $u_{x2}(t)$ пронормовані таким чином, що їхні енергії на інтервалі однакові. Розкривши в останньому співвідношенні дужки, одержимо з врахуванням того, що $C_1 = C_2$:

$$\int_0^{T_x} u_{x2}^2(t) dt - \int_0^{T_x} u_y(t) u_{x2}(t) dt < \int_0^{T_x} u_{x1}^2(t) dt - \int_0^{T_x} u_y(t) u_{x1}(t) dt.$$

Перші доданки кожної частини нерівності, що є енергіями корисних сигналів, взаємно знищуються. Правило рішення в цьому випадку перетворюється в порівняння вихідних сигналів взаємо-кореляційних фільтрів:

$$\int_0^{T_x} u_y(t) u_{x2}(t) dt < \int_0^{T_x} u_y(t) u_{x1}(t) dt.$$

Відповідна схема зображена на рис. 9.37:

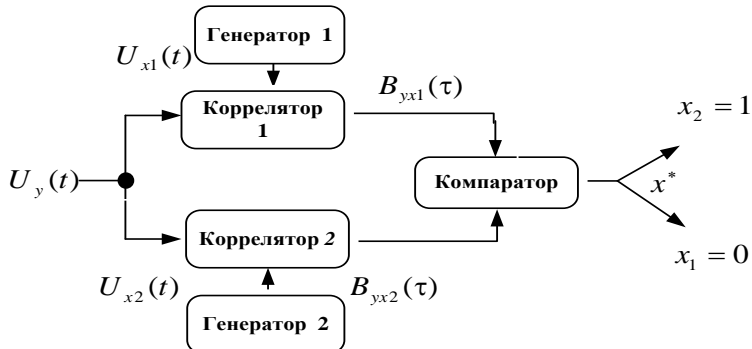


Рис. 9.37. Схема кореляційного приймача з визначеним порогом прийняття рішення на тлі завад

Основні висновки

Будь-яку ІВС можна розділити на джерело, перетворювач, канал передачі, накопичувач і пристрій, що відображає, інформацію, а інформаційні процеси, що відбуваються в цих пристроях, представити в загальному випадку у вигляді процесу передачі інформації по каналу зв'язку.

Якщо стиснення проводиться так, що за стислими даними можна абсолютно точно відновити початкову інформацію, кодування називається неруйнівним. Неруйнівне кодування використовується при передачі (або зберіганні) текстової інформації, числових даних, комп'ютерних файлів тощо, тобто там, де недопустимі навіть щонайменші відмінності початкових і відновлених даних.

Кодування в каналі, або завадостійке кодування, це спосіб обробки переданих даних, що забезпечує зменшення кількості помилок, що виникають в процесі передачі по каналу з завадами.

Узагальнені характеристики інформаційних каналів: час передачі сигналу; ширина його частотного спектру; енергетична характеристика — середня потужність.

Дискретний канал в загальному вигляді представляється сукупністю дискретного модулятора на вході, безперервного каналу і дискретного демодулятора на виході. Моделі дискретних каналів передачі даних поділяються на: двійковий симетричний канал без пам'яті; двійковий несиметричний канал без пам'яті; симетричний канал без пам'яті.

Швидкість передачі інформації сигналами з обмеженою середньою потужністю через канал, у якому діє білий гаусів шум, виявляється максимальною при повній подібності між сигналом і завадою;

Максимальна швидкість передачі інформації буде забезпечена, якщо як фізичний носія інформації застосовувати стаціонарний випадковий процес у вигляді білого гаусового шуму.

Задачею синтезу елементів інформаційних систем є визначення алгоритму функціонування інформаційних систем по заданому критерію якості і інтерпретація цього алгоритму за допомогою технічних засобів.

Задачею аналізу елементів інформаційних систем є розрахунок робочих характеристик та блок/структурних схем інформаційних систем.

Розрізняють наступні основні типи задач статистичного синтезу: виявлення сигналу на фоні завад; розрізнення сигналів на фоні завад; сумісне виявлення (розрізнення) сигналів і оцінка їх параметрів на фоні завад; виділення сигналів на фоні завад.

Теорія потенційної завадостійкості полягає в тому, що для зменшення впливу флуктуаційних завад існує найкращий (ідеальний) приймач, що володіє найбільшою (потенційною) завадостійкістю для даного методу передачі.

Підвищення функціональної надійності або завадостійкості систем передачі інформації досягають збільшенням часової, частотної, енергетичної надлишковості.

Запитання для самоконтролю

1. Які елементи входять до моделі системи передачі інформації?
 2. Що Ви розумієте під поняттям “Канал зв’язку”?
 3. Розкрийте поняття “об’єм сигналу” та “ємність каналу”.
 4. Назвіть основні характеристики дискретних каналів.
 5. Розкрийте відмінність між двійковим симетричним та несиметричним каналами без пам’яті.
 6. Сформулюйте теорему Шеннона для дискретного каналу з завадами та без завад.
 7. Виведіть вираз, який визначає дійсний характер залежності пропускної здатності каналу від ширини смуги пропускання та інших параметрів каналу.
 8. Назвіть статистичні критерії виявлення сигналів на тлі завад і зазначте в яких випадках вони реалізуються.
 9. Визначте структурну схему ідеального приймача Котельникова.
 10. Яким шляхом досягається збільшення завадостійкості систем передачі інформації?
-

Conclusions

Any information system could be divided into source, converter, transmission channel, storage device and device, displaying information, while information processes, occurring in these devices, could be presented in general case in form of information transmission via communication channel.

If compression is performed in the way that compressed data can be exactly restored as initial information, the coding is called nondestructive. Nondestructive coding is used in transmission (or storing) of text information, numerical data, computer files, etc., i.e. in cases where even the smallest differences between primary and recovered data are not allowed.

Coding in channel or interference immune coding is a way of transmitted data processing, capable of reducing the number of errors that occur during transmission

via noisy channel.

Generalized characteristics of information channels: signal transmission time; frequency spectrum width; energy characteristic - average power.

Digital channel is generally represented by set of discrete modulator at the input, continuous channel and discrete demodulator at the output.

Discrete data channels models are subdivided into following types: binary symmetric channel without memory; binary asymmetric channel without memory; symmetric channel without memory.

Information transmission speed by signals with limited average power via channel where white Gaussian noise is functioning is maximal when signal and noise are identical.

Maximum information transmission speed is ensured, if the stationary random process in the form of white Gaussian noise is used as the physical data carrier.

Synthesis task of information systems elements is to define an algorithm of information systems functioning within predetermined quality criteria and to interpret this algorithm using technical means.

The task of analyzing information systems elements is to calculate operational characteristics and structural schemes of information systems.


There are following basic types of statistical synthesis tasks: identifying signal on background noise; signals distinction on background noise; simultaneous signals detection (distinction) and estimation of their parameters on background noise; signals selection on background noise

Theory of potential interference immunity says that to reduce the impact of fluctuation interference there is the best (ideal) receiver, which has the greatest (potential) interference immunity for given transmission method.

Increasing functional reliability, or interference immunity, information transmission systems can be reached by increasing time, frequency and energy redundancy.

Ключові слова

Українською	Англійською
канал зв'язку	communication channel
кодер джерела	coder of the source
декодер джерела	decoder of the source
дискретний канал	discrete channel
безперервний канал	continuous channel
оптимальний приймач	optimum receiver



ЗАХИСТ ВИМІРЮВАЛЬНОЇ ІНФОРМАЦІЇ ВІД НЕСАНКЦІОНОВАНОГО ДОСТУПУ

10

- 10.1. Методи несанкціонованого доступу до ресурсів інформаційних систем
- 10.2. Засоби захисту від несанкціонованого доступу
- 10.3. Моделювання систем і процесів захисту інформації
- 10.4. Протидія мережевому несанкціонованому доступу

10.1. Методи несанкціонованого доступу до ресурсів інформаційних систем

Проблема несанкціонованого доступу (НСД) до ресурсів інформаційних систем загострювалась у міру розвитку інформаційних технологій і тотального використання комп'ютерних мереж у всіх сферах діяльності суспільства.

Вирішення задач розроблення та вибору відповідних ефективних методів і засобів захисту інформації значною мірою залежить від низки чинників, пов'язаних із самим процесом несанкціонованого доступу. З урахуванням того, що спектр несанкціонованих дій достатньо різномірний, то основою класифікації можуть бути базові ознаки (рис. 10.1).

Базові ознаки несанкціонованого доступу				
1	2	3	4	5
Автоматизація	Взаємодія з політикою безпеки	Дистанційність	Зовнішні прояви	Ініціалізаційна умова
мануальний; автоматизований; автоматичний (вірусний)	постполітизаційний; деполітизаційний	локальний; віддалений	пасивний; активний	умовний; безумовний
6	7	8	9	
Інструментальний засіб	Наявність зворотного зв'язку	Порушення характеристик безпеки	Природа взаємодії	
прграмний; апаратний; нетиповий	зі зворотнім зв'язком; без зворотного зв'язку	К-дієвий (порушення конфіденційності ресурсів); Ц-дієвий (порушення цілісності ресурсів); Д-дієвий (порушення доступності ресурсів)	фізичний; логічний	
10	11	12	13	14
Реляційні ознаки	Специфіка реалізації	Спрямованість результату	Ступінь складності	Семирівнева еталонна модель
монономонний; поліномонний; монополічний; поліполічний	фрагментований; без замовчувань; прихований; пігібекінговий; маскарадний; непрямий; соціотехнічний; криптоаналітичний; неспецифічний	розширюючий; викривляючий; розповсюджуючий; розкрадаючий; перевантажуючий; інформаційний; утримуючий; знищуючий	простий; складний; системний	ISO-00 ISO-01 ... ISO-7E ISO-7F

Рис. 10.1. Класифікація несанкціонованого доступу за базовими ознаками

Мануальний НСД (підглядання, збирання сміття, вилучення інформаційних носіїв, підміна положень вмикачів режимів тощо) реалізується за прямою участю людини без використання будь-яких спеціальних засобів.

Автоматизований НСД здійснюється з постійною участю оператора з використанням широкого спектру програмних і апаратних засобів та пов'язаний з суперзапінгом, снупінгом, сніфінгом, підключенням додаткових терміналів, використанням мережевих аналізаторів тощо. Приклад автоматизованого НСД зображений на рис. 10.2, неавторизована сторона реалізовує несанкціонований доступ, за допомогою апаратних та програмних засобів (наприклад, шляхом використання утилітів віддаленого керування), на робочу станцію при прямому підключенні до мережі.

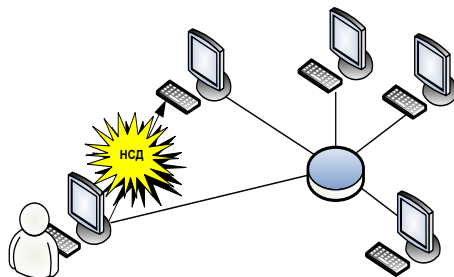


Рис. 10.2. Приклад автоматизованого НСД

Автоматичний НСД реалізується без участі людини, як правило з використанням спеціалізованих програмних засобів, функціонування яких базується на вірусних технологіях. Прикладом автоматичного НСД може бути, зараження комп'ютера вірусом під час підключення до глобальної мережі Internet, без участі людини (рис. 10.3).

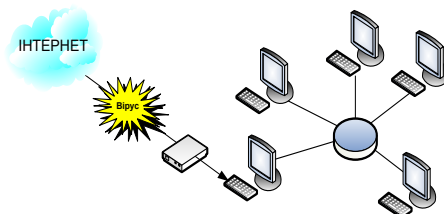


Рис. 10.3. Приклад автоматичного НСД

Постполітизаційний НСД базується на використанні недоліків у вже реалізованій політиці безпеки. Такими недоліками можуть бути невірно побудовані правила розмежування доступу, використання програмних і апаратних засобів з недостатнім рівнем захищеності, прорахунки при блокуванні каналів витоку інформації з обмеженим доступом тощо. Наприклад, якщо користувачу та адміністратору мережі надані однакові права доступу до сервера, то неавторизована сторона, що одержала права доступу користувача 1, зможе реалізувати НСД до інформаційних ресурсів на рівні прав адміністратора (рис. 10.4,а).

Деполітизаційний НСД пов'язаний з помилками і недбалістю, які мають місце при реалізації заходів із забезпеченням вже існуючої політики безпеки. Це в першу чергу, пов'язане з людським чинником і залежить від достатньої адміністративної підтримки, коректності виконання функцій захисту, своєчасного реагування на нештатні ситуації тощо. Наприклад, якщо Користувач залишає своє робоче

Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу

місце, не дотримуючись політики безпеки (не закінчує сеанс роботи чи не блокує комп'ютер), то неавторизована сторона за час його відсутності має доступ до інформаційних ресурсів через не коректність виконання користувачем функцій захисту (рис. 10.4,б).

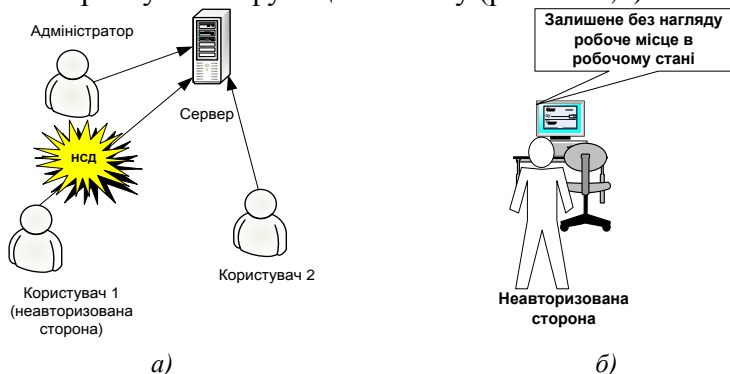


Рис. 10.4. Приклад постполітизаційного і деполітизаційного НСД

Якщо НСД на ресурс здійснюється в локалізованій області його розташування (локальна обчислювальна мережа, робоча станція, принтер, носій інформації, ОС, додатки тощо), то він називається *локальним*, а у протилежному випадку – *віддаленим*. Наприклад, локальний НСД може бути реалізований всередині сегменту (фізичного об'єднання станцій за допомогою комутаційних пристроїв не вище каналного рівня) при цьому джерело НСД і ресурс, що піддався несанкціонованій дії, будуть знаходитись в межах одного сегмента. Приклади реалізацій локального та віддаленого НСД відповідно зображені на рис. 10.5,а і 10.5,б. Локальний НСД реалізується неавторизованою стороною в одному сегменті мережі, а відділений несанкціонований доступ відбувається через мережу Інтернет.

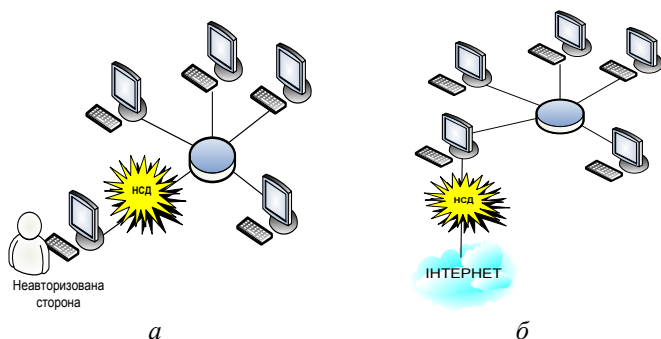
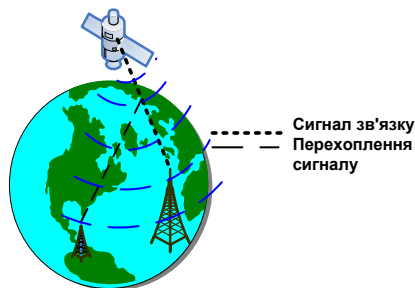


Рис. 10.5. Реалізація локального та віддаленого НСД

Міжсегментний НСД може бути прикладом віддаленого НСД, коли джерело і ресурс знаходяться в різних сегментах комп'ютерної мережі, наприклад, кампусної. В першому випадку ініціатором НСД зазвичай буває легальний користувач, який, наприклад, із файл-сервера через робочу станцію здійснює копіювання конфіденційних даних на позаштатний носій інформації.

В результаті своєї дії (наприклад, візуального перегляду даних з терміналу) *пасивний НСД* не здійснює безпосередній вплив на ресурси і може не порушити їх характеристики безпеки, наприклад, при перехопленні зашифрованих даних.

В результаті *активного НСД* на ресурси здійснюється безпосередній вплив (зміна конфігурації, порушення функціональності і т. ін.) і порушуються їх характеристики безпеки. Практично всі типи віддаленого НСД є активними, особливістю якого у порівнянні з пасивним є принципова можливість його виявлення, тому, що в результаті безпосереднього впливу здійснюються певні зміни. На відміну від активного, пасивний НСД не залишає слідів втручання.



Наприклад, перехоплення супутникового сигналу (рис. 10.6) без його модифікації відноситься до пасивного НСД.

Перехоплення повідомлення неавторизованою стороною з подальшою його модифікацією (рис. 9.7) відноситься до активного НСД.

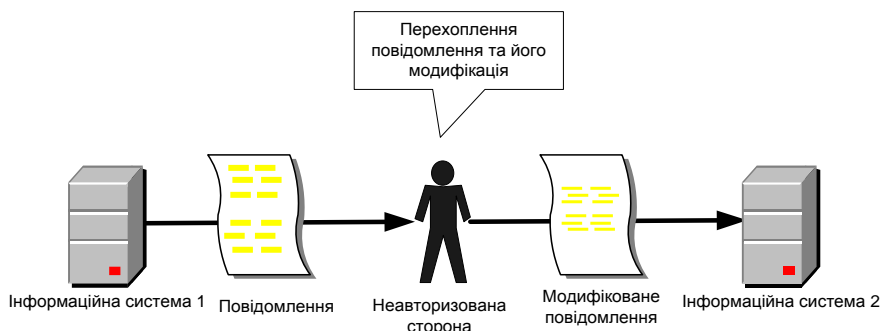


Рис. 10.7. Реалізація активного НСД

Умовний НСД ініціалізується у випадку виникнення певної події (механізм логічної бомби) і в свою чергу можуть поділятися на

пасивний і активний. **Прикладом** ініціалізації пасивного умовного НСД може бути передача від потенційної цілі запиту певного типу, який і буде умовою початку атаки. У цьому випадку, наприклад, такою умовою можуть бути DNS- і ARP-запити в стеку протоколів TCP/IP.

Активний умовний НСД здійснює постійний моніторинг стану окремих ресурсів і при його певній зміні формується сигнал ініціалізації. Прикладом такого стану може бути подія, пов'язана з перериванням сеансу роботи користувача з сервером без стандартної команди, наприклад, LOGOFF. Момент ініціалізації безумовного НСД не супроводжується певною зміною стану ресурсів і визначається джерелом атаки.

Програмний НСД ґрунтується на спеціальних мікро- або макрокодованих засобах (наприклад, суперзапінгових утилітах, внутрішніх командах, сценаріях автоматизації тощо), які функціонують в межах інформаційних систем для реалізації своїх функцій.

Наприклад, віддалений комп'ютер (див. рис. 10.8) відправляє повідомлення на робочу станцію, до якої несанкціоновано підключається неавторизована сторона за допомогою портативного комп'ютера і реалізує зараження цієї станції вірусом, який модифікує це повідомлення.

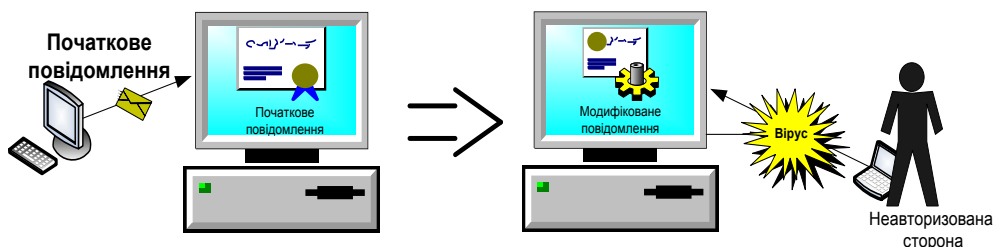


Рис. 10.8. Реалізація програмного НСД

Апаратний НСД засновується на різноманітних механічних, електричних, електромеханічних, електронних, електронно-механічних та інших пристроях, які використовуються автономно чи в поєднанні з іншою апаратурою для виконання відповідних функцій.

Нетиповий НСД реалізується на підґрунті таких засобів, які не відносяться до апаратних або програмних, наприклад, вибухівка, радіоактивні матеріали, кислоти, луги, комахи, гризуни тощо.

В процесі реалізації *НСД зі зворотнім зв'язком* здійснюється отримання нею від ресурсу, що піддався несанкціонованим діям, відповіді на свої дії, які необхідні, наприклад, для подальшого продовження зазначеного процесу на більш ефективному рівні, який досягається завдяки аналізу реакцій об'єкта НСД на певні зміни.

Прикладом НСД з зворотнім зв'язком є сканування портів спеціальним програмним забезпеченням. Під час сканування на порти сервера відправляється пакет синхронізації SYN. Якщо на це повідомлення (рис. 10.9) приходить відповідь (зворотній зв'язок) у вигляді пакету SYN/ACK, то це означає, що сканований порт знаходиться у стані очікування і можна виконувати наступну дію.

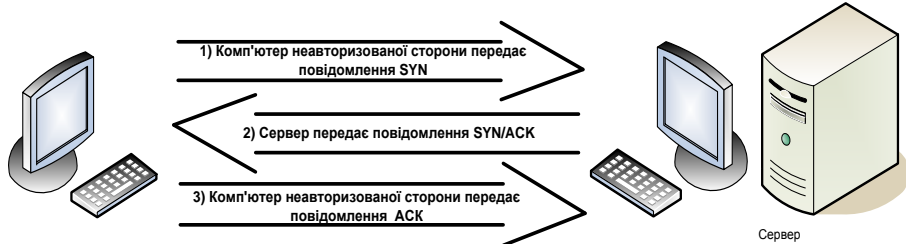


Рис. 10.9. Реалізація НСД зі зворотнім зв'язком

НСД без зворотного зв'язку реалізує свої дії незалежно від реакції ресурсу, що піддався несанкціонованим діям. Прикладом таких несанкціонованих дій є відмова в обслуговуванні.

Наприклад, блокування маршрутизатора (рис. 10.10), що приводить до відмови в обслуговуванні.

При реалізації НСД здійснюється порушення основних характеристик безпеки ресурсів інформаційних систем:

конфіденційності (характеристика безпеки ресурсів, що відображає їх властивість невиявленості й доступності без відповідних повноважень. Фактично, ресурси не можуть бути доступні або розкриті неавторизованій стороні, тобто для неї їх нібито немає. В свою чергу, авторська сторона (наприклад, обслуговуючий персонал, користувачі, програми і т. ін.), якій надано відповідні повноваження, має повний доступ до ресурсів);

цілісності (характеристика безпеки ресурсів, що відображає їх властивість протистояти несанкціонованій зміні. Наприклад, користувач, що накопичує інформацію, має право очікувати, що вміст його файлів залишиться незмінним, незважаючи на цілеспрямовані впливи, відмови програмних або апаратних засобів. За цією характеристикою ресурси не зазнають будь-якої зміни з боку неавторизованої сторони);

доступності (характеристика безпеки ресурсів, що відображає їх властивість, яка полягає в можливості їх використання у заданий момент часу відповідно до пред'явлених повноважень. Фактично

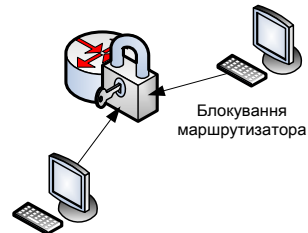


Рис. 10.10. Реалізація НСД без зворотного зв'язку

Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу

авторська сторона, за потреби, відразу, в будь-який момент часу отримує необмежений доступ до необхідного ресурсу).

У цьому контексті за типом порушення зазначених характеристик НСД буває:

К-дієвий (порушення конфіденційності ресурсів);

Ц-дієвий (порушення цілісності ресурсів);

Д-дієвий (порушення доступності ресурсів).

Якщо в процесі НСД порушуються різні характеристики безпеки, то результуючий тип буде комбінований на базі основних, наприклад, КЦД-дієвий – НСД, що порушує конфіденційність, цілісність і доступність ресурсів.

За природою взаємодії з ресурсами інформаційної системи НСД поділяється на *фізичний* та *логічний*. Для першого характерна фізична форма взаємодії, яка проявляється у вигляді різного роду прямих блокувань, пошкоджень, проникнень, крадіжок тощо, наприклад, розмикання електричних з'єднань, пошкодження носіїв інформації, розукомплектування, подолання фізичного кордону захисту, підслуховування, перехоплення побічних електромагнітних випромінювань і наведень тощо.

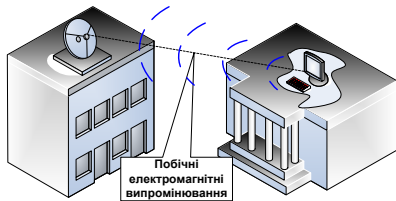


Рис. 10.11. Реалізація фізичного НСД

Для другого не властива пряма фізична взаємодія з ресурсами і вона в основному пов'язана з логікою подій, наприклад, аналізом протоколів, перевантаженням, визначення паролів, захопленням сеансів

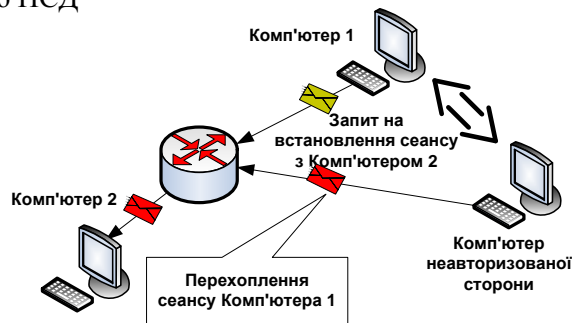


Рис. 10.12. Приклад логічного НСД

тощо.

Прикладом логічної форми НСД, може бути перехоплення сеансу зв'язку (рис. 10.12).



Рис. 10.13. Приклад монономного НСД

Монономний НСД реалізується із одного джерела на один конкретний ресурс.

Часто такий НСД називають нерозподіленим, **наприклад**, монономний НСД можна реалізувати за допомогою сканування портів комп'ютера з визначеною IP-адресою (рис. 10.13).

Поліномний НСД здійснюється одночасно з декількох джерел (два і більше) на один ресурс і націлений на досягнення однієї конкретної мети. Такий НСД часто називають *розподіленим* (рис. 10.14).

Монополічний НСД реалізується із одного джерела одночасно на множину ресурсів (два і більше) і направлений на досягнення конкретної мети (рис. 10.15).

Приклад такого типу НСД може базуватися на ширококомовному передаванні повідомлення від джерела на всі комп'ютери сегменту, адреси яких знаходяться під одною маскою підмережі.

Поліполічний НСД поєднує у собі *поліномну* і *монополічну* технології за якими множина джерел здійснює НСД на множину ресурсів з метою досягнення однієї конкретної мети (рис. 10.16).

Фрагментований НСД базується на принципі декомпозиції і поетапної реалізації, наприклад, на основі використання механізму розбиття IP-пакетів (на множину більш мілких) і подальшого їх передавання (рис. 10.17). Такий підхід дозволяє обходити системи виявлення атак, які не розраховані

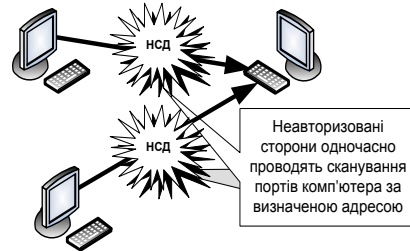


Рис. 9.14. Приклад поліномного НСД

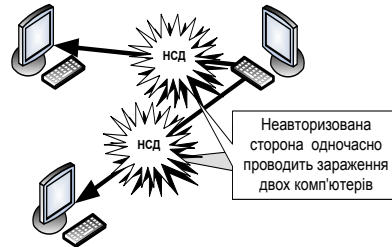


Рис. 10.15. Приклад монополічного НСД

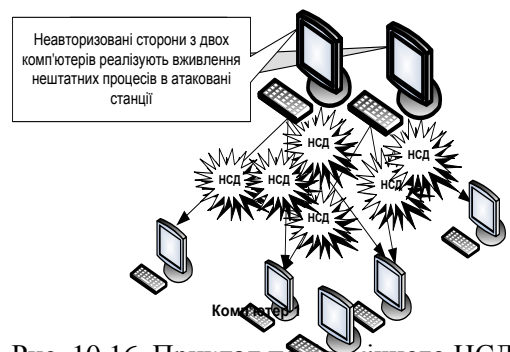


Рис. 10.16. Приклад поліполічного НСД

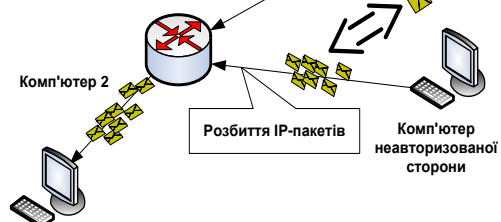


Рис. 10.17. Приклад фрагментованого НСД

на протидію декомпозиційним технологіям.

НСД, який реалізується без використання значень за замовчуванням (наприклад, програмна закладка BackOffice за замовчуванням використовує порт 31337, але це значення може бути змінено наприклад на 31336), орієнтований на подолання систем виявлення атак, які засновуються на сигнатурних (шаблонних) технологіях за аналогією з антивірусними програмами, призначеними для захисту від сигнатурних вірусів.

Приховані атаки використовують різноманітні заходи (підміна контрольних сум, перехоплення різноманітних даних, модифікація ядра операційної системи, використання стандартних або схожих до стандартних імен тощо), які дозволяють залишатися невиявленими в локалізованій області атакованого ресурсу. Технологія *прихованого НСД* за своєю ідеологією подібна до технології *стелс-вірусів*.

Пігібекінговий НСД заснований на несанкціонованому одержанні доступу до тимчасово неконтрольованого ресурсу, наприклад, шляхом проникнення до інформаційної системи в результаті тимчасової відсутності або після некоректного завершення сеансу роботи легального користувача.

Маскарадний НСД заснований на формуванні такої поведінки порушника, яка дозволяє йому видати себе за легальне джерело, наприклад, шляхом обману (spoofing) атакувати обчислювальну мережу (з протоколом TCP/IP) привласнюючи IP-адресу за допомогою якої обходиться система захисту.

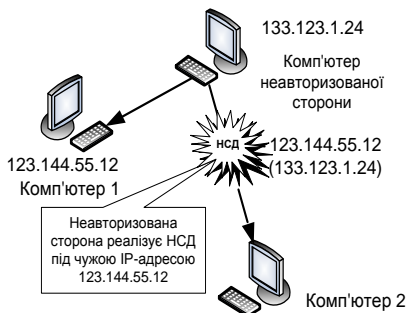


Рис. 10.18. Приклад непрямого НСД



Рис. 10.19. Приклад соціотехнічного НСД

Непрямий НСД заснований на тому, що напад здійснюється через третю особу (посередника), а істинне джерело нападу залишається невідомим, при цьому часто використовуються маскарадні технології (рис. 10.18). Наприклад, скориставшись варіантом нападу з перенаправленням несанкціонованих дій, для виключення можливості виявлення реального джерела, порушник проводить або перенаправляє свій трафік через чужий комп'ютер, який для ресурсу (що піддався несанкціонованим діям) і буде вихідним джерелом.

Соціотехнічний (соціоінжиніринговий) НСД пов'язаний з

отриманням даних (наприклад, імен користувачів, паролів, телефонних номерів віддаленого доступу тощо) від атакованих людей у процесі інформаційного обміну (рис. 10.19).

Криптоаналітичний НСД заснований на використанні широкого спектру криптоаналітичних методів та засобів для зламування ресурсів, захищених різними криптографічними засобами.

Розширюючий НСД (рис. 10.20) орієнтований на одержання більших повноважень на права доступу до ресурсу, наприклад, на вхід в локальну обчислювальну мережу з правами адміністратора,

отримання доступу на запис до полів баз даних, зміну атрибутів файлів тощо.



Рис. 10.20. Приклад розширюючого НСД

Викривлюючий НСД пов'язаний зі здійсненням будь-яких прямих змін в цільовому ресурсі, наприклад, підробка полів баз даних, підміна інформаційних носіїв, зміна часу і дати тощо.

Розповсюджуючий НСД спрямований на отримання доступу до ресурсу та його розкриття без відповідних повноважень, наприклад, несанкціоноване одержання файла даних з пароллями і їх публікація на хакерських сайтах або розсилання серед абонентів мережі.

Розкрадаючий НСД направлений на використання ресурсу без нанесення прямого збитку, наприклад, без зниження якості обслуговування користувачів здійснити тимчасове вилучення частини пам'яті (для розширення можливостей іншої системи), завантаження телекомунікаційних каналів, використання робочої станції або мережевого сервісу тощо.

Перевантажуючий НСД спрямований на завантаження ресурсу до такого рівня, що він втрачає властивість його використання. Результатом таких несанкціонованих дій можуть бути неможливість



Рис. 10.21. Приклад перевантажуючого НСД

Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу

використання, перевантаження, перешкоджання використанню (відмова в обслуговуванні) тощо.

Прикладом такого НСД (рис. 10.21) може бути перевантаження маршрутизатора.

Інформаційний НСД пов'язаний зі збиранням необхідних даних (як правило для реалізації подальших дій) і не передбачає здійснення прямого НСД до ресурсу.

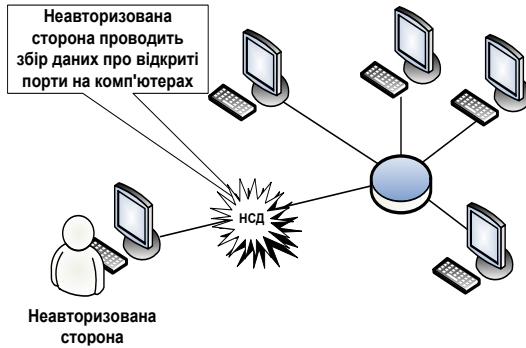


Рис. 10.22. Приклад інформаційного НСД

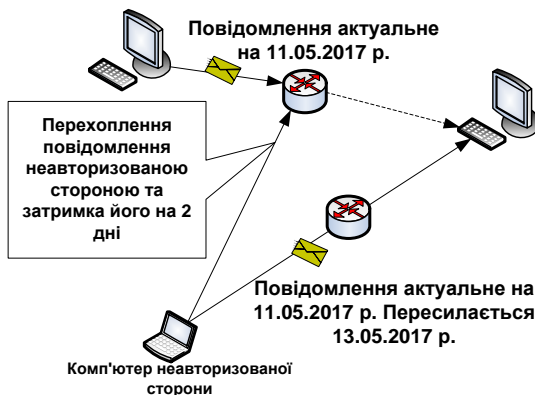


Рис. 10.23. Приклад утримуючого НСД



Рис. 10.24. Приклад знищуючого НСД

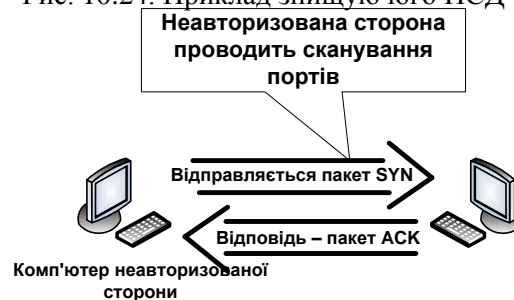


Рис. 10.25. Приклад простого НСД

Наприклад, одержання інформації в результаті аналізу публікацій, використання системних утиліт для виявлення активних робочих станцій, сервісів тощо (рис. 10.22).

Утримуючий НСД призначений для тимчасової затримки ресурсу з метою зниження його актуальності.

Наприклад, притримування шифрограми на проміжному вузлі при її передачі телекомунікаційними каналами загального користування (рис. 10.23).

Знищуючий НСД орієнтований на безповоротну ліквідацію ресурсу, наприклад, вилучення файлу, дроблення інформаційного носія, низькорівневе форматування жорсткого диску з метою знищення даних тощо (рис. 10.24).

Простий НСД (рис. 10.25) є нескладними в реалізації діями, направленними на виконання окремих процедур, наприклад, сканування портів, аналіз трафіку, пошук

активних робочих станцій, віддалене управління тощо.

Складний НСД є комбінацією простого, призначеного для реалізації низки необхідних функцій, наприклад, виявлення активної робочої станції і здійснення віддаленого управління нею.

Системний НСД будується на основі сформованого системного підходу з багатокроковою комбінацією дій та використанням простого НСД для ефективної реалізації спеціально направленою комплексу функцій, наприклад, пошук активних робочих станцій, моніторинг трафіку, сканування, зламування ОС і робочих додатків, несанкціоноване копіювання даних та замітання слідів.

Міжнародна організація стандартизації (ISO) запропонувала *семирівневу еталонну модель* з метою розмежування функцій різних протоколів у процесі передачі інформації від одного абонента іншому. Таких класів функцій виділено сім. Вони отримали назву рівнів, кожний з яких виконує певні задачі в процесі передачі блоку інформації, причому відповідний рівень зі сторони приймача виконує перетворення, зворотні тим, що зроблені на тому самому рівні на передавальній стороні (джерелі). У цьому зв'язку *НСД за семирівневою еталонною моделлю* можна визначити на наступних рівнях:

- фізичний, що кодується параметром $K0(2) = 20$;
- канальний, що кодується параметром $K1(2) = 21$;
- мережевий, що кодується параметром $K2(2) = 2^2$;
- транспортний, що кодується параметром $K3(2) = 23$;
- сеансовий, що кодується параметром $K4(2) = 24$;
- представлення даних, що кодується параметром $K5(2) = 25$;
- прикладний, що кодується параметром $K6(2) = 2^6$.

На *фізичному рівні* забезпечуються необхідні механічні, електричні, функціональні і процедурні характеристики для встановлення, підтримки і розмикання фізичного з'єднання. На *канальному рівні* забезпечуються функціональні і процедурні засоби для встановлення, підтримки і вивільнення ліній передавання даних між абонентами мережі (наприклад, терміналами і вузлами мережі). На *мережевому рівні* забезпечуються функціональні і процедурні засоби для обміну службовою інформацією між двома об'єктами транспортного рівня мережі (тобто пристроями, що підтримують протоколи на транспортному рівні за допомогою мережевого з'єднання). Гарантує незалежність поведінки об'єктів транспортного рівня від схеми маршрутизації і комутації. На *транспортному рівні* забезпечується оптимізація комутаційного обслуговування (підтримуваного реалізацією більш низьких рівнів зв'язку) шляхом забезпечення прозорих передач даних між абонентами в рамках сеансу.

На *сеансовому рівні* здійснюється забезпечення на логічному рівні обслуговування двох «зв'язаних» на рівні представлення даних об'єктів мережі і керування введенням діалогу між ними шляхом синхронізації повідомлень.

На *рівні представлення даних* забезпечується сукупність службових операцій, які можна вибрати на прикладному рівні для інтерпретації переданих і одержуваних даних. Ці службові операції містять вміщують керування інформаційним обміном, відображення даних і керування структурованими даними. Службові операції цього рівня являють собою основу всієї семирівневої моделі і дозволяють зв'язувати в одне ціле термінали і засоби обчислювальної техніки будь-яких типів.

На *прикладному рівні* забезпечується безпосередня підтримка прикладних процесів і програм кінцевого користувача і керування взаємодії цих програм з різними об'єктами мережі передачі даних. Визначення класу несанкціонованої дії за даною ознакою наступне – ISO-K, де $K = K_{(2)0} + K_{(2)1} + K_{(2)2} + K_{(2)3} + K_{(2)4} + K_{(2)5} + K_{(2)6}$ – двійковий код. Для зручності K будемо представляти у шістнадцятковому коді ($K_{(16)}$).

Наприклад, якщо за цією класифікацією несанкціонована дія записана як ISO-3A (де $3A_{(16)} = 0111010 = 2^5 + 2^4 + 2^3 + 2^1 = K5 + K4 + K3 + K1$), то вона інтерпретує реалізацію на рівні каналному, транспортному, сеансовому і представлення даних, а, наприклад, запис ISO-00 означає, що клас несанкціонованої дії не підтримується семирівневою еталонною моделлю.

Не важко помітити, що несанкціоновані дії, які класифікуються за ознаковим принципом, можуть у кожному конкретному випадку при визначенні загального класу містити не тільки одну, але і більше компонент кожної з ознак. З появою нових методів і засобів реалізації несанкціонованих дій, ознаки запропонованої класифікації можуть бути розширені.

При практичному використанні класифікації, наприклад, такий НСД, як сканування портів, можна визначити як: автоматизований, постполітизаційний, віддалений, пасивний, безумовний, програмний, зі зворотним зв'язком, K-дієвий, логічний, мономоніторний, неспецифічний, інформаційний, простий, ISO-36.

За допомогою даної класифікації можна здійснювати (відповідну їй) формалізацію можливостей систем протидії для підвищення ефективності їхнього вибору і формуванні вимог при їхній розробці.

З НСД пов'язані такі поняття, як *доступ* та *перехоплення*. Під *доступом* (*access*) розуміють взаємодію між ресурсами інформаційних систем, що забезпечує передавання інформації між такими ресурсами,

а в процесі доступу до інформації (access to information) зокрема реалізується її копіювання, модифікація, знищення, ініціалізація тощо. Розрізняють *несанкціонований* та *санкціонований доступ*. Якщо доступ до ресурсів системи здійснюється, наприклад, з порушенням або поза правилами, розмежування доступу, то такий доступ – несанкціонований. Одною з базових дій, що породжує НСД, є перехоплення (intercept), під яким розуміють несанкціоноване одержання інформації незаконним підключенням до каналів зв'язку (наприклад, пряме перехоплення) чи візуально (наприклад, підглядання), або за допомогою радіотехнічних засобів (наприклад, непряме перехоплення).

Перехоплення за дією на інформацію можна розділити на *активне* і *пасивне*, а за типом підключення – на *пряме* і *непряме*.

Активне перехоплення (active eavesdropping). Перехоплення, під час якого у супротивника є можливість не тільки перехоплювати повідомлення, але й впливати на нього, наприклад, затримувати або вилучати сигнали, що передаються каналами зв'язку.

Пасивне перехоплення (passive tapping). Отримання інформації з можливістю тільки спостерігати за обміном повідомленнями (наприклад, з метою виявлення різної системної інформації в обчислювальній мережі (ОМ), не справляючи на нього ніякого впливу.

Пряме перехоплення (direct eavesdropping). Перехоплення інформації безпосереднім підключенням (наприклад, додаткового терміналу) до лінії зв'язку. Пряме перехоплення можна виявляти перевіркою лінії зв'язку.

Непряме перехоплення (indirect eavesdropping). Перехоплення інформації (наприклад, індуктивних хвиль) без використання безпосереднього підключення до лінії зв'язку (threat). Пасивне перехоплення важко виявити, оскільки немає безпосереднього приєднання термінального обладнання до лінії зв'язку.

Руйнівні програмні впливи. *Руйнівний програмний вплив* – це програмний код або його частини, за допомогою яких здійснюється загроза хоча б одній характеристиці безпеки певних ресурсів інформаційних систем. Руйнівні впливи можна поділити на такі групи: комп'ютерні віруси (віруси), логічні бомби, потайні ходи і лазівки; програми розкриття паролів, реплікатори, мережеві програмні аналізатори, суперзапінгові утиліти, троянські коні.

Вірус. Програма, яка здатна до багаторазового самовільного створення свого тіла і зазвичай модифікує (заражає) інші програми, записані у файлах чи системних областях, для подальшого відтворення нового тіла та одержання керування з метою модифікації записів,

знищення файлів, завантаження ресурсів і виконання інших руйнівних впливів у інформаційній системі.

Логічні бомби. Програма, що ініціюється з виникненням різних подій, наприклад, відкриття певного файлу, обробка заданих записів та інші дії з метою порушення характеристик безпеки ресурсів інформаційних систем. Використовуються, наприклад, для розкрадання за допомогою змінювання певним чином (на свою користь) коду програми, що реалізує фінансові операції.

Потайний хід. Уразливість в системі, що навмисно створює її розробник або яка виникла випадково і фактично є додатковим способом проникнення в систему.

Програми розкриття паролів. Програми зазвичай призначені для вгадування паролів (наприклад, архівованих файлів) через перебирання варіантів, можливих для використання символів або проникнення в систему за допомогою словників. Програми, що засновані на останньому методі, здійснюють зламування системи парольного захисту через перебирання елементів одного або декількох словникових файлів, складеного спеціально чи взятого із серверів або жорстких дисків локальних станцій.

Реплікатори. Програми, що при виконанні створюють декілька своїх копій в інформаційній системі. Наприклад, коли реплікатор створює тільки одну і після цього виконує її, то пам'ять системи швидко переповнюється, що обмежує доступ до певних компонентів системи.

Мережеві аналізатори. Програмно-апаратні засоби (в окремих випадках програми, що запускається з робочої станції, підключеної до мережі), призначені для зчитування будь-яких параметрів потоку даних в інформаційній системі.

Суперзапінг. Руйнівний вплив, що пов'язаний з несанкціонованим використанням утиліт для модифікації, знищення, копіювання, розкриття, вставлення, застосування або заборони застосування даних інформаційної системи.

Троянські коні. Спеціалізована програма, яка, як правило, виступає від імені інших програм і дозволяє дії, відмінні від визначених в специфікації, що використовуються програмним забезпеченням.

Із часом цей перелік може доповнюватися новими складовими, оскільки рівень зростання програмного і апаратного забезпечення настільки інтенсивний, що його навіть важко прогнозувати.

Деякі з руйнівних програмних впливів, наприклад логічні бомби або троянські коні, можуть бути реалізовані у вигляді програмних закладок (жучків), які навмисно впроваджуються в тіло певних програм з метою реалізації руйнуючих дій на характеристики безпеки ресурсів інформаційних систем. Наприклад, в командний процесор операційної системи можна встановити програмну закладку, яка є резидентною, виконує функції логічної бомби (активізується під час запуску DISKREET.EXE з нортоновських утиліт) і призначається для запису в певне місце на диску (зазвичай сектор, помічений як збійний) паролів, що вводяться з клавіатури. У цьому разі, така закладка порушує не тільки цілісність командного процесора, але і конфіденційність даних, що передаються до потрібної утиліти з клавіатури.

Класифікація комп'ютерних вірусів. Одним з найбільш поширених руйнівних програмних впливів є комп'ютерні віруси. У світі створені тисячі вірусів, а їх різновиди (особливо чисельність) постійно збільшуються. З урахуванням аналізу зазначених руйнівних програмних впливів їх найзручніше класифікувати за такими ознаками (рис. 10.26): *середовищем поширення; способом маскування в середовищі; інфікованими об'єктами; способом інфікування об'єкта; способом розміщення в інфікованому об'єкті; руйнівним впливом; здатністю до зміни; стилем написання; типом коду.*

За *середовищем розповсюдження* віруси можна поділити на такі: DOSівські; WINDOWSівські; UNIXівські; NETWAREвські; INTERNETівські та ін. Ці віруси функціонують відповідно в середовищі DOS, WINDOWS, UNIX, NETWARE, INTERNET та інших системах, функції і програмні додатки яких використовують віруси. Назву вірусам дають за іменем середовища, в якому вони розповсюджуються.

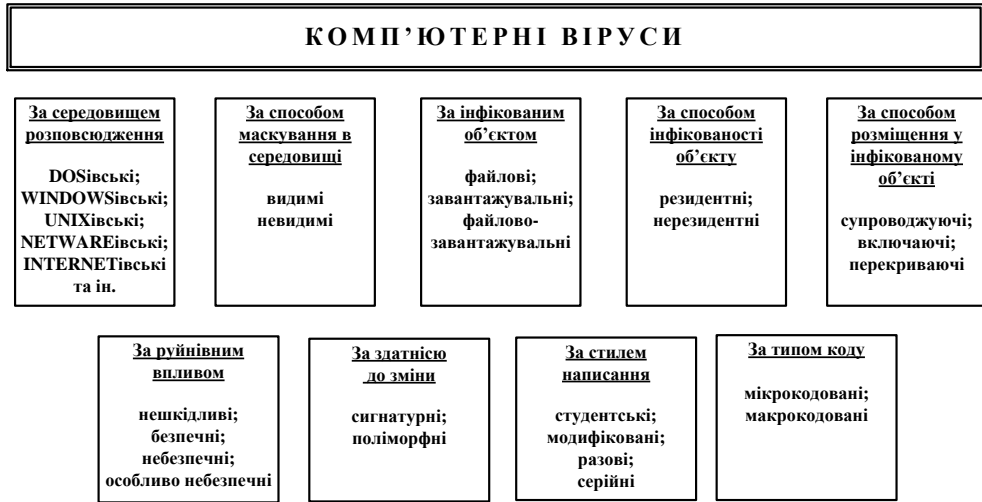


Рис. 10.26. Класифікація комп'ютерних вірусів

За *способом маскування* в середовищі віруси розділяють на видимі і невидимі. *Видимі віруси* досить легко виявити, використовуючи найпростіші засоби. Наприклад, за допомогою команди DOS dir можна побачити зміну розміру інфікованого файлу або, запустивши простий шістнадцятковий редактор, знайти, наприклад, сигнатуру вірусу і т. ін. *Невидимі віруси* (stealth; стелс-віруси; віруси-невидимки) містять майстерні алгоритми, що дозволяють їм маскуватися в середовищі розповсюдження. Такі віруси, знаходячись в системі, перехоплюють її звернення до інфікованих об'єктів і замінюють їх інфіковані ділянки на оригінальні. Після таких дій, наприклад за командою dir, не можна ідентифікувати зміну розміру файлу, а шістнадцятковий редактор, що використовує в своїй роботі функції ОС, не надасть можливості знайти, наприклад, сигнатуру вірусу і т. ін.

За *типом інфікованих об'єктів* серед вірусів розрізняють файлові, завантажувальні, файлово-завантажувальні. *Файлові віруси* розміщуються у файлах різних форматів (COM, EXE, SYS, DOC і т. ін.) і, як правило, ініціалізуються першими під час їх обробки. *Завантажувальні віруси* розміщуються в Boot-секторах дисків або в секторі вінчестера із системним завантажувачем і активізуються під час початкового завантаження ОС. *Файлово-завантажувальні віруси* (поєднання файлових і завантажувальних) інфікують як файли, так і зазначені сектори, а алгоритм їх роботи значно ускладнюється з урахуванням бінарної дії.

За *способом інфікування* об'єкта серед вірусів виділяють резидентні, нерезидентні. *Резидентні віруси* після запуску залишають в оперативній пам'яті комп'ютера свою резидентну частину, яка перехоплює звернення системи до підлягаючих інфікуванню об'єктів і заражає їх. Такий вірус залишається в пам'яті аж до вимкнення комп'ютера або його перезавантаження. Іноді для того, щоб інфікувати всі файли, наприклад, всі виконувані EXE-файли поточного каталога диска А:, досить вставити незахищену дискету з файлами в дисковод А і виконати команду `dir A:`. *Нерезидентні віруси* ініціалізуються в період обробки системою інфікованого об'єкта і не є активними (тобто не заражають пам'ять) після цього періоду. Тому нерезидентний вірус не може інфікувати інші файли, якщо не буде ініційований в момент інфікування файл-носії.

За *способом розміщення в інфікованому об'єкті* віруси поділяють на супровідні, включальні, перекривальні. *Супровідні віруси* трапляються дуже рідко; вони інфікують EXE-файли непрямо, тобто для зазначеного файлу створюється новий з таким самим іменем, але з СОМ-розширенням, куди і вміщується тіло вірусу. У момент запуску файлу з іменем першим активізується СОМ-файл (тобто вірус), який виконує свої функції і далі запускає однойменний EXE-файл. *Включальні віруси* вмонтовуються в початок, кінець або у внутрішню частину файлу, при цьому межі останнього розширюються і розміри файлів відповідно збільшуються на розмір впровадженого тіла вірусу, а сам інфікований об'єкт залишається неушкодженим і зберігає свою працездатність. *Перекривальні віруси* безповоротно пошкоджують інфікований об'єкт, оскільки у разі його зараження тіло вірусу накладається на код об'єкта. Файли, уражені перекривальним вірусом, не виліковуються і тому зазвичай вилучаються.

За *руйнівним впливом* серед вірусів розрізняють нешкідливі, безпечні, небезпечні, особливо небезпечні. *Нешкідливі віруси* відповідно до їх алгоритму не завдають прямої шкоди інформаційній системі, за винятком зайнятого дискового простору і частини оперативної пам'яті (особливо, якщо вірус резидентний). Але будь-який вірус, навіть нешкідливий, може виявити себе інакше в нових умовах (наприклад, у разі зміни версії операційної системи або середовища, характерного для розповсюдження вірусів, форматів дисків і т. ін.) і призвести до непередбачуваних наслідків. *Безпечні віруси* характеризуються проявами різних звукових та відео ефектів і, аналогічно нешкідливим вірусам, зменшують розмір вільної пам'яті. *Небезпечні віруси*, крім того, що зменшують розмір вільного дискового простору і оперативної пам'яті, ще і призводять до серйозних збоїв в

роботі системи. *Особливо небезпечні віруси* завдають найбільшої шкоди користувачеві та системі. Вони знищують програми, дані, різні записи на дисках й іншу інформацію, що може призвести до безповоротних втрат і серйозних відмов роботи системи.

За *здатністю до змінювання* віруси поділяють на сигнатурні та поліморфні. *Сигнатурні віруси* завжди мають постійний код або, принаймні, в тілі вірусу можна виділити незмінний код (сигнатуру), за яким його можна розпізнати. *Поліморфні віруси* дуже складно виявити, оскільки вони не мають сигнатур. Іншими словами, складно знайти два однакові тіла вірусу. Такий ефект досягається через трансформацію його тіла, наприклад, за допомогою нескладного шифрування (зазвичай з використанням операції XOR), а також змінюванням програми шифрування. Зазначені віруси називають *привидами* і досить часто при їх побудові використовують стелс-технологію, що робить їх невидимими у відповідному середовищі.

За *стилем написання* серед вірусів виокремлюють студентські, модифіковані, разові та серійні. *Студентські віруси* містять зазвичай багато помилок, побудовані за нескладним алгоритмом і є першими спробами авторів у цій області програмування. *Модифіковані віруси* створюють шляхом зміни коду вірусу, який раніше створили інші автори. Разові віруси створюють здебільшого досвідчені програмісти після копіткої роботи зазвичай з метою перевірки своїх професійних здібностей. *Серійні віруси* пишуть автори або група авторів і виконують їх в єдиному стилі. Досліджуючи їх, можна побачити, як зростає майстерність виробника.

За *типом коду* віруси поділяють на мікрокодовані та макрокодовані. *Мікрокодовані віруси* утворюються, як правило, внаслідок роботи компілятора і є бінарним кодом, що приєднується до інфікованого об'єкта. *Макрокодовані віруси* зазвичай обробляються інтерпретатором макрокоду і є послідовністю макрокодів, що зберігаються переважно в файлах (наприклад, з розширенням DOC), створених прикладними програмами, які мають можливість створення макросів (наприклад, MS WORD for WINDOWS).

Такі розгорнуті класифікації дають можливість створення більш ефективних комплексних систем захисту інформації.

10.2. Засоби захисту від несанкціонованого доступу

Індустрія сучасних засобів захисту інформації від НСД визначається широким номенклатурним арсеналом. Такі засоби з

практичного погляду можна розподілити на такі класи: апаратні, програмні, програмно-апаратні, криптографічні, стеганографічні, організаційні, законодавчі та морально-етичні.

Апаратні засоби – різноманітні механічні, електричні, електромеханічні, електронні, електронно-механічні та інші пристрої і системи (наприклад, джерела безперебійного живлення, криптографічні обчислювачі та ЗВІС-процесори, електронні ідентифікатори та ключі, пристрої для виявлення жучків, генератори шумів тощо), що функціонують автономно або вбудовуються чи з'єднуються з іншою апаратурою з метою блокування дій дестабілізуючих чинників та розв'язання інших задач захисту інформації.

Програмні засоби – спеціальні програми (наприклад, антивіруси, шифрування даних, реалізації алгоритмів цифрового підпису, розмежування доступу, оцінки ризиків, визначення рівня безпеки, організації експертиз тощо), які функціонують в межах інформаційних систем для розв'язання задач захисту інформації.

Програмно-апаратні засоби – взаємопов'язані апаратні і програмні засоби (наприклад, банківські системи електронних платежів, комплексні інформаційні системи конфіденційного зв'язку, автоматизовані системи контролю доступу персоналу та транспортних засобів в режимних зонах тощо), які функціонують автономно чи в складі інших систем з метою вирішення задач захисту інформації.

Криптографічні засоби – засоби, призначені для захисту інформації шляхом криптографічного перетворення інформації (шифрування, дешифрування), яке реалізується за допомогою асиметричних або симетричних криптографічних систем. Асиметричні криптографічні системи ґрунтуються на криптографії з відкритим ключем. Наприклад, найвідомішими практичними реалізаціями цього типу є системи Діффі-Хеллмана, RSA та Ель-Гамала. Симетричні криптографічні системи ґрунтуються на криптографії із секретним ключем, найбільш відомими практичними реалізаціями яких є, наприклад, DES, ГОСТ тощо.

Практичне використання сучасних криптографічних засобів тісно пов'язано з фундаментальними дослідженнями в цій галузі і здійснюється через відповідні апаратні, програмні і апаратно-програмні засоби, які побудовані на їх основі, наприклад, системи Тессера, Кліппера, Кріптона тощо. Слід зазначити, що з цим класом засобів тісно пов'язаний криптоаналіз, який ефективно використовується для випробування надійності криптографічних систем.

Стеганографічні засоби орієнтовані на приховування інформації в такій формі, коли сам факт її наявності не очевидний, наприклад, приховування даних у звукових або графічних файлах, які входять до складу ОС Windows.

Організаційні засоби захисту інформації це множина процесів та дій (наприклад, контроль за утилізацією носіїв інформації з обмеженим доступом, планування заходів щодо відновлення втраченої інформації, аудит систем захисту, реалізація експертиз тощо), що здійснюються протягом усіх технологічних етапів (проектування, виготовлення, модифікація, експлуатація, утилізація тощо) існування відповідних ресурсів інформаційних систем і ведуть до створення, удосконалення, упорядкування і узгодженості взаємозв'язків та взаємодії їх компонентів з метою розв'язання задач захисту інформації. Розробляючи організаційні засоби, необхідно враховувати, щоб у загальній множині механізмів захисту вони могли самостійно або в комплексі з іншими засобами вирішувати задачі захисту, забезпечувати ефективне використання засобів інших класів, а також раціонально об'єднувати всі засоби в єдину цілісну систему захисту. Слід зазначити, що множина всіх потрібних і потенційно можливих організаційних засобів невизначена і не існує формальних методів формування їх переліку і змісту. Виходячи з цього, основними методами формування організаційних засобів можна вважати лише неформально-евристичні.

Законодавчі засоби захисту інформації є множиною нормативно-правових актів (конвенції, закони, укази, постанови, нормативні документи тощо), що діють у певній державі і забезпечують юридичну підтримку для розв'язання задач захисту інформації. Загалом, за допомогою законодавчих засобів визначаються права, обов'язки та відповідальність щодо правил взаємодії з інформацією, порушення яких може вплинути на стан її захищеності. У світовій практиці основу зазначених засобів складають патентне та авторське право, національні закони про державну таємницю та обробку інформації в інформаційних системах, ліцензування, страхування, сертифікація, класифікаційні нормативні документи тощо (розділ 10).

Морально-етичні засоби – моральні норми та етичні правила, що склалися в суспільстві, колективі та об'єкті інформаційної діяльності, порушення яких ототожнюється з недотриманням загальноприйнятих дисциплінарних правил та професійних ідеалів. Прикладом таких засобів може бути кодекс честі, етикет, етика хакера тощо.

Базові вимоги до будь-якого засобу захисту інформаційних систем можна розбити на п'ять категорій (рис. 10.27).

КАТЕГОРІЇ ЗАГАЛЬНИХ ВИМОГ				
Функціональність	Надійність	Адаптованість	Ергономічність	Економічність
Вирішення необхідної сукупності задач захисту	Здатність вчасно, правильно і коректно виконувати всі передбачені функції захисту	Адаптація при зміні структури, технологічних схем і умов функціонування інформаційної системи	Вимоги до зручності адміністрування, експлуатації і мінімізації перешкод користувачам	Мінімізація фінансових і ресурсних витрат

Рис. 10.27. Базові вимоги до засобів захисту інформаційних систем

Категорії засобів захисту програмного забезпечення. Як показує практика, найдорожчою складовою інформаційної системи є її програмне забезпечення. Тому розробка програмних засобів захисту інформації для захисту програмного забезпечення з практичного погляду – найбільш привабливі задачі. Питання захисту програмного забезпечення від хакерів (стосовно його копіювання і динамічних та статичних методів дослідження) найгостріше стоїть перед його розробниками і власниками. З системних позицій, захист програмного забезпечення забезпечується цілим комплексом засобів, що починається законодавчими актами і закінчується конкретними апаратурними розробками. Відомі такі категорії засобів захисту програмного забезпечення (рис. 10.28): власні, в складі інформаційної системи, із запитом інформації, активні та пасивні.

Засоби власного захисту визначають елементи захисту, що містяться в самому програмному забезпеченні або супроводжують його. До них належать: документація, розповсюдження продукції у вигляді виконуючих модулів, супроводження програм розробником, обмеження застосування, проектування на замовлення, вбудовані ідентифікаційні мітки власника та авторське право.



Рис. 10.28. Категорії засобів захисту програмного забезпечення

До *засобів захисту в складі інформаційних систем* відносять: захист магнітних дисків, захисні механізми самих пристроїв інформаційних систем, замки захисту доступу та зміну функцій в системі.

Серед механізмів захисту магнітних дисків потрібно виділити дві групи:

1. Перешкода безпосередньому копіюванню програм з диска на диск.

2. Захист програм від реасемблювання і налагоджувачів.

Реасемблери та *налагоджувачі* дозволяють представити програму в формі, більш доступній для сприйняття, а також допомагають вивчити логіку захисту і оперативно здійснити модифікацію програм.

Перша група захищає програму від несанкціонованого відтворення, друга - від несанкціонованої ревізії. Ці групи не взаємопов'язані, оскільки програми, призначені для вільного розповсюдження, можуть бути захищені від налагоджувачів і реасемблерів, і навпаки, різні програмні продукти можуть бути захищені не тільки від копіювання. Найбільш надійний захист забезпечується одночасним застосуванням двох груп, оскільки її подолання в програмі, скопійованій із захищеного диска, буде пов'язане з подоланням захисту другої групи. Отже, програми, які захищені від налагоджувачів, реасемблерів і якщо знаходяться на стандартному диску, можуть бути легко скопійовані, а ті що розташовані на захищеному диску і які не містять другої групи, можуть бути легко вивчені та модифіковані з метою подолання механізмів захисту.

Програми, що використовують *засоби захисту із запитом інформації* для подальшого виконання призначених функцій, здійснюють запит на введення додаткової інформації, представлені, наприклад, у вигляді ключових слів. До структури зазначених засобів входять паролі, криптографічні шифри, сигнатури на основі унікальних характеристик інформаційних систем і апаратура захисту.

Засоби активного захисту поділяють на зовнішні і внутрішні; вони активізуються у разі виникнення певних (нештатних) обставин, що підпадають під контроль, наприклад, таких, як неправильно введені паролі, вичерпаний час користування, вичерпана кількість запусків, неправильна контрольна сума ділянок програми тощо.

Засоби пасивного захисту включають методи ідентифікації програм, пристрої контролю подій, водяні знаки, що перешкоджають створенню копії і психологічні заходи. За допомогою цих засобів здійснюється підтвердження справжності, контроль доступу, пошук доказів несанкціонованого копіювання тощо.

Програмні засоби захисту є найважливішою і необхідною частиною механізму захисту сучасних інформаційних систем і можуть бути як автономними, так і входити до складу різного програмного забезпечення. Це насамперед пов'язано з такими їх властивостями, як універсальність, гнучкість, надійність, простота реалізації, можливість модифікації та удосконалення. З універсальністю пов'язана можливість вирішення програмними засобами широкого кола задач захисту. Гнучкість асоціюється з тим, що вони можуть адаптуватися до конкретних умов функціонування інформаційних систем, а також у структурі інформаційних систем, наприклад, бути складовою частиною операційної системи (ОС), функціонувати як самостійні пакети програм захисту, розподілятися між окремими компонентами системи і т. ін. З надійністю пов'язана висока програмна стійкість у разі тривалої неперервної роботи і задоволення високих вимог до достовірності впливів управління за наявності різних дестабілізуючих факторів. Простота реалізації програмних засобів захисту очевидна порівняно з можливістю реалізації будь-яких інших засобів. Можливості їх модифікації і удосконалення визначаються самою природою.

Основними недоліками їх використання є те, що вони додатково завантажують процесор, що призводить до збільшення часу реагування на запити і, відповідно, до зменшення ефективності роботи, а також до зменшення ємності доступної і зовнішньої оперативної пам'яті. Слід зазначити, що у разі випадкової чи зловмисної модифікації програми вона може втратити здатність виконувати функції захисту, а також стати додатковим каналом несанкціонованого отримання інформації. Крім цього,

Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу

жорстка орієнтація на архітектуру певних типів ПК, а також залежність від ОС доповнює зазначені недоліки.

Перелік програмних засобів захисту може бути будь-яким, а виконуваними ними функції можуть бути такі: перевірка прав доступу (за простим паролем, за складним паролем, за разовими паролями); розпізнавання користувачів за різними ідентифікаторами та компонентів програмного забезпечення і елементів баз даних; розмежування доступу до захищених даних за матрицею повноважень, рівнем секретності та іншими ознаками; керування доступом до задач, програм і елементів баз даних за спеціальними мандатами; реєстрація звернень до системи, задач, програм і елементів захищених даних; підготовка до видачі конфіденційних документів (формування і нумерація сторінок, визначення і присвоєння грифа обмеження доступу, реєстрація виданих документів тощо); перевірка адресата перед видачею захищених даних у канали зв'язку; керування видачею даних у канали зв'язку; криптографічне перетворення даних; контроль процесів обробки і видачі захищених даних; знищення залишкової інформації у постійному запам'ятовувальному пристрої після виконання запитів користувачів; сигналізація спроб несанкціонованих дій; блокування роботи користувачів, які порушують правила захисту інформації; організація псевдороботи з порушником з метою відволікання його уваги; забезпечення комплексних засобів і систем захисту; комплексний захист від вірусів та інших програмних руйнівних впливів; забезпечення проведення експертиз і сертифікації; керування ризиком та організація дій у кризових ситуаціях тощо.

Слід зазначити, що індустрії цього напрямку властива стихійність розвитку програм захисту, що, з одного боку, не дає гарантій повноти наявних засобів, а з другого боку – не виключає дублювання тих самих задач захисту. Виходячи з вищерозглянутого, можна відокремити три принципово важливі *вимоги до формування програмних засобів захисту: функціональна повнота, гнучкість і уніфікованість використання.*

Що стосується першої вимоги, то неважко переконатися, що за наведеним вище переліком можна створити комплексний продукт, який найповніше охоплює всі класи захисту.

Задоволення інших двох вимог залежить від форм і способів представлення програм захисту. Аналіз показав, що найповніше вимоги гнучкості й уніфікованості задовольняє така сукупність принципів: наскрізна модульна побудова, повна структуризація, представлення машинезалежною мовою.

Принцип наскрізної модульної побудови полягає в тому, що кожному з програм будь-якого рівня (обсягу) потрібно представляти у вигляді системи модулів, кожен з яких має бути цілком автономним і мати

стандартні вхід і вихід, що забезпечують комплексування з будь-якими іншими модулями.

Представлення машинезалежною мовою означає, що програмні модулі мають бути такими, щоб їх з мінімальними зусиллями можна було включати до складу програмного забезпечення будь-якої інформаційної системи. Повну універсалізацію представлення модулів можна забезпечувати поданням їх у вигляді блок-схеми, деталізованої до такого ступеня, щоб кожний блок можна було реалізовувати малою кількістю операторів найбільш розповсюджених мов високого чи низького рівня.

10.3. Моделювання систем і процесів захисту інформації

У процесі розвитку теорії і практики інформаційної безпеки сформувалися *емпіричний, теоретичний і теоретико-емпіричний методологічні підходи* до оцінки уразливості інформації.

Суть *емпіричного* підходу полягає в тому, що на основі тривалого збору й обробки даних реальних проявів загроз інформації і розмірів нанесеного збитку, виключно емпіричним шляхом встановлюються залежності між потенційно можливим збитком і коефіцієнтами, що характеризують частоту прояву відповідної загрози і значення розміру збитку.

Вихідною посилкою при розробці моделей є припущення, що при порушенні захищеності інформації наноситься деякий збиток, а забезпечення захисту інформації пов'язане з витратами. Очікувана вартість захисту може бути визначена сумою витрат на захист і збитком від її порушення. Зазначена залежність графічно представлена на рис. 10.29. Очевидно, що оптимальним рішенням є виділення на захист інформації засобів у розмірі B_{opt} , оскільки саме при цьому забезпечується мінімізація загальної вартості захисту інформації.

Для того щоб скористатися даним підходом до розв'язання проблеми, необхідно

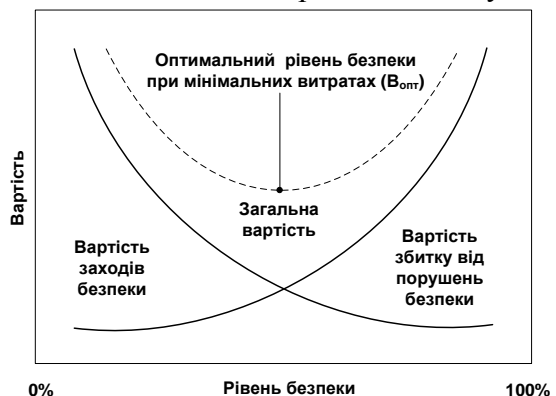


Рис. 10.29. Оцінка вартості захисту інформації

Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу

визначити очікувані втрати від порушення захищеності інформації і залежність між рівнем захищеності і витратами, на її захист. Витрати, що забезпечує необхідний рівень захищеності, залежать від повної множини загроз інформації, потенційної небезпеки від реалізації кожної з них та розмірів витрат, необхідних для їх нейтралізації.

Оскільки оптимальний рівень витрат на захист (рис. 10.31) відповідає рівню очікуваних збитків від порушень безпеки, то достатньо визначити тільки рівень збитку. Очікувані збитки від i -ї загрози інформації можна визначити за відомою емпіричною залежністю:

$$R_i = 10^{(S_i + V_i - 4)}, \text{ де } S_i \text{ та } V_i - \text{ коефіцієнти, що}$$

характеризують можливу частоту виникнення відповідної загрози та значення можливого збитку при її виникненні, а значення цих коефіцієнтів були визначені на основі суджень експертів:

Очікувана частота появи загрози	S_i	Значення можливого збитку при прояві загрози (USD)	V_i
Майже ніколи	0	1	0
1 раз у 1000 років	1	10	1
1 раз у 100 років	2	100	2
1 раз у 10 років	3	1000	3
1 раз у рік	4	10000	4
1 раз на місяць	5	100000	5
2 рази в тиждень	6	1000000	6
3 рази в тиждень	7	10000000	7

Сумарний збиток визначається за виразом: $R_i = \sum_{\forall_i} \alpha_i V_i$, де α_i -

ваговий коефіцієнт. Розглянута модель є достатньо грубою, а підвищення її адекватності приводить до складних аналітичних викладень, заснованих на методах теорії імовірностей і прийняття рішень.

Модель з повним перекриттям. Розвитком моделей оцінки загроз інформаційних систем є моделі їх нейтралізації, тобто моделі захисту, найбільш загальною з яких є модель *системи з повним перекриттям*.

Побудова даної моделі засновується на тому, що в механізмі захисту повинен міститися принаймні один засіб для перекриття будь-якого потенційно можливого каналу витоку інформації. Відома наступна методика формального опису такої системи:

1. Складається повний перелік об'єктів O системи, що підлягають захисту;
2. Складається повний перелік потенційно можливих загроз T інформації;

3. Складені таким чином множини поєднуються в двочастковий граф з дотриманням умови: ребро $\langle t_i, o_j \rangle$ існує тільки тоді, коли загроза t_i є реальною для об'єкта o_j .

4. Для кожного ребра в графі визначається кількісна міра відповідної загрози для відповідного об'єкта;

5. Формується множина M засобів захисту інформації в обчислювальній системі;

6. Визначається кількісна міра можливості протидії кожного засобу захисту кожної з загроз. Якщо можливість протидії перевищує рівень загрози, то відповідне ребро графа виключається.

Якщо множина M така, що усуваються всі ребра графа, то така система є **системою з повним перекриттям**.

Одною з різновидів **теоретично** строгих моделей є моделі систем розмежування доступу до ресурсів інформаційних систем.

У загальному випадку суть цих моделей може бути представлена наступним чином. Інформаційна система є *системою множинного доступу, тобто до тих самих її ресурсів мають права доступу деяке число користувачів (процесів)*. Якщо будь-які із зазначених ресурсів захищаються, то доступ до них здійснюється лише за наявності відповідних повноважень, система розмежування доступу є механізмом, що регулює такий доступ. За цим механізмом не повинен бути дозволений доступ користувачам (процесам), що не мають на це повноважень, і не повинно бути відмовлення в доступі користувачам (процесам), що мають відповідні повноваження.

Розглянемо **теоретико-емпіричний** підхід до оцінки уразливості інформації.

За визначенням, базовим є показник уразливості інформації в одному структурному ресурсів інформаційної системи щодо одного дестабілізуючого чинника і щодо одного порушника однієї категорії (для чинників, пов'язаних зі злочинними діями людей).

Для визначення базових показників уразливості різних видів застосовуються аналітичні моделі, що дозволяють визначати шукані величини (у даному випадку - показники уразливості інформації) шляхом проведення обчислень за заздалегідь установленими (виведеними) залежностями.

Наведемо деякі аналітичні моделі.

1. Порушення фізичної цілісності інформації. Введемо наступні позначення: $P_{iex}^{(u)}$ – імовірність того, що на вхід i -го структурного компонента надходить інформація з порушеною цілісністю; $P_{ijk}^{(u)}$ – імовірність того, що цілісність інформації, що знаходиться

Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу

(оброблюваної, збереженої, переданої) у i -му структурному компоненті, буде порушена під впливом j -го дестабілізуючого чинника i (у випадку злочинних дій людей) щодо одного порушника k -ї категорії. Для тих дестабілізуючих чинників, що не пов'язані зі злочинними діями людей, індекс “ k ” ігнорується, тобто значення $P_{ijk}^{(u)}$

для всіх k однакові і залежать тільки від i та j ; $P_{ijk \text{ вих}}^{(u)}$ - імовірність того, що цілісність вихідної з i -го компонента інформації порушена під впливом j -го дестабілізуючого чинника i (у випадку злочинних дій людей) щодо одного порушника k -ї категорії (щодо індексу “ k ” справедливо висловлене вище зауваження).

Відповідно до теореми множення ймовірностей випадкових подій величина $P_{ijk \text{ вих}}^{(u)}$ (порушення цілісності вихідної інформації) може бути

виражена наступною залежністю:
$$P_{ijk \text{ вих}}^{(u)} = 1 - (1 - P_{i \text{ вх}}^{(u)}) \cdot (1 - P_{ijk}^{(u)}) .$$

Розкривши дужки і зробивши тотожні перетворення, остаточно одержуємо

$$P_{ijk \text{ вих}}^{(u)} = P_{i \text{ вх}}^{(u)} + [1 - P_{i \text{ вх}}^{(u)}] \cdot P_{ijk}^{(u)} . \tag{10.1}$$

Графічно дана залежність представлена сімейством кривих на рис. 10.30.

Як бачимо на цю модель є певні обмеження. Тут передбачається, що різні дестабілізуючі чинники впливають на інформацію незалежно один від одного, а процес порушення цілісності в ресурсів інформаційної системи не залежить від того, чи порушена цілісність інформації, що надходить на його вхід.

Тому правомірність використання такої моделі повинна бути обґрунтована.

Події $P_{ijk}^{(u)}$ – є складними з тієї точки зору, що для їхнього здійснення необхідне одночасне проявлення відповідного дестабілізуючого чинника і власне порушення цілісності інформації при його дії. Якщо імовірність

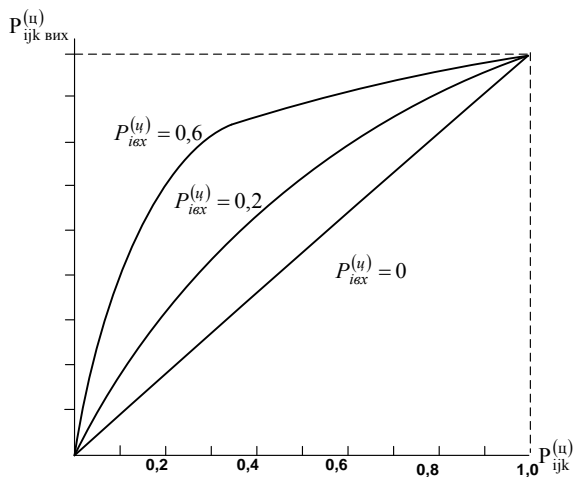


Рис. 10.30. Графік значень базової ймовірності порушення фізичної цілісності інформації

першої події позначити через $P_{ijk}^{(y,n)}$, а другої – через $P_{ijk}^{(y,u)}$, то $P_{ijk}^{(y)} = P_{ijk}^{(y,n)} \cdot P_{ijk}^{(y,u)}$.

Підставляючи це значення в (10.1), одержуємо

$$P_{ijk\text{вих}}^{(y)} = P_{i\text{вх}}^{(y)} + [1 - P_{i\text{вх}}^{(y)}] \cdot P_{ijk}^{(y,n)} \cdot P_{ijk}^{(y,u)}$$

Значення $P_{ijk\text{вих}}^{(y)}$ і є базовим показником уразливості з погляду порушення цілісності вихідної інформації. Однак для його використання необхідні значення $P_{ijk}^{(y)}$ для всіх структурних компонентів і всіх дестабілізуючих інформаційної системи. Слід зазначити, що формування всієї множини значень цих величин пов'язано з певними труднощами.

2. Несанкціоноване одержання інформації. З погляду несанкціонованого одержання інформації головну небезпеку представляють злочинні дії людей. Введемо такі позначення: $P_{ikl}^{(n,k)}$ – імовірність доступу порушника k -ї категорії в l -у зону i -го ресурсу інформаційної системи; $P_{ijl}^{(n,k)}$ – імовірність наявності (прояву) j -го каналу несанкціонованого отримання інформації в l -й зоні i -го ресурсу інформаційної системи; $P_{ijkl}^{(n,u)}$ – імовірність доступу порушника k -ї категорії до j -му каналу несанкціонованого отримання інформації в l -й зоні i -го компонента за умови доступу порушника в зону; $P_{ijl}^{(n,u)}$ – імовірність наявності інформації, що захищається, у j -му каналі несанкціонованого отримання інформації у l -й зоні i -го компонента в момент доступу туди порушника. Зазначимо, що зони є вкладеними і, наприклад, $l=1,5$ (рис. 10.31), то перша зона є зовнішньою (перший рубіж), а п'ята – внутрішньою (останній рубіж).

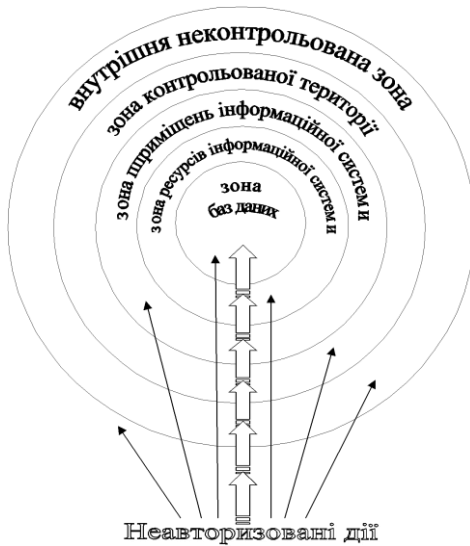


Рис. 10.31. Приклад схеми можливих несанкціонованих дій в інформаційній системі

Класичні мандатні моделі цілісності і розмежування доступу. Модель мандатного розмежування доступу, що одержала назву моделі Белла-Лападула (БЛМ), дотепер впливає на дослідження і розробки в

області комп'ютерної безпеки. Ідеї, закладені в БЛМ, можуть бути використані при побудові різних політик безпеки.

Принципи, що лежать в основі БЛМ, беруть походження від паперового документування інформації з обмеженим доступом, а Белл і Лападула перенесли відомі підходи забезпечення безпеки в інформаційні технології. Тут основною тезою є те, що всі суб'єкти і об'єкти асоціюються з рівнями безпеки, які варіюються від низьких рівнів (некласифікованих) до високих (цілком секретних). Крім того для запобігання витоку інформації до неуповноважених суб'єктів з низькими рівнями безпеки їм не дозволяється читати інформацію з об'єктів з високими рівнями безпеки. Ця теза веде до першого правила БЛМ.

Проста властивість безпеки, також відома як правило “заборона читання вгору” (NRU), говорить, що суб'єкт із рівнем безпеки x_s може читати інформацію з об'єкта з рівнем безпеки x_o , тільки якщо x_s переважає над x_o . Це означає, що якщо в системі, що задовольняє правила моделі БЛМ, суб'єкт із рівнем доступу секретний спробує прочитати інформацію з об'єкта, класифікованого як цілком секретний, то такий доступ не буде дозволений.

Белл і Лападула також використали тезу, що суб'єктам не дозволяється розміщати інформацію чи записувати її в об'єкти, що мають більш низький рівень безпеки. Наприклад, коли цілком секретний документ переміщується в некласифіковане сміттєве відро, то може відбутися витік інформації і це призвело до другого правила БЛМ.

Властивість відома як правило “заборона запису вниз” (NWD), говорить, що суб'єкт безпеки x_s може записати інформацію в об'єкт із рівнем безпеки x_o тільки якщо x_o має перевагу над x_s . Це означає, що якщо в системі, що задовольняє правилам моделі БЛМ, суб'єкт із рівнем доступу цілком секретний спробує записати інформацію в некласифікований об'єкт, то такий доступ не буде дозволений. Така властивість, наприклад, розв'язує проблему троянських коней, тому що запис інформації на більш низький рівень безпеки (типова дія троянських коней) заборонений.

Правило заборони за записом є великим спрощенням деяких реалізацій БЛМ. Так, деякі описання включають більш детальне поняття типу доступу (наприклад такі, як додавання і виконання).

Правила заборони за записом і читанням БЛМ відповідають інтуїтивним поняттям того, як запобігти витоку інформації до неуповноважених (неавторизованих) джерел.

Незважаючи на всі переваги, виявилось, що при використанні БЛМ у контексті практичного проектування і розробки реальних інформаційних систем виникає низка технічних питань, які є логічним наслідком переваг БЛМ - її простоти. Проблеми виникають при розгляді питань побудови політик безпеки для конкретних типів систем, тобто на менш абстрактному рівні розгляду. При даному розгляді системний компонент моделі ускладнюється, що може привести до неадекватності БЛМ у її класичній

формі, що ініціювало широку полеміку з приводу застосовності БЛМ для побудови безпечних систем.

Мандатна модель контролю цілісності (Модель Біба). Відомо, що в БЛМ важливість чи чутливість суб'єктів і об'єктів зростає з підвищенням в ієрархії рівнів безпеки. При розгляді моделей контролю цілісності запис вгору може бути загрозою в тому випадку, коли суб'єкт із низьким рівнем безпеки модифікує чи знищує дані в об'єкті, що лежить на більш високому рівні. Тому, виходячи з задач цілісності, існує вимога заборони такого запису. Дотримуючись подібних аргументів, можна розглядати читання знизу як потік інформації, що йде від об'єкта нижнього рівня і порушує цілісність суб'єкта високого рівня. Тому достатньо ймовірно, що і таке читання необхідно заборонити.

Таких два спостереження зробив у середині сімдесятих минулого століття Кен Біба. Вони були послідовно внесені в модель безпеки і з тих пір називається моделлю цілісності Біба (чи просто моделлю Біба), який відобразив її у такий спосіб, як і БЛМ, за тим виключенням, що правила його моделі є повною протилежністю правилам БЛМ. Зазвичай розглядаються три варіації моделі Біба: мандатна модель цілісності; модель зниження рівня суб'єкта; модель зниження рівня об'єкта, а загальний термін “модель Біба” використовується для позначення будь-якої чи відразу всіх трьох моделей. Для мандатної моделі контролю цілісності відомий формальний опис і її часто називають інверсією БЛМ. Це досить точна назва, оскільки основні правила цієї моделі є обернені правилам БЛМ і відображаються як “заборона читання знизу” (NRD) та “заборона запису вгору” (NWU). Визначимо їх у термінах суб'єктів, об'єктів, і нового типу рівнів безпеки – *рівнів цілісності*, над якими введене відношення переваги.

Правило NRD мандатної моделі цілісності Біба визначається як заборона суб'єктам на читання інформації з об'єкта з більш низьким рівнем цілісності. Правило NRD є повною протилежністю NRU БЛМ, за винятком того, що тут використовуються рівні цілісності (а не безпеки, як у БЛМ), а правило NWU (заборона суб'єктам на запис інформації в об'єкт із більш високим рівнем цілісності) є повною протилежністю правилу NWD БЛМ для того ж випадку рівнів цілісності.

Однією з переваг цієї моделі є те, що вона успадкувала багато важливих характеристик БЛМ, включаючи її простоту і інтуїтивність. Це значить, що проектувальники реальних систем можуть легко зрозуміти суть цих правил і використовувати їх для прийняття рішень при проектуванні. Крім того, оскільки мандатна модель цілісності Біба, подібно БЛМ, заснована на простій ієрархії, її легко пояснити і відобразити користувачам системи.

З іншого боку, модель є очевидним протиріччям з правилами NRU і NWD. Це значить, що коли необхідно побудувати систему, яка запобігає загрозам секретності і цілісності, то одночасне використання правил моделей БЛМ і Біба може призвести до ситуації, у якій рівні безпеки і цілісності будуть використовуватися протилежними способами.

Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу

Розглянемо формальний опис моделі Біба. Для цього опишемо прості математичні конструкції, що допоможуть описати різні правила, які складають цю мандатну модель.

Подамо множини суб'єктів і об'єктів, де рівні цілісності суб'єкта чи об'єкта x позначаються як рівень (x), і для них введено відношення *переваги*. Використовуючи ці визначення, сформулюємо правила NRD і NWU мандатної моделі цілісності Біба в термінах булевої функції *ДОЗВОЛИТИ*:

NRD: $\forall s \in \text{суб'єкти}, o \in \text{об'єкти}$:

ДОЗВОЛИТИ (s, o , читання), тоді і тільки тоді, якщо рівень (o) *переважає* рівень (s).

Даний тип визначення передбачає умови, при яких функція *ДОЗВОЛИТИ* приймає значення істина. Визначення стверджує, що для всіх визначених суб'єктів і об'єктів операція читання дозволена тільки в тому випадку, коли виконується умова переваги. Правило NWU просто є оберненим до використання відношення переваги, як це показано в наступному визначенні:

NWU: $\forall s \in \text{суб'єкти}, o \in \text{об'єкти}$:

ДОЗВОЛИТИ (s, o , запис) $\Leftrightarrow \text{clerence}(s) \geq \text{classification}(o)$.

Це визначення стверджує, що для всіх суб'єктів і об'єктів операція запису дозволяється тільки в тому випадку, коли виконується умова переваги. Подібність визначення цих двох правил правилам моделі БЛМ може надати зручний спосіб для розробників системи передбачити можливість переконфігурування правил БЛМ таким чином, щоб підтримувати мандатну модель цілісності Біба.

Модель зниження рівня суб'єкта (друга модель Біба) пов'язана з невеликим ослабленням правила читання знизу. Ця модель не дозволяє суб'єктам з високою цілісністю читати інформацію із об'єктів з більш низькою цілісністю. Таке правило гарантує, що інформація із об'єкта з низькою цілісністю не порушить цілісність суб'єкта. Але в цій моделі суб'єкту дозволяється здійснювати читання знизу в результаті якого рівень цілісності суб'єкта знижується до рівня цілісності об'єкта.

Модель зниження рівня об'єкта (третя модель Біба) пов'язана з послабленням правил запису вгору, тобто замість повної заборони запису вгору модель дозволяє такий запис, але знижує рівень цілісності об'єкта до рівня цілісності суб'єкта, що здійснює запис.

10.4. Протидія міжмережевому несанкціонованому доступу

Основою сучасних інформаційних систем є комп'ютерні мережі, на які також поширюються дії раніше описаних методів НСД. В загальному випадку мережа організації фактично є неоднорідною низкою комп'ютерів, неоднакової конфігурації і призначення,

керованих різними операційними системами і зв'язаних між собою за допомогою мережевого обладнання. У таких умовах здійснити надійний захист від несанкціонованих дій при взаємодії із зовнішніми мережами можна лише за допомогою спеціалізованих програмно-апаратних комплексів, що забезпечують відповідний захист. Такі комплекси називають міжмережевими екранами, міжмережевими фільтрами, екрануючими фільтрами або брандмауерами або системи FireWall, які встановлюються на стику між внутрішньою і зовнішньою мережами і функції протидії беруть на себе.

Функції міжмережевого екранування. Як вже зазначалося для захисту від несанкціонованого міжмережевого доступу брандмауер розташовується між мережею, що підлягає захисту (внутрішня) і мережею, з якої можливий НСД (зовнішня) (рис. 10.32), при цьому всі міжмережеві взаємодії здійснюються тільки через екрануючий фільтр, який організаційно входить до складу внутрішньої мережі.

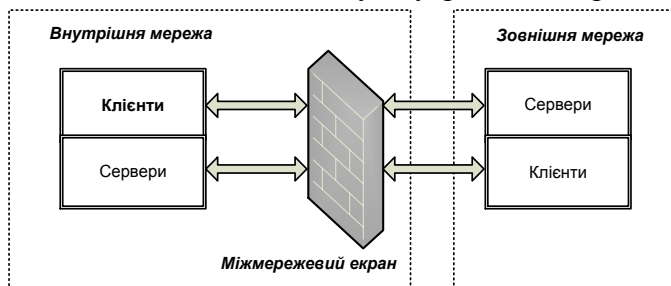


Рис. 10.32. Схема підключення брандмауера

Міжмережевий екран повинен враховувати протоколи інформаційного обміну, які покладені в основу функціонування внутрішньої і зовнішньої мереж, а якщо протоколи відрізняються, то брандмауер повинен підтримувати багатопрокольний режим роботи, забезпечуючи протокольне перетворення рівнів відповідно до моделі OSI для взаємодії відкритих мереж.

Для міжмережевого екрану окремо задаються правила, які обмежують доступ із внутрішньої мережі в зовнішню та навпаки і у загальному випадку його робота заснована на динамічному виконанні двох груп функцій: фільтрація інформаційних потоків та посередництво у міжмережевій взаємодії.

При цьому прості міжмережеві екрани орієнтовані на виконання тільки однієї з даних функцій, а комплексні – забезпечують спільне їх виконання. Для прийняття управляючих рішень щодо використовуваних сервісів брандмауер повинен отримувати, запам'ятовувати, вибирати й обробляти інформацію, одержану від усіх комунікаційних рівнів і додатків. Повнота і правильність управління вимагають, щоб комплексний брандмауер мав можливість аналізу і використання низки чинників:

Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу

Інформація про з'єднання – інформація від усіх семи рівнів у пакеті.

Історія з'єднань – інформація, отриманої від попередніх з'єднань.

Стан рівня додатка – інформація про стан, отримана з інших додатків.

Агрегуючий елементі – обчислення різноманітних залежностей, оснований на всіх вище перерахованих чинниках.

При екрануванні окремого комп'ютера підтримується доступність мережесервісів та зменшується навантаження, ініційоване зовнішньою активністю, що знижує вразливість захищених внутрішніх сервісів комп'ютера, оскільки спочатку неавторизована сторона повинна подолати механізм захисту екрануючого фільтра.

Базова задача брандмауера – фільтрація трафіку, що пов'язано з вибіркоким пропусканням даних через екрануючий фільтр (іноді з виконанням деяких перетворень).

Фільтрація здійснюється на основі низки правил, що завантажуються в екран, і мережесервісів, що є вираженням, прийнятої політики безпеки. З цих позицій брандмауер зручно представляти як послідовність фільтрів (рис. 10.33), що обробляють інформаційні потоки на основі інтерпретації окремих правил фільтрації шляхом:

аналізу інформації за заданими в інтерпретуючих правилах критеріями, наприклад, за адресами одержувача і або відправника, за типом додатка, для якого призначена інформація тощо.

прийняття на основі інтерпретуючих правил одного з наступних рішень:

- 1) не пропустити дані;
- 2) обробити дані від імені одержувача і повернути результат відправнику;
- 3) передати дані на наступний фільтр для продовження аналізу;
- 4) пропустити дані, ігноруючи наступні фільтри.

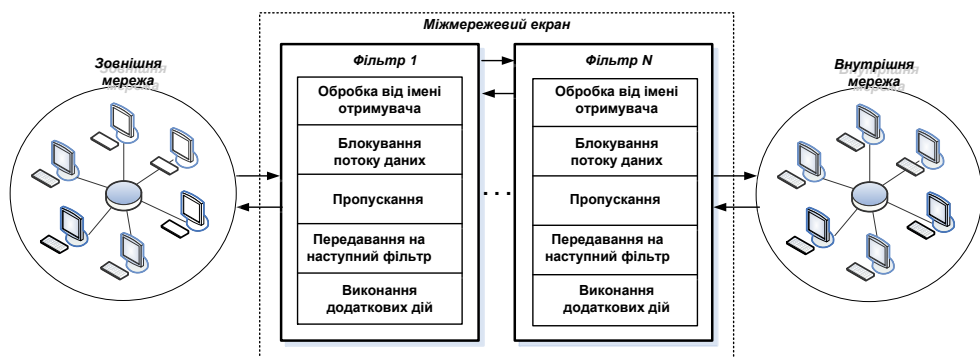


Рис. 10.33. Логічна структура міжмережевого екрану

При фільтрації можуть задаватися і додаткові дії (посередництва), наприклад, перетворення даних, реєстрація подій тощо і тоді відповідні

правила фільтрації визначають перелік умов, за якими з використанням зазначених критеріїв аналізу здійснюються:

дозвіл чи заборона подальшої передачі даних;

виконання додаткових захисних функцій.

Критеріями аналізу інформаційного потоку можуть використовуватися:

службові поля пакетів повідомлень, що містять мережеві адреси, ідентифікатори, адреси інтерфейсів, номери портів і інші важливі дані;

безпосередній зміст пакетів повідомлень (наприклад, на наявність комп'ютерних вірусів), що підлягають перевірці;

зовнішні характеристики інформаційного потоку, наприклад, часові та частотні характеристики, обсяг даних і тощо.

Використовувані критерії аналізу залежать від рівнів моделі OSI, на яких здійснюється фільтрація, а у загальному випадку діє правило: чим вище рівень моделі OSI, на якому брандмауер фільтрує пакети, тим вище і забезпечуваний ним рівень захисту.

Ще однією базовою функцією брандмауера є виконання посередництва, яку він здійснює за допомогою спеціальних екрануючих програм (агентів) – програмам-посередників, які є резидентними і контролюють (забороняють, дозволяють) передачу пакетів між зовнішньою і внутрішньою мережею.

В разі доступу з одної мережі в іншу, спочатку встановлюється логічне з'єднання з програмою-посередником, яка перевіряє допустимість запиту міжмережевої взаємодії і при його дозволі встановлює з'єднання з необхідним комп'ютером. Надалі обмін між комп'ютерами мереж здійснюється через програмного посередника, що може виконувати фільтрацію трафіку та здійснювати інші захисні функції.

Зазначимо, що функції фільтрації міжмережевий екран може виконувати без програм-посередників, забезпечуючи прозору взаємодію між внутрішньою і зовнішньою мережею, програмні посередники в свою чергу можуть і не здійснювати відповідну фільтрацію.

У загальному випадку екрануючі агенти блокуючи прозору передачу можуть виконувати ще ряд функцій (рис. 10.34).



Рис. 10.34. Базові функції екрануючих агентів

Ідентифікація і аутентифікація **користувачів** необхідна для забезпечення високого ступеня безпеки при їх доступі в мережі, при цьому для запобігання перехоплення паролів не передається у відкритому вигляді через загальні комунікації. Найбільш ефективним способом аутентифікації є використання одноразових паролів, а також застосування цифрових сертифікатів, виданих довірчими органами (наприклад, центром розподілу ключів).

Перевірка дійсності одержуваних і переданих даних є особливо актуальною функцією не тільки для аутентифікації електронних повідомлень, але і мігруючих програм (наприклад, Java), стосовно яких може бути виконана підробка. Ця функція пов'язана з контролем цифрових підписів, при цьому також можуть застосовуватися цифрові сертифікати. Розмежування доступу до ресурсів внутрішньої або зовнішньої мережі при звертанні до міжмережевому екрану реалізується за допомогою функцій ідентифікації і аутентифікації користувачів. Функція розмежування доступу базується на одному з таких підходів: дозвіл доступу тільки за визначеними адресами у зовнішній мережі; фільтрація запитів за оновлюваними списками недопустимих адрес і блокування пошуку з небажаними ключовими словами; накопичування і поновлення санкціонованих інформаційних ресурсів зовнішньої мережі в масиві пам'яті брандмауера і повна заборона доступу в зовнішню мережу.

Фільтрація і перетворення потоку повідомлень виконується посередником на основі заданого набору правил, при цьому розрізняють: екрануючі агенти, орієнтовані на аналіз потоку повідомлень для визначених видів сервісу (наприклад, FTP, HTTP, Telnet тощо) та універсальні екрануючі агенти, що обробляють всі повідомлення, наприклад, з метою пошуку і знешкодження комп'ютерних вірусів, прозорого шифрування даних тощо.

Трансляція внутрішніх мережевих адрес реалізується для всіх пакетів, що слідує із внутрішньої мережі в зовнішню. Для цих пакетів здійснюється автоматичне перетворення IP-адрес комп'ютерів-відправників в одну IP-адресу, асоційовану із брандмауером, з якого передаються усі вихідні пакети, що виключає прямий контакт між внутрішньою і зовнішньою мережею, а IP-адреса брандмауера стає єдиною активною IP-адресою, яка потрапляє в зовнішню мережу.

Це дозволяє приховати топологію внутрішньої мережі, а також мати усередині мережі власну систему адресації, що ефективно вирішує проблему розширення адресного простору.

Аналіз та реєстрація подій, реагування на визначені події, а також аналіз зареєстрованої інформації і складання звітів дозволяє виявити спроби несанкціонованих дій. Ефективність брандмауера в значній мірі залежить від ефективності виконання цієї функції, що дозволяє формувати попереджувальні сигнали при виявленні нападу.

Система реєстрації, збору й аналізу статистики опрацьовує адреси клієнтів і сервера, ідентифікатори користувачів, час сеансів та з'єднань, кількість переданих та прийнятих даних, дії адміністратора і користувачів тощо, та надають адміністраторам докладні звіти та оповіщають про певні події в режимі реального часу.

Кешування даних, запитуваних при доступі користувачів внутрішньої мережі до інформаційних ресурсів зовнішньої здійснюється шляхом мережі накопичування інформації на просторі жорсткого диску брандмауера (технологія проху-сервера). Тому, якщо при запиті потрібна кешована інформація, то вона надається без звертання до зовнішньої мережі, що істотно прискорює доступ.

При такій технології всі санкціоновані інформаційні ресурси зовнішньої мережі накопичуються й обновляються адміністратором на проху-сервері, а користувачам внутрішньої мережі дозволяється доступ тільки до кешованих інформаційних ресурсів.

Екрануючі агенти надійніші від звичайних фільтрів і забезпечують високий ступінь захисту, але вони знижують продуктивність обміну даними між мережами і не мають достатньої прозорості для додатків і кінцевих користувачів.

Особливості міжмережевого екранування. Брандмауери підтримують безпеку міжмережевої взаємодії на різних рівнях еталонної моделі OSI, де функції захисту істотно відрізняються. У цьому зв'язку комплексний брандмауер представляють у вигляді сукупності неподільних екранів орієнтованих на окремий рівень моделі OSI. Зазвичай комплексний екран функціонує на мережевому, сеансовому і прикладному рівнях еталонної моделі, тому відповідно розрізняють екрануючий маршрутизатор та шлюзи сеансового і прикладного рівня (рис. 10.35).

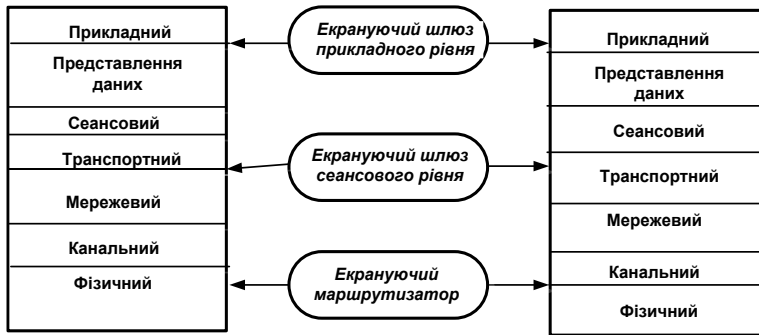


Рис. 10.35. Типи міжмережєвих екранів, що функціонують на окремих рівнях моделі OSI

Використовувані в мережах протоколи не завжди однозначно відповідають моделі OSI і тому зазначені екрани можуть відображати і сусідні рівні еталонної моделі, наприклад, шлюз прикладного рівня може зашифрувати повідомлення при їх передачі та розшифрувати прийняті дані, тоді він функціонує на рівні представлення даних і прикладному.

Міжмережєві екрани кожного з типів мають свої переваги і недоліки, але надійний захист забезпечують тільки комплексні системи, які поєднують всі види екранування.

Екрануючий шлюз прикладного рівня (пакетний фільтр) призначений для фільтрації пакетів повідомлень і забезпечує прозору взаємодію між внутрішньою і зовнішньою мережами. Він функціонує на мережевому рівні моделі OSI, а також може охоплювати і транспортний. Рішення про пропускання пакета приймається незалежно за заданими правилами фільтрації на основі аналізу пакетів мережевого і транспортного рівнів (рис. 10.36).

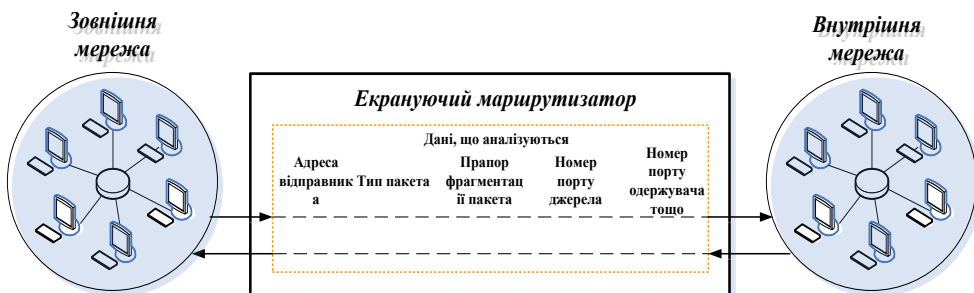


Рис. 10.36. Схема функціонування пакетного фільтра

Адреси відправника й одержувача можуть бути IP-адресами, які визначаються при формуванні пакета і залишаються незмінними при передачі

його через мережу.

Тип пакета містить код протоколу, що відповідає мережевому або транспортному рівню. Прапор фрагментації пакета визначає наявність або відсутність фрагментації. Номера портів джерела й одержувача однозначно ідентифікують додаток-відправник, а також додаток, для якого призначений цей пакет. Для можливості фільтрації пакетів за адресами портів необхідне знання прийнятих у мережі угод щодо виділення номерів портів протоколам високого рівня.

Екрануючі маршрутизатори не забезпечують високий ступінь безпеки, оскільки перевіряють тільки заголовки пакетів і не підтримують такі функції як аунтентифікація кінцевих вузлів, шифрування пакетів, перевірка їх цілісності тощо. Ці маршрутизатори уразливі для таких атак, як підробка вихідних адрес, несанкціонована модифікація вмісту пакетів повідомлень, а обхід таких екранів реалізується на основі формування заголовків пакетів, що задовольняють правилам фільтрації.

Екрануючий шлюз сеансового рівня призначений для контролю віртуальних з'єднань і трансляції адрес (наприклад, IP-адрес) при взаємодії з зовнішньою мережею. Він функціонує на сеансовому рівні моделі OSI, охоплює транспортний і мережевий рівні, а захисні механізми відносяться до функцій посередництва.

Контроль віртуальних з'єднань полягає в контролі квітування зв'язку та передачі інформації з установлених віртуальних каналів.

При контролі квітування здійснюється контроль за установленням віртуального з'єднання між вузлами внутрішньої і зовнішньої мережі шляхом визначення допустимості зв'язку, на основі інформації, що міститься в заголовках пакетів сеансового рівня, наприклад, протоколу TCP. Однак якщо пакетний фільтр при аналізі заголовків перевіряє тільки номери портів джерела й одержувача, то екрануючий шлюз сеансового рівня аналізує інші поля, що відносяться до процесу квітування зв'язку.

Після встановлення з'єднання з комп'ютером зовнішньої мережі шлюз діє від імені клієнта стежить за виконанням квітування зв'язку, наприклад, за протоколом TCP, а прикладом базового критерію фільтрації може бути можливість DNS-сервера визначити IP-адресу клієнта і асоційоване з нею ім'я.

Розглянемо процедуру квітування зв'язку на прикладі обміну TCP-пакетами, де прапором SYN позначається синхронізація, а ACK – підтвердження (рис. 10.37).

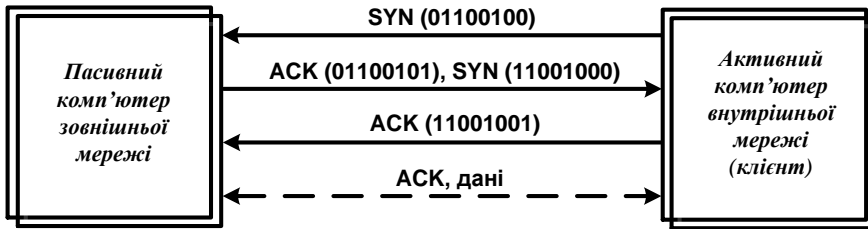


Рис. 10.37. Схема квітуння зв'язку за протоколом TCP

Перший пакет сеансу TCP, позначений флагом SYN і який вміщує довільне число, наприклад 01100100, є запитом клієнта на відкриття сеансу. Комп'ютер зовнішньої мережі, що отримав цей пакет, посилає відповідь, позначену флагом ACK і містить число, на одиницю більше, ніж у прийнятому пакеті (у нашому випадку 01100101), підтверджуючи, таким чином, прийом пакета SYN від клієнта. Крім того, здійснюючи зворотну процедуру, комп'ютер зовнішньої мережі посилає також клієнту пакет SYN, але вже з порядковим номером першого байта переданих даних (наприклад, 11001000), а клієнт підтверджує його одержання передачею пакета ACK, що містить число 11001001. На цьому процес квітуння зв'язку завершується.

Для екрануючого шлюзу сеансового рівня (рис. 10.38) сеанс вважається допустимим тоді, коли при квітунні зв'язку прапори SYN і ACK та числа заголовків TCP-пакетів виявляються логічно зв'язаними між собою. Після того як шлюз визначив, що комп'ютер внутрішньої і зовнішньої мережі є авторизованими учасниками сеансу TCP, і перевірів допустимість даного сеансу, то він встановлює з'єднання.



Рис. 10.38. Схема функціонування шлюзу сеансового рівня

Для контролю віртуальних з'єднань у таких шлюзах використовуються спеціальні програми (каналні посередники) встановлюють між внутрішньою та зовнішньою мережами віртуальні канали та контролюють передачу.

В основному екрануючі шлюзи сеансового рівня працюють в комплекті з екрануючими шлюзами прикладного рівня.

З погляду практичної реалізації шлюз сеансового рівня достатньо простою і надійною програмою і доповнює екрануючий маршрутизатор функціями контролю віртуальних з'єднань і трансляції внутрішніх адрес (наприклад, IP-адрес).

Недоліки шлюзу сеансового рівня майже ті самі, що і у екрануючого маршрутизатора тому він застосовується як доповнення до екрануючого шлюзу прикладного рівня.

Екрануючий шлюз прикладного рівня функціонує на прикладному рівні моделі OSI, охоплюючи також рівень представлення даних і забезпечує найбільш надійний захист міжмережевих взаємодій. Його захисні функції відносяться до функцій посередництва, але на відміну від екрануючого шлюзу сеансового рівня він виконує більшу кількість функцій захисту, до яких відносяться ті, що реалізуються екрануючими агентами (рис. 10.39) і використовуються по одному для кожного обслуговуваного прикладного протоколу.

Прикладний шлюз, так само, як і шлюз сеансового рівня, перехоплює за допомогою відповідних екрануючих агентів вхідні і вихідні пакети та копіює і перенаправляє інформацію через шлюз, і функціонує як сервер-посередника, крім прямих з'єднань між внутрішньою і зовнішньою мережею.

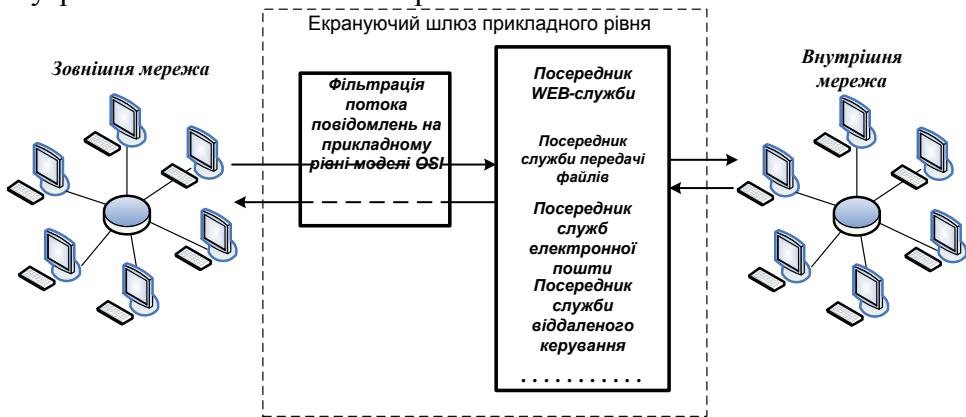


Рис. 10.39. Схема функціонування екрануючого шлюзу прикладного рівня

Посередники, що використовуються прикладним шлюзом, мають важливі відмінності від каналних посередників шлюзів сеансового рівня. Посередники прикладного шлюзу пов'язані з конкретними додатками, а також вони можуть фільтрувати потік повідомлень на прикладному рівні моделі OSI (рис. 10.40).

ЕКРАНЮЮЧИЙ ШЛЮЗ ПРИКЛАДНОГО РІВНЯ	
ПЕРЕВАГИ	НЕДОЛІКИ
<ul style="list-style-type: none"> ❶ За рахунок можливості виконання великої кількості функцій посередництва забезпечує найбільш високий рівень захисту локальної мережі; ❷ Захист на рівні додатків дозволяє здійснювати велику кількість додаткових перевірок, зменшуючи тим самим імовірність проведення успішних атак, заснованих на недоліках програмного забезпечення; ❸ При порушенні працездатності прикладного шлюзу блокується наскрізне проходження пакетів між поділюваними мережами, що не знижує безпеку захищеної мережі у випадку відмов 	<ul style="list-style-type: none"> ❶ Досить велика складність самого брандмауера, а також процедур його установки і конфігурування; ❷ Високі вимоги до продуктивності і ресурсоемності комп'ютерної платформи; ❸ Відсутність "прозорості" для користувачів і зниження пропускової здатності при реалізації міжмережєвих взаємодій

Рис. 10.40. Переваги і недоліки екранюючого шлюзу прикладного рівня

Прикладні шлюзи як посередників використовують програмні сервери конкретних служб (передача файлів, електронної пошти, віддаленого керування тощо), що функціонують на брандмауері в резидентному режимі і реалізують функції захисту.

Як і у випадку екранюючого шлюзу сеансового рівня, для зв'язку між комп'ютерами внутрішньої і зовнішньої мережі відповідний посередник прикладного шлюзу утворює два з'єднання: від комп'ютерами (внутрішньої мережі) до брандмауера і від брандмауера до місця призначення. Але, на відміну від каналних посередників, посередники прикладного шлюзу пропускають тільки пакети, сгенеровані тими додатками, які їм доручено обслуговувати. Наприклад, програма-посередник Web-служби може обробляти лише трафік, генерований цією службою. Якщо в мережі працює прикладний шлюз, то вхідні і вихідні пакети можуть передаватися лише для тих служб, для яких є відповідні посередники.

Під час налагоджування прикладного шлюзу й опису правил фільтрації повідомлень використовуються такі параметри, як назва сервісу, припустимий часовий діапазон його використання, обмеження на зміст повідомлень, зв'язаних з даним сервісом, комп'ютери, з яких можна користуватися сервісом, ідентифікатори користувачів, схеми

аунтентифікації тощо. На рис 9.43 показані основні переваги і недоліки екрануючого шлюзу прикладного рівня.

Для ефективного захисту міжмережевої взаємодії брандмауер повинен бути правильно встановлений і зконфігурований, для цього необхідно:

- розробити політику міжмережевої взаємодії;
- визначити схеми підключення міжмережевого екрана;
- налагодження параметрів функціонування брандмауера.

Розроблення політики міжмережевої взаємодії. Політика міжмережевої взаємодії є тією частиною політики безпеки в організації, що визначає вимоги до безпеки інформаційного обміну із зовнішнім оточенням. Дані вимоги обов'язково повинні відобразити два аспекти:

- політику доступу до мережевих сервісів;
- політику роботи міжмережевого екрана.

Політика доступу до мережевих сервісів визначає правила надання, а також використання всіх можливих сервісів захищеної комп'ютерної мережі. Відповідно в межах даної політики повинні бути задані всі сервіси, що надаються через міжмережевий екран, визначені допустимі адреси клієнтів для кожного сервісу, зазначені правила (описують, коли і які користувачі яким сервісом і на якому комп'ютері можуть скористатися) для користувачів визначаються правила аунтентифікації комп'ютерів і користувачів, а також умови роботи останніх поза локальною мережею.

Політика роботи міжмережевого екрана задає базовий принцип управління міжмережевим взаємодією, покладений в основу функціонування брандмауера. Може бути обраний один із двох таких принципів:

- заборонено усе, що не дозволено;
- дозволено усе, що не заборонено.

У першому випадку міжмережевий екран конфігурується так, щоб блокувати будь-які не дозволені міжмережеві взаємодії. Такий підхід дозволяє адекватно реалізувати принцип мінімізації привілеїв і з погляду безпеки він є кращим.

У другому випадку міжмережевий екран налагоджується таким чином, щоб блокувати тільки заборонені міжмережеві взаємодії. У цьому випадку підвищується зручність використання мережевих сервісів користувачами, але знижується безпека міжмережевої взаємодії, коли адміністратор може врахувати не всі дії, що заборонені користувачам.

Визначення схеми підключення міжмережевого екрана. Для підключення міжмережевих екранів використовуються різні схеми, що залежать від умов функціонування, а також кількості мережевих інтерфейсів брандмауера.

Брандмауери з одним мережевим інтерфейсом (рис. 10.41) не досить ефективні і з погляду безпеки, і з позицій зручності конфігурування. Фізично не розмежовується внутрішню і зовнішню мережі, а відповідно не можуть забезпечувати надійний захист міжмережевих взаємодій. Налагодження таких екранів, а також сполучення з ними маршрутизаторів досить складна задача, вартість вирішення якої перевищує вартість брандмауера з двома чи трьома мережевими інтерфейсами. Тому розглядатимемо лише дві останні схеми підключення, при цьому захищаємо локальну мережу представимо як сукупність закритої і відкритої підмереж, де відкрита підмережа є та, доступ до якої з боку зовнішньої мережі може бути цілком чи частково відкритим.

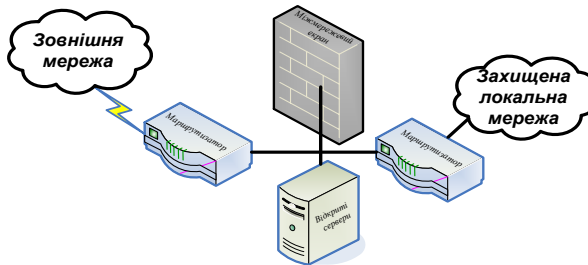


Рис. 10.41. Захист локальної мережі брандмауером з одним мережевим інтерфейсом

- Серед множини можливих схем підключення брандмауерів типовими є:
 - схема єдиного захисту локальної мережі;
 - схема із захищеною закритою і незахищеною відкритою підмережами;
 - схема з розподіленим захистом закритої і відкритої підмереж.

Схема єдиного захисту локальної мережі є найбільш простим рішенням (рис. 10.42), при якому брандмауер цілком екранує локальну мережу від зовнішньої, а між маршрутизатором і брандмауером мається тільки один шлях для трафіку.

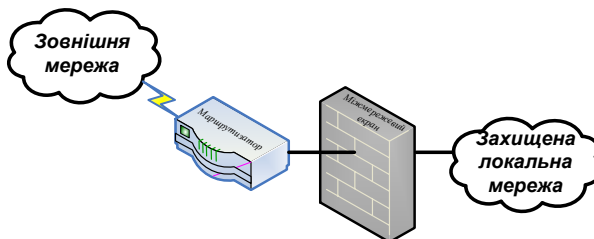


Рис. 10.42. Схема єдиного захисту локальної мережі

При цьому маршрутизатор налагоджується таким чином, що брандмауер є єдиним видимим зовні вузлом. Сервери, що входять у локальну мережу, також будуть захищені міжмережевим екраном, але об'єднання серверів, доступних із зовнішньої мережі, разом з іншими ресурсами захищеної локальної мережі, істотно знизить безпеку міжмережєвих взаємодій. Тому дану схему підключення брандмауера можна використовувати лише при відсутності в локальній мережі відкритих серверів або коли вони є доступними з зовнішньої мережі тільки для обмеженого числа довірених користувачів. За наявності у складі локальної мережі загальнодоступних відкритих серверів їх доцільно винести у відкриту підмережу до міжмережєвого екрана (рис. 10.43). Такий спосіб має більш високу захищеність закритої частини локальної мережі, але забезпечує знижену безпеку відкритих серверів. Деякі брандмауери дозволяють розмістити ці сервери на собі, але таке рішення не є кращим з погляду завантаження комп'ютера і безпеки самого брандмауера, тому таку схему доцільно використовувати лише при невисоких вимогах на безпеку до відкритої підмережі.



Рис. 10.43. Схема з захищеною і незахищеною відкритою під мережами

Якщо до безпеки відкритих серверів пред'являються підвищені вимоги, то необхідно використовувати схему з розподілим захистом закритої і відкритої підмереж. Така схема може бути побудована на основі одного брандмауера з трьома мережевими інтерфейсами (рис. 10.44) або двох

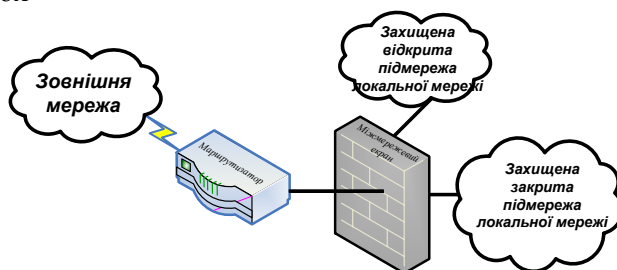


Рис. 10.44. Схема з розподілим захистом закритої і відкритої підмереж на основі одного брандмауера з трьома мережевими інтерфейсами

брандмауерів із двома мережевими інтерфейсами (рис. 10.45). В обох випадках доступ до відкритої і закритої підмереж локальної мережі можливий тільки через міжмережєвий екран, при цьому доступ до відкритої

Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу

підмережі не дозволяє здійснити доступ до закритої. З останніх двох схем великий ступінь безпеки міжмережєвих взаємодій забезпечує схема з двома брандмауерами, кожний з яких утворює окремий ешелон захисту закритої підмережі, а захищена відкрита підмережа тут виступає як екрануюча підмережа, що конфігурується таким чином, щоб забезпечити доступ до комп'ютерів підмережі як із зовнішньої мережі, так і з закритою локальною. Але прямий обмін інформаційними пакетами між зовнішньою мережею і закритою підмережею неможливий.

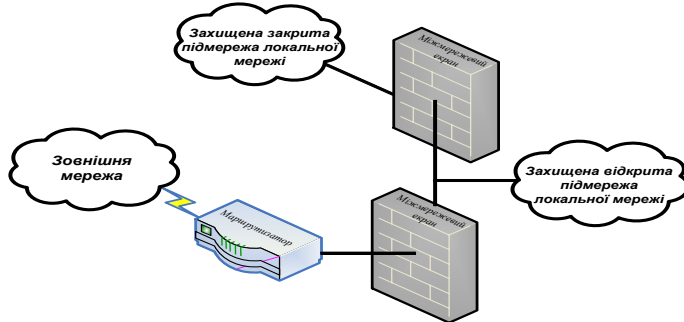


Рис. 10.45. Схема з роздільним захистом закритої і відкритої підмереж на основі двох брандмауерів з двома мережевими інтерфейсами

При атаці системи з екрануючою підмережею необхідно подолати, принаймні, дві незалежні лінії захисту. Засоби моніторингу стану міжмережєвих екранів як правило знайдуть подібну спробу атаки і адміністратор системи зможе вчасно почати необхідні дії щодо запобігання НСД.

Налагодження параметрів функціонування брандмауера.

Міжмережєвий екран є програмно-апаратним засобом захисту, що складається з комп'ютера, а також функціонуючих на ньому операційної системи і спеціального програмного забезпечення, яке часто також називають брандмауером.

Комп'ютер брандмауера повинен бути досить потужним і фізично захищеним (наприклад, знаходитися в спеціально відведеному охороняемому приміщенні) і мати засоби захисту від завантаження операційної системи з несанкціонованого носія.

Операційна система брандмауера також повинна задовольняти низці вимог:

- мати засоби розмежування доступу до ресурсів системи;
- блокувати доступ до комп'ютерних ресурсів в обхід існуючого програмного інтерфейсу;
- забороняти привілейований доступ до своїх ресурсів з локальної мережі;
- містити засоби моніторингу (аудита) будь-яких адміністративних дій.

Після встановлення брандмауера здійснюється налагодження параметрів, що включає наступні етапи:

формування правил роботи міжмережевого екрана відповідно до розробленої політики міжмережевої взаємодії і описання правил в інтерфейсі брандмауера;

перевірка заданих правил на несуперечність;

перевірка відповідності параметрів налагодження брандмауера розробленій політиці міжмережевої взаємодії.

Зформована на першому етапі база правил роботи міжмережевого екрана є формалізованим відображення розробленої політики міжмережевої взаємодії, а компонентами правил є об'єкти захисту, користувачі і сервіси.

У число таких об'єктів можуть входити комп'ютери з одним мережевим інтерфейсом, шлюзи (комп'ютери з декількома мережевими інтерфейсами), маршрутизатори, мережі тощо. Об'єкти, кожен з яких має низку атрибутів (мережева адреса, маска підмережі тощо) можуть об'єднуватися в групи. Частина таких атрибутів необхідно задати вручну, а впливають автоматично з інформаційних баз. Важливо звертати увагу на необхідність повного опису об'єктів для переконання в коректності заданих правил екранування, що можливе тільки тоді, коли визначені всі мережеві інтерфейси шлюзів і маршрутизаторів.

При описі правил роботи міжмережевого екрана користувачам надаються вхідні імена, які поєднуються в групи. Для користувачів зазначаються допустимі вихідні і цільові мережеві адреси, діапазон дат і часу роботи, а також схеми і порядок аунтентифікації.

Визначення набору і властивостей використовуваних сервісів здійснюється на основі вбудованої в дистрибутив брандмауера бази даних, а для нестандартних сервісів параметри можуть задаватися вручну за допомогою спеціальних атрибутів.

Після формування бази правил здійснюється її перевірка на несуперечність. Це дуже важливий момент, особливо для розвинених, багатокомпонентних мережевих конфігурацій із складною політикою міжмережевої взаємодії. Перевірка сформованих правил на несуперечність виконується автоматично, а виявлені неоднозначності усувають шляхом редагування суперечливих правил. Більшість брандмауерів після формування бази правил виконують процес остаточного налагодження автоматично.

Перевірка відповідності параметрів налагодження брандмауера розробленій політиці міжмережевої взаємодії може виконуватися на основі аналізу протоколів роботи міжмережевого екрана. Однак найбільша результативність такої перевірки буде досягнута при використанні спеціалізованих програмних засобів аналізу захищеності мережі (мережеві сканери).

Після пошуку слабких місць у конфігурації міжмережевого екрану надаються рекомендації щодо їх корекції. Пошук слабких місць здійснюється на основі перевірки реакції міжмережевих екранів на різні типи спроб порушення безпеки, при цьому виконується сканування всіх мережевих сервісів, доступ до яких здійснюється через міжмережевий екран, а для

постійної підтримки високого ступеня безпеки мережі сканери вмонтувуються в міжмережеві екрани.

При налагодженні міжмережевого екрана слід враховувати, що він не може захистити від некомпетентності адміністраторів і користувачів, наприклад, через вибір пароля, що легко вгадується, неконтрольований канал зв'язку тощо.

Загальні вимоги до міжмережевих екранів. Відповідно до вимог засобів захисту інформаційних систем міжмережеві екрани повинні задовольняти наступним групам вимог.

За цільовими якістьми – забезпечення безпеки внутрішньої захищеної мережі і повний контроль над зовнішніми підключеннями і сеансами зв'язку. Наявність засобів авторизації доступу користувачів через зовнішні підключення (приклад: типовою є ситуація, коли частина персоналу організації повинна виїжджати у відрядження, і в процесі роботи їм потрібен доступ до деяких ресурсів внутрішньої комп'ютерної мережі організації. Брандмауер повинен надійно розпізнавати таких користувачів і надавати їм необхідні види доступу).

За керованістю і гнучкістю – наявність потужних і гнучких засобів управління для повного впровадження політики безпеки організації. Забезпечення простої реконфігурації системи при зміні структури мережі. Якщо в організації є кілька зовнішніх підключень, у тому числі й у віддалених філіях, то система керування екранами повинна мати можливість централізовано забезпечувати для них проведення єдиної політики міжмережевої взаємодії.

За продуктивністю і прозорістю – достатньо ефективне функціонування і забезпечення в реальному часі обробку всього вхідного і вихідного трафіку при максимальному навантаженні. Це необхідно для того, щоб брандмауер не можна було перевантажити великою кількістю викликів, що призвели б до порушення його роботи. Міжмережевий екран повинен працювати непомітно для користувачів локальної мережі і не ускладнювати виконання ними легальних дій. У протилежному випадку користувачі будуть намагатися будь-якими способами обійти встановлені рівні захисту.

За самозахищеністю – наявність механізмів самозахисту від будь-яких несанкціонованих впливів. Оскільки міжмережевий екран є і ключем і дверима до конфіденційної інформації в організації, то він повинен блокувати будь-які спроби несанкціонованої зміни його параметрів налагодження, а також включати розвинуті засоби самоконтролю свого стану і сигналізації, які повинні забезпечувати своєчасне повідомлення служби безпеки при виявленні будь-яких

несанкціонованих дій та порушення працездатності міжмережевого екрана.

Основні висновки

Конфіденційність – характеристика безпеки ресурсів, що відображає їх властивість невиявленості й доступності без відповідних повноважень. Фактично, ресурси не можуть бути доступні або розкриті неавторизованій стороні, тобто для неї їх нібито немає. В свою чергу, авторська сторона (наприклад, обслуговуючий персонал, користувачі, програми і т. ін.), якій надано відповідні повноваження, має повний доступ до ресурсів.

Цілісність – характеристика безпеки ресурсів, що відображає їх властивість протистояти несанкціонованій зміні. Наприклад, користувач, що накопичує інформацію, має право очікувати, що вміст його файлів залишиться незмінним, незважаючи на цілеспрямовані впливи, відмови програмних або апаратних засобів. За цією характеристикою ресурси не зазнають будь-якої зміни з боку неавторизованої сторони.

Доступність – характеристика безпеки ресурсів, що відображає їх властивість, яка полягає в можливості їх використання у заданий момент часу відповідно до пред'явлених повноважень. Фактично авторська сторона, за потреби, відразу, в будь-який момент часу отримує необмежений доступ до необхідного ресурсу.

Активне перехоплення – перехоплення, під час якого у супротивника є можливість не тільки перехоплювати повідомлення, але й впливати на нього, наприклад, затримувати або вилучати сигнали, що передаються каналами зв'язку. Пасивне перехоплення – отримання інформації з можливістю тільки спостерігати за обміном повідомленнями (наприклад, з метою виявлення різної системної інформації в обчислювальній мережі (ОМ), не справляючи на нього ніякого впливу.

Пряме перехоплення (direct eavesdropping) – перехоплення інформації безпосереднім підключенням (наприклад, додаткового терміналу) до лінії зв'язку. Пряме перехоплення можна виявляти перевіркою лінії зв'язку. Непряме перехоплення (indirect eavesdropping) – перехоплення інформації (наприклад, індуктивних хвиль) без використання безпосереднього підключення до лінії зв'язку (threat). Пасивне перехоплення важко виявити, оскільки немає безпосереднього приєднання термінального обладнання до лінії зв'язку.

Апаратні засоби – різноманітні механічні, електричні, електромеханічні, електронні, електронно-механічні та інші пристрої і системи (наприклад, джерела безперебійного живлення, криптографічні обчислювачі та ЗВІС-процесори, електронні ідентифікатори та ключі, пристрої для виявлення жучків, генератори шумів тощо), що функціонують автономно або вбудовуються чи з'єднуються з іншою апаратурою з метою блокування дій дестабілізуючих чинників та розв'язання інших задач захисту інформації.

Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу

Програмні засоби – спеціальні програми (наприклад, антивіруси, шифрування даних, реалізації алгоритмів цифрового підпису, розмежування доступу, оцінки ризиків, визначення рівня безпеки, організації експертиз тощо), які функціонують в межах інформаційних систем для розв'язання задач захисту інформації.

Програмно-апаратні засоби – взаємопов'язані апаратні і програмні засоби (наприклад, банківські системи електронних платежів, комплексні інформаційні системи конфіденційного зв'язку, автоматизовані системи контролю доступу персоналу та транспортних засобів в режимних зонах тощо), які функціонують автономно чи в складі інших систем з метою вирішення задач захисту інформації.

Організаційні засоби захисту інформації – це множина процесів та дій (наприклад, контроль за утилізацією носіїв інформації з обмеженим доступом, планування заходів щодо відновлення втраченої інформації, аудит систем захисту, реалізація експертиз тощо), що здійснюються протягом усіх технологічних етапів (проекування, виготовлення, модифікація, експлуатація, утилізація тощо) існування відповідних ресурсів інформаційних систем і ведуть до створення, удосконалення, упорядкування і узгодженості взаємозв'язків та взаємодії їх компонентів з метою розв'язання задач захисту інформації.

Законодавчі засоби захисту інформації – є множиною нормативно-правових актів (конвенції, закони, укази, постанови, нормативні документи тощо), що діють у певній державі і забезпечують юридичну підтримку для розв'язання задач захисту інформації. Загалом, за допомогою законодавчих засобів визначаються права, обов'язки та відповідальність щодо правил взаємодії з інформацією, порушення яких може вплинути на стан її захищеності.

Руйнівний програмний вплив – це програмний код або його частини, за допомогою яких здійснюється загроза хоча б одній характеристиці безпеки певних ресурсів інформаційних систем. Руйнівні впливи можна поділити на такі групи: комп'ютерні віруси (віруси), логічні бомби, потайні ходи і лазівки; програми розкриття паролів, реплікатори, мережеві програмні аналізатори, суперзапінгові утиліти, троянські коні.

Вірус – програма, яка здатна до багаторазового самовільного створення свого тіла і зазвичай модифікує (заражає) інші програми, записані у файлах чи системних областях, для подальшого відтворення нового тіла та одержання керування з метою модифікації записів, знищення файлів, завантаження ресурсів і виконання інших руйнівних впливів у інформаційній системі.

Логічні бомби – програма, що ініціюється з виникненням різних подій, наприклад, відкриття певного файлу, обробка заданих записів та інші дії з метою порушення характеристик безпеки ресурсів інформаційних систем. Використовуються, наприклад, для розкрадання за допомогою змінювання

певним чином (на свою користь) коду програми, що реалізує фінансові операції.

Потайний хід – уразливість в системі, що навмисно створює її розробник або яка виникла випадково і фактично є додатковим способом проникнення в систему.

Програми розкриття паролів – програми зазвичай призначені для вгадування паролів (наприклад, архівованих файлів) через перебирання варіантів, можливих для використання символів або проникнення в систему за допомогою словників.

Реплікатори - програми, що при виконанні створюють декілька своїх копій в інформаційній системі. Наприклад, коли реплікатор створює тільки одну і після цього виконує її, то пам'ять системи швидко переповнюється, що обмежує доступ до певних компонентів системи.

Мережеві аналізатори – програмно-апаратні засоби (в окремих випадках програми, що запускається з робочої станції, підключеної до мережі), призначені для зчитування будь-яких параметрів потоку даних в інформаційній системі.

Суперзапінг – руйнівний вплив, що пов'язаний з несанкціонованим використанням утиліт для модифікації, знищення, копіювання, розкриття, вставлення, застосування або заборони застосування даних інформаційної системи.

Троянські коні – спеціалізована програма, яка, як правило, виступає від імені інших програм і дозволяє дії, відмінні від визначених в специфікації, що використовуються програмним забезпеченням.

Екрануючий шлюз прикладного рівня (пакетний фільтр) – призначений для фільтрації пакетів повідомлень і забезпечує прозору взаємодію між внутрішньою і зовнішньою мережами.

Екрануючий шлюз сеансового рівня – призначений для контролю віртуальних з'єднань і трансляції адрес (наприклад, IP-адрес) при взаємодії з зовнішньою мережею.

Політика роботи міжмережевого екрана – задає базовий принцип управління міжмережевим взаємодією, покладений в основу функціонування брандмауера.

Запитання для самоконтролю

1. *Як класифікуються НСД за автоматизацією?*
2. *Як класифікуються НСД за ініціалізаційною умовою?*
3. *Як реалізується НСД зі зворотнім зв'язком?*
4. *Чи можливо реалізувати монономний НСД з декількох джерел?*
5. *На чому заснований пігібекінговий НСД?*
6. *Що можна віднести до неспецифічних категорій НСД?*
7. *Як класифікуються засоби захисту від НСД?*
8. *Наведіть класифікацію комп'ютерних вірусів.*

9. *В чому полягає сутність емпіричного підходу?*
10. *Розкрийте зміст моделі системи з повним перекриттям.*
11. *Якими показниками може бути оцінено якість розподілу кодів доступу?*
12. *В чому полягає сутність теоретико-емпіричного підходу при побудові моделей захисту інформації?*
13. *Наведіть і розкрийте основні функції міжмережевого екранування.*
14. *Які є типи міжмережевих екранів?*
15. *Що потрібно для ефективного захисту міжмережевої взаємодії?*
16. *Розкрийте основну суть моделей Белла-Лападула і Біба.*

Conclusions

Privacy is security characteristic of resources that reflects their characteristic of inaccessibility without appropriate authority. In fact, resources cannot be available or disclosed to unauthorized parties, i.e. they do not exist for them. In turn, the author's side (for example, service staff, users, programs, etc.), which received an appropriate authority, has full access to resources.

Integrity is security characteristic of resources that reflects their feature to resist unauthorized change.

Availability is security characteristic of resources that reflects their feature to use them in a given moment of time according to granted authority. In fact, the author's side, when needed, immediately at any time gets unlimited access to necessary resource.

Active interception is interception, when enemy has a chance not only to intercept the message but also to affect it, for example to delay or remove signals that are transmitted by communication channels. Passive interception is access to information with the only option to observe connection exchange in order to identify system information in computing network without affecting it. Passive interception is difficult to identify because there is no direct connection of terminal equipment to communication link.

Direct interception is information interception by direct connection of additional terminal to communication line. Direct interception could be detected by communication line test. Indirect interception is information interception without use of direct connection to communication line.

Hardware is a variety of mechanical, electrical, electromechanical, electronic, electronic-mechanical and other devices and systems (UPS, cryptographic calculators, electronic identifiers and keys, "bugs" detecting devices, noise generators, etc.) functioning autonomously (embedded or linked with other equipment) in order to block the action of destabilizing factors and solving other tasks of information security.

Software – special programs (antiviruses, data encoding, digital signature algorithms implementation, access separation, risk estimation, security level evaluation, etc.) that function within information systems to meet information protection challenges.

Software-hardware means - interrelated hardware and software (banking electronic payment systems, complex information system of confidential communication, automated access control systems for personnel and vehicles in secure areas, etc.) that operate autonomously or as part of other systems to meet information security challenges.

Organizational means of information protection – is a set of processes and actions (control of limited access storage devices disposal, planning of recovery activities for lost information, audit of protection systems, expertise implementation, etc.), carried out throughout all technological stages (design, production, modification, maintenance, disposal, etc.) of existence of appropriate information systems resources and leading to creation, improvement, streamline and harmonization of linkages and interactions of its components in order to meet information security challenges.

Destructive software impact is program code or its part that threatens at least one characteristic of information systems certain resources security. Destructive effects could be divided into following groups: computer viruses (viruses), logic bombs, backdoors and loopholes; passwords disclosure programs, replicators, networking software analyzers, super-zapping utilities, trojans.

Virus is a program capable of creating its reusable arbitrary body that usually modify (infect) other programs, stored in files or system areas, for further reproduction of new body and getting control for records modification, files destruction, resources download and performing other destruction influences in information system.

Logic bomb is a program that is initiated with emergence of various events, such as particular file opening, specified records processing and other actions in order to violate characteristics of information systems resource security. They are used, for example, for, stealing by changing in certain way (in their favor) program code that implements financial transactions.

Backdoor is a system vulnerability that is deliberately created by developer or randomly emerged and in fact is an additional way to break into system.

Passwords disclosure program is a program designed for unraveling passwords (for example, archived files) through enumeration of variants (brute force method), all possible symbols or penetration into the system with the help of dictionaries.

Replicators – programs, which, while executing, create several self-copies in information system. When replicator creates self-copy and executes it, system memory becomes overloaded and it limits access to some system components.

Network analyzers – software-hardware means (in some cases – programs that are launching from workstation, connected to network), designed for reading any parameter of data stream in information system.

Super-zapping – devastating impact associated with unauthorized use of utilities for modification, deleting, copying, disclosure, insertion, use or prohibition of use of information system data.

Розділ 10. Захист вимірювальної інформації від несанкціонованого доступу

Trojan is a specialized program that usually acts on behalf of other programs and allows actions different from that determined in specification used by the software.

Shielding application level gateway (package filter) - a device designed for messages packages filtering, providing transparent interaction between internal and external networks.

Shielding circuit level gateway – a device designed for virtual connection control and address translation (for example, IP-address) while interacting with external network.

Firewall operational policy – policy that defines basic principle of networking interaction management, and is a basis for firewall functioning.

Ключові слова

Українською	Англійською
несанкціонований доступ	unauthorized access
комп'ютерна мережа	computer network
неавторизована сторона	unauthorized side
засоби захисту	security facilities



МЕТРОЛОГІЧНЕ ЗАБЕЗПЕЧЕННЯ ІВС

11

- 11.1. Зміст метрологічного забезпечення ІВС
- 11.2. Основні метрологічні характеристики ІВС
- 11.3. Методи визначення метрологічних характеристик ІВС

11.1. Зміст метрологічного забезпечення ІВС

Сучасний розвиток вимірювальної техніки, зокрема ІВС, ускладнення вимірювальних завдань і умов експлуатації засобів вимірювань, їх системне застосування висуває нові вимоги до описання властивостей засобів вимірювань. Прилади, що призначені для самостійного використання, для яких клас точності однозначно визначає комплекс нормованих метрологічних характеристик, практично непридатні для синтезу вимірювальних каналів ІВС. Комплекс метрологічних характеристик повинен вибиратися таким чином, щоб за деякою сукупністю метрологічних характеристик засобів вимірювальної техніки, засобів обчислювальної техніки і характеристик інших пристроїв, що утворюють вимірювальний канал, можна було визначити метрологічні характеристики всього вимірювального каналу. Інтелектуалізація ІВС, включення в їх склад обчислювального компоненту з метою опрацювання результатів вимірювань, оцінювання їх характеристик точності, управління процедурою вимірювань призводить до зростаючого значення метрологічного аспекту розроблення і використання алгоритмів і програм опрацювання даних.

Метрологічне забезпечення ІВС - система наукової, технічної, правової та організаційної діяльності, спрямована на досягнення єдності процесів перетворення інформації, що здійснюється в ІВС і необхідної точності результатів її функціонування.

Єдність процесів перетворення інформації - якість процесів перетворення інформації, при якому їх результати, визначені з використанням метрологічних характеристик інформаційних засобів, можна порівняти. Необхідною умовою порівнянності результатів є однаковість метрологічних характеристик.

Точність результату функціонування ІВС - якісна характеристика системи, що відображає близькість дійсного результату функціонування до істинного.

Метрологічна експертиза (складова метрологічного забезпечення ІВС) - аналіз та оцінювання оптимальних науково-технічних рішень, пов'язаних із забезпеченням єдності процесів перетворення інформації, що здійснюються в ІВС.

Розглянемо результати аналізу особливостей ІВС і пов'язані з цим проблеми їх метрологічного забезпечення (табл. 1).

Об'єктами метрологічного забезпечення ІВС є вимірювальні канали та їх складові - вимірювальні та обчислювальні компоненти.

Таблиця 1

Багатофункціональність	Забезпечення одночасного вимірювання ряду фізичних величин; побудова узагальнених оцінок на основі великого числа параметрів; обчислення комплексних параметрів
Наявність в складі ЕОМ	Оцінювання метрологічних характеристик алгоритмів і програм обробки результатів вимірювань
Багатоканальність	Оцінка, зменшення або виключення впливу каналів один на одного
Нерозривний зв'язок спеціалізованих ІВС з об'єктом, на якому вони експлуатуються	Проблеми проведення метрологічного обслуговування в умовах неможливості прив'язки до ідеалу шляхом переміщення ІВС до місця знаходження зразка. Неможливість комплектної повірки вимірювальних каналів за умовами установки датчиків на об'єкті
Складність опису об'єктів і їх моделювання	Складність обліку впливу об'єктів на точність вимірювання в умовах невизначеності апіорі інформації
Агрегатний спосіб побудови	Можливість дослідження ІВС як закінченого цілого тільки на об'єкті
Розподіленість компонентів і складових частин ІВС в просторі	Облік впливу на метрологічні характеристики різних умов експлуатації компонентів ІВС
Можливість зміни складу ІВС в процесі експлуатації	Складність регламентації вимог до систем на момент їх випуску
Наявність динамічних режимів вимірювання	Необхідність дослідження динамічних властивостей ІВС і узгодження їх з об'єктом

Основними завданнями метрологічного забезпечення ІВС є:
 аналіз і оцінювання технічних рішень на відповідність нормам точності і вибору параметрів, що підлягають вимірюванню, а також забезпеченості методами і засобами вимірювань;
 встановлення єдиного порядку передачі розміру одиниць фізичних величин ІВС;

Розділ 11. Метрологічне забезпечення ІВС

встановлення єдиних вимог до номенклатури, нормування та поданням метрологічних характеристик вимірювальних каналів ІВС;

визначення дійсних значень метрологічних характеристик вимірювальних каналів ІВС і їх відповідності встановленим нормам;

визначення похибок обчислювальних компонентів вимірювальних каналів ІВС і їх відповідності встановленим нормам;

встановлення повноти робіт з метрологічного забезпечення ІВС і відповідності вимогам нормативних документів з метрології;

встановлення вимог до метрологічного нагляду над ІВС і можливості його здійснення;

встановлення повноти робіт з метрологічного забезпечення ІВС і відповідності вимогам нормативної документації з метрології.

Метрологічне забезпечення вимірювальних компонентів вимірювальних каналів ІВС, що випускаються серійно, здійснюється відповідно до нормативної документації з метрології з урахуванням вимог, що стосуються експлуатаційної сумісності.

Державного метрологічного контролю і нагляду підлягають ІВС, які використовуються:

як робочі еталони для повірки засобів вимірювальної техніки (ЗВТ);

для проведення державних випробувань стосовно затвердження типу ЗВТ;

у сфері охорони здоров'я;

при виробництві і контролі продуктів харчування;

в ході здійснення заходів щодо захисту навколишнього природного середовища, охорони праці та техніки безпеки;

при геодезичних і гідрометеорологічних роботах;

при проведенні торгових операцій, включаючи всі види діяльності, пов'язані з визначенням кількості і якості товарів;

для надання послуг зв'язку;

для державного обліку матеріальних ресурсів (електричної і теплової енергії, газу, нафтопродуктів і т.д.);

для проведення податкових, митних та поштових операцій;

при проведенні випробувань і контроль якості продукції державними або іншими уповноваженими органами для визначення відповідності продукції обов'язковим вимогам державних стандартів;

при обов'язковій сертифікації продукції;

для судових, арбітражних, слідчих та інших експертиз, які проводяться від імені державних органів;

на транспорті для забезпечення безпеки руху;

для реєстрації національних і міжнародних рекордів.

Склад робіт з метрологічного забезпечення ІВС. Склад робіт залежить від стадії «життєвого циклу» ІВС (рис. 11.1). Метрологічне забезпечення ІВС здійснюється на всіх етапах їх життєвого циклу.

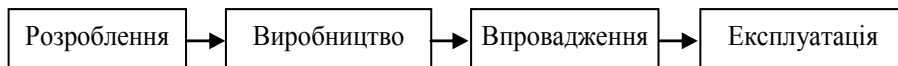


Рис. 11.1. Схема «життєвого циклу» ІВС

При розробленні ІВС виконують такі роботи з метрологічного забезпечення:

- метрологічну експертизу технічного завдання на розробку ІВС, конструкторської та програмної документації;

- розроблення ІВС, програм і методик приймальних випробувань, програм і методик метрологічної атестації ІВС, програм експериментальних досліджень ІВС, програм метрологічної атестації обчислювальних компонентів, методик перевірки ІВС;

- державні приймальні випробування ІВС;

- метрологічну атестацію ІВС одиничного виробництва.

В ході виробництва ІВС виконують:

- метрологічну експертизу технологічної документації на ІВС, призначених для серійного виробництва;

- первинну перевірку (калібрування) ІВС серійного виробництва.

На етапі впровадження ІВС виконують:

- метрологічну атестацію ІВС, монтаж і налагодження яких здійснюють безпосередньо на об'єкті експлуатації;

- метрологічну атестацію ІВС в умовах, які відрізняються від умов, для яких нормовані їх метрологічні характеристики;

- метрологічну атестацію обчислювальних компонентів, вимірювальних каналів ІВС.

Під час експлуатації ІВС виконують:

- метрологічний супровід обчислювальних компонентів вимірювальних каналів ІВС;

- періодичну перевірку (калібрування) ІВС.

Зміст робіт з метрологічного забезпечення ІВС. Основними роботами з метрологічного забезпечення ІВС є:

- встановлення єдиних вимог до метрологічних характеристик систем;

- розроблення методів і засобів контролю метрологічних характеристик систем;

Розділ 11. Метрологічне забезпечення ІВС

метрологічна експертиза технічної документації;
забезпечення єдності і достовірності результатів функціонування ІВС шляхом проведення випробувань з метою затвердження типу ІВС або їх одиничних екземплярів;

аналіз стану метрологічного забезпечення ІВС і розроблення на його основі комплексних програм розвитку метрологічного забезпечення систем;

організація і здійснення державного метрологічного контролю і нагляду за станом і застосуванням ІВС;

організація і проведення перевірочних та калібрувальних робіт;

організація і проведення робіт з атестації алгоритмів опрацювання інформації, що застосовуються при роботі ІВС.

Аналіз стану метрологічного забезпечення ІВС проводять з метою встановлення: можливості здійснення постійного контролю метрологічної справності систем, що знаходяться в експлуатації; відповідності розроблюваної документації ІВС, що виготовляються і знаходяться в експлуатації, вимогам нормативної документації і розроблення на цій основі заходів вдосконалення ІВС і їх метрологічного забезпечення.

Державний метрологічний контроль і нагляд, відомчий контроль встановлюють:

наявність нормативних документів, що регламентують вимоги до організації і порядку проведення розроблення, виробництва і експлуатації ІВС, а також вимоги до точності результату функціонування ІВС, порядок і правила перевірки ІВС;

ефективність проведення метрологічної експертизи конструкторської, технологічної, проектної документації;

правильність експлуатації ІВС і організації контролю їх стану;

наявність необхідних еталонів для здійснення перевірок;

правильність проведення перевірок ІВС в процесі експлуатації і дотримання міжповіркових інтервалів.

Мета метрологічної експертизи технічної документації - забезпечення ефективності метрологічного забезпечення, виконання загальних і конкретних вимог до метрологічного забезпечення найбільш раціональними методами і засобами. Конкретні цілі метрологічної експертизи визначаються призначенням і змістом технічної документації.

Метрологічна експертиза технічної документації на вітчизняні ІВС проводиться метрологічними службами організацій (підприємств), які розробляють, виготовляють і експлуатують ІВС, головними і базовими

організаціями метрологічної служби в галузях, а також органами державної метрологічної служби та державними науковими метрологічними центрами, в тому числі акредитованими як державних центрів випробувань засобів вимірювань в установленому порядку.

Основним змістом метрологічної експертизи технічного завдання (ТЗ) на розроблення (проектування) ІВС, що містить вихідні дані для розроблення, є перевірка достатності вихідних вимог, наведених в проекті ТЗ, для раціональної регламентації метрологічних характеристик вимірювальних каналів ІВС на етапі їх розроблення і побудови ефективного способу їх метрологічного забезпечення на наступних етапах життєвого циклу системи.

Основним змістом метрологічної експертизи технічних умов, конструкторської, технологічної, проектної та експлуатаційної документації є перевірка відповідності закладених в зазначеній документації комплексу метрологічних характеристик вимірювальних каналів та їх компонентів, методів і засобів їх визначення, контролю та (або) розрахунку, відповідності вихідним вимогам ТЗ, а також дотримання метрологічних правил, вимог і норм, що регламентуються в нормативній документації. Зокрема, перевіряють:

наявність в технічних умовах і експлуатаційної документації вичерпного переліку вимірювальних каналів і метрологічних вимог до них;

придатність до контролю конструкції ІВС;

наявність у проектній документації, призначеній для монтажу та наладки ІВС на об'єкті, вимог до параметрів і характеристик, необхідних для контролю якості монтажу ІВС;

наявність і зміст матеріалів (протоколів, актів, журналів, звітів і т.п.) попередніх випробувань, що стосуються метрологічних властивостей ІВС.

Атестація алгоритмів опрацювання інформації, що застосовуються при роботі ІВС проводиться для визначення, якою мірою алгоритм обчислень відповідає функції, що зв'язує вимірювану величину з результатами прямих вимірювань (зі значеннями величини на вході вимірювальних компонентів ІВС). Зазвичай виникають невідповідності викликані можливостями обчислювальної техніки і вимушеними спрощеннями алгоритмів обчислень (лінеаризацією функцій, їх дискретними уявленнями і т.п.).

Випробування для затвердження типу та саме затвердження типу проводяться для ІВС, що підлягають застосуванню в сферах поширення державного метрологічного контролю і нагляду. Випробування для добровільної "сертифікації відповідності" і

Розділ 11. Метрологічне забезпечення ІВС

"сертифікації відповідності" проводяться для ІВС, що не підлягають застосуванню в вищевказаних сферах. Якщо в сфері поширення державного метрологічного контролю і нагляду застосовується тільки частина із загальної кількості вимірювальних каналів ІВС, а інша частина - поза цією сферою, то випробуванням з метою затвердження типу ІВС підлягає тільки перша частина вимірювальних каналів.

Повірки і калібруванню піддаються вимірювальні канали ІВС, що застосовуються в сферах поширення державного метрологічного контролю і нагляду. Зміст робіт з повірки (калібрування) визначається документами на методику повірки ІВС.

При проведенні метрологічної атестації обчислювальних компонентів вимірювальних каналів ІВС до розгляду надають:

програмну документацію на обчислювальний компонент;

інформацію про діапазони зміни вхідних даних і їх похибки;

програму і методику метрологічної атестації обчислювального компоненту;

вихідний текст програми на електронних носіях;

відомості про технічні характеристики і параметри засобів обчислювальної техніки;

проект програми і методики метрологічного супроводу обчислювального компоненту.

Програма і методика метрологічної атестації обчислювальних компонентів може бути представлена окремим документом або складовою частиною програми метрологічної атестації ІВС.

Під час проведення метрологічної атестації обчислювальних компонентів визначають:

оцінки похибок обчислювальних компонентів і зіставлення їх з вимогами ТЗ на розробку;

придатність обчислювальних компонентів для використання в складі ІВС.

Під час метрологічного супроводу обчислювальних компонентів визначається:

оцінка похибок обчислювальних компонентів після модифікації і відповідність їх ТЗ завдання на розробку;

ступінь впливу модифікації обчислювальних компонентів на їх метрологічні властивості;

доцільність проведення метрологічної атестації обчислювальних компонентів після модифікації;

придатність обчислювального компонента для використання в складі ІВС після модифікації.

Метрологічний супровід обчислювальних компонентів проводять відповідно до програми і методики супроводу.

11.2. Основні метрологічні характеристики ІВС

Комплекси метрологічних характеристик ІВС регламентуються в нормативній, технічній та проектній документації, а саме: стандартах загальних технічних вимог (умов), стандартах технічних вимог (умов), технічних завданнях, технічних умовах, конструкторській, технічній, експлуатаційній та іншій документації.

Основні принципи регламентування та нормування метрологічних характеристик ІВС. Метрологічні характеристики відображають властивості засобів вимірювань (вимірювальних каналів), які мають визначальний вплив на результат і точність вимірювань.

Нормування метрологічних характеристик - це встановлення комплексу метрологічних характеристик і способів їх подання.

Загальні принципи нормування, застосовувані до ІВС:

можливість порівняння і вибору засобів вимірювання;

можливість визначення похибки (невизначеності) вимірюваної величини з урахуванням нормованих метрологічних характеристик практично реалізованими методами;

можливість визначення оцінки похибки (невизначеності) результатів вимірювань, проведених за допомогою даного засобу вимірювання.

Комплекс метрологічних характеристик повинен включати в себе характеристики, інваріантні до умов застосування і режиму роботи. Способи подання метрологічних характеристик повинні забезпечувати простоту їх контролю.

Особливості ІВС обумовлюють специфіку регламентації їх метрологічних властивостей, при цьому слід враховувати: особливості випуску та комплектації систем; умови їх експлуатації з урахуванням просторового розподілу компонентів; використання в складі систем складних обчислювальних пристроїв.

Розглянемо докладніше цей вплив.

1. Якщо ІВС випускається і комплектується як єдине ціле, то в нормативній документації встановлюються метрологічні характеристики вимірювального каналу в цілому і методи їх контролю. Якщо ІВС будується за агрегатним принципом, то регламентуються розрахункові метрологічні характеристики і методи розрахунку метрологічних характеристик вимірювального каналу системи за

Розділ 11. Метрологічне забезпечення ІВС

метрологічними характеристиками, які входять до їх складу компонентів, або нормують індивідуальні характеристики вимірювального каналу і методи їх експериментального визначення.

2. Якщо всі компоненти ІВС знаходяться в однакових зовнішніх умовах, то функції впливу величин та додаткові складові похибки (невизначеності), що виникають під їх впливом, нормуються як для звичайних засобів вимірювань. Якщо вони знаходяться в різних зовнішніх умовах, то нормування і визначення величин впливу є складним завданням, для вирішення якого вдаються до планування експерименту. Якщо функція впливу однієї величини впливу на метрологічні характеристики залежить від інших величин впливу, необхідно унормувати і визначати багатовимірні функції впливу і проводити багатofакторний експеримент.

3. Якщо використовуються аналогові обчислювальні пристрої, то їх розглядають як звичайний вимірювальний компонент системи і відповідним чином нормують його метрологічні характеристики. Якщо використовують складні обчислювальні пристрої (включаючи ЕОМ), то виникає завдання регламентації метрологічних характеристик алгоритмів і програм обчислень.

Основною структурною одиницею будь ІВС є вимірювальний канал - послідовний ланцюг вимірювальних компонентів поєднаних каналами зв'язку. Початком вимірювального каналу є або вхід системи, або вихід проміжного пристрою, від якого розходяться кілька таких послідовних ланцюгів; кінцем - вихід системи або вхід системи розгалуження.

Таким чином, щоб досить повно описати метрологічні властивості системи, необхідно унормувати метрологічні характеристики:

вимірювальних каналів системи;

вимірювальних компонентів системи;

аналогових обчислювальних компонентів;

алгоритмів обчислення, що реалізуються цифровими обчислювальними компонентами.

Для ІВС регламентуються метрологічні характеристики вимірювальних каналів в цілому, метрологічні характеристики компонентів, які входять в вимірювальні канали у вигляді нормованих значень, які визначені експериментально і (або) за результатами відповідних розрахунків. Регламентація метрологічних характеристик вимірювальних каналів ІВС не виключає нормування метрологічних характеристик їх компонентів, що забезпечує відповідність

метрологічних властивостей вимірювальних каналів регламентованим і взаємозамінність компонентів.

Для вимірювальних каналів ІВС, як правило, регламентуються такі характеристики, які можуть застосовуватися як окремо, так і в комбінаціях:

- а) нормовані метрологічні характеристики і методи їх контролю;
- б) значення метрологічних характеристик, одержані розрахунковим шляхом і методи їх розрахунку за характеристиками компонентів;
- в) перелік метрологічних характеристик, визначених експериментально, і методи їх експериментального визначення;
- г) перелік або значення розрахункових характеристик похибки вимірювань, які виконуються за допомогою вимірювальних каналів ІВС, і методи їх розрахунку за метрологічними характеристиками компонентів і значеннями показників умов експлуатації.

Для ІВС, до складу яких входить обчислювальний компонент, нормуються характеристики точності алгоритмів опрацювання результатів вхідних вимірювань для отримання кінцевих результатів вимірювань, які реалізуються обчислювальним компонентом відповідно до заданої програми. Метрологічні характеристики повинні врахувати вплив алгоритму опрацювання даних на методичні похибки вимірювань і трансформацію похибок вихідних вимірювань в процесі опрацювання.

Остаточний вибір доцільного способу регламентації метрологічних характеристик ІВС є складним багатоваріантним завданням і здійснюється на основі аналізу особливостей конкретних видів ІВС.

Метрологічні характеристики вимірювальних каналів ІВС, які нормуються або визначаються розрахунковим шляхом. Комплекс метрологічних характеристик для конкретних вимірювальних каналів ІВС, які нормуються або визначаються розрахунковим шляхом за нормованими характеристиками компонентів, вибираються із загального переліку, до якого входять такі метрологічні характеристики вимірювальних каналів ІВС:

- 1) номінальна статична характеристика перетворення вимірювального каналу ІВС, який закінчується вимірювальним перетворювачем або вимірювальним приладом зі шкалою, проградуєваною в одиницях, відмінних від одиниць вимірювання вхідного сигналу вимірювального каналу;

Розділ 11. Метрологічне забезпечення ІВС

2) вихідний код, число розрядів коду, номінальна ціна одиниці найменшого розряду коду вимірювального каналу ІВС, призначеного для видачі результатів у цифровому коді;

3) характеристики систематичної складової похибки вимірювального каналу ІВС (межа $\Delta_{\text{с.доп}}$ допустимого значення систематичної складової похибки вимірювального каналу ІВС). При необхідності допускається нормувати найбільшу допустиму зміну систематичної складової на заданому інтервалі часу або зміну в часі меж допустимої систематичної складової похибки. Також додатково до межі $\Delta_{\text{с.доп}}$ може бути унормовано середньоквадратичне відхилення $\sigma(\Delta_{\text{с}})$ систематичної складової похибки, яке розглядається як випадкова величина, властива безлічі вимірювальних каналів ІВС даного типу;

4) характеристики випадкової складової похибки вимірювального каналу ІВС:

а) межа допустимого значення середньоквадратичного відхилення випадкової складової похибки $\sigma_{\text{доп}}(\Delta)$ вимірювального каналу ІВС;

б) межа допустимого значення середньоквадратичного відхилення і номінальна нормалізована (виражена в частках дисперсії) автокореляційна функція $r_{\Delta}(\tau)$ або номінальна спектральна щільність $S_{\Delta}(\omega)$ і межі припустимих відхилень від наведених номінальних функцій випадкової складової похибки вимірювального каналу ІВС;

5) характеристика похибки Δ вимірювального каналу ІВС (межа $\Delta_{\text{доп}}$ допустимого значення похибки вимірювального каналу ІВС);

б) межа допустимого значення варіації вихідного сигналу $b_{\text{доп}}$ або показань вимірювального каналу ІВС, які закінчуються вимірювальним перетворювачем або приладом;

7) характеристики, що відображають взаємодію вимірювального каналу ІВС з об'єктом;

8) динамічні характеристики вимірювальних каналів ІВС:

а) номінальні функції зв'язку між змінними в часі вхідним і вихідним сигналами (передавальна функція, імпульсна вагова функція, перехідна характеристика, амплітудно-фазова характеристика і т.д.) і межі допустимих відхилень від номінальних функцій;

б) час встановлення показань або вихідного сигналу t_y , вимірювального каналу ІВС; (для конкретного вимірювального каналу

ІВС динамічна характеристика з наведеного списку вибирається з урахуванням практичної можливості її нормування).

9) номінальна функція впливу $\psi(\xi)$ і межі допустимого відхилення від номінальної функції;

10) найбільші допустимі зміни $\Delta L(\xi)$ метрологічних характеристик вимірювального каналу ІВС, викликані відхиленням зовнішніх факторів, що впливають і неінформативних параметрів вхідного сигналу від номінальних значень;

11) динамічні функції впливу вимірювального каналу ІВС. Динамічні функції впливу нормуються для ІВС, зі швидкими змінами впливають величин (наприклад, для ІВС, встановлених на рухомих об'єктах), що призводить до виникнення перехідних процесів в вимірювальних ланцюгах вимірювальних каналів ІВС, які суттєво впливають на похибку вимірювань. Для конкретного вимірювального каналу динамічна характеристика вибирається з урахуванням практичної можливості її нормування;

12) характеристики лінії зв'язку, якщо вона не входить до складу комплектуючих компонентів ІВС на момент випуску та з'являється як компонент системи тільки при монтажі на об'єкті.

Розглянемо деякі рекомендації, які слід враховувати при нормуванні метрологічних характеристик вимірювальних каналів ІВС.

Нормування метрологічних характеристик компонентів вимірювального каналу ІВС здійснюється з урахуванням тих же чинників, що і для вимірювальних каналів. Нормування метрологічних характеристик аналогових обчислювальних компонентів здійснюється аналогічно нормуванню метрологічних характеристик вимірювальних компонентів.

Найбільші допустимі межі, функція впливу або динамічна функція впливу нормуються окремо для кожної величини впливу; зазначені характеристики нормуються також для спільних впливів цих величин, якщо функція впливу кожної величини істотно залежить від інших величин впливу.

Якщо довжина вимірювального каналу ІВС така, що одна і та ж величина впливу для різних частин вимірювального каналу може одночасно прийняти істотно різні значення, то для такого вимірювального каналу нормується $\psi(\xi)$ або $\Delta L(\xi)$ для кожної з зазначених частин вимірювального каналу окремо і для різних його частин встановлюються різні нормальні і робочі умови.

Метрологічні характеристики вимірювальних каналів ІВС, які визначаються експериментально або розрахунково-експеримен-

Розділ 11. Метрологічне забезпечення ІВС

гальним шляхом. Комплекс метрологічних характеристик для конкретних вимірювальних каналів ІВС, які визначаються експериментально, або розрахунковим шляхом за метрологічними характеристиками компонентів, отриманих експериментально, вибирається із загального переліку метрологічних характеристик, до якого входять такі метрологічні характеристики вимірювальних каналів ІВС:

1) градувальна характеристика вимірювального каналу ІВС;
2) таблиці поправок вимірювального каналу для ІВС, які закінчуються показуючим або реєструючим вимірювальним приладом зі шкалою, проградуйованою в одиницях вхідного сигналу каналу;

3) характеристики систематичної складової похибки вимірювального каналу ІВС - верхня $\Theta_{\text{в}}$ і нижня $\Theta_{\text{н}}$ межі не виключеної систематичної складової похибки вимірювального каналу ІВС та ймовірність $P(\Theta_{\text{с}})$, з якою не виключена систематична складова похибки знаходиться в даних межах. При необхідності може визначатися найбільша зміна систематичної складової похибки на заданому інтервалі часу;

4) характеристики випадкової складової похибки вимірювального каналу ІВС:

а) середньо квадратичне відхилення $\sigma(\Delta)$, його довірчі межі й відповідна довірна ймовірність P_{σ} ;

б) середньо квадратичне відхилення $\sigma(\Delta)$, його довірчі межі й відповідна довірна ймовірність P_{σ} та нормалізовані автокореляційна функція $r_{\Delta}(\tau)$ або спектральна щільність $S_{\Delta}(\omega)$ випадкової складової похибки з зазначенням довірчих меж цих функцій і відповідних їм довірчих ймовірностей P_r або P_s ;

5) варіація h вихідного сигналу вимірювального каналу ІВС, який закінчується вимірювальним перетворювачем, або варіація показань вимірювального каналу, який закінчується вимірювальним приладом, довірчі межі її визначення і відповідна їм довірна ймовірність;

6) характеристики, які відображають взаємодію вимірювального каналу ІВС з об'єктом вимірювання і похибки їх експериментального визначення;

7) неінформативні параметри вихідного сигналу вимірювального каналу ІВС, який закінчується вимірювальним перетворювачем і похибка їх експериментального визначення;

9) динамічні характеристики вимірювального каналу ІВС:

а) функція зв'язку між змінними в часі вхідними та вихідними сигналами (передавальна функція, імпульсна вагова функція, перехідна характеристика, амплітудно-фазова характеристика і т.д.) і похибка її експериментального визначення;

б) час встановлення показань або вихідного сигналу вимірювального каналу і похибка його експериментального визначення; (для конкретного вимірювального каналу ІВС динамічна характеристика з наведеного вище переліку вибирається з урахуванням практичної можливості її експериментального визначення);

10) функція впливу вимірювального каналу ІВС і похибка її експериментального визначення;

11) динамічні функції впливу вимірювального каналу ІВС. Для конкретного вимірювального каналу динамічна характеристика вибирається з урахуванням практичної можливості її експериментального визначення.

Регламентування метрологічних характеристик компонентів проводиться з урахуванням тих же чинників, які передбачені і для вимірювальних каналів.

Основні принципи контролю та визначення метрологічних характеристик вимірювальних каналів ІВС. Контроль і визначення метрологічних характеристик вимірювальних каналів проводять з метою:

а) забезпечення якості розроблення та виготовлення ІВС в частині дотримання метрологічних вимог до них на підприємствах, які розробляють і виробляють ІВС;

б) перевірки якості монтажу та наладки ІВС (включаючи дотримання вимог до параметрів ліній зв'язку, які впливають на метрологічні характеристики вимірювальних каналів) в процесі і після установки ІВС на об'єкті; можливості проведення дослідної експлуатації і передачі в постійну експлуатацію в частині дотримання метрологічних вимог, закладених в проєкті;

в) перевірки та збереження метрологічних характеристик ІВС в процесі експлуатації.

Контроль і визначення метрологічних характеристик здійснюється експериментально «наскрізним методом», тобто шляхом подачі на вхід вимірювального каналу ІВС зразкового сигналу (імітує вимірювану величину) і зняття його вихідного сигналу (результату вимірювання). Отримані в результаті експерименту значення метрологічних характеристик служать безпосередньо як визначені або ж порівнюються з нормованими значеннями.

Розділ 11. Метрологічне забезпечення ІВС

Якщо експеримент проводився в конкретно реалізованих умовах застосування ІВС, відмінних від тих (зазвичай нормальних), для яких нормувалися або визначалися метрологічні характеристики вимірювального каналу, то отримані в процесі експерименту значення порівнюються з розрахованими значеннями тих же характеристик, обчислених для конкретно реалізованих умов за нормованими або визначеними метрологічними характеристиками компонентів, які входять до складу вимірювального каналу.

Необхідними умовами використання «наскрізного методу» визначення або контролю метрологічних характеристик є:

а) наявність доступу до входу вимірювального каналу (обмеження доступу може бути викликано конструкцією або способами установки сенсорів, температурними умовами або наявністю небезпечного середовища в місцях їх розташування і т.д.);

б) можливість завдання необхідного набору всіх значень величин впливу, характерних для умов експлуатації;

в) наявність еталонів і засобів завдання вимірюваних величин.

Значення метрологічних характеристик можуть бути отримані розрахунковим (розрахунково-експериментальним) методом. Він використовується в тих випадках, коли для вимірювального каналу не виконуються умови «наскрізного методу» контролю або визначення його метрологічних характеристик. Метод полягає в наступному:

а) у вимірювальному каналі виділяється та його частина, яка складається з компонентів (або їх сукупності) з нормованими або експериментально визначеними метрологічними характеристиками, для якої застосовується «наскрізний метод» контролю її метрологічних характеристик (доступна частина). Бажано, щоб в доступну частину вимірювального каналу входила максимальна кількість його компонентів, що дозволить дослідити метрологічні характеристики лінії зв'язку, що нормуються і функціональні перетворювачі, пристрої зв'язку з об'єктом, обчислювальні пристрої. Для сучасних ІВС як доступну частину доцільно використовувати вимірювальні канали інтегрованих компонентів, на основі яких створюються ІВС;

б) здійснюється контроль або визначення метрологічних характеристик доступної частини вимірювального каналу;

в) обчислюються значення метрологічних характеристик вимірювального каналу, значення в цілому за нормованими або експериментально визначеними метрологічними характеристиками доступної частини і нормованим або приписаним (за результатами раніше проведених експериментальних визначень метрологічних

характеристик) недоступної частини вимірювального каналу. Якщо недоступна частина вимірювального каналу вносить основний вклад в результуючу похибку вимірювального каналу, то даний спосіб стає практично чисто розрахунковим, оскільки не дивлячись на експериментальну перевірку доступної частини, її похибка становить незначну частину в сумарній похибці вимірювального каналу.

При визначенні метрологічних характеристик число експериментів (розмір вибірки) і максимально допустима похибка вимірювань в даній точці діапазону вимірювального каналу ІВС встановлюються в документації таким чином, щоб для експериментально визначених характеристик забезпечувалися довірчий інтервал і довірча ймовірність не гірші, ніж задані.

Характеристики точності алгоритмів опрацювання результатів вимірювань, що реалізуються обчислювальним компонентом, підлягають одноразовому контролю (визначенню) експериментальними, аналітичними або методами імітаційного моделювання. При контролі або визначенні метрологічних характеристик аналогових обчислювальних компонентів використовуються ті ж принципи, що і для вимірювальних компонентів ІВС.

11.3. Методи визначення метрологічних характеристик ІВС

У попередніх розділах було показано, що будь-яка ІВС складається з вимірювальних каналів, відповідно процедура визначення метрологічних характеристик ІВС зводиться до визначення метрологічних характеристик вимірювальних каналів. Методи визначення їх метрологічних характеристик (МХ) можна звести до трьох основних груп, які представлені на рис. 11.2.



Рис. 11.2. Методи визначення МХ вимірювальних каналів ІВС

Розділ 11. Метрологічне забезпечення ІВС

Кожен з наведених методів має свої специфічні вимоги до апріорної інформації про ІВС, складових процедури визначення метрологічних характеристик, умов проведення процедури тощо

Метод визначення метрологічних характеристик систем за метрологічними характеристиками компонентів вимірювальних каналів. Необхідність застосування розрахункових методів визначення метрологічних характеристик ІВС за метрологічними характеристикам компонентів обумовлена агрегатним принципом їх конструкції, тобто наявністю універсального ядра (інформаційно-обчислювального комплексу), на основі якого з використанням датчиків різних фізичних величин конструюють ІВС.

Оскільки розрахункові методи передбачають ідеалізацію властивостей системи і вимагають великого обсягу апріорної інформації, їх використання повинно бути обґрунтовано техніко-економічними причинами.

Методи поширюються на вимірювальні канали, що складаються з послідовно включених лінійних аналогових компонентів, а також на вимірювальні канали, що містять дискретні компоненти, впливом дискретності яких на похибку або невизначеність результатів вимірювань ІВС можна знехтувати.

Для правильного отримання вихідних даних для розрахунків у вигляді функціональних залежностей, які пов'язують метрологічні характеристики з вхідним (вихідним) сигналом, необхідно вибрати математичну модель компонента. Як правило, в нормативній документації відсутні дані, необхідні для побудови моделі. Тому при використанні розрахункових методів необхідно провести додаткові дослідження, пов'язані з побудовою моделі і перевіркою її адекватності.

У загальному випадку розрахунок номінальної функції перетворення, характеристик похибки (невизначеності) показань ІВС ґрунтується на послідовному приведення до виходу каналу складових похибки ІВС з подальшим їх підсумовуванням.

Розглянемо метрологічні характеристики, що підлягають розрахункам:

а) для нормальних умов експлуатації ІВС в статичному режимі розрахунку підлягають:

номінальні статичні характеристики перетворення вимірювальних каналів $f_n(x)$;

систематична складова похибки каналів $\Delta_{\text{сист.}}$;

випадкова складова похибки вимірювальних каналів $\overset{\circ}{\Delta}$;

похибки вимірювальних каналів Δ ;

б) для робочих умов експлуатації ІВС в статичному режимі розраховуються:

найбільші допустимі зміни метрологічних характеристик вимірювальних каналів викликаних змінами зовнішніх впливають величин і неінформативних параметрів вхідного сигналу $\Delta I(\xi)$;

функції впливу на метрологічні характеристики вимірювальних каналів $\psi(\xi)$;

в) для нормальних умов експлуатації вимірювальних каналів в стаціонарному динамічному режимі розраховуються:

амплітудно- і фазочастотні характеристики $A_n(\omega)$ і $\phi_n(\omega)$;

спектральна щільність випадкової складової похибки $S_{\overset{\circ}{\Delta}}(\omega)$.

Для *робочих умов* експлуатації в *стаціонарному динамічному* режимі розраховують функції впливу на метрологічні характеристики.

Як вихідні дані для розрахунку метрологічних характеристик вимірювальних каналів використовуються нормовані метрологічні характеристики компонентів, що входять до їх складу.

1. Для розрахунку метрологічних характеристик вимірювальних каналів в *статичному* режимі в *нормальних умовах* експлуатації використовуються наступні метрологічні характеристики компонентів, що складають вимірювальний канал:

$f_{ni}(x)$ – номінальна статична характеристика перетворення i -го компоненту в вигляді формули;

характеристики систематичної складової $\Delta_{\text{сист}}$ похибки:

$M[\Delta_{ci}(x)]$ – математичне очікування систематичної складової похибки компонентів у вигляді лінійної функції вхідного сигналу;

$\sigma[\Delta_{ci}(x)]$ – середньоквадратичне відхилення систематичної складової похибки компонентів даного типу у вигляді функції вхідного сигналу;

характеристики випадкової похибки $\overset{\circ}{\Delta}$ ($\sigma_{i\text{доп}}(\overset{\circ}{\Delta}) = \xi_i$ – межа допустимого значення середньоквадратичного відхилення випадкової складової похибки компонентів даного типу).

У розрахункових формулах, отриманих з урахуванням взаємного впливу каналів багатоканальних компонентів, які входять до складу ІВС, використовуються наступні метрологічні характеристики цих компонентів:

Розділ 11. Метрологічне забезпечення ІВС

$f_{n1}(x), f_{n2}(x), \dots$ – набір номінальних статичних характеристик перетворення для кожного вимірювального каналу кожного багатоканального компонента;

$M(\Delta_{c1}), M(\Delta_{c2}), \dots$ – набір математичних очікувань систематичних складових похибки;

$\sigma(\Delta_{c1}), \sigma(\Delta_{c2}), \dots$ – набір середньоквадратичних відхилень систематичних складових похибки для кожного вимірювального каналу кожного багатоканального компоненту.

2. Для розрахунку метрологічних характеристик каналів в статичному режимі в робочих умовах експлуатації на додаток до вищезазначених метрологічних характеристик використовуються найбільші допустимі зміни перерахованих в п. 1 метрологічних характеристик, викликані зміною зовнішніх величин впливу і неінформативних параметрів вхідного сигналу або їх функції впливу.

3. Вихідними даними для розрахунку метрологічних характеристик в *стаціонарному динамічному* режимі вимірювань в *нормальних* умовах експлуатації є:

$A_{ni}(\omega)$ – номінальна амплітудно-частотна характеристика (АЧХ) компонента, що входить до вимірювального каналу;

$\phi_{ni}(\omega)$ – номінальна фазочастотна характеристика (ФЧХ) компоненту;

$\Delta A_i(\omega)$ – найбільше допустиме відхилення АЧХ від номінального;

$\Delta \phi_i(\omega)$ – найбільше допустиме відхилення ФЧХ від номінального;

$S_{\Delta i}(\omega)$ – спектральна щільність випадкової складової похибки

даного типу.

4. У *стаціонарному динамічному* режимі вимірювання в *робочих* умовах експлуатації вихідними даними поряд з характеристиками, зазначеними в п. 3, є функції впливу компонентів, які входять в канал:

$\psi \Delta A_i(\xi_1, \xi_2, \dots)$ – функція впливу на АЧХ;

$\psi \Delta \phi_i(\xi_1, \xi_2, \dots)$ – функція впливу на ФЧХ;

$\psi S_{\Delta i}(\xi_1, \xi_2, \dots)$ – функція впливу на спектральну щільність

випадкової складової похибки.

Вихідні дані, необхідні для розрахунку метрологічних характеристик компонентів, отримують з нормативно-технічної документації, а в тих випадках, коли характеристики не вказані або представлені неповно, вони повинні бути визначені шляхом

експериментальних досліджень компонентів з використанням спеціальних методик.

Далі, для прикладу, розглянемо методику розрахунку метрологічних характеристик ІВС тільки в стаціонарному режимі і без урахування взаємного впливу вимірювальних каналів.

Розрахунок метрологічних характеристик вимірювальних каналів в статичному режимі при нормальних умовах експлуатації:

1. Номінальна статична характеристика вимірювального каналу може бути представлена як

$$f_H(x) = A^0 x + a, \quad (11.1)$$

де $A^k = \prod_{j=k+1}^N A_j$, $k=1,2,\dots,N-1$, $A^N = 1$, $A^0 = \prod_{i=1}^N A_i$, $a = \sum_{k=1}^N A^k a_k$; N –

кількість компонентів в каналі; A_i і a_i – відповідно мультиплікативна і адитивна складові номінального i -го компонента, номінальна статична характеристика якого задається в вигляді лінійних функцій вхідного сигналу.

Наприклад, для вимірювального каналу, що складається з трьох послідовно з'єднаних компонентів (рис. 12.3) адитивні і мультиплікативні складові яких позначимо відповідно, розрахунок номінальної функції проводиться в такий спосіб.

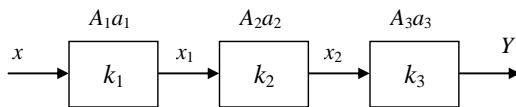


Рис. 11.3. Проста лінійна модель вимірювального каналу

Сигнал на виході компонента k_1 можна подавити як $x_1 = A_1 x + a_1$. Цей сигнал є вхідним для k_2 . вихідний сигнал якого можна отримати з виразу $x_2 = A_2 x_1 + a_2 = A_2 (A_1 x + a_1) + a_2 = A_1 A_2 x + A_2 a_1 + a_2$, тоді сигнал на виході компонента k_3 можна подати у вигляді $x_3 = A_3 x_2 + a_3 = A_3 (A_1 A_2 x + A_2 a_1 + a_2) + a_3 = A_1 A_2 A_3 x + A_2 A_3 a_1 + A_3 a_2 + a_3$, що відповідає формулі (11.1).

2. Розрахунок характеристик систематичної складової похибки вимірювального каналу у вигляді функцій вхідного сигналу.

В результаті розрахунків оцінюються $M[\Delta_{\text{сист}}(x)]$ – математичне очікування систематичної складової похибки і $\sigma[\Delta_{\text{сист}}(x)]$ – середньоквадратичне відхилення систематичної складової похибки вимірювального каналу даного типу.

Вихідними даними для розрахунку є:

N – кількість компонентів в каналі;

Розділ 11. Метрологічне забезпечення ІВС

$f_{ni}(x)$ – номінальна статична характеристика компонента даного типу у вигляді лінійної залежності від вхідного сигналу $f_{ni}(x) = A_i x + a_i$, ($i=1,2,\dots,N$);

$M(\Delta_{ci}(x)) = B_i x + b_i$ – математичне очікування систематичної складової похибки компонента даного типу у вигляді лінійної функції вхідного сигналу;

$\sigma(\Delta_{ci}(x))$ – середньоквадратичне відхилення систематичної складової похибки компонентів даного типу (для всіх компонентів), ($i=1,2,\dots,N$). Для розрахунку $\sigma[\Delta_{ci}(x)]$ задається в вигляді функції вхідного сигналу $\sigma(\Delta_{ci}(x)) = \sqrt{(G_i^2 x^2 + g_i^2)}$, де G_i^2 і g_i^2 – дисперсії мультиплікативної та адитивної складових систематичної похибки відповідно.

За вихідними даними обчислюються допоміжні величини:

$$B^k = \prod_{j=k+1}^N (B_j + A_j), \quad k=0,1,\dots,N-1; \quad (11.2)$$

$$Q_i^2 = \prod_{j=i+1}^N ((B_j + A_j)^2 + G_j^2), \quad i=0,1,\dots,N-1; \quad (11.3)$$

$$B^N = A^N = Q_N^2 = 1; \quad A^k = \prod_{j=k+1}^N A_j, \quad k=0,1,\dots,N-1;$$

$$b = \sum_{i=1}^N B^i (b_i + a_i) - \sum_{i=1}^N A^i a_i; \quad (11.4)$$

$$B = B^0 - A^0. \quad (11.5)$$

Підставляючи їх в наступні формули, знаходять оцінки характеристик систематичної складової похибки вимірювального каналу:

$$M[\Delta_{\text{сист}}(x)] = Bx + b, \quad (11.6)$$

$$\sigma[\Delta_{\text{сист}}(x)] = \sqrt{x^2 (Q_0^2 - (B^0)^2) + \sum_{i=1}^N (b_i^2 (Q_i^2 - (B^i)^2) + g_i^2 Q_i^2)}. \quad (11.7)$$

3. Розрахунок межі допустимого значення систематичної складової похибки вимірювального каналу даного типу.

Вихідними даними для розрахунку є:

N – кількість компонентів, які входять до каналу;

$[0\dots q]$ – діапазон зміни вхідного сигналу;

$f_{ni}(x)$ – номінальна статична характеристика компоненти даного типу у вигляді лінійної залежності від вхідного сигналу $f_{ni}(x) = A_i x + a_i$,

$(i=1,2\dots N)$;;

$\Delta_{\text{с.доп}i}$ – межа допустимого значення систематичної складової похибки компонента даного типу, що входить в вимірювальний канал.

За початковими даними для обчислення допоміжних величин позначимо $\lambda_i = \Delta_{\text{с.доп}i}$, $q_1 = q$, тоді:

$$\begin{aligned} q_i &= q_{i-1}A_{i-1}, i=2,3,\dots,N-1; \\ B^N &= 1; B^{i-1} = B^i(A_i + 2\lambda_i/q_i), i=2,3,\dots,N-1. \\ \Delta_{\text{с.доп.}} &= \sum_{i=1}^N B^i \lambda_i. \end{aligned} \quad (11.8)$$

4. Розрахунок межі допустимого значення середньоквадратичного відхилення випадкової складової похибки вимірювального каналу даного типу.

Вихідними даними для розрахунку є:

N – кількість компонентів, які входять в ізмерительний канал;

$[0\dots q]$ – діапазон зміни вхідного сигналу;

$f_{ni}(x)$ – номінальна статична характеристика компонента даного типу у вигляді залежності від вхідного сигналу $f_{ni}(x) = A_i x$, $(i=1,2\dots N)$;

$\Delta_{\text{с.доп}i}$ – межа допустимого значення систематичної складової похибки компонента даного типу;

$\sigma_{\text{доп.}i}(\Delta)$ – межа допустимого значення середньоквадратичного відхилення випадкової складової похибки компонента даного типу.

Для обчислення $\sigma_{\text{доп.}}(\Delta)$ позначимо: $q_1 = q$, $\lambda_i = \Delta_{\text{с.доп}i}$, $\chi_i^2 = \sigma_{\text{доп.}i}^2(\Delta)$ при $i=1,2,\dots,N$. Обчислюються такі допоміжні величини:

$$q_i = q_{i-1}A_{i-1}, i=2,3\dots,N; B^N = S_N = 1; \quad (11.9)$$

$$B^{i-1} = B^i \left(A_i + \frac{2\lambda_i}{q_i} \right), i=1,2\dots,N; \quad (11.10)$$

$$S_{i-1}^2 = S_i^2 \left(\left(A_i + \frac{2\lambda_i}{q_i} \right)^2 + \frac{2\lambda_i^2}{q_i} + \frac{4\chi_i^2}{q_i^2} \right) i=1,2\dots,N; \quad (11.11)$$

$$C_{ij+1} = 1; C_{ij} = \prod_{k=i+1}^{j-1} \left(A_k + \frac{2\lambda_k}{q_k} \right), j > i+1. \quad (11.12)$$

Межа допустимого значення середньоквадратичного відхилення випадкової складової похибки вимірювального каналу

$$\sigma_{\text{доп}}(\Delta) = \sqrt{\sum_{i=1}^N \left(\chi_i^2 S_i^2 + \lambda_i^2 \left(S_i^2 - (B^i)^2 \right) \right)} + 2 \sum_{i=1}^{N-1} \lambda_i \sum_{j=i+1}^N S_j^2 \lambda_j C_{ij} \frac{\chi_j^2}{q_j^2}. \quad (15.13)$$

5. Для розрахунку межі допустимого значення похибки $\Delta_{\text{доп}}$ вимірювального каналу в якості вихідних даних використовують такі характеристики:

N – кількість компонентів, які входять в канал;

$[0...q]$ – діапазон зміни вхідного сигналу;

$f_{\text{ни}}(x)$ – номінальну статичну характеристику (для всіх компонентів, які входять в вимірювальний канал): $f_{\text{ни}}(x) = A_i x$;

$\Delta_{\text{доп}i}$ – межу допустимого значення похибки компонентів даного типу (для всіх компонентів, які входять в вимірювальний канал).

Примітка. Під межею допустимої похибки вимірювального каналу розуміється найбільше значення величини $[\Delta_{\text{сист}} + \gamma \sigma(\Delta)]$ в діапазоні зміни вхідного сигналу (γ - постійна величина, яка залежить від закону розподілу ймовірностей похибки).

Введемо позначення

$$\lambda_i = \alpha_i \delta_i, \quad \chi_i^2 = (1 - \alpha_i)^2 \frac{\delta_i^2}{\gamma^2}, \quad i = 1, 2, \dots, N, \quad (11.14)$$

де α_i – параметр перебору, який при кожному $i = 1, 2, \dots, N$ приймає значення 0 або 1.

Розрахунок $\Delta_{\text{доп}}$ відповідає завданню пошуку $\max V(\alpha_1, \dots, \alpha_N)$ за умови, що α_i може приймати одне з двох значень $\alpha_i = 1$ або $\alpha_i = 0$.

За формулами (11.9) – (11.12) обчислюються допоміжні величини B_i, S_i^2, C_{ij}, q_i . Введемо функцію $V(\alpha_1, \alpha_2, \dots, \alpha_N)$ від величини α_i :

$$V(a_1, a_2, \dots, a_N) = \sum_{i=1}^N B^i \lambda_i + \gamma \sqrt{\sum_{i=1}^N \left(\chi_i^2 S_i^2 + \lambda_i^2 \left(S_i^2 - (B^i)^2 \right) \right)} + 2 \sum_{i=1}^N \lambda_i \sum_{j=i+1}^N S_j^2 \lambda_j C_{ij} \frac{\chi_j^2}{q_j^2} \quad (11.15)$$

Наприклад, при $N = 3$ необхідно обчислити вісім значень $V(0,0,0), V(0,0,1), V(0,1,0), V(1,0,0), V(1,0,1), V(0,1,1), V(1,1,0), V(1,1,1)$ і вибрати з них найбільше.

Розрахунок метрологічних характеристик вимірювальних каналів ІВС в статичному режимі в робочих умовах експлуатації:

1. Розрахунок функцій впливу на характеристики систематичної складової похибки каналів дає можливість отримати:

$\Psi_{M(\Delta_{\text{сист}})}(\xi_1, \xi_2, \dots)$ – функцію впливу на $M(\Delta_{\text{сист}})$ каналу при спільному зміні величин впливу;

$\Psi_{\sigma(\Delta_{\text{сист}})}(\xi_1, \xi_2, \dots)$ – функцію впливу на $\sigma(\Delta_{\text{сист}})$ каналу при спільному зміні величин впливу.

У набір величин впливу $(\xi_1, \xi_2, \dots, \xi_p)$ входять величини для кожного компоненту вимірювального каналу. Взаємний вплив вимірювальних каналів не враховується. Для кожного компонента, що входить до складу вимірювального каналу, задаються $f_{ni}(x)$, $\mathbf{M}(\Delta_{ci}(x))$, $\sigma(\Delta_{ci}(x))$.

Функція впливу на при спільному зміні впливають величин визначається як $\Psi_{M(\Delta_{ci})} = a_{1i}(\xi_1, \xi_2, \dots, \xi_p) + a_{2i}(\xi_1, \xi_2, \dots, \xi_p)x$, де $\beta_{1i}(\xi_1, \xi_2, \dots, \xi_p)$, $\beta_{2i}(\xi_1, \xi_2, \dots, \xi_p)$ функції впливу.

Обчислення проводяться за наступним алгоритмом.

Обчислюються допоміжні величини:

$$B_{i\xi} = B_i + \alpha_{2i}(\xi_1, \dots, \xi_p); \quad b_{i\xi} = b_i + \alpha_{1i}(\xi_1, \dots, \xi_p);$$

$$G_{i\xi}^2 = G_i^2 + \frac{1}{2}\beta_{2i}^2(\xi_1, \dots, \xi_p); \quad g_{i\xi}^2 = g_i^2 + \frac{1}{2}\beta_{1i}^2(\xi_1, \dots, \xi_p).$$

Підставляючи їх у формули (11.2) і (11.3), обчислюють $B_{i\xi}^i$, $Q_{i\xi}^2$ для $i = 1, 2, \dots, N$.

Ці величини використовують для обчислень $\mathbf{M}[\Delta_{\text{сист}}(x)]$ і $\sigma[\Delta_{\text{сист}}(x)]$, підставляючи їх у формули (11.4) – (11.7) замість B , Q_i^2 , b . За формулами (11.2) – (11.9) визначаються $\sigma[\Delta_{\text{сист}}(x)]$ і $\mathbf{M}[\Delta_{\text{сист}}(x)]$. І як результат, обчислюються шукані функції впливу:

$$\Psi_{M(\Delta_{\text{сист}})}(\xi_1, \xi_2, \dots, \xi_p) = \mathbf{M}[\Delta_{\text{сист}}(x)]_{\xi} - \mathbf{M}[\Delta_{\text{сист}}(x)],$$

$$\Psi_{\sigma(\Delta_{\text{сист}})}(\xi_1, \xi_2, \dots, \xi_p) = \sigma[\Delta_{\text{сист}}(x)]_{\xi} - \sigma[\Delta_{\text{сист}}(x)].$$

3. Розрахунок найбільшої допустимої зміни $\sigma_{\text{доп}}(\overset{\circ}{\Delta})$ вимірювального каналу. Результат розрахунку – значення середньоквадратичного відхилення $\sigma_{\text{доп}}(\overset{\circ}{\Delta})(\xi_1, \dots, \xi_p)$. Вихідними даними для розрахунку є:

N – кількість компонентів у вимірювальному каналі;

$[0 \dots q]$ – діапазон зміни параметра вхідного сигналу.

Для кожного компонента, що входить до вимірювальний канал,

задаються $f_{ni}(x) = A_i x$; $\Delta_{\text{с доп } i} = \lambda_i$; $\sigma_{\text{доп } i}(\overset{\circ}{\Delta}) = \Theta_i$;

$$\Delta(\Delta_{\text{с доп } i}(\xi_1, \xi_2, \dots, \xi_p)) = K_i \lambda_i,$$

де p – кількість впливають величин; $\Delta\sigma_{\text{доп } i}(\Delta)(\xi_1, \xi_2, \dots)$ – найбільша допустима зміна $\sigma_{\text{доп } i}(\Delta)$ при спільній зміні величин впливу $\Delta\sigma_{\text{доп } i}(\Delta)(\xi_1, \xi_2, \dots, \xi_p) = l_i \Theta_i$; $l_i > 0$.

Обчислюють величини $\lambda_{i\xi} = (1 + k_i)\lambda_i$, $\chi_{i\xi} = (1 + l_i)\chi_i$.

Використовуючи $\lambda_{i\xi}$, $\chi_{i\xi}$ замість λ_i , χ_i , за формулами (11.9) - (11.13) обчислюється $\sigma_{\text{доп}}(\Delta)_\xi$. Шукану характеристику знаходять як $\Delta\sigma_{\text{доп}}(\Delta)(\xi_1, \xi_2, \dots, \xi_p) = l\sigma_{\text{доп}}(\Delta)$, де $l = (\sigma_{\text{доп}}(\Delta)_\xi - \sigma_{\text{доп}}(\Delta)) / \sigma_{\text{доп}}(\Delta)$.

4. Розрахунок найбільшої допустимої зміни $\Delta_{\text{доп}}$ вимірювального каналу при спільній зміні величин впливу $\Delta(\Delta_{\text{доп}}(\xi_1, \xi_2, \dots, \xi_p))$.

Результатом розрахунку є $\Delta(\Delta_{\text{доп}})$, виражене в одиницях $\Delta_{\text{доп}}$.

Далі для кожного компонента, що входить до складу вимірювального каналу, задають $f_{ni}(x) = A_i x$; кількість величин впливу p ; $\Delta_{\text{доп } i} = \delta_i$.

Обчислюються допоміжні величини $\delta_{i\xi} = (1 + r_i)\delta_i$. За формулами (11.14), (11.15) обчислюють $\Delta_{\text{доп}}$. Використовуючи $\delta_{i\xi}$ замість δ_i , визначають δ_i за тими ж формулами. Шукана характеристика $\Delta(\Delta_{\text{доп}}) = r\Delta_{\text{доп}}$, де $r = (\Delta_{\text{доп } \xi} - \Delta_{\text{доп}}) / \Delta_{\text{доп}}$.

Методи експериментального визначення характеристик похибки вимірювальних каналів ІВС. *Загальні відомості.* Методи експериментального визначення характеристик похибки вимірювальних каналів, як правило, однакові для всіх точок діапазону вимірювання, тому будемо розглядати їх для однієї точки діапазону. Для спрощення будемо вважати, що вимірювальний канал є простим (тобто здійснює пряме вимірювання величини або її носія). Вибір кількості точок в діапазоні вимірювань проводиться з міркувань лінійності номінальної характеристики.

Для ІВС з лінійною номінальною статичною характеристикою:

якщо нульове значення вимірюваної величини розташоване в середині діапазону вимірювання ІВС, то число експериментальних точок приймається 11 (по п'ять точок на позитивній та негативній частинах діапазону вимірювання і одна точка поблизу нульового значення);

якщо нульове значення вимірюваної величини розташоване на краю діапазону вимірювання ІВС, то число експериментальних точок приймається рівним п'яти.

Для ІВС з нелінійною статичною характеристикою число експериментальних точок в діапазоні вимірювання ІВС вибирається не меншим, ніж для ІВС з лінійною характеристикою, а їх розташування може бути нерівномірним.

До експериментальних включають точки, в яких очікуються найбільші значення характеристик похибки, а також точки, які відповідають нижньому і верхньому межам діапазону вимірювання.

Ці точки вибираються з урахуванням характеру не лінійності.

Наприклад:

для аналогових ІВС при використанні в них лінійної або кусково-лінійної апроксимації нелінійних залежностей вхідного сигналу ІВС від вимірюваної величини доцільно вибирати значення вимірюваної величини (тобто експериментальні точки) в інтервалах з найбільшою нелінійністю;

для аналого-цифрових ІВС, в які входять АЦП, що мають суттєві диференціальну і інтегральну нелінійності, до складу експериментальних включаються "критичні" точки (в яких відбувається втрата або деформація окремих кодів).

Вказані методи визначення характеристик похибки базуються на тому, що похибка вимірювального каналу в точці, яка перевіряється, складається з систематичної і випадкової складових, а варіація така мала, що нею можна знехтувати. При цьому характеристиками похибки є: значення систематичної складової похибки; середньоквадратичне відхилення випадкової складової похибки вимірювання; межі інтервалу, в якому з достовірністю P лежить похибка.

Вихідними даними для знаходження їх точкових та інтервальних оцінок використовуються отримані експериментально вибіркові значення похибки D_i , де $i=1,2,\dots,n$, які опрацьовуються відповідно наведених нижче розрахункових методів.

При визначенні характеристик похибки простих вимірювальних каналів на їх вхід подають від сигнали від еталону або необхідні значення прямо вимірюваної величини, або необхідні значення носія прямо вимірюваної величини, відповідно точкам перевірки. Сигнали, які подають на вхід і отримують на виході вимірювального каналу, виражають в одиницях прямо вимірюваної величини або відповідного її носія.

При практичному використанні описаних нижче методів визначення похибки може знадобитися перерахунок характеристик сигналів, виражених в значеннях прямо вимірюваної величини, в значенні носія або поданого на вхід каналу або одержаного на його виході. Для перерахунку повинні використовуватися номінальні прямі і зворотні функції сенсорів і вимірювального каналу в цілому, нормовані у відповідній документації.

Розділ 11. Метрологічне забезпечення ІВС

Під номінальною прямою функцією $F_{\text{ном}}$ розуміється номінальна залежність значення Y показання або носія вимірювального каналу від значення X прямо вимірюваної величини або вхідного носія $Y = F_{\text{ном}}(X)$, в окремому випадку можливо $Y = X$.

Під номінальною оберненою функцією $F_{\text{ном}}^{-1}$ розуміється номінальна залежність значення X вхідної величини або вхідного носія від значення Y показання або носія $Y = F_{\text{ном}}^{-1}(X)$, в окремому випадку можливо $X = Y$. Очевидно, що повинна виконуватися умова $Y = F_{\text{ном}}^{-1}(F_{\text{ном}}(X))$.

Для розрахунку інтервальних оцінок похибки або її характеристик при наявності істотної випадкової складової похибки рекомендується використовувати сучасний робастний метод опрацювання спостережень. Цей метод не заперечує використання традиційних методів опрацювання, заснованих на припущенні про нормальність закону розподілу випадкової складової похибки і є одним з можливих окремих випадків робастного методу lp -оцінок (коли $p = 2$). Вказівки про допустимість використання традиційних методів даються при викладі методик визначення похибки або її характеристик.

Вибір методу визначення похибки. Вибір експериментального методу визначення похибки або її характеристик залежить від багатьох факторів, які визначають постановку і проведення експерименту. До основних факторів, які впливають на вибір методу, відносяться:

наявність або відсутність апріорної інформації про склад і властивості похибок вимірювального каналу;

мета експерименту;

вид і склад вимірювального каналу.

Відповідно складу і властивостей похибки вимірювальні канали підрозділяють на дві групи: з істотною і дуже малою, якою нехтують, випадковими складовими похибки.

Метою експерименту може бути визначення або контроль характеристик похибки вимірювального каналу. У процесі визначення знаходяться точкові і (або) інтервальні оцінки дійсних значень характеристик. У процесі контролю перевіряється відповідність дійсних значень характеристик регламентованим для них нормам.

Апріорна інформація про склад і суттєвість складових похибки вимірювальних каналів можуть бути отримані з нормативної та технічної документації на ІВС або на основі експериментальних досліджень складу та властивостей похибок каналів.

При відсутності апріорних відомостей про склад похибки і суттєвість її складових проводять, керуючись наведеними нижче

вказівками, дослідження похибок каналів, типових для ІВС даного типу.

Для досліджень вибирають вибірки (не менше п'яти) однотипних вимірювальних каналів. Для кожної вибірки визначають не менше чотирьох точок, які перевіряються, розподілених рівномірно в межах діапазону вимірювань. Для кожної точки, визначають толерантні межі і межі довірчих інтервалів для систематичної складової похибки і для середньоквадратичного відхилення випадкової складової похибки.

Зіставляючи результати, отримані для всіх перевірених точок у вибірці, визначають найгірші (найбільші) значення границь інтервалів:

$D_{\text{сист гр}}$ – для довірчого інтервалу систематичної складової похибки;

$S_{\text{гр}}$ – для довірчого інтервалу середнє відхилення випадкової складової похибки;

$D_{\text{гр}}$ – для толерантного інтервалу похибки.

Вибрані найбільші значення кордонів інтервалів далі розглядаються як параметри сукупності однотипних вимірювальних каналів.

На основі отриманих даних роблять висновок про можливість визнання випадкової складової похибки нехтує малою або навпаки суттєвою. Для цього обчислюють значення інтервального коефіцієнта

$$K_p = D_{\text{гр}} / \sqrt{D_{\text{сист гр}}^2 / 3 + S_{\text{гр}}^2 + q^2 / 12},$$

де q - номінальна ступінь квантування (ціна одиниці найменшого розряду вихідного коду); для аналогових та аналого-цифрових каналів приймають $q = 0$;

Перевіряється виконання нерівності $S_{\text{гр}}^2 \leq A(2-A)D_0^2 / K_p^2$, де D_0 – граничне значення (межа допустимих значень) похибки каналу в реальних умовах його експлуатації (якщо воно невідоме, замість нього в формулу рекомендується підставити);

A – критерій малості складової похибки (відносна похибка розрахунку, обумовлена нехтуванням однією або декількома складовими похибки) рекомендується для практичних розрахунків вибирати $0,1 \leq A \leq 0,2$.

Якщо нерівність виконується, то слід вважати, що випадкова складова похибки дуже мала. В іншому випадку випадкова складова не може бути визнана такою. На підставі отриманих даних вибирають один з наведених нижче методів визначення похибки досліджуваних каналів.

Методи визначення похибки. Описані тут методи визначення характеристик похибки можна застосувати й у тому випадку, коли крім

Розділ 11. Метрологічне забезпечення ІВС

систематичної і випадкової складових похибки вимірювальний канал має істотну варіацію, яка є додатковою до вже вказаних характеристик. Тому експеримент відповідно описаних методів здійснюють двічі: спочатку при підході до точки діапазону, що визначається, з боку менших, а потім – з боку більших значень (або навпаки). При цьому отримують дві вибірки значень і похибки обсягом $n/2$ кожна, які є вихідними даними для знаходження точкових та інтервальних оцінок характеристик і опрацьовуються відповідно наведених розрахункових методів.

Метод визначення похибки аналогових і цифро-аналогових вимірювальних каналів для випадку нехтування малою випадковою складовою похибки. Структурна схема контролю похибки приведена на рис. 11.4.

Еталон 1 задає на вході каналу значення X прямо вимірюваної величини або її носія, відповідні точкам діапазону вимірювань, які перевіряються (при визначенні похибки цифро-аналогових каналів як еталон 1 використовується довільний задатчик коду). СІТ-еталон вимірює значення носія Y на виході вимірювального каналу.

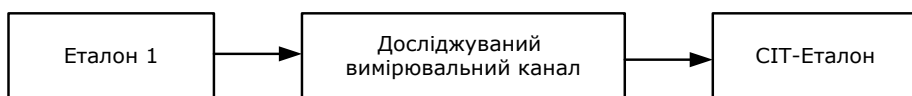


Рис. 11.4. Схема контролю похибки

Якщо точка X , що перевіряється, діапазону вимірювань задана в одиницях прямо вимірюваної величини або її носія, то за допомогою еталона 1 встановлюють значення вхідного сигналу X , зчитують і реєструють показання СІТ-еталону і розраховують значення абсолютної похибки, виражене в одиницях сигналу, за формулою $D = Y - F_{\text{ном}}(X)$, де $F_{\text{ном}}(X)$ – значення сигналу, обчислене для досліджуваної точки X за номінальної функції перетворення.

Якщо точка Y , що перевіряється, діапазону вимірювання задана в одиницях носія, то на еталоні 1 встановлюють таке значення вхідного сигналу, при якому показання СІТ-еталону дорівнюють Y . Значення абсолютної похибки обчислюється в одиницях вхідного сигналу вимірювального каналу $D = X - F_{\text{ном}}^{-1}(Y)$.

Метод визначення характеристик похибки аналогових і цифро-аналогових вимірювальних каналів для випадку істотної випадкової складової похибки. У кожній точці, що перевіряється, проводиться не менше $n=10$ вимірювань D_i ($i=1,2,\dots,n$) похибки вимірювального

каналу. За раніше розглянутою методикою обчислюються точкові оцінки характеристик систематичної і випадкової складових похибки і межі довірчих інтервалів для характеристик складових похибки і (або) толерантні границі похибки.

Якщо не потрібно великої точності експерименту або є підстави вважати, що закон розподілу випадкової складової похибки є нормальним, можна для спрощення розрахунків прийняти параметр $p = 2$. Інакше доцільно застосувати робастну методику.

Метод визначення похибки аналого-цифрових вимірювальних каналів для випадку нехтування малою випадковою складовою похибки. Структурна схема визначення похибки наведена на рис. 11.5.



Рис. 11.5. Схема визначення похибки

Еталон 1 задає на вході вимірювального каналу значення прямо вимірюваної величини або її носія, які відповідають точкам, що перевіряються, діапазону вимірювань. На виході каналу отримують код (показання), який може бути знятий експериментатором або автоматичним пристроєм. Всі сигнали і коди слід виражати або в одиницях вимірюваної величини, або в одиницях носія вимірюваної величини, яка надходить на вхід досліджуваного вимірювального каналу.

Точки, які перевіряються, діапазону вимірювань задають установленням значення N_0 вихідного коду або показання вимірювального каналу. Регулюючи вхідний сигнал еталону 1 (рівень регуляції повинен бути не більше 0,25 від номінальної ступені квантування q вимірювального каналу), встановлюють на вході каналу таке значення прямо вимірюваної величини або її носія, при якому на виході каналу або спостерігається перехід від коду (показання) $(N_0 - q)$ до заданого коду N_0 точки, що перевіряється, чи настає приблизно рівночасне чередування кодів $(N_0 - q)$ і N_0 . Значення похибки вимірювального каналу при вхідному коді обчислюють за формулою:

$$|D| = |N_0 - 0,5 \cdot q - X_m| + 0,5 \cdot q.$$

Метод визначення характеристик похибки аналого-цифрових вимірювальних каналів для випадку істотної випадкової складової похибки. Структурна схема визначення похибки приведена на рис. 11.5. Метод застосовується в тому випадку, коли середньоквадратичне відхилення випадкової складової похибки перевищує $0,25q$, тобто при будь-якому значенні вимірюваної величини в межах будь-якого рівня квантування чергуються випадковим чином не менше двох значень вихідного коду (показань) вимірювального каналу. Точки, що

Розділ 11. Метрологічне забезпечення ІВС

перевіряються, діапазону вимірювань задають установкою значення X_0 прямо вимірюваної величини або її носія.

На вхід каналу, що перевіряється, подають від еталону 1 значення X_0 вимірюваної величини, відповідне досліджуваній точці діапазону вимірювань. Зчитують і реєструють $n \geq 10$ значень N_i вихідного коду (показань) вимірювального каналу. Обчислюють значення похибок каналу $D_i = N_i - X_0$.

Метод розрахунку толерантних границь похибки і довірчих інтервалів характеристик складових похибки незалежно від виду закону розподілу. Вихідними даними для розрахунку толерантних границь похибки і меж довірчих інтервалів характеристик її складових в конкретній точці є знайдені експериментально одним з вище описаних вище методів вибіркові значення D_i похибки, де $i=1,2,\dots,n$, при $5 \leq n \leq 250$. Розрахунок проводиться за такою методикою:

1) Визначаються:

$$\text{середнє значення похибки } D_a = \sum_{i=1}^n D_i / n;$$

$$\text{оцінка другого центрального моменту } M_2 = \sum_{i=1}^n (D_i - D_a)^2 / n;$$

$$\text{оцінка четвертого центрального моменту } M_4 = \sum_{i=1}^n (D_i - D_a)^4 / n;$$

$$\text{оцінка вибіркового ексцесу } E_c = M_4 / M_2^2;$$

$$\text{незміщена оцінка ексцесу } E_x = \frac{[n^2 - 2n + 3]E_c - 3(2n - 3)}{n^2 - 3n + 3 - (n - 1)E_c}.$$

Якщо виконується нерівність $1,8 < E_x \leq 6$, то обчислюють значення параметру p за наближеною формулою $p = [4,2 / (E_x - 1,8)]^{0,5886}$. (15.16)

Якщо $E_x > 6$, то необхідно або збільшити число відліків похибки в експерименті і повторити розрахунок, або прийняти $p = 1$.

Якщо $E_x \leq 1,8$, то також необхідно збільшити число відліків похибки в експерименті або прийняти $p = 15$.

Якщо повторення експерименту неможливо, то в особливо відповідальних випадках значення параметру p знаходять з точного рівняння $E_c = \Gamma(1/p)\Gamma(5/p)/[\Gamma(3/p)]^2$, де $\Gamma(x)$ – гамма-функція.

Якщо отримано значення $p=1$, це свідчить про можливу наявність грубого промаху в вимірюваннях, і за lp -оцінку систематичної складової похибки приймають значення вибіркової медіани.

Якщо отримано значення $p \geq 15$, це може свідчити про наявність варіації або двохмодальний закон розподілу похибки.

2) Розраховують lp -оцінку систематичної складової (математичного очікування) похибки з умови

$$D_{p \text{ сист}} = \arg \min_f \sum_{i=1}^n |D_i - f|^p. \quad (15.17)$$

Іншими словами, потрібно знайти таке значення змінної f , при якому наведена сума буде мінімальною; це значення змінної f приймається за оцінку систематичної складової похибки. Рішення може бути знайдено ітеративним методом, при цьому рекомендується прийняти для першої ітерації $f = D_a$.

Якщо $f = D_a$ то $D_{p \text{ сист}} = D_a$.

3) Обчислюють lp - оцінку середнє відхилення випадкової складової похибки $S_p = (p/(n-1))^{1/p} \cdot (\Gamma(3/p)/\Gamma(1/p))^{1/p} \cdot \sqrt[p]{\sum_{i=1}^n |D_i - f|^p}$.

Якщо $p=2$, оцінка середньоквадратичного відхилення дорівнює

$$S_p = \sqrt{\sum_{i=1}^n (D_i - D_a)^2 / (n-1)}.$$

4) Знаходять межу довірчого інтервалу для систематичної складової похибки $D_{p \text{ сист}}$ при довірчій ймовірності $P=0,95$:

$$D_{p \text{ сист}} - t \frac{S_p}{\sqrt{n-1}} \leq D_{\text{сист}} \leq D_{p \text{ сист}} + t \frac{S_p}{\sqrt{n-1}}, \quad (11.18)$$

$$\text{де } t = \frac{0,4446 + 1,1146(a-n)}{1 + 0,57(a-n)} + 0,154 \frac{p-2}{p-0,6266}; \quad a = \frac{2,357(P-2)}{P+0,316}.$$

5) Знаходять кордону довірчого інтервалу для середньоквадратичного відхилення випадкової складової похибки при довірчій ймовірності $P=0,95$:

$$s_p \sqrt{(n-1)/\chi_2^2} \leq s \leq s_p \sqrt{(n-1)/\chi_1^2},$$

де $\chi_1^2 = (n-1)[1 - (1,2 + a_1n)]/(1 - b_1n)$; $\chi_2^2 = (n-1)([1 - (c_0 + c_1n)])/(1 - c_2n)$;

$a_1 = (0,0189 - 0,00316p)/(1 + 0,068p)$; $b_1 = (0,0388 - 0,266p)/(1 + 2,27p)$;

$c_0 = (4,93 - 0,464p)/(1,16p - 1)$; $c_1 = (0,0024p - 0,1255)/(1 - 1,474p)$;

$c_2 = (0,431 + 0,095p)/(1,414p - 1)$.

6) Знаходять толерантні межі похибки для довірчої ймовірності

Розділ 11. Метрологічне забезпечення ІВС

$P=0,95: (D_{p,\text{сист}} - \chi \cdot s_p) \leq D \leq (D_{p,\text{сист}} + \chi \cdot s_p)$, де

$$\chi = \frac{a_0 + a_1 n}{1 + b_1 n}; \quad a_0 = \frac{2,787 + 1,8244 p}{1 + 0,03007 p}; \quad a_1 = \frac{0,8282 + 0,576 p}{1 + 0,106 p}; \quad b_1 = \frac{0,264 + 0,286 p}{1 + 0,072 p}.$$

Визначення метрологічних характеристик вимірювального каналу ІВС, до складу якого входить обчислювальний компонент. При метрологічній атестації обчислювальних компонентів ІВС, алгоритмів опрацювання вимірювальної інформації досліджуються три основних групи показників: показники точності, стійкості і складності.

Показники точності характеризують точність результатів, отриманих за допомогою алгоритму з дотриманням загальноприйнятої моделі вхідних даних.

Показники стійкості (надійності) характеризують стійкість по відношенню до спотворень початкових даних, перешкод, надмірних похибок тощо.

Показники складності визначають трудомісткість рішення задачі при використанні даного алгоритму (кількість елементарних операцій опрацювання даних).

Мета атестації алгоритму – вибір оптимального алгоритму для вирішення конкретного завдання (на етапі розробки системи).

Під метрологічними характеристиками програми обчислень розуміють характеристики тих властивостей програми, які впливають на результат вимірювань і можуть привести до додаткових втрат вимірювальної інформації.

Втрати вимірювальної інформації можуть бути обумовлені:

застосуванням наближених методів обчислень (недосконалість методів або алгоритмів);

недостовірністю експериментальних даних, які надходять на вхід обчислювального компоненту (спадкова втрата вимірювальної інформації);

похибкою округлення результатів обчислень.

У разі суттєвих обмежень обчислень (за часом розрахунків, кроком дискретизації, числом операцій) доцільно досліджувати функції зв'язку між метрологічними характеристиками і цими обмеженнями. Отримавши в процесі метрологічної атестації залежність, наприклад методичної похибки від кроку дискретизації можна вибрати значення параметру обмеження, при якому ця похибка буде мінімальною.

Оцінка метрологічних характеристик програм обчислень може бути отримана за допомогою обчислювальних експериментів, організація яких асоціюється з розглянутими вище методом "зразкової міри" або методом "зразкового приладу".

У першому випадку на вхід обчислювального компоненту подають цифрові сигнали, які імітують роботу аналогової частини системи (рис. 11.6,а). Необхідна послідовність числових даних може утворюватися шляхом

зчитування з пристрою пам'яті або відтворюватися програмними засобами ЕОМ відповідно заданого алгоритму. Результати опрацювання a_i^{ok} числових даних $N(t_i)$ і очікувані результати їх опрацювання a_i^t , які надходять на вхід пристрою порівняння, повинні бути відомі з необхідною точністю. За значенням отриманої різниці приймається висновок щодо метрологічних характеристик обчислювального компонента.



Рис. 11.6. Методи визначення метрологічних характеристик обчислювального компоненту IBC\

Для організації експерименту за методом "зразкового приладу" (рис.11.6,б) використовується імітація результатів прямих вимірювань (з урахуванням відомих законів розподілу даних і відомої інформації про вимірювані величини). Результат застосування програми обчислення a_i^{ok} порівнюється з "ідеальним" значенням a_i^{et} , визначеним з урахуванням апріорної інформації про досліджувану систему (наприклад, за допомогою "еталонної" ЕОМ). Однак через складність цей експеримент у багатьох випадках недоцільний.

Метрологічна атестація програм обчислень проводиться тільки для атестації програм обчислень, які розробляються або не пройшли метрологічну атестацію. Повторна атестація проводиться тільки при зміні обчислювального компонента.

Далі розглянемо метод визначення метрологічних характеристик обчислювального компоненту IBC послідовно-паралельної структури, що складається з вимірювальних компонентів і обчислювального компоненту. Вимірювальний компонент організований як сукупність послідовно з'єднаних

Розділ 11. Метрологічне забезпечення ІВС

вимірювальних компонентів з нормованими метрологічними характеристиками, який вимірює одну фізичну величину. На обчислювальному компоненті виконуються програми, призначені для розрахунків за результатами прямих вимірювань оцінок результатів непрямих або сукупних вимірювань.

Як модель таких вимірювальних каналів приймають багатополосник, що має k входів і один вихід. Вхідними величинами каналу є електричні сигнали, пропорційні вимірюваним величинам, а вихідною величиною – значення величини, обчислене за допомогою програми, яка виконується обчислювальним компонентом.

Метод визначення характеристик похибки вимірювального каналу ґрунтується на моделюванні електричних сигналів вимірювальних компонентів, які після перетворення в код надходять в обчислювальний компонент для програмного опрацювання результатів. Під час визначення характеристик похибки вимірювального каналу за характеристиками вимірювальних компонентів моделюють вимірювальні сигнали, які надходять на вхід обчислювального компонента, тобто електричні сигнали заміняють масивами числових даних. Моделювання похибки вимірювальних компонентів здійснюють, як правило, за допомогою спеціальних програм-генераторів випадкових чисел, призначених для формування чисел із заданою функцією розподілу і необхідними статистичними характеристиками.

Розглядаються такі характеристик похибки: систематична складова похибки $\Delta_{\text{сист}}$, середньоквадратичне відхилення $\bar{\sigma}(\Delta)$ похибки та похибка каналу Δ .

Процедура визначення характеристик похибки вимірювального каналу складається з етапів, представлених на рис. 11.7.



Рис. 11.7. Етапи визначення характеристик похибки вимірювального каналу

Вимоги до початкової інформації. Інформація про вимірювальних компонентах, які входять в вимірювальний канал має містити такі

відомості:

значення верхнього та нижнього меж діапазону вимірювань кожного з вимірювальних компонентів;

значення математичного очікування і середньоквадратичного відхилення систематичної складової основної похибки вимірювальних компонентів.

Інформація про програму, реалізовану в обчислювальному компоненті повинна включати:

кількість і назву вхідних величин, порядок і спосіб введення їх значень;

назву величини і порядок виведення її значень;

вид і властивості функції програми.

Підготовка вхідних даних до експерименту. Вибір експериментальних точок $X_r[i]$ $r=\overline{1, K}$, для яких проводиться визначення характеристик похибки вимірювального каналу, здійснюють для кожної вхідної величини X_r на основі інформації про вимірювальний компонент і вид та властивості функції перетворення програми. Кількість експериментальних точок за r -ою вхідною величиною X_r , позначимо як N_r , $r=\overline{1, K}$. Мінімальна кількість таких точок – три: на початку, в кінці і всередині діапазону вимірювань допустимих значень вхідної величини. Якщо функція програми має розриви, то точки розривів додатково включають в число експериментальних точок.

Вибір кількості експериментів в експериментальних точках для кожної вхідної величини здійснюють на підставі інформації про закон розподілу похибки вимірювальних компонентів і її статистичні характеристики, а також про якість застосовуваних генераторів випадкових чисел, за допомогою котрих формується масив вхідних даних в кожній експериментальній точці $X_r[i]$, де $i=\overline{1, N_r}$.

Генератор випадкових чисел (програмний або фізичний) повинен формувати масиви чисел $\xi_r[j]$, $j=\overline{1, M_r}$, розподілених за тим же законом розподілу, що і похибка r -го вимірювального компоненту (з тими ж статистичними характеристиками). Кількість масивів вхідних даних в кожній експериментальній точці за r -ою вхідною величиною X_r вважають однаковим і позначають як M_r , $r=\overline{1, K}$.

Масиви значень формують за формулою

$$X_r[i, j] = X_r[i] \cdot (1 + \xi_r[j]), \quad (11.19)$$

де $\xi_r[j]$ – j -е значення r -ої випадкової величини, сформованої за

допомогою r -го генератора випадкових чисел.

У кожній експериментальній точці $X_r[i]$ формують стільки масивів вхідних даних $X_r[i, j]$, скільки необхідно для перевірки статистичної гіпотези про розподіл випадкової величини ξ за заданим законом із необхідними статистичними характеристиками.

Підготовка програми до проведення експериментів полягає в заміні автоматизованого введення вхідних величин (від вимірювальних компонентів) на ручний – з клавіатури або на зчитування підготовлених масивів вхідних даних з носія, в залежності від потреб дослідника. Така заміна повинна проводитися кваліфікованим програмістом (бажано розробником цієї програми) і не повинна порушувати основний алгоритм опрацювання програми.

Послідовність визначення характеристик похибки вимірювального каналу полягає в наступному:

1. Проводять перетворення кожного з масивів вхідних даних $X_r[i, j]$, де $r = \overline{1, K}$, $i = \overline{1, N_r}$, $j = \overline{1, M_r}$ за допомогою досліджуваного програмного забезпечення і фіксують результати спостережень $Y[i, j]$, $i = \overline{1, N_r}$, $j = \overline{1, M_r}$.

2. При необхідності в кожній експериментальній точці оцінюють аномальність результатів за відомими критеріями (див. розд. 5). Виправлені результати в i -й експериментальній точці в j -му експерименті позначимо $Z[i, j]$.

3. Проводять оцінювання систематичної складової похибки вимірювального каналу в i -й експериментальній точці за формулою

$$\bar{\Delta}_{\text{сист}}[i] = (\bar{Z}[i] - Z[i]), \quad (11.20)$$

де $\bar{Z}[i]$ – середнє арифметичне значення скоригованих результатів спостережень в i -й експериментальній точці; $Z[i]$ – дійсне (справжнє) значення результату спостережень в i -й експериментальній точці (при вхідних даних з нульовою похибкою).

Величину $\bar{Z}[i]$ визначають за формулою
$$\bar{Z}[i] = \frac{1}{M} \sum_{j=1}^M Z[i, j].$$
 (11.21)

4. Проводять оцінювання систематичної складової похибки вимірювального каналу
$$\bar{\Delta}_{\text{сист гр}} = \max \bar{\Delta}_{\text{сист}}[i].$$
 (11.22)

5. Проводять оцінювання дисперсії випадкової складової похибки

каналу в i -й експериментальній точці

$$\bar{D}(\overset{\circ}{\Delta}[i]) = \frac{1}{M-1} \cdot \sum_{j=1}^M (Z[i, j] - \bar{Z}[i])^2. \quad (11.23)$$

6. Проводять оцінювання середньоквадратичного відхилення випадкової складової похибки каналу в i -й точці

$$\bar{\sigma}(\overset{\circ}{\Delta}[i]) = \sqrt{\bar{D}(\overset{\circ}{\Delta}[i])}. \quad (11.24)$$

7. Проводять оцінювання середньоквадратичного відхилення випадкової складової похибки вимірювального каналу:

обчислюють відношення найбільшого значення дисперсії (в діапазоні) до суми всіх значень дисперсій (критерій Кохрена, див. розд. 5)

$$G = \left(\max \bar{D}(\overset{\circ}{\Delta}[i]) \right) / \sum_{i=1}^N \bar{D}(\overset{\circ}{\Delta}[i]); \quad (11.25)$$

обчислене значення G порівнюють з критичним значенням $G_{кр}$. Якщо $G \leq G_{кр}$, то слід вважати, що дисперсії однорідні, в іншому випадку – неоднорідні;

якщо дисперсії неоднорідні, то за оцінку середньоквадратичного відхилення випадкової складової похибки каналу приймають найбільше значення

$$\bar{\sigma}(\overset{\circ}{\Delta}) = \max \bar{\sigma}(\overset{\circ}{\Delta}[i]); \quad (11.26)$$

якщо дисперсії однорідні, оцінкою середньоквадратичного відхилення випадкової складової похибки каналу є середнє значення

$$\bar{\sigma}(\overset{\circ}{\Delta}) = \frac{1}{N} \sum_{i=1}^N \bar{\sigma}(\overset{\circ}{\Delta}[i]). \quad (15.27)$$

8. Оцінюють межі інтервалу, в якому з довірчою ймовірністю P знаходиться похибка вимірювального каналу:

обчислюють в кожній експериментальній точці кордону довірчого інтервалу $\Delta_v[i] = \bar{\Delta}_{сист}[i] + t \cdot \bar{\sigma}(\overset{\circ}{\Delta}[i])$ і $\Delta_n[i] = \bar{\Delta}_{сист}[i] - t \cdot \bar{\sigma}(\overset{\circ}{\Delta}[i])$, де t – квантиль розподілу Стюдента.

межі інтервалу, в якому з довірчою ймовірністю P знаходиться похибка вимірювального каналу $\Delta_v = \max \bar{\Delta}_v[i]$, $\Delta_n = \min \bar{\Delta}_n[i]$.

Основні висновки

Такі особливості ІВС, як багатофункціональність, багатоканальність,

Розділ 11. Метрологічне забезпечення ІВС

наявність обчислювального компонента, нерозривний зв'язок спеціалізованих ІВС з об'єктом, на якому вони експлуатуються, агрегатний спосіб побудови, розподіленість компонентів і складових частин ІВС в просторі, зумовлюють основні відмінності їх метрологічного забезпечення від метрологічного забезпечення інших засобів вимірювальної техніки. До проблем, які виникають при розробці метрологічного забезпечення, можна віднести: забезпечення одночасного вимірювання ряду фізичних величин; оцінювання метрологічних характеристик алгоритмів і програм обробки результатів; оцінювання впливу вимірювальних каналів один на одного; проведення метрологічного обслуговування в умовах неможливості прив'язки до ідеалу; облік впливу на метрологічні характеристики різних умов експлуатації компонентів ІВС і ін.

Об'єктами метрологічного забезпечення ІВС є вимірювальні канали та їх складові, тобто вимірювальні і обчислювальні компоненти.

До завдань метрологічного забезпечення ІВС відносяться: аналіз і оцінювання технічних рішень на відповідність нормам точності і вибору параметрів; встановлення єдиних вимог до номенклатури, нормування та поданням метрологічних характеристик вимірювальних каналів ІВС; визначення дійсних значень метрологічних характеристик вимірювальних каналів і обчислювальних компонентів і їх відповідності встановленим нормам; встановлення повноти робіт з метрологічного забезпечення ІВС і відповідності вимогам НД з метрології.

Склад робіт з метрологічного забезпечення визначається стадією «життєвого циклу» ІВС - розробкою, виробництвом, впровадженням і експлуатацією ІВС.

Основними роботами з метрологічного забезпечення ІВС є: встановлення єдиних вимог до метрологічних характеристик систем; розробка методів і засобів контролю метрологічних характеристик; метрологічна експертиза технічної документації; аналіз стану метрологічного забезпечення ІВС; організація і здійснення державного метрологічного контролю і нагляду за станом і застосуванням ІВС; організація і проведення повірочних робіт; організація і проведення робіт з атестації алгоритмів обробки інформації.

Метрологічні характеристики відображають властивості засобів вимірювань (вимірювальних каналів), які надають визначальний вплив на результат і точність вимірювань. Нормування метрологічних характеристик - це встановлення комплексу метрологічних характеристик і способів їх подання. Комплекс метрологічних характеристик повинен включати в себе характеристики, інваріантні до умов застосування і режиму роботи. Способи подання метрологічних характеристик повинні забезпечувати простоту їх контролю. При регламентації метрологічних властивостей ІВС слід враховувати особливості випуску та комплектації систем, умови їх експлуатації з урахуванням просторової розподіленості компонентів, використання в їх складі складних обчислювальних пристроїв.

Методи визначення метрологічних характеристик ІВС підрозділяються

на розрахункові, експериментальні і методи визначення метрологічних характеристик обчислювального компонента.

Розрахунок метрологічних характеристик вимірювальних каналів може проводитися як для статичного, так і для динамічного режимів роботи в нормальних і робочих умовах експлуатації.

Експериментальні методи включають методи визначення характеристик аналогових, цифро-аналогових та аналого-цифрових вимірювальних каналів при наявності нехтує малою і суттєвою випадковою складовою похибки.

Під метрологічними характеристиками обчислювального компонента розуміють характеристики тих властивостей програм, які впливають на результат вимірювань і можуть привести до додаткових втрат вимірювальної інформації, які можуть бути обумовлені застосуванням наближених методів обчислень (недосконалість методів або алгоритмів), недостовірність експериментальних даних, які надходять на вхід обчислювального компонента (спадкова втрата вимірювальної інформації), похибкою округлення результатів обчислень.

Метрологічна атестація програм обчислень проводиться тільки при атестації програм обчислень, які розробляються або не пройшли метрологічну атестацію. Повторна атестація проводиться при зміні обчислювального компонента.

Запитання для самоконтролю

- 1. Що розуміється під метрологічним забезпеченням ІВС?*
- 2. Перерахуйте основні роботи з метрологічного забезпечення ІВС.*
- 3. На яких етапах життєвого циклу здійснюється метрологічне забезпечення ІВС?*
- 4. Охарактеризуйте спільні цілі метрологічної експертизи технічної документації на ІВС різних видів.*
- 5. У чому полягає складність у здійсненні державного метрологічного контролю і нагляду по відношенню до інформаційно-вимірювальних систем?*
- 6. Проаналізуйте, які проблеми в області метрологічного забезпечення виникають в зв'язку з основними особливостями ІВС.*
- 7. У чому полягають загальні принципи нормування метрологічних характеристик ІВС?*
- 8. Які особливості ІВС в першу чергу обумовлюють специфіку регламентації їх метрологічних характеристик?*
- 9. Які характеристики вимірювальних каналів належать до динамічних?*
- 10. Охарактеризуйте основні проблеми і специфічні особливості експериментальних досліджень метрологічних властивостей ІВС.*
- 11. Як враховується вплив впливають величин при визначенні метрологічних характеристик ІВС?*
- 12. Сформулюйте основні принципи, які використовуються при визначенні метрологічних характеристик програм обчислень.*
- 13. Від чого залежить вибір експериментальних і розрахункових методів*

Розділ 11. Метрологічне забезпечення ІВС

визначення і контролю метрологічних характеристик вимірювальних каналів ІВС?

14. Назвіть етапи визначення характеристик похибки вимірювального каналу.

15. Охарактеризуйте методи визначення метрологічних характеристик обчислювальних компонентів ІВС.

16. Назвіть методи експериментального визначення характеристик похибки вимірювальних каналів.

17. Назвіть класифікацію методів визначення метрологічних характеристик вимірювальних каналів ІВС.

Conclusions

IMS features such as multifunctionality, multichannel, computational component presence, specialized IMS inextricable link with object on which they operate, aggregate method of construction, distribution of IMS components and parts in space cause main differences between their metrological assurance and metrological assurance of other means of measuring techniques. Amongst problems arising during development of metrological assurance are: ensuring simultaneous measurement of a number of physical quantities, evaluation of metrological characteristics of algorithms and software for results processing, evaluation of measurement channels' influence on each other, providing metrological service in conditions of impossibility of binding to the standard, accounting for effects of various IMS components operating conditions, etc. on the metrological characteristics

IMS metrological assurance objects are measuring channels and their measuring and computing components.

IMS metrological assurance tasks: analysis and evaluation of technical solutions to meet the standards of accuracy and parameters choice, establishment of uniform standards for nomenclature, standardization and presenting of IMS measuring channels' metrological characteristics, determination of metrological characteristics valid values of measuring channels and computing components and their regulatory compliance, establishment of work completeness on IMS measurement assurance and compliance with metrology regulatory documents.

The scope of work on measurement assurance is determined by the stage of IMS "life cycle" - development, production, implementation and IMS operation.

IMS metrological assurance major work: establishing of uniform requirements for systems metrological characteristics; methods and means development of metrological characteristics control; metrological examination of technical documentation; analysis of IMS metrological assurance state; organization and implementation of state metrological control and supervision of IMS condition and use; organization and conduction of verification (calibration); organization and conduction of data processing algorithms certification.

Metrological characteristics reflect properties of measuring means (measurement channels) that have determining influence on measurement result and accuracy. Rationing of metrological characteristics - is establishing of complex metrological characteristics and methods of its presentation. Metrological characteristics complex should include characteristics invariant to operating conditions and operation mode. Ways of metrological characteristics presenting should provide control simplicity.

When regulating IMS metrology properties, system production and configuration particularities, operating conditions considering spatial components distribution, and use of complex computing devices should be considered.

Methods for IMS metrological characteristics determination are divided into calculation methods for determining, by metrological characteristics, of components that are part of measurement channel, experimental methods and methods of determination of computational component metrological characteristics.

Calculation of measuring channels metrological characteristics can be performed for both static and dynamic modes, and in normal and actual operating conditions.

Experimental methods include methods of characteristics determination of analog, digital-to-analog and analog-to-digital measurement channels in the presence of negligible and substantial random error component.

Metrological characteristics of computational component are characteristics of those programs properties that affect measurement result and could lead to additional losses of measurement information that might be caused by application of approximate calculation methods (imperfect methods or algorithms), unreliability of experimental data that are coming to input of computational component (hereditary loss of measurement information), rounding error of calculation results.

Metrological attestation of calculations programs is undertaken only for attesting computations programs under development or failed to pass metrological attestation. Re-attestation is possible after the change of computational component.

Ключові слова

метрологічне забезпечення	metrology assurance
метрологічна атестація	metrology attestation
метрологічні характеристики	metrology characteristics

СПИСОК ЛІТЕРАТУРИ

1. *Бабак В.П.* Обробка сигналів: Підручник / / В.П. Бабак, В.С. Хандецький, Е. Шрюфер. – К.: Либідь, 1999. - 392 с.
2. *Бабак В.П.* Теоретичні основи захисту інформації: Підручник. – Книжкове вид-во НАУ, 2008. – 752 с.
3. *Бабак В.П.* Теорія ймовірностей, випадкові процеси та математична статистика: Підручник / В.П. Бабак, Б.Г. Марченко, М.Є. Фриз. – К.: Техніка, 2004. - 288 с.
4. *Баскаков С.И.* Радиотехнические цепи и сигналы: Учебник. – М.: Высш. школа, 2000. – 462 с.
5. *Бронштейн И.Н.* Справочник по математике для инженеров и учащихся втузов. Изд-е 13-е / И.Н. Бронштейн, К.А. Семендяев. - М.: Наука, 1986. – 544с.
6. Датчики: [справочное пособие]; под общ. ред. В.М. Шарапова и Е.С. Полищука. – М.: Техносфера, 2012. – 624 с.
7. Информационно-измерительные системы. Применение интеллектуальных модулей, методов и средств повышения точности физических измерений: Монография / А.А. Зори, С.И. Клевцов, В.Д. Коренев и др. – Донецк: ДВНЗ «ДонНТУ», 2011. – 206 с.
8. *Кузнецов В.П.* Метрологические характеристики измерительных систем. – М.: Машиностроение, 1984. – 412 с.
9. *Кузнецов Б.Ф.* Стохастические модели и методы анализа информационно-измерительных систем АСУ ТП. – Ангарск: Ангарская гос. техн. академия, 2007. – 180 с.
10. *Кузьмин И.В.* Основы теории информации и кодирования: Учебник / И.В. Кузьмин, В.А. Кедрус. – К.: Вища школа, 1986. - 238 с.
11. *Куликовский К.Л.* Методы и средства измерений / К.Л. Куликовский, В.Я. Купер. – М.: Энергоатомиздат, 1986. – 448 с.
12. *Кухарчук В.В.* Основы метрології та електричних вимірювань: Підручник / В.В. Кухарчук, В.Ю. Кучерук, Є.Т. Володарський, В.В. Грабко. – Херсон: Олді-плюс, 2013.–538 с.
13. *Новопашенный Г.Н.* Информационно-измерительные системы: Учебное пособие. – М.: Высш. школа, 1977. – 208 с.
14. *Орнатский П.П.* Теоретические основы информационно-измерительной техники: Учебник. – К.: Вища школа, 1983. - 455 с.
15. Основы метрології та вимірювальної техніки: Підручник у 2 т./М. Дорожовець, В. Мотало, Б. Стадник, В. Василюк, Р. Борек, А. Ковальчик. – Львів: Видавництво Національного університету «Львівська політехніка», 2005. – 1300 с.

16. *Поліщук Є.С.* Засоби та методи вимірювань неелектричних величин: Підручник / Є.С. Поліщук, М.М. Дорожовець, Б.І. Стадник та ін.; за ред. Є.С. Поліщука. – Львів: Вид-во Бескид Біт. 2012. – 618 с.
17. *Раннев Г.Г.* Измерительные информационные системы: учебное пособие. – М.: Академия, 2010. – 336 с.
18. *Рабинер Л.* Теория и применение цифровой обработки сигналов / Л. Рабинер, Б. Гоулд Пер. с англ. – М.: Мир, 1978. – 848 с.
19. *Рубичев Н.А.* Измерительные информационные системы: учебное пособие. – М.: Дрофа, 2010. – 334 с.
20. Русско-украинско-английский словарь терминов по информационным технологиям / В. П. Бабак, О. Г. Байбуз, А. П. Приставка. – К.: НАУ, 2006. – 252 с.
21. *Скляр Б.* Цифровая связь. Теоретические основы и практическое применение / Пер. с англ. – М.: Изд. дом Вильямс, 2004. – 1104 с.
22. *Темников Ф.Е.* Теоретические основы информационной техники: Учеб. Пособие / Ф.Е. Темников, В.А. Афонин, В.И. Дмитриев. – М.: Энергия, 1979. – 512 с.
23. Теоретические основы информационно-измерительных систем: Учебник / В.П. Бабак, С.В. Бабак, В.С. Еременко и др., под ред. чл.-кор. НАН Украины В.П. Бабака / – К.: София-А, 2014. – 832с.
24. *Томпкинс У., Уебстер Дж.* Сопряжение датчиков и устройств ввода данных с компьютерами IBM PC. – М.: Мир, 1992. – 592 с.
25. *Фишер-Крипс А. С.* Интерфейсы измерительных систем. Справочное руководство. – М.: Изд. Дом «Технологии», 2006. – 336 с.
26. *Цапенко М.П.* Измерительные информационные системы: Структуры и алгоритмы, системотехническое проектирование: Учебн. пособие. – М.: Энергоатомиздат, 1985. – 440 с.
27. *Ярочкин В.И.* Информационная безопасность: Учебник. – М.: Академпредпроект: Трикста, 2005. – 544 с.
28. *Applied Measurement System / Edited by Md. Zachurul Had.* – In Tech, USA, 2012.
29. *Belkhamza Z.* Measuring Organizational Information Systems Success: New Technologies and Practices / Z. Belkhamza, A. Wafa. – Universiti Malaysia Sabah, Malaysia, 2012.
30. *Herold R.* Encyclopedia of Information Assurance / R. Herold, M. Robers. – Indiana, USA, 2010.
31. *Kalsi H.S.* Electronic instrumentation; edition by H.S. Kalsi. – Third edition. – New Delhi: Tata McGraw-Hill Education, 2012. – 829 p.

The textbook systematically presents basic concepts and principles of information-measurement systems construction and studies, based on mathematical apparatus for conversion of measuring signals, measurement technologies, transmission and processing of signals and data, interference immune coding, the use of modern information and communication channels for information transmission, methods of metrology characteristics studies. After each chapter main conclusions, self-control questions and key words in English and Ukraine can be found.

Textbook is recommended for technical specialties students of higher educational institutions, graduate students, researchers, engineers and technicians, measurement technology professionals.

Навчальне видання

Бабак Віталій Павлович,
Бабак Сергій Віталійович, Єременко Володимир Станіславович,
Куц Юрій Васильович, Щербак Леонід Миколайович

ТЕОРЕТИЧНІ ОСНОВИ ІНФОРМАЦІЙНО-ВИМІРЮВАЛЬНИХ СИСТЕМ

ВИДАННЯ ДРУГЕ, ПЕРЕРОБЛЕНЕ ТА ДОПОВНЕНЕ

Підручник

За редакцією чл.-кор. НАН України В. П. Бабака

Підп. до друку 20.04.2017. Формат 60×80/16. Папір офс. №1.
Гарн. Таймс. Умов. друк. арк. 31,0. Обл.-вид. арк. 32,4.
Наклад 300 пр.

Віддруковано у друкарні ТОВ «Наш формат»
01042, м. Київ, пр. Миру, 7, оф. 45.