
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БІОРЕСУРСІВ І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ
ІНСТИТУТ ЕНЕРГЕТИКИ, АВТОМАТИКИ І ЕНЕРГОЗБЕРЕЖЕННЯ

МЕТОДИЧНІ ВКАЗІВКИ

по виконанню самостійних робіт з дисципліни

ТЕОРІЯ ЕЛЕКТРОМЕХАНІЧНОГО ПЕРЕТВОРЕННЯ ЕНЕРГІЇ

галузь знань - 14 «Електрична інженерія»
спеціальність - 141 «Електроенергетика, електротехніка та
електромеханіка»
рівень вищої освіти – третій (освітньо-науковий рівень)

КИЇВ 2024

Методичні вказівки присвячені вивченню сучасних підходів до фізичної трактовки всіх явищ та процесів, які притаманні електромеханічним системам. Показана роль та місце математичних моделей, які використовуються на різних етапах розроблення та експлуатації електромеханічних перетворювачів, розкриті відповідні цілі та задачі їх створення, в узагальненому вигляді розглянуті ефективні методи формування базису (аргументів) математичної моделі, а саме, системи первісних параметрів, якими визначаються споживчі властивості. Суттєва увага при цьому приділена й найсучаснішим методам, що дозволяють одночасно використовувати як кількісні (об'єктивні), так і якісні (суб'єктивні) характеристики матеріальний об'єкт.

Методичні вказівки орієнтовані на аспірантів технічних спеціальностей, а особливо на аспірантів, які проводять наукові дослідження в галузі енергетичних технологій і технологій аграрного виробництва. Подані матеріали можуть бути використані для самостійної роботи, формування практичних навичок, а також як базис для написання наукових публікацій і дисертаційних робіт.

Рецензенти:

доктор. техн. наук, професор С.А.Шворов (НУБІП України. кафедра автоматизації та робототехнічних систем ім. акад. І.І. Мартиненка)

доктор. техн. наук, професор А.І. Чміль (НУБІП України. кафедра електротехніки, електромеханіки та електротехнологій)

Розробник: Заблудський М.М., професор кафедри електротехніки, електро-механіки та електротехнологій, д.т.н., професор.

Київ: НУБІП України. 2024. 78 С.

©НУБІП України

ПЕРЕЛІК ВВЕДЕНИХ СКОРОЧЕНЬ

БШП – багатошаровий перцептрон

ГА – генетичний алгоритм;

ЕМП – електромагнітні пристрої, електромеханічні перетворювачі енергії;

ЕРС – електрорушійна сила

ЕСЗ – еквівалентна схема заміщення;

МО – матеріальний об'єкт;

ОМ – об'єкт моделювання;

ПБР – поведінка бджолиного рою;

ШН – штучний нейрон;

ШНМ – штучна нейронна мережа.

ПЕРЕДМОВА

Характерна риса сучасності – стрімке зростання видів та кількості матеріальних об'єктів, які створені людиною та оточують її у повсякденному житті. Є всі підстави вважати, що ця тенденція не тільки збережеться, але й отримуватиме в найближчому майбутньому додатковий імпульс розвитку.

Будь-який виріб, апарат, система, процес (надалі, узагальнено – матеріальний об'єкт – МО) має набір властивостей, які у сукупності визначають його здатність задовольняти ті чи інші потреби споживача. Такі властивості прийнято називати *споживчими*, підкреслюючи тим самим їх корисний для визначеного споживача зміст. Для різних МО ці властивості безумовно різні. Але за будь-яких умов головним є *проявлення МО своїх споживчих властивостей в повному обсязі й на протязі часу, який визначено як час його експлуатації*. Тобто спільною та найважливішою характеристикою будь-якого створеного людиною МО є *якість*, точніше – *споживча якість*.

Споживчу якість можна визначити як властивість МО забезпечувати визначені відповідними нормативно-технічними документами технічні, економічні, естетичні, екологічні та інші характеристики на протязі визначеного терміну експлуатації у визначених умовах. Причому вказані у нормативно-технічних документах значення характеристик МО не можуть бути абстрактними, а *мають відповідати потребам конкретних споживачів*. Ці потреби зазвичай встановлюються на підставі ретельних досліджень можливого ринку збуту МО ще до початку формування його технічних характеристик.

Розглядаючи збереження екосистеми як загальну та необхідну характеристику будь-якого МО, до його споживчих властивостей слід відносити й безпечну для навколишнього середовища утилізацію після завершення терміну експлуатації. Особливе значення ця характеристика набуває під час проектування, виробництва та експлуатації енергетичних об'єктів та їх окремих складових. Адже саме ці об'єкти, функціонування яких пов'язане з перетворенням значної кількості енергії, суттєво впливають на стан навколишнього середовища, й при помилках та/або прийнятті невдалих технічних рішень в процесі проектування, виробництва, експлуатації та утилізації здатні привести до порушення екологічної рівноваги.

Серед вимог до якості МО, зазвичай, прийнято називати показники, що визначають його технічний рівень (функціональні параметри, надійність, ергономічність тощо) та ринкові можливості (собівартість, обсяг виробництва, ринкова ціна тощо). Всі ці показники входять до складу системи споживчих властивостей, якими наділяються сучасні МО. Беззаперечно, що до такої системи показників мають увійти й показники, які характеризуватимуть гуманітарні, соціальні, екологічні та ін. наслідки виробництва, експлуатації, утилізації того чи іншого МО.

Якість споживчих властивостей МО поняття *системне*, визначається значною кількістю найрізноманітніших факторів, серед яких і виконання відповідних робіт з проектування, виробництва, технічного обслуговування, і властивості матеріалів, комплектуючих, напівфабрикатів, і ефективність менеджменту, і глибина маркетингових досліджень тощо. За таких умов, для забезпечення якості МО дотримуються спеціальної методології, якою визначається весь комплекс робіт, тим чи іншим чином пов'язаних зі створенням, експлуатацією та утилізацією цього МО. Одним, чи не найважливішим складником цієї методології є розроблення та ефективне використання моделей МО. Наприклад, в залежності від поставлених цілей, результатом роботи з відповідною моделлю МО може бути інформація, яка на стадії проектування дозволяє оптимізувати цей МО з метою отримання заданих споживчих характеристик, на стадії виробництва – виявити причини відхилення споживчих властивостей від заданих нормативно-технічними документами, на стадії експлуатації – оптимізувати процес технічного обслуговування МО та, за умови непрацездатності МО, виявити пошкоджені складові частини тощо.

Функціонування більшості МО пов'язане з перетворенням одних видів енергії в енергію інших видів. Процес такого перетворення базується на визначених фізичних (та/або хімічних) явищах, завдяки яким саме й формуються корисні для людини споживчі властивості відповідного МО. Ці фізичні явища завжди мають конкретний характер, що надає можливість розділити всі МО на відповідні класи. Кожен такий клас об'єднуватиме МО, функціонування яких має однакові базові фізичні засади. Таке розділення суттєво спрощує дослідження властивостей МО, їх проектування виробництво, експлуатацію та утилізацію.

РОЗДІЛ 1

ЗАГАЛЬНІ ПРИНЦИПИ МОДЕЛЮВАННЯ МАТЕРІАЛЬНИХ ОБ'ЄКТІВ НА РІЗНИХ СТАДІЯХ ЇХ ЖИТТЄВОГО ЦИКЛУ

*Оточуючу реальність ми пізнаємо
через нами ж створені моделі, доступні
при даному стані наших знань*

NN

Шлях будь-якого матеріального об'єкта (елементарного комплектуючого, пристрою, апарату, системи, технологічного процесу тощо) від початку проектування (точніше, від реалізації заходів, що призводять до формування технічних вимог) до зняття з експлуатації, демонтажу та утилізації як непридатного або морально застарілого, називають його *життєвим циклом* цього об'єкту.

В укрупненому вигляді життєвий цикл будь-якої технічної системи (будь-якого МО) включає чотири узагальненні стадії — *проектування, виробництво, експлуатацію, утилізацію* (див. рис. 1.1). При цьому кожна з наведених стадій життєвого циклу при детальнішому розгляді характеризується сукупністю декількох етапів [1-3].

Стадія проектування передбачає формування технічних вимог та розроблення технічного завдання; проведення відповідних науково-технічних досліджень, розроблення віртуальних та/або матеріальних моделей, розроблення схем (структурних, функціональних, принципівих, монтажних, кінематичних та ін.), конструкції, технологічних процесів тощо.

Стадія виробництва включає підготовку та підтримання у працездатному стані відповідного технологічного обладнання та оснащення, проведення визначених організаційних заходів (зокрема навчання та/або перенавчання персоналу), заходів щодо захисту навколишнього середовища, придбання матеріалів та комплектуючих, виготовлення деталей, збирання та електричний монтаж виробів, виготовлення тари та ін. Крім того ця стадія життєвого циклу МО включає процеси налаштування МО, контроль його характеристик та (за необхідністю) пошук несправних складових у випадку, коли МО втрачає працездатність на лінії випробувань.

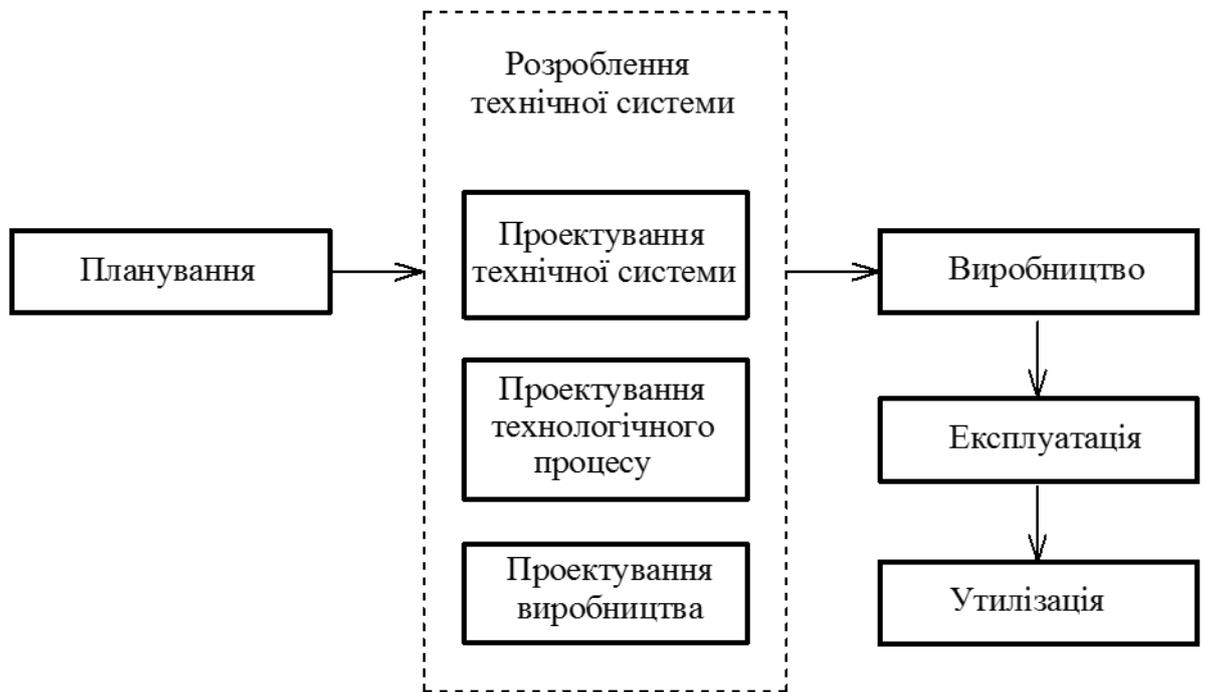


Рис. 1.1. Основні стадії життєвого циклу технічної системи

Стадія експлуатації складається з сукупності робіт і організаційних заходів, спрямованих на використання МО за призначенням та його підтримку в працездатному стані; до таких робіт відносяться — зберігання, транспортування, підготовка до використання (передпродажна підготовка або роботи, що передбачені регламентом вводу до експлуатації тощо), використання за призначенням, підготовка відповідного технічного персоналу, технічне обслуговування, пошук несправних складових та ремонт для МО, що втратили працездатність та ін.

Стадія утилізації включає виведення МО з експлуатації, його демонтаж, знищення або перероблення складових частин методами та засобами, що безпечні для персоналу та навколишнього середовища.

Послідовність наведених етапів, звичайно, може бути іншою, окремі етапи (при експлуатації конкретних видів МО) можуть бути відсутні, тоді як інші неодноразово повторюватимуться.

Життєвий цикл МО протікає у визначеному середовищі. Це середовище характеризується сукупністю кліматичних, механічних, організаційних, науково-технічних, економічних, екологічних, соціальних та ін. складників, які з точки зору функціонування МО визначаються як *зовнішні впливи*.

Наприклад, на будь-якій стадії життєвого циклу вказане середовище включає такі фактори як кваліфікація відповідного (науково-технічного, виробничого, адміністративного) персоналу, матеріально-технічне забезпечення, визначені організаційні засади, професіоналізм колективу та соціальний фон у ньому тощо.

Крім того, середовище, в якому протікає життєвий цикл МО, з повним

правом можна (і слід) характеризувати рівнем розвитку знань, на базі яких розробляються відповідні алгоритми та програми для проведення необхідних розрахунків, що супроводжують процеси проектування МО та його технічного обслуговування (включаючи, звичайно, і пошук несправностей).

Спроможність МО задовольняти ті чи інші потреби споживача характеризується *якістю* цього МО. З цієї точки зору якість МО можна подати, як набір показників (так би мовити, *споживчих властивостей*), що у сукупності складають його *споживчу цінність* та *споживчу вартість*.

Рівень якості МО, зазвичай, визначається відповідністю його вихідних параметрів (*показників якості*) нормативно-технічній документації та формується на всіх стадіях його життєвого циклу.

Слід зазначити, що під час визначення якості МО (особливо наукоємних), зазвичай, вводять критерії, які характеризують науково-технічний рівень, тобто ступінь використання для реалізації споживчих характеристик даного МО найсучасніших досягнень у відповідних галузях знань [1, 3]. Адже такі критерії (які часто називають критеріями інноваційності) характеризують життєздатність МО, тобто можливий строк його виробництва та використання до досягнення рівня морального старіння.

Для забезпечення необхідного рівня якості МО протягом його життєвого циклу, прийняті на визначених етапах технічні рішення щодо схеми, конструкції, матеріалів, технології виготовлення, зовнішнього вигляду та ін. можуть піддаватися (і, часто, піддаються) перегляду на наступних етапах. При цьому слід враховувати, що доступна кількість вказаних змін та їх вартість суттєво залежатиме від стадії життєвого циклу МО, на якому такі зміни відбуваються [2].

На рис. 1.2 представлені залежності від окремих стадій життєвого циклу (проектування, виробництво, експлуатація) вартості C проведення змін (схемотехнічних, конструктивних, технологічних та ін.) в МО та кількості N можливостей, які є для цього в розпорядженні підприємства-виробника або організації, в якій цей МО розробляється.

Практикою доведено, що витрати на корегування схем, конструкції, технології виготовлення будь-якого МО збільшуються у 10 разів при переході від однієї стадії життєвого циклу до наступної (див. рис. 1.3).

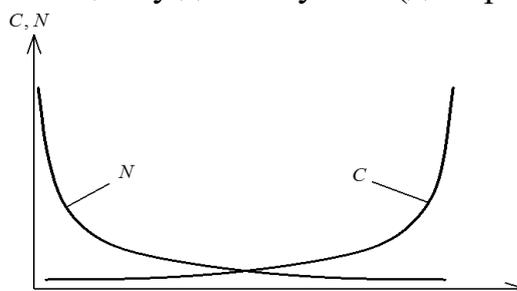


Рис. 1.2. Кількість (N) можливостей проведення змін та їх вартість (C) в залежності від стадії життєвого циклу МО

Отже, навіть за умови виявлених недоліків в схемних, конструктивних, технологічних, екологічних рішеннях на підставі відповідного контролю якості МО після завершення відповідної стадії його життєвого циклу, виправлення цих недоліків потребуватиме значних зусиль та матеріальних витрат, може призвести до порушення терміну введення МО в експлуатацію або до порушення регламенту робіт, що відбуваються на відповідній стадії, що само по собі є суттєвим фактором зниження якості.

Таким чином, визначення споживчих властивостей МО, зв'язків цих властивостей зі схемними, конструктивними, технологічними та ін. рішеннями на ранніх стадіях життєвого циклу МО є єдино можливий та економічно ефективний шлях досягнення бажаного рівня якості цього МО.

Зазначимо, що визначення умов експлуатації МО буде повним та адекватним лише за умови врахування *всіх* факторів, які впливають на його характеристики. Адже визначення умов експлуатації, наприклад, без урахування таких суттєвих факторів як кваліфікація персоналу, матеріально-технічні можливості та організаційні засади підприємства, де виробляється, експлуатується чи утилізується МО (тобто тих характеристик, які стосуються роботи підприємств, пов'язаних з МО на відповідному етапі його життєвого циклу) часто стає причиною суттєвих відрізень функціонування одного й того ж типу МО у різних споживачів. Під впливом факторів оточуючого середовища параметри на виходах складових частин МО (а, відповідно, і визначені цими параметрами вихідні характеристики МО в цілому) приймають відповідні значення. Саме ці параметри і забезпечують можливість проявлення МО притаманних йому споживчих властивостей

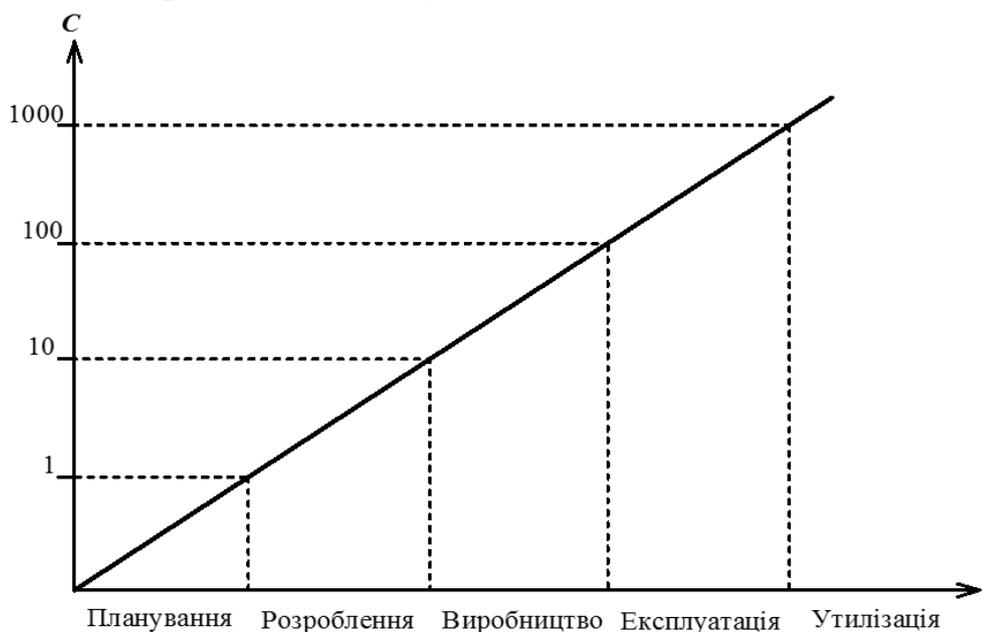


Рис. 1.3. Схема зростання витрат на усунення дефектів, виявлених після закінчення виконання відповідної стадії життєвого циклу

Сукупність значень вказаних параметрів визначає *технічний стан* (або просто *стан*) матеріального об'єкта.

1.1. ЦІЛІ, ЗАДАЧІ ТА ФУНКЦІЇ МОДЕЛЮВАННЯ МАТЕРІАЛЬНИХ ОБ'ЄКТІВ

Найважливіші принципи можуть та мають бути гнучкими

А. Лінкольн¹

Дослідження будь-якого МО, з метою прогнозування його поведінки у визначених умовах, зазвичай, починається з опису цього об'єкту, який можна виконати в найрізноманітніших формах: математичній, графічній, текстовій, фізичній (матеріальній). Такий опис об'єкту називають *моделлю* (математичною, графічною, фізичною тощо). За будь-яких умов процес побудови моделі починається з визначеної ідеалізації об'єкту, тобто з введення тих чи інших припущень щодо його структури, взаємодії окремих складових частин та ін. З цією метою, виділяють *суттєві для розв'язання конкретної задачі* характеристики МО і відкидають другорядні (*несуттєві для даного конкретного дослідження*).

1.1.1. Функції моделі на різних етапах життєвого циклу МО

Під математичною моделлю прийнято розуміти сукупність співвідношень (рівнянь, нерівностей, логічних умов, операторів і т.п.), які визначають характеристики стану об'єкту моделювання, а через них і вихідні значення - реакції, в залежності від параметрів об'єкта-оригінала, вхідних впливів, початкових та граничних умов, часу тощо.

Один і той же МО може бути описаний різними моделями, в залежності від цілей побудови моделі та методів, спрямованих на досягнення цих цілей. В визначеному сенсі можна казати, що *моделі – це посередники, які відображають деякі аспекти реальних явищ*. Такі посередники не несуть нового знання, а лише сприяють його розширенню та конкретизації, стають важливим компонентом узагальненої схеми проектування, виробництва та експлуатації об'єкту.

Цілі, що мають бути досягнуті за результатами моделювання МО, суттєво різняться в залежності від стадії життєвого циклу, на якому таку модель передбачається використовувати.

На стадії проектування - це одержання інформації, яка необхідна для прийняття (та підтвердження ефективності) схемотехнічних, конструкторських та технологічних рішень, завдяки яким досягаються

¹ *Abraham Lincoln* - шістнадцятий президент США

відповідні споживчі характеристики МО, забезпечується можливість використання ефективних методів контролю його параметрів в процесі виробництва та технічного обслуговування під час експлуатації, використання безпечних методів та засобів утилізації тощо.

На цій стадії життєвого циклу математичні моделі забезпечують

- синтез всіх схем (структурної, функціональної, принципової, монтажної, кінематичної та ін.);

- аналіз схем, з метою перевірки забезпечення вихідних характеристик у визначених умовах використання;

- вибір комплектуючих;

- компоновання об'єкту;

- розроблення складових частин;

- розроблення алгоритмів діагностування тощо.

На стадії виробництва - це контроль технологічного процесу та управління ним, визначення відповідності вихідних (споживчих) характеристик МО нормативно-технічній документації, набуття інформації для прийняття оптимальних рішень щодо оперативного внесення необхідних корективів в схеми, конструкцію, технологію за результатами накопиченого досвіду виробництва та експлуатації тощо.

На стадії виробництва модель МО використовується для отримання інформації щодо впливу параметрів технологічних процесів на вихідні характеристики МО, з метою прийняття рішень щодо корекції режимів вказаних процесів. Відмітимо, що цей тип моделей, зазвичай, заснований на глибокому аналізі фізичних та хімічних явищ, що відбуваються у складових частинах МО під час технологічного процесу.

На стадії експлуатації - це підтримання та прогнозування працездатності МО, визначення складових частин, що призвели до втрати МО своїх споживчих властивостей, отримання, накопичення та систематизація інформації, яка має стати основою прийняття рішень щодо модернізації даного виду МО або розроблення більш ефективних об'єктів наступного покоління тощо.

На стадії експлуатації модель МО використовується для контролю його поточного стану з метою визначення місця та характеру відмов, з метою прогнозування рівня споживчих характеристик МО на визначений відрізок часу, з метою отримання необхідної інформації для прийняття рішень щодо здійснення відповідних заходів під час виникнення нештатних ситуацій. За допомогою моделі МО на цій стадії життєвого циклу здійснюють імітацію зовнішніх впливів (завад, характеристик оточуючого середовища, деградації параметрів окремих складових частин, дій персоналу тощо), які можуть виникнути під час функціонування МО. Така імітація є необхідною у випадках, коли штучне створення таких умов на реальному МО утруднене або неможливе (наприклад, може призвести до небажаних чи катастрофічних наслідків). Вказані дії нерідко обумовлені необхідністю визначення

можливості експлуатації МО в умовах, що відрізняються від обумовлених нормативно-технічною документацією.

В загальному випадку модель МО – це деяка допоміжна штучна (математична, віртуальна) або матеріальна (якщо передбачається експериментальне вивчення взаємодії окремих складових МО) система, яка піддається теоретичному та/або практичному вивченню замість вивчення самого МО.

Вказана система має відповідати визначеним вимогам, а саме:

- знаходиться у визначеній об'єктивній відповідності з МО, що вивчається (тобто має бути адекватною МО);

- здатна замінити МО у визначеному сенсі (здатна імітувати «поведінку» МО у визначених умовах виробництва, експлуатації, утилізації тощо);

- здатна забезпечити отримання потрібної інформації щодо самого МО та його складових частин.

Головні функції моделі – спрощення процесу отримання інформації щодо властивостей МО у визначених умовах або, взагалі, забезпечення можливості реалізації такого процесу, управління об'єктом, його оптимізація, прогнозування та діагностика стану та ін.

Модель характеризується відповідною множиною параметрів складових та споживчими (вихідними) характеристиками. Оригінал і модель схожі за одними параметрам та характеристиками й різняться за іншими. Заміщення одного з цих об'єктів іншим правомірне за умови, що властиві їм характеристики, з числа тих, які цікавлять дослідника, визначаються однотипними підмножинами параметрів і пов'язані з цими параметрами однаковими залежностями.

Процес дослідження МО на їх моделях, рівно як і процес побудови моделей реально існуючих МО з метою пояснення та/чи передбачення зміни їх станів називають *моделюванням*. Теорія моделювання – це взаємопов'язана сукупність положень, визначень, методів та засобів створення моделей. Самі моделі є предметом *теорії моделювання*.

На сьогодні не існує єдиної класифікації видів моделювання, що пояснюється багатозначністю самого поняття «модель». Класифікувати види моделювання можна за характером моделей, за характером МО, що моделюються, за областю використання моделі (в техніці, в фізичних науках, в соціології тощо). Наприклад, можна виділити інформаційне, комп'ютерне, математичне, статистичне, структурне, імітаційне та ін. моделювання.

Для формальної класифікації моделей часто використовується класифікаційна ознака, що характеризує математичні засоби, які використовуються для опису стану та взаємодії складових МО. При цьому, наприклад, виділяють моделі лінійні та нелінійні, аналітичні та чисельні, зосереджені та розподілені, детерміновані та стохастичні, статичні та динамічні, дискретні та безперервні тощо. Звичайно, можливі й змішані типи

моделей.

За способом подання МО розрізняють моделі *структурні* та *функціональні*. При цьому *структурні моделі* подають МО як *систему* з відповідною структурою, характер та обсяг якої відображає суттєві для цілей дослідження властивості та взаємні зв'язки окремих складових частин цього МО. В той же час, *функціональні моделі* відображають лише характер реакції МО на зовнішні впливи, подають цей МО як «чорний ящик».

За будь-яких умов *математична модель МО має відповідати конкретним потребам та умовам*, тобто має бути пристосована для розв'язання конкретних задач, наприклад, для опису фізичних процесів, що відбуваються в МО і забезпечують його функціональні можливості; для оцінювання відповідності МО визначеним нормам; для пошуку несправних складових частин; для контролю та керування тощо. Звичайно, для розв'язання різних задач моделі одного й того ж МО будуть суттєво відрізнятися.

Наприклад, для розрахунку показників безвідмовності за відсутністю резервування математичну модель МО можна подати у вигляді послідовного з'єднання складових частин, для яких відомі інтенсивності відмов, тоді як при наявності резервування математична модель цього ж МО подається у вигляді послідовно-паралельного з'єднання відповідних складових частин; для аналізу та/або синтезу МО необхідна модель, що відображає (звичайно, з певною мірою припущень) фізичні процеси, які відбуваються у складових частинах під час роботи цього МО у визначених умовах експлуатації; для пошуку несправної складової в МО при побудові моделі можна обмежитись бінарним відображенням вихідних характеристик його складових і т.п.

1.1.2. Загальні принципи моделювання матеріальних об'єктів

Процес моделювання будь-якого матеріального об'єкту, в узагальненому вигляді, можна подати як сукупність визначених етапів.

Перший етап побудови моделі передбачає наявність визначених знань щодо МО (об'єкту-оригіналу). Пізнавальні можливості моделі визначаються тим, що модель відображає (імітує) якісь суттєві риси об'єкту-оригіналу. Питання щодо необхідної та достатньої міри збіжності оригіналу з моделлю потребує конкретного аналізу. Звичайно, що моделювання втрачає сенс, якщо модель (мається на увазі модель матеріальна) тотожна оригіналу (при цьому вона перестає бути моделлю), або у випадку надмірного та суттєвого (відповідно до поставлених цілей) відрізнення від оригіналу. Отже, зазвичай, в процесі моделювання *вивчення одних сторін МО здійснюється за рахунок відмови від вивчення інших сторін*.

Модель заміщує оригінал лише в строго обмеженому сенсі. З цього випливає, що для одного МО може бути побудовано декілька

«спеціалізованих» моделей, кожна з яких «концентрує увагу» на визначених сторонах (властивостях) цього МО, або характеризує його з різним ступенем деталізації.

На *другому етапі* модель виступає як самостійний об'єкт (артефакт) дослідження. Однією з форм такого дослідження є проведення «модельних» експериментів, під час яких змінюють умови функціонування моделі та систематизують дані щодо її «поведінки». За результатами таких дій отримується інформація щодо властивостей моделі у тих чи інших умовах.

На *третьому етапі* здійснюється перенесення інформації, отриманої за допомогою моделі на оригінал. Тобто формується сукупність інформаційних даних щодо МО. Дані, отримані на другому етапі, при необхідності мають бути кореговані з урахуванням тих властивостей об'єкту-оригіналу, що не знайшли відображення, або були свідомо спотворені (спрощені) під час побудови моделі.

Четвертий етап полягає в практичній перевірці отриманої за допомогою моделі інформації та її використання для побудови узагальненої теорії МО, для його модифікації в потрібному напрямку, для пошуку несправностей (у разі втрати МО працездатності), для цілей керування цим МО тощо.

Моделювання - циклічний процес. Це означає, що за першим чотирьохетапним циклом може слідувати другий, третій і т.д. При цьому інформація щодо МО уточнюється та розширюється, а сама модель удосконалюється. Недоліки, що виявлені після «проходження» першого циклу моделювання та обумовлені малим обсягом інформації щодо МО та/або помилками в побудові моделі, виправляються на наступних циклах.

Для визначення математичної (віртуальної) моделі МО використовують ідеалізацію «вхід - стан – вихід». При цьому математичну модель МО визначають як абстрактне уявлення, згідно якого за допомогою визначеного набору змінних, характеризуються «входи», «виходи» та «стани» об'єкту, а для опису їх взаємодії використовують відповідний математичний апарат.

В загальному випадку процес математичного моделювання МО пов'язаний з необхідністю виконання ряду процедур, які передують безпосередньому застосуванню відповідних математичних процедур моделювання, а саме:

- визначення сукупності параметрів, які виступатимуть аргументами математичної моделі (тобто формування області визначення майбутнього математичного артефакту);

- встановлення оптимальної інформативної сукупності параметрів, які мають визначатися на підставі моделювання (тобто формування області можливих значень майбутнього математичного артефакту);

- визначення критеріїв, вибір методів і засобів порівняння параметрів, що належать до вказаної інформативної сукупності.

Після реалізації вказаних процедур процес математичного моделювання зводиться до встановлення функціонального зв'язку між вибраними параметрами-аргументами та параметрами, які мають визначатися за результатами моделювання. Зазначимо, що параметри, які визначаються за результатами моделювання, як правило, формують інформацію щодо технічного стану МО.

1.2. ВИБІР ПАРАМЕТРІВ МАТЕРІАЛЬНОГО ОБ'ЄКТА ДЛЯ ЦІЛЕЙ МОДЕЛЮВАННЯ

Значущих факторів небагато, а факторів тривіальних величезна множина — лише поодинокі дії приводять до важливих результатів.

NN

Кожен МО характеризується визначеною множиною Y параметрів. До складу цієї множини входять параметри комплектуючих, конструктивних елементів, вихідні параметри складових частин різного ступеню, вихідні параметри МО в цілому. Параметри, що входять до множини Y , можна класифікувати (розділити) за ступенем узагальнення інформації про технічний стан МО. При цьому виділяють [6-8] параметри *первинні*, *вторинні* та *проміжні*.

Первинні параметри мають найнижчу ступінь узагальнення і, по суті, є параметрами комплектуючих та конструктивних елементів МО. Інформація про ці параметри, не надає можливості безпосередньо (лише за наявності такої інформації) визначити стан МО в цілому.

Вторинні параметри мають найвищу ступінь узагальнення інформації. На їх підставі можна безпосередньо зробити висновки щодо поточного стану МО. Це, як правило, вихідні параметри МО. Для визначення цих параметрів використовують терміни: *параметри якості*, *споживчі характеристики*, *споживчі властивості* і т.п.

Проміжні параметри – це параметри, ступінь узагальнення інформації якими має проміжний (між первинними і вторинними параметрами) характер. Як проміжні параметри зазвичай визначають параметри складових частин МО більшого чи меншого ступеня інтеграції. На підставі цих параметрів, як правило, не вдається безпосередньо і в повному обсязі охарактеризувати технічний стан МО. Однак, вони можуть у визначеній мірі використовуватися для характеристики стану МО замість вихідних параметрів, особливо у тих випадках, коли безпосереднє визначення вихідних параметрів надто витратне або, взагалі, неможливе.

Із сукупності всіх параметрів Υ , які характеризують ті чи інші властивості даного МО, виділяють підмножину $\Upsilon_{\text{вих}}$ вихідних параметрів цього МО (або його визначених складових частин). Серед параметрів, що входять до складу множини $\Upsilon_{\text{вих}}$ є сукупність параметрів $\Upsilon_{\text{визн}}$, яка мінімально необхідна для однозначного визначення наскільки правильно даний МО виконує свої функції (теж саме, звичайно, можна сказати й про складові частини МО). Такі параметри, зазвичай, називають *визначальними параметрами*, або *комплексними вихідними параметрами* даного МО (або, відповідно, його складової частини).

Визначальні параметри по відношенню до інших параметрів МО можна вважати параметрами найвищого рівня (найвищого рангу) з точки зору інформативності для розроблення математичної моделі МО. Визначальний параметр може являти собою не тільки поодинокую чисельну величину, а бути набором чисельних величин, функцією, функціоналом, оператором інших параметрів МО.

Виявлення визначальних параметрів іноді потребує проведення суттєвих теоретичних та експериментальних досліджень МО. Адже проектування МО, налаштування процесу його виробництва, контроль технічного стану в процесі експлуатації за умови використання визначальних параметрів МО, зазвичай, виявляється найінформативнішим з технічної точки зору та найефективнішим з точки зору економічних показників.

Для математичного моделювання МО важливим є питання встановлення *мінімальної достатньої кількості параметрів*, за якими необхідно вести відповідні розрахунки та/або спостереження (тобто вимірювати, контролювати, оцінювати та ін.). Так, для пошуку відмов в МО необхідна детальна інформація про стан його складових частин, що, за звичай, потребує контролю більшої кількості параметрів, ніж при визначенні факту працездатності МО. Для проектування МО необхідно мати інформацію про первинні та (за необхідністю про проміжні) параметри. Для прогнозування працездатності МО на визначений подальший відрізок часу необхідно мати інформацію не лише про параметри складових (а то і комплектуючих) МО, але й про закони зміни цих параметрів протягом часу.

Існує чимало методів вибору необхідної сукупності параметрів МО для цілей моделювання [6 – 15]. За всяких умов схема розв'язання задачі вибору визначальних параметрів ґрунтується на виконанні відповідних робіт, до яких, серед іншого, слід віднести:

- аналіз структурної, функціональної, принципової, кінематичної та ін. схем усього МО та відібраних за відповідними критеріями його складових частин;
- визначення ознак станів нормального функціонування МО та/або його складових частин; при цьому під ознакою стану нормальної роботи МО

розуміють визначену сукупність значень параметрів, які відповідають такому стану;

- вибір визначальних параметрів на підставі проведеного аналізу схем та ознак нормальної роботи МО з урахуванням інформативності цих параметрів щодо ідентифікації стану МО;

- аналіз можливих відхилень визначальних параметрів від їх номінальних значень під впливом змінення умов експлуатації, виробничих факторів, природного старіння матеріалів та комплектуючих тощо; при цьому враховуються економічні фактори, пов'язані з виміром того чи іншого визначального параметра, можливість здійснення його контролю за принципом «відповідає нормативним документам – не відповідає нормативним документам» (тобто за принципом «так – ні»); за необхідністю відслідковувати повільне змінення значення відповідного визначального параметра в межах поля допуску з реєстрацією моменту досягнення значення цього параметру граничної межі, враховується методична та технічна можливість реєстрації вказаного моменту з достатньою достовірністю тощо.

З наведеного переліку складових процесу вибору визначальних параметрів видно, що універсалізація цього процесу можлива лише на методологічному рівні, адже різноманіття видів МО робить таку універсалізацію неможливою.

З усієї сукупності існуючих на сьогодні методів вибору визначальних параметрів МО розглянемо принцип Парето, метод експертного дослідження та формування зведеного показника якості МО, які визнані найефективнішим та знайшли найширше використання на практиці.

1.2.1. Принцип Парето

На методологічному рівні до наведених вище складових процесу вибору визначальних параметрів слід додати ще одне універсальне правило, яке на сьогодні широко використовується для розв'язання задач пошуку оптимальних стратегій. Адже саме до таких задач відноситься задача вибору визначальних параметрів для математичного моделювання МО. Таким правилом є *принцип Парето*.

Принцип Парето – це універсальне емпіричне правило (введене італійським соціологом Вільфредо Парето у 1987 році), згідно якого *базовий принцип оптимізації будь-якої діяльності полягає в правильному виборі мінімуму найважливіших дій*.

Сам Парето на підставі спостереження соціальних та економічних процесів встановив співвідношення «важливих» і «неважливих» факторів як 1:4. Звичайно, вказане співвідношення в конкретних випадках потерпає відхилень (доречи не надто великих), але сутність принципу Парето залишається беззаперечним, доведеним практикою фактом.

В контексті вибору визначальних параметрів МО найважливіші

положення, що впливають з принципу Парето, можна подати як:

- параметрів, які надають більшу частину інформації щодо стану МО, за звичай, не багато, тоді як параметрів тривіальних, малоінформативних, малозначущих для визначення стану цього МО значно більше; це стосується і сукупності вихідних параметрів МО, які часто, виявляються пов'язаними один з одним визначеною функціональною залежністю, або визначаються, по більшості, вихідними параметрами одних і тих же складових частин;

- вихідні параметри МО, в окремих випадках, можуть нести менше інформації щодо визначення стану цього МО ніж вихідні параметри його складових частин; наприклад, для визначення працездатності МО, часто, взагалі необов'язково вимірювати його вихідні параметри, а достатньо визначити загальне споживання енергії цим МО, адже відмова електрокомпонента (а тим більше складової частини з помітним споживанням електроенергії), за звичай, супроводжується зміненням режиму споживання енергії МО від джерела, при цьому, звичайно, для більш точного визначення стану МО (місце, причина відмови, характер пошкоджень тощо) будуть необхідні додаткові дії щодо отримання відповідної інформації;

- за результатами дослідження реального МО нерідко виникає необхідність корегування рішень, що приймаються на підставі дослідження його віртуальної моделі; зазвичай це відбувається внаслідок впливу на реальний МО сукупності факторів оточуючого середовища, які не враховані моделлю цього об'єкту; особливо це проявляється коли реакція МО залежить не тільки від дії безпосередньо того чи іншого з окремого фактора, а визначається певною сукупністю та співвідношеннями цих факторів (що на моделі, зазвичай, неможливо реалізувати в повному обсязі).

В загальній постановці задача вибору визначальних параметрів МО полягає у розв'язанні багатокритеріальної задачі максимізації отриманої інформації щодо стану МО при мінімізації ризику «пропустити» значущі фрагменти цієї інформації. Розв'язанням такої задачі оптимізації з визначеною сукупністю критеріїв є *множина Парето*. В цьому контексті, загальна кількість критеріїв дорівнює кількості умов щодо параметрів МО, які розглядаються при виборі визначальних параметрів.

Множина Парето визначається сукупністю умов, які для випадку n критеріїв можна подати у вигляді:

елемент x співвідноситься з елементом y як

- «краще»;

ця категорія означає, що для всіх i з інтервалу $[1, n]$ існує таке k з інтервалу $[1, n]$, що з умови $x_i > y_i$ випливає $x_k \succ y_k$ (тобто за вказаної умови вибір x_k має перевагу, пріоритет перед вибором y_k);

- «гірше»;

ця категорія означає, що для всіх i з інтервалу $[1, n]$ існує таке k з

інтервалу $[1, n]$, що з умови $x_i < y_i$ випливає $x_k \prec y_k$ (тобто за вказаної умови вибір y_k має перевагу, пріоритет перед вибором x_k);

- «не гірше», але й «не краще»,

така категорія означає, що існує таке k з інтервалу $[1, n]$, що $x_k \approx y_k$ (тобто існують умови, за яких не можна визначити пріоритетність якогось одного з цих параметрів, тобто жодному з параметрів x_k, y_k не можна віддати перевагу).

Наведене дозволяє сформулювати правило:

якщо X – є множина всіх можливих параметрів (варіантів) x , то x^ ($x^* \in X$) називають оптимальним за Парето параметром (варіантом), якщо не існує такого $x \in X$, який був би «кращим» за x^* .*

Методологію пошуку множини Парето, тобто оптимальних за Парето елементів, розглянемо на прикладі.

Приклад

Припустимо, що на підставі розгляду схемотехнічних та конструктивних рішень відібрано для проведення моніторингу стану деякого МО шість параметрів. Під час відбирання цих параметрів враховувалась, серед іншого, їх інформативність відносно стану нормального функціонування МО. Виявилось, що деякі з відібраних параметрів мають однакову інформативність. Тобто кількість параметрів з точки зору їх інформативності можна було б зменшити.

Введемо додаткову умову, наприклад, час вимірювання параметру, що безпосередньо пов'язано з економічною ефективністю реалізації процесу пошуку несправної складової МО.

Обидва критерії (інформативність і час) подамо у вигляді дискретних значень (рангів) в діапазоні від 1 до 3. За цих умов кожен з шести відібраних параметрів можна подати як вектор у двомірному просторі, а саме:

(1, 2), (2, 2), (1, 3), (2, 1), (1, 1), (2, 2).

Застосовуючи наведені вище правила, простим перебиранням параметрів можна сформулювати множину Парето:

(2, 2), (1, 3), (2, 2).

Звичайно, в більш складних випадках з великою кількістю параметрів та/або критеріїв метод простого перебору є неприйнятним і для розв'язання задачі пошуку множини Парето використовуються відповідні математичні методи розв'язання багатокритеріальних задач оптимізації.

Відмітимо, що пошук множини Парето або використання будь-якого іншого методу розв'язання багатокритеріальної задачі з вибору визначальних параметрів МО, можливе лише за умови наявності знань щодо кількості

інформації, яку містить той чи інший параметр МО відносно відповідного технічного стану останнього, та після детального аналізу схемних та конструкторських рішень, що прийняті під час проектування МО. Зазвичай це не проста задача, яка потребує визначених і глибоких досліджень МО, але від її розв'язання суттєво залежатиме складність та техніко-економічна ефективність математичної моделі МО, а отже і всіх рішень, що приймаються на її основі.

1.2.2. Метод експертного дослідження

Експертний метод оцінювання МО [16, 17] полягає в реалізації процедури отримання відповідної оцінки (результату) на основі групової думки спеціалістів (експертів). Така комплексна, узагальнена думка, звичайно, точніша (тобто «об'єктивніша») за індивідуальну думку кожного окремого спеціаліста.

В кваліметрії експертний метод використовується для визначення пріоритетності окремих показників якості матеріальних об'єктів або величини вагових коефіцієнтів узагальненого критерію (показника) якості (споживчої якості). Однак експертний метод не належить виключно кваліметрії. Він використовується в техніці (комісії, анкетування), в медицині (консиліуми), в мистецтві та спорті (журі), в соціально-політичній сфері (референдуми), в державному управлінні (колегіальність) та ін.

Весь процес експертного оцінювання зазвичай розділяється на окремі етапи, кожен з яких характеризується відповідною специфікою. Найчастіше виділяється сім таких етапів:

- формулювання цілей;
- вибір форми опитування;
- підготування інформаційних матеріалів, бланків анкет, модератора (менеджера) процедури тощо;
- підбір експертів;
- проведення експертизи;
- оброблення та аналіз результатів;
- підготування звіту за результатами експертного оцінювання.

На методологічному рівні експертне оцінювання полягає у створенні колективного розуму (мультирозуму), що володіє більшими можливостями у порівнянні з можливостями розуму окремої людини. Джерелом надможливостей такого «комплексного» мультирозуму є пошук слабких асоціацій та припущень, заснованих на досвіді окремого спеціаліста. Саме за рахунок створення мультирозуму експертний метод володіє суттєвими можливостями щодо розв'язання задач, які не піддаються розв'язанню звичайними аналітичними або чисельними методами. До таких задач, наприклад, відносяться задачі з

- вибору найкращого варіанту рішення серед наявних;

- визначення вагових коефіцієнтів під час формування узагальнених показників якості МО;

- прогнозування розвинення визначеного процесу, об'єкту;

- пошуку, вибору стратегії розв'язання складних задач;

Перед початком експертного дослідження необхідно чітко визначити його мету та сформулювати відповідне питання експерту. Це є чи не найважливішим та найскладнішим етапом усього процесу експертного дослідження. При цьому рекомендується додержуватися наступних правил:

- час, умови, зовнішні та внутрішні обмеження щодо розв'язання поставленої задачі, мають визначатися чітко та *не припускати* подвійного тлумачення;

- питання мають формулюватися так, щоб потрібна точність відповіді на них відповідала, була доступна людському досвіду;

- пріоритетність має надаватися формулюванню питань у вигляді якісних стверджень; за умови необхідності чисельної оцінки слід запобігати значної кількості градацій (зазвичай використовується не більше 5 – 7 таких градацій).

На всіх етапах експертного дослідження слід пам'ятати, що експерти лише *оцінюють можливі варіанти, а не будують закінчений план дій* щодо реалізації шляхів розв'язання поставлених задач.

Незалежно від цілей та задач використання експертного методу передбачає додержання відповідних загальних умов:

- експертне оцінювання має здійснюватися лише у випадку, коли для розв'язання поставленої задачі не можна використати більш об'єктивні методи (наприклад, аналітичні, експериментальні);

- в роботі експертної комісії не повинно бути факторів, котрі могли б в будь-якій мірі вплинути на судження експертів;

- думки експертів мають бути незалежними;

- питання, що поставлені перед експертами, не повинні припускати різного тлумачення;

- експерти мають бути компетентними у питаннях, що розглядаються, мати досвід практичної роботи у відповідній галузі та, бажано, досвід експертної діяльності;

- кількість експертів має бути оптимальною;

- формулювання питань має передбачати, стимулювати надання експертами однозначних відповідей та забезпечувати можливість їх оброблення математичними методами.

Якісний склад експертної комісії — важлива умова ефективності експертного методу. Вочевидь, що у всіх без винятку випадках експертиза має проводитись грамотними, висококваліфікованими, компетентними в питаннях, які розглядаються, спеціалістами, бажано, з досвідом виконання подібної роботи. За будь-яких умов всі залучені для експертизи проходять спеціальний інструктаж.

На завершальному етапі формування експертної групи доцільно провести

серед експертів відповідне тестування щодо визначення адекватності розуміння ними поставлених цілей та задач, особистого та взаємного оцінювання, узгодженості думок тощо. Таке тестування полягає в розв'язанні експертами задач, подібних реальним, але із заздалегідь відомими (звичайно, не залученим експертам) відповідями. На підставі результатів такого тестування встановлюється компетентність і профпридатність експертів.

Наприклад, з метою визначення особистої оцінки (самооцінки) кожному з експертів пропонується в *умовах обмеженого часу* відповісти на питання спеціальної анкети. За результатами такого опитування перевіряється оцінка кожним експертом своїх професійних знань та ділових рис. Досвід показує, що, не зважаючи на суб'єктивний характер такої оцінки, експертні групи з високими показниками самооцінки експертів в процесі роботи помиляються менше ніж групи з низькими показниками самооцінки.

Тестування експертів щодо їх оцінювання один одного (взаємне оцінювання) також надзвичайно показове та корисне для прийняття остаточних рішень, але може бути реалізоване, звичайно, лише за умов наявності у експертів досвіду спільної роботи.

За умов наявності відомостей щодо результатів роботи експерта в інших експертних групах критерієм його кваліфікації може стати *показник або ступінь надійності*, який визначається відношенням кількості випадків, коли думка експерта співпала з результатами експертизи, до загальної кількості експертиз, в яких цей експерт приймав участь.

Під час реалізації експертного оцінювання кожен експерт надає одне із значень відліку, котрий згідно основному постулату метрології визначається випадковим числом. Порядок і правила подальших дій визначаються прийнятою методикою проведення експертизи.

Слід мати на увазі, що однократне вимірювання експертним методом потребує наявності великого об'єму апріорної інформації. Так, наприклад, під час визначення естетичних показників якості МО суттєве значення матиме художній смак, при визначенні ергономічних показників – власні звички експерта, психологічні особливості характеру, фізіологічні властивості організму тощо. Саме тому, однократне вимірювання використовується рідко, адже суб'єктивні особливості експерта при цьому виступатимуть в якості факторів, що постійно діють та важко піддаються виключенню, компенсації або врахуванню.

Багатократне вимірювання однієї величини (фізичної або будь-якої іншої) постійного розміру або показника якості МО може бути організоване з наступним усередненням експериментальних даних за часом (якщо вимірювання виконуються одним експертом) або за множиною (якщо вимірювання здійснюється одночасно декількома експертами). За такий спосіб вказані величини виступають як випадкові і нівелюються при усередненні. Адже відлік, отриманий групою експертів, подається множиною окремих

значень або законом розподілення ймовірності. За умови великої кількості окремих значень відліку за правилом «трьох сигм» легко виявляються та усуваються хибні значення. Якщо значення відліку визначається нормальним законом розподілу імовірності, то його середнє арифметичне при кількості експертів

$n > 30 \dots 40$ також визначається нормальним законом, а при меншій їх кількості — законом розподілу ймовірностей Стьюдента.

1.3. ТЕОРІЯ ЧОТИРИПОЛЮСНИКІВ. МАТЕМАТИЧНІ МОДЕЛІ «ЧОРНОГО ЯЩИКА»

Відкрити «чорний ящик» може бути не лише складно, але й небезпечно... А чи варто взагалі намагатися його відкрити?

NM

Поняття чотириполюсника виникло як зручне узагальнення характерних властивостей електричних кіл [18].

Чотириполюсником зветься частина електричного кола, яка містить два вхідних виводи (полюси) для підключення джерела енергії (сигналу) та два вихідних виводи для підключення навантаження.

В науково-технічній літературі чотириполюсник іноді зветься 2x2-полюсником, що підкреслює наявність двох виводів на вході та двох – на виході. В іноземній літературі чотириполюсник зветься 2-*port* пристроєм (пристроєм з двома портами, тобто з двома входами, плечима), зважаючи на те, що кожен «*port*» відображає наявність двох виводів (саме з цим пов'язане те, що при перекладі на українську мову чотириполюсник іноді зветься 2-полюсником - двополюсником).

За умови відповідного вибору контрольних перерізів у вхідній та вихідній лініях, які приєднані, відповідно, до входу та виходу чотириполюсника, рівняння, які зв'язуватимуть однакові за змістом величини у вказаних перерізах (тобто струми, напруги, електричні чи магнітні поля), за формою виявляються тотожними рівнянням чотириполюсника.

Причина вказаної тотожності полягає в тому, що рівняння чотириполюсника описують загальні фізичні закономірності, пов'язані з процесами проходження, відбиття, поглинання енергії в лінійних системах, крізь які саме енергія джерела живлення (вхідна енергія) передається до споживача або приймача (вихідна енергія). Такі закономірності виявляються незмінними у всьому освоєному на сьогодні діапазоні частот електричних

коливань, а отже їх можна використовувати в тому діапазоні частот і в тих термінах (категоріях), які в даному діапазоні частот є загальноприйнятими та найзручнішими. Саме завдяки наведеним властивостям чотириполюсники (як і їх узагальнення – багатополюсники) вважають універсальними (у визначеному вище змісті) математичними моделями – математичними моделями «чорного ящика».

1.3.1. Базові поняття та визначення чотириполюсника

Чотириполюсники на сьогодні настільки широко використовуються в процесі розв’язання найрізноманітніших задач, що можна впевнено говорити про формування самостійного розділу аналізу (моделювання) пристроїв з чотирма виводами – теорію чотириполюсників, сутність якої можна подати у вигляді сукупності відповідних базових принципів.

1. Користуючись деякими узагальненими параметрами чотириполюсника, можна визначити напруги та струми на його вході та виході, не розглядаючи процеси та явища, що відбуваються всередині чотириполюсника (тобто чотириполюсник являє собою «чорний ящик» - об’єкт, використання якого не передбачає знання його внутрішньої структури).

2. Складне електричне коло, як правило, можна подати у вигляді сукупності об’єднаних визначеним чином відносно простих за структурою складових – чотириполюсників. Теорія чотириполюсників дозволяє розрахувати параметри такого складного електричного кола на підставі інформації щодо параметрів чотириполюсників, які входять до його структури.

3. Отримані на підставі теорії чотириполюсників значення електричних величин на вході й виході електричного кола дозволяють оцінити режим роботи каналу передачі енергії (сигналів) в цілому. При цьому узагальнений характер параметрів чотириполюсника дозволяє співставляти та адекватно оцінювати передавальні властивості електричних кіл, що різняться за структурою та процесами.

4. На підставі теорії чотириполюсників в окремих випадках вдається на підставі заданих вхідних/вихідних параметрів знайти структуру й значення елементів чотириполюсника (розкрити структуру «чорного ящика»), тобто розв’язати задачу синтезу електричних кіл.

Класифікація чотириполюсників

Чотириполюсники класифікують за різними ознаками. Зазвичай виокремлюють чотириполюсники *активні* та *пасивні*, *симетричні* та *несиметричні*, *лінійні* та *нелінійні*, *зворотні* та *незворотні*, *еквівалентні* та *нееквівалентні*.

Активні чотириполюсники — це чотириполюсники, які містять вітки з

некомпенсованими джерелами електричної енергії. За таких умов на розімкнених виводах чотириполюсника напруга не дорівнюватиме нулю. Тобто активний чотириполюсник можна уявляти як деяке джерело електричної напруги (струму).

Пасивні чотириполюсники або не містять джерел електричної енергії, або сумарна дія наявних в структурі чотириполюсника таких джерел взаємно компенсується всередині чотириполюсника таким чином, що напруга на його вхідних і вихідних виводах дорівнює нулю.

В подальшому розглядатимемо виключно теорію пасивних чотириполюсників. Зазначимо, що будь-який активний чотириполюсник можна замінити еквівалентним йому пасивним чотириполюсником та винесеним за межі цього чотириполюсника еквівалентними джерелами ЕРС.

Симетричний чотириполюсник характеризується тим, що взаємна заміна місць підключення його вхідних і вихідних виводів не змінює розподілення струмів та напруг у зовнішньому (по відношенню до цього чотириполюсника) електричному колі.

В *лінійних* чотириполюсниках зв'язок між струмами та напругами має лінійну залежність. До складу таких чотириполюсників не входять нелінійні елементи, тобто елементи вольт-амперні характеристики яких мають нелінійний характер.

Зворотні чотириполюсники – це чотириполюсники, для яких має силу принцип взаємності, який полягає в тому, що відношення напруги на вході чотириполюсника до струму на його виході не залежить від того, яка пара виводів цього чотириполюсника вибрана як вхідна.

Еквівалентні чотириполюсники – це такі чотириполюсники, при

взаємній заміні яких струми та напруги на відповідних вхідних та вихідних полюсах не змінюються. При цьому, звичайно, не змінюватимуться напруги та струми й у всіх елементах електричного кола, за умови, що такі чотириполюсники є його складовою частиною.

Зазначимо, що внутрішня структура еквівалентних чотириполюсників може бути суттєво різною, але їх заміна один на другий ніяк не впливає на реакцію зовнішніх (по відношенню до клем приєднання чотириполюсників) елементів електричного кола, тобто «не відчувається» електричними елементами, що приєднуються до входів/виходів таких чотириполюсників.

Чотириполюсники також можна підрозділяти за схемами з'єднання елементів, що входять до їх складу. При цьому виділяють *T*-, *Π*- та *Γ*-образні чотириполюсники, мостові чотириполюсники та ін.

1.3.2. Основні рівняння чотириполюсника

Для виводу основних рівнянь, що пов'язують електричні величини (категорії) на вході та виході чотириполюсника, а також встановлення

адекватних співвідношень між цими рівняннями слід уважно відноситися до вибраних напрямків вказаних електричних категорій (напруги та струмів).

На рис. 1.6 представлений чотириполіусник з прийнятими в даному навчальному посібнику напрямками струмів і напруги на виводах чотириполіусника за умови прямого (рис. 1.6а) та зворотного (рис. 1.6б) передавання енергії.

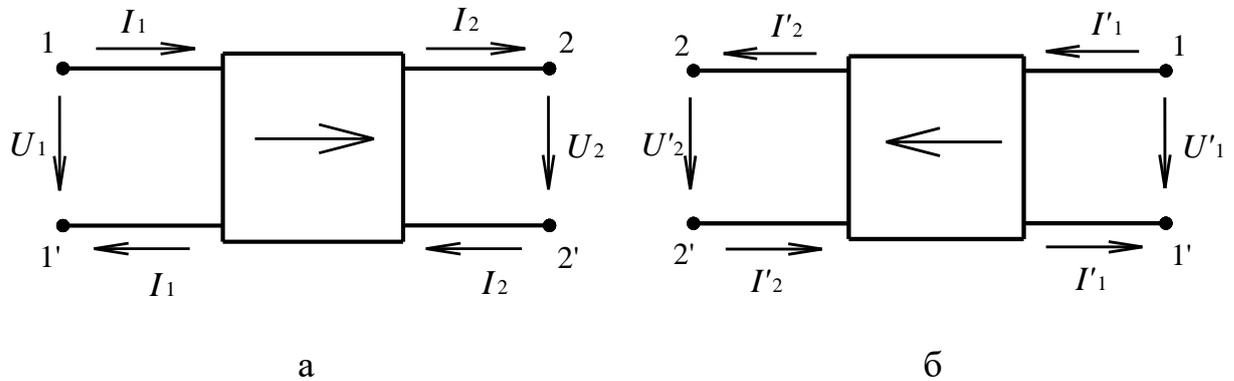


Рис. 1.6. Напрямки струмів та напруги на виводах чотириполіусника за умови передачі енергії в прямому (а) та зворотному (б) напрямках

На сьогодні найбільше поширення отримали три види запису рівнянь, що пов'язують напругу U_1 та струм I_1 на вході чотириполіусника з напругою U_2 та струмом I_2 на його виході.

Зазначимо, що коефіцієнти при величинах U_1, U_2, I_1, I_2 у вказаних рівняннях зазвичай зветься параметрами чотириполіусника. За умови лінійності чотириполіусника ці параметри не залежатимуть від величини напруги та струму на його полюсах.

Перша система рівнянь встановлює зв'язок напруги та струму на вході чотириполіусника з напругою та струмом на його виході, а саме

$$\begin{cases} U_1 = a_{11}U_2 + a_{12}I_2 \\ I_1 = a_{21}U_2 + a_{22}I_2 \end{cases} \quad (1.9)$$

Легко бачити, що у співвідношеннях (1.9)

- величини a_{11}, a_{22} - безрозмірні,
- величина a_{12} має розмірність опору,
- величина a_{21} має розмірність провідності.

Про чотириполіусник, властивості якого описуються співвідношеннями (1.9) зазвичай говорять, як про чотириполіусник заданий a -параметрами.

Слід зазначити, що нерідко (особливо у зарубіжній науково-технічній літературі) співвідношення (1.9) подають у вигляді

$$\begin{cases} U_1 = AU_2 + BI_2 \\ I_1 = CU_2 + DI_2 \end{cases} \quad (1.10)$$

За таких умов говорять, що чотириполюсник заданий $ABCD$ -параметрами.

Зв'язок напруги на вході та виході чотириполюсника зі струмами на його вході та виході подається системою рівнянь

$$\begin{cases} U_1 = z_{11}I_1 + z_{12}I_2 \\ U_2 = z_{21}I_1 + z_{22}I_2 \end{cases} \quad (1.11)$$

Як видно, у співвідношеннях (1.11) коефіцієнти $z_{11}, z_{12}, z_{21}, z_{22}$ при струмах I_1, I_2 мають розмірність опору. Про чотириполюсник, властивості якого описуються співвідношеннями (1.11) зазвичай говорять, як про чотириполюсник заданий z -параметрами (заданий у термінах опору).

Зв'язок струмів на вході та виході чотириполюсника з напругами на його вході та виході подається системою рівнянь

$$\begin{cases} I_1 = y_{11}U_1 + y_{12}U_2 \\ I_2 = y_{21}U_1 + y_{22}U_2 \end{cases} \quad (1.12)$$

Як видно, у співвідношеннях (1.12) коефіцієнти $y_{11}, y_{12}, y_{21}, y_{22}$ при напругах U_1, U_2 мають розмірність провідності. Про чотириполюсник, властивості якого описуються співвідношеннями (1.12) зазвичай говорять, як про чотириполюсник заданий y -параметрами (заданий у термінах провідності).

Звичайно, що параметри чотириполюсника в загальному випадку є комплексними величинами.

Сучасна теорія багатополюсників і, зокрема, чотириполюсників будується на основі використання математичного апарату теорії матриць. В матричному вигляді співвідношення (1.9), (1.11), (1.12) набувають вигляду

$$\begin{bmatrix} U_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \times \begin{bmatrix} U_2 \\ I_2 \end{bmatrix} = [a] \times \begin{bmatrix} U_2 \\ I_2 \end{bmatrix}. \quad (1.13)$$

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \times \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = [z] \times \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}. \quad (1.14)$$

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \times \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = [y] \times \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}. \quad (1.15)$$

Матриці параметрів чотириполюсника, що входять до співвідношень (1.13) – (1.15), мають відповідну назву, а саме

$[a]$ - матриця передачі або матриця a , або a -матриця;

$[z]$ - матриця опорів або матриця z , або z -матриця;

$[y]$ - матриця провідностей або матриця y , або y -матриця.

Зв'язок між елементами окремих матриць чотириполюсника неважко встановити, розв'язуючи ту чи іншу систему рівнянь відносно відповідних вхідних та вихідних категорій (напруг та струмів). Цей зв'язок встановлюється співвідношеннями

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \frac{1}{z_{21}} \begin{bmatrix} z_{11} & -|z| \\ 1 & -z_{22} \end{bmatrix} = \frac{1}{y_{21}} \begin{bmatrix} -y_{22} & 1 \\ -|y| & y_{11} \end{bmatrix}, \quad (1.16)$$

$$\begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \frac{1}{a_{21}} \begin{bmatrix} a_{11} & -|a| \\ 1 & -a_{22} \end{bmatrix} = \frac{1}{|y|} \begin{bmatrix} y_{22} & -y_{12} \\ -y_{21} & y_{11} \end{bmatrix}, \quad (1.17)$$

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \frac{1}{|z|} \begin{bmatrix} z_{22} & -z_{12} \\ -z_{21} & z_{11} \end{bmatrix} = \frac{1}{a_{12}} \begin{bmatrix} a_{22} & -|a| \\ 1 & a_{11} \end{bmatrix}, \quad (1.18)$$

де

$$|y| = \frac{1}{|z|} = y_{11}y_{22} - y_{12}y_{21} = -\frac{a_{21}}{a_{12}},$$

$$|z| = \frac{1}{|y|} = z_{11}z_{22} - z_{12}z_{21} = -\frac{a_{12}}{a_{21}},$$

$$|a| = a_{11}a_{22} - a_{12}a_{21} = -\frac{z_{12}}{z_{21}} = -\frac{y_{12}}{y_{21}}.$$

Наведені вище рівняння (1.16) – (1.18) відповідають передачі енергії в прямому напрямку, тобто згідно рис. 1.6а - зліва направо. За умови передачі енергії у зворотному напрямку (див. рис. 1.6б) вхідні та вихідні виводи чотириполюсника міняються місцями. При цьому слід враховувати змінення напрямку струмів на виводах чотириполюсника. Позначивши величини напруги та струмів, що відповідають зворотному напрямку передачі енергії, індексом «'», виконавши в співвідношеннях (1.16) – (1.17) заміни

$$U'_2 \rightarrow U_1, U'_1 \rightarrow U_2, -I'_2 \rightarrow I_1, -I'_1 \rightarrow I_2,$$

після незначних перетворень, спрямованих на приведення відповідних рівнянь до прийнятої (канонічної) форми запису, отримуємо наступні матричні рівняння, які відповідають зворотному напрямку передачі енергії чотириполіусником:

$$\begin{bmatrix} U'_1 \\ I'_1 \end{bmatrix} = \frac{1}{|a|} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \times \begin{bmatrix} U'_2 \\ I'_2 \end{bmatrix}, \quad (1.19)$$

$$\begin{bmatrix} U'_1 \\ U'_2 \end{bmatrix} = \begin{bmatrix} -z_{22} & -z_{21} \\ -z_{12} & -z_{11} \end{bmatrix} \times \begin{bmatrix} I'_1 \\ I'_2 \end{bmatrix}, \quad (1.20)$$

$$\begin{bmatrix} I'_1 \\ I'_2 \end{bmatrix} = \begin{bmatrix} -y_{22} & -y_{21} \\ -y_{12} & -y_{11} \end{bmatrix} \times \begin{bmatrix} U'_1 \\ U'_2 \end{bmatrix}. \quad (1.21)$$

1.3.3. З'єднання чотириполіусників

Під час математичного моделювання (аналізу) електричних кіл чотириполіусники можна з'єднувати один з одним найрізноманітнішими способами. Найпоширеніші три способи такого з'єднання, які надають можливості подати складне електричне коло як сукупність з'єднаних чотириполіусників. При цьому виникає задача визначення матриці отриманого за результатом з'єднання чотириполіусника на підставі матриць окремих чотириполіусників, з яких саме вказане з'єднання й отримане.

Каскадне (ланцюгове) з'єднання чотириполіусників

Каскадний спосіб з'єднання чотириполіусників показаний на рис. 1.7.

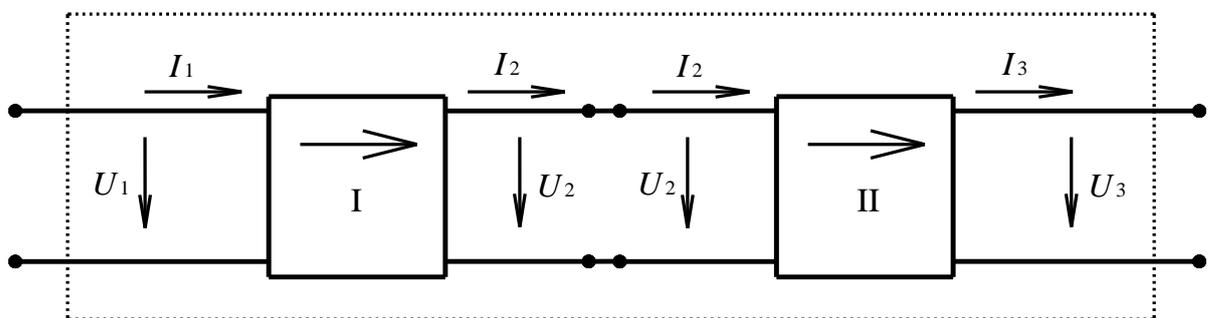


Рис. 1.7. Каскадне з'єднання двох чотириполіусників

За такий спосіб вихід одного чотириполіусника приєднується до входу другого, утворюючи каскад (ланцюг).

Каскадне з'єднання чотириполіусників зазвичай аналізується в термінах

a -матриці (що приводить до найпростішого алгоритму) . Отже задача полягає в тому, щоб на підставі відомих матриць передачі кожного з чотириполіусників, визначити матрицю передачі усього з'єднання.

Враховуючи, що за умови з'єднання чотириполіусників вихідні струми та напруги першого чотириполіусника дорівнюють вихідним струмам та напругам другого чотириполіусника (див. рис. 1.7), відповідні матричні рівняння можна подати як

$$\begin{bmatrix} U_1 \\ I_1 \end{bmatrix} = [a]_1 \times \begin{bmatrix} U_2 \\ I_2 \end{bmatrix}, \quad (1.22)$$

$$\begin{bmatrix} U_2 \\ I_2 \end{bmatrix} = [a]_2 \times \begin{bmatrix} U_3 \\ I_3 \end{bmatrix}. \quad (1.23)$$

Співвідношення для результуючого чотириполіусника встановлюються простою підстановкою матриці стовпця $\begin{bmatrix} U_2 \\ I_2 \end{bmatrix}$ із співвідношення (1.23) до співвідношення (1.22). Як результат отримуємо

$$\begin{bmatrix} U_1 \\ I_1 \end{bmatrix} = [a]_1 \times [a]_2 \times \begin{bmatrix} U_3 \\ I_3 \end{bmatrix}. \quad (1.24)$$

Звичайно, така закономірність утворення матриці каскадного з'єднання зберігатиметься і при збільшенні кількості поєднаних чотириполіусників. Тобто при каскадному з'єднанні чотириполіусників матриця $[a]$ усієї системи дорівнюватиме добутку матриць $[a]$ чотириполіусників, які цю систему утворюють. Отже, при з'єднанні у систему k чотириполіусників маємо

$$[a] = \prod_{i=1}^k [a]_i. \quad (1.25)$$

Зазначимо, що у співвідношенні (1.25) матриці-співмножники мають розміщуватися в тому самому порядку, в якому відповідні чотириполіусники включені до ланцюга в напрямку передачі енергії (адже добуток матриць не підкоряється закону переміщення).

Паралельне з'єднання чотириполіусників

Паралельне з'єднання двох чотириполіусників наведене на рис. 1.8.

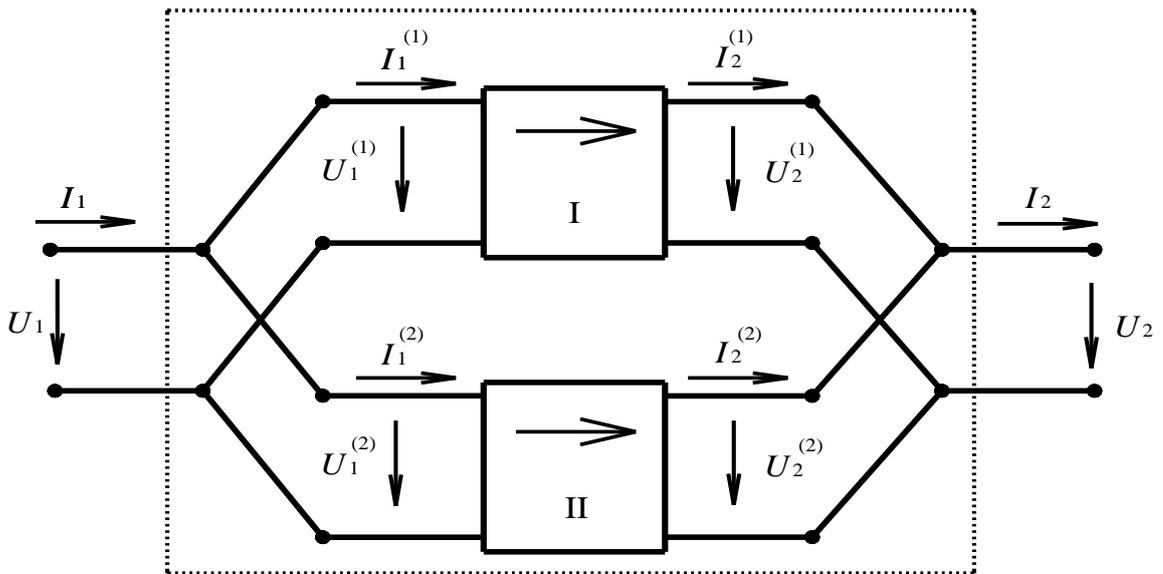


Рис. 1.8. Паралельне з'єднання двох чотириполіусників

При такому способі з'єднання чотириполіусників їх входи та виходи з'єднуються паралельно, причому таке з'єднання має відбуватися саме так, як це показано на рис. 1.8, тобто зі збереженням напрямків напруги та струму на полюсах.

Паралельне з'єднання чотириполіусників зазвичай аналізується в термінах y -матриці (саме це приводить до найпростішого алгоритму). Отже задача полягає в тому, щоб на підставі відомих матриць провідності кожного з чотириполіусників, визначити матрицю провідності усього з'єднання.

Припустимо, що матриці провідності кожного з чотириполіусників визначаються як

$$\begin{bmatrix} I_1^{(1)} \\ I_2^{(1)} \end{bmatrix} = [y]_1 \times \begin{bmatrix} U_1^{(1)} \\ U_2^{(1)} \end{bmatrix}, \quad (1.26)$$

$$\begin{bmatrix} I_1^{(2)} \\ I_2^{(2)} \end{bmatrix} = [y]_2 \times \begin{bmatrix} U_1^{(2)} \\ U_2^{(2)} \end{bmatrix}. \quad (1.27)$$

За умови паралельного з'єднання чотириполіусників вхідні та вихідні напруги першого, другого та результуючого чотириполіусників дорівнюють одна одній, а струми на вході та виході результуючого чотириполіусника дорівнюють сумі відповідних струмів першого та другого чотириполіусника, тобто

$$\begin{aligned}
 U_1^{(1)} = U_1^{(2)} = U_1; \quad U_2^{(1)} = U_2^{(2)} = U_2; \\
 I_1 = I_1^{(1)} + I_1^{(2)}; \quad I_2 = I_2^{(1)} + I_2^{(2)}.
 \end{aligned}
 \tag{1.28}$$

Враховуючи (1.26) – (1.28), отримуємо співвідношення, яким визначається матриця провідності з'єднання двох чотириполіусників, а саме

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = ([y]_1 + [y]_2) \times \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}.
 \tag{1.29}$$

За умови паралельного з'єднання k чотириполіусників усі наведені міркування щодо взаємовідношення струмів та напруг на входах/виходах складових чотириполіусників та результуючої системи чотириполіусників, звичайно зберігатимуться.

Отже для чотириполіусника, отриманого за результатом паралельного з'єднання k чотириполіусників, матимемо

$$[y] = \sum_{i=1}^k [y]_i.
 \tag{1.30}$$

Послідовне з'єднання чотириполіусників

Послідовне з'єднання двох чотириполіусників наведене на рис. 1.9.

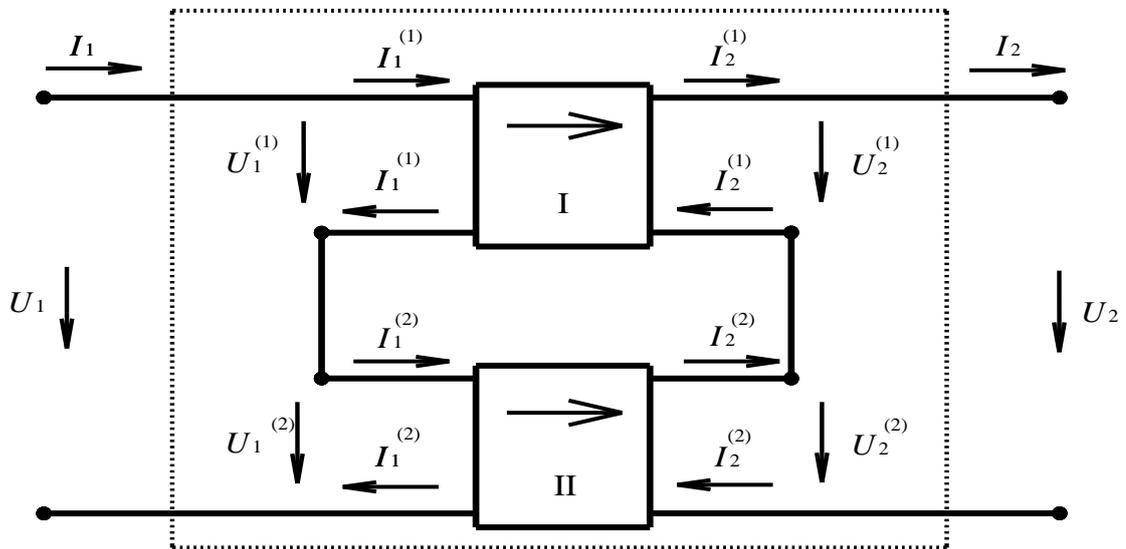


Рис. 1.9. Послідовне з'єднання двох чотириполіусників

При такому способі з'єднання чотириполіусників їх входи та виходи з'єднуються послідовно, а саме так, як це показано на рис. 1.9. тобто з узгодженням напрямків напруги та струму на полюсах.

Послідовне з'єднання чотириполіусників зазвичай аналізується в термінах z -матриці (саме це приводить до найпростішого алгоритму). Отже задача полягає в тому, щоб на підставі відомих матриць опорів кожного з чотириполіусників, визначити матрицю опору усього з'єднання.

Припустимо, що матриці опорів кожного з чотириполіусників визначаються як

$$\begin{bmatrix} U_1^{(1)} \\ U_2^{(1)} \end{bmatrix} = [z]_1 \times \begin{bmatrix} I_1^{(1)} \\ I_2^{(1)} \end{bmatrix}, \quad (1.31)$$

$$\begin{bmatrix} U_1^{(2)} \\ U_2^{(2)} \end{bmatrix} = [y]_2 \times \begin{bmatrix} I_1^{(2)} \\ I_2^{(2)} \end{bmatrix}. \quad (1.32)$$

За умови послідовного з'єднання чотириполіусників напруга на вході та виході результуючого чотириполіусника дорівнюватиме сумі напруг, відповідно, на вході та виході першого та другого чотириполіусників, а струми на вході та виході результуючого чотириполіусника дорівнюють відповідним струмам на вході та виході першого та другого чотириполіусника, тобто

$$\begin{aligned} I_1^{(1)} = I_1^{(2)} = I_1; \quad I_2^{(1)} = I_2^{(2)} = I_2; \\ U_1 = U_1^{(1)} + U_1^{(2)}; \quad U_2 = U_2^{(1)} + U_2^{(2)}. \end{aligned} \quad (1.33)$$

Враховуючи (1.31) – (1.33), отримуємо співвідношення, яким визначається матриця опорів з'єднання двох чотириполіусників, а саме

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = ([z]_1 + [z]_2) \times \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}. \quad (1.34)$$

За умови послідовного з'єднання k чотириполіусників усі наведені міркування щодо взаємовідношення струмів та напруг на входах/виходах складових чотириполіусників та результуючого чотириполіусника, звичайно зберігатимуться. Отже для чотириполіусника, отриманого за результатом послідовного з'єднання k чотириполіусників, матимемо

$$[z] = \sum_{i=1}^k [z]_i. \quad (1.35)$$

Зазначимо, що співвідношення (1.30), (1.35) залишаються справедливими лише за умови, що відповідне з'єднання чотириполіусників не змінює розподілення струмів і напруги в їх структурах.

1.3.4. Вхідний опір чотириполіусника. Режими холостого ходу та короткого замикання

Визначення вхідного опору чотириполіусника за умови прямої та зворотної передачі енергії, а також його аналіз в режимах холостого ходу та короткого замикання надає чимало корисної інформації як для використання системи, що моделюється цим чотириполіусником, так і для визначення параметрів самого чотириполіусників (точніше, для визначення елементів відповідних матриць, якими цей чотириполіусник задається).

Вхідний опір $z_{\text{ВХ}}$ чотириполіусника визначається як відношення напруги до струму на його вхідних полюсах.

За умови, що чотириполіусник включений в режимі прямої передачі енергії (див. рис. 1.6а) відповідно до (1.13) матимемо

$$z_{\text{ВХ}} = \frac{U_1}{I_1} = \frac{a_{11}z_{\text{Н}} + a_{12}}{a_{21}z_{\text{Н}} + a_{22}}, \quad (1.36)$$

де

$$z_{\text{Н}} = \frac{U_2}{I_2} - \text{опір навантаження.}$$

За умови включення чотириполіусника в режимі зворотної передачі енергії (див. рис. 1.6б), використовуючи співвідношення (1.19), отримуємо

$$z'_{\text{ВХ}} = \frac{U'_1}{I'_1} = \frac{a_{11}z'_{\text{Н}} + a_{12}}{a_{21}z'_{\text{Н}} + a_{22}}, \quad (1.37)$$

де

$$z'_{\text{Н}} = \frac{U'_2}{I'_2} - \text{опір навантаження при зворотній передачі енергії.}$$

За умови реалізації режиму холостого ходу ($z_{\text{Н}} \rightarrow \infty$) або короткого замикання ($z_{\text{Н}} = 0$) із (1.36) безпосередньо впливає

$$z_{\text{ХХ}} = \frac{a_{11}}{a_{21}}; \quad z_{\text{КЗ}} = \frac{a_{12}}{a_{22}}. \quad (1.38)$$

При включенні чотириполіусника в режимі зворотної передачі енергії із (1.37) впливає

$$z'_{\text{ХХ}} = \frac{a_{22}}{a_{21}}; \quad z'_{\text{КЗ}} = \frac{a_{12}}{a_{11}}. \quad (1.39)$$

Порівнюючи співвідношення (1.38) і (1.39), неважко помітити, що параметри $z_{\text{ХХ}}$, $z'_{\text{ХХ}}$, $z_{\text{КЗ}}$, $z'_{\text{КЗ}}$ пов'язані між собою співвідношенням

$$\frac{z_{xx}}{z'_{xx}} = \frac{z_{kz}}{z'_{kz}} = \frac{a_{11}}{a_{22}}. \quad (1.40)$$

Співвідношення (1.40), зокрема, означає, що параметри холостого ходу та короткого замикання, визначені в режимах прямої та зворотної передачі енергії, непридатні для опису незворотного чотириполюсника.

Використання режимів холостого ходу та короткого замикання надає можливість встановити фізичний зміст елементів матриць $[a]$, $[z]$, $[y]$, які використовуються для опису його характеристик.

За умови реалізації режиму холостого ходу при прямій передачі енергії вихідні полюси розімкнені, а отже $I_2 = 0$. При цьому згідно (1.9) та (1.11) матимемо

$$a_{11} = \frac{U_1}{U_2}, \quad a_{21} = \frac{I_1}{U_2}, \quad z_{11} = \frac{U_1}{I_1}, \quad z_{21} = \frac{U_2}{I_1}. \quad (1.41)$$

Як видно із співвідношень (1.41) елементи матриць $[a]$ та $[z]$ чотириполюсника, а саме елементи a_{11} , a_{21} , z_{11} , z_{21} , визначають його *властивості в режимі холостого ходу та короткого замикання при прямій передачі енергії*.

Тобто

z_{11} - характеризує вхідний опір чотириполюсника, тобто $z_{11} = z_{xx}$;

a_{11} - характеризує передачу напруги чотириполюсником (визначає коефіцієнт передачі чотириполюсника за напругою;

z_{21} та $a_{21} \left(z_{21} = \frac{1}{a_{21}} \right)$ - визначають відношення напруги на виході до

струму на вході (або навпаки: напруги на вході до струму на виході).

Зазначимо, що параметри типу $\frac{U_2}{I_1}$, $\frac{U_1}{I_2}$, $\frac{I_2}{U_1}$, $\frac{I_1}{U_2}$ зуться, відповідно,

опорами або *провідностями зв'язку* чотириполюсника

За умови, що в режимі прямої передачі енергії, вихідні полюси чотириполюсника короткозамкнені ($U_2 = 0$), тобто реалізується режим короткого замикання, неважко отримати (використовуючи процедуру аналогічну наведеній вище)

$$a_{12} = \frac{U_1}{I_2}, \quad a_{22} = \frac{I_1}{I_2}, \quad y_{11} = \frac{I_1}{U_1}, \quad y_{21} = \frac{I_2}{U_1}. \quad (1.42)$$

Таким чином, елемент y_{11} - вхідна провідність чотириполюсника за

умови короткозамкнених вихідних полюсів; а a_{22} - характеризує передачу струму чотириполюсником (визначає коефіцієнт передачі за струмом); елементи u_{21} , a_{12} - провідність (опір) зв'язку чотириполюсника.

Щоб з'ясувати зміст інших елементів матриць, якими описуються чотириполюсники, достатньо розглянути режими холостого ходу та короткого замикання при зворотній передачі енергії.

Реалізуючи режим холостого ходу при зворотній передачі енергії, тобто прийнявши, що у співвідношеннях (1.20) $I'_2 = 0$, знаходимо

$$z_{22} = -\frac{U'_1}{I'_1}, \quad z_{12} = -\frac{U'_2}{I'_1}. \quad (1.43)$$

З (1.43) видно, що z_{22} є вхідним опором (зі знаком «мінус») чотириполюсника, який працює в режимі холостого ходу при зворотній передачі енергії, тобто $z_{22} = -z_{xx}$, а z_{12} - опір зв'язку (зі знаком мінус) за тих же умов.

1.3.5. Зворотні (взаємні) чотириполюсники

Серед пасивних лінійних чотириполюсників виокремлюють дві основні групи, а саме *зворотні* та *незворотні* чотириполюсники.

Зворотні чотириполюсники підкоряються принципу взаємності, який можна сформулювати як

«якщо деяке джерело ЕРС знаходиться у визначеній вітці деякого електричного кола і викликає визначений струм у другій вітці цього кола, то те саме джерело, розміщене у другій вітці кола, викликати в першій вказаній вітці той самий струм, що і в першому випадку»

Застосування цього принципу до чотириполюсника, вихідні полюси якого короткозамкнені (див. рис. 1.10), призводить до висновку, що відношення напруги на вході чотириполюсника до струму на його короткозамкненому виході буде однаковим для прямого та зворотного напрямку передачі енергії цим чотириполюсником.

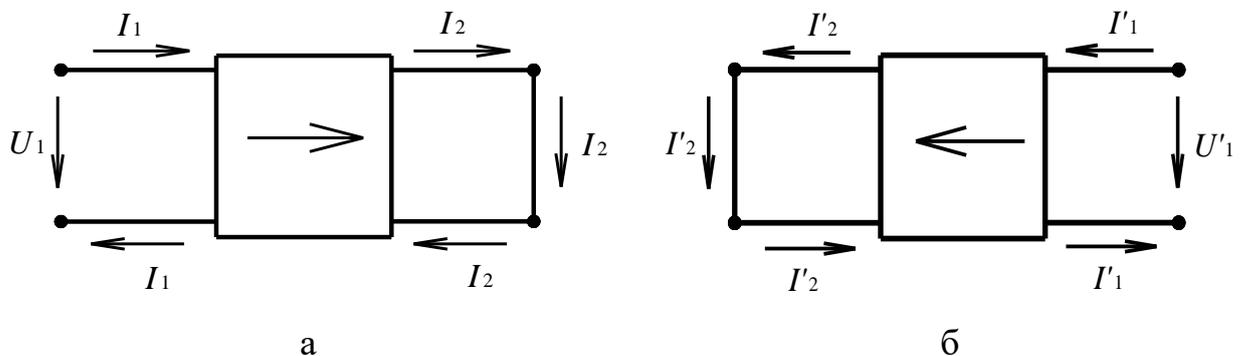


Рис. 1.10. До пояснення властивості зворотності чотириполюсника

З наведеного випливає, що у зворотного чотириполюсника провідності зв'язку для режимів прямої та зворотної передачі енергії однакові, тобто

$$y_{21} = -y_{12}. \quad (1.44)$$

Зі співвідношення (1.18) випливає, що виконання (1.44) можливе лише за умови

$$|a| = 1 \text{ та } z_{21} = -z_{12}.$$

Таким чином, умова зворотності чотириполюсника може бути записана в трьох рівноправних варіантах, а саме

$$y_{21} = -y_{12}, |a| = 1, z_{21} = -z_{12}. \quad (1.45)$$

Зазначимо, що будь-який незворотний чотириполюсник можна подати як каскадне з'єднання зворотної та незворотної частин.

Дійсно, припустимо, що деякий незворотний чотириполюсник описується матрицею

$$[a] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \text{ причому } |a| \neq 1.$$

Ту саму матрицю $[a]$ можна подати як

$$[a] = \begin{bmatrix} \frac{a_{11}}{\sqrt{|a|}} & \frac{a_{12}}{\sqrt{|a|}} \\ \frac{a_{21}}{\sqrt{|a|}} & \frac{a_{22}}{\sqrt{|a|}} \end{bmatrix} \times \begin{bmatrix} \sqrt{|a|} & 0 \\ 0 & \sqrt{|a|} \end{bmatrix}. \quad (1.46)$$

Перша з матриць добутку (1.46) описує зворотну частину чотириполюсника (адже, як легко переконатися, детермінант цієї матриці дорівнює одиниці - $|a| = 1$), а друга – незворотну (детермінант цієї матриці дорівнюватиме $|a| \neq 1$).

Зазначимо, що елемент електричного кола, матриця $[a]$ якого подається як

$$[a] = \begin{bmatrix} \sqrt{|a|} & 0 \\ 0 & \sqrt{|a|} \end{bmatrix}$$

зветься *ідеальним перетворювачем потужності*. Такий перетворювач потужності не впливає на вхідний опір електричного кола і зосереджує в собі виключно незворотні (невзаємні, вентиляльні) властивості цього кола.

1.3.6. Симетричні та антисиметричні чотириполюсники

В групі зворотних (взаємних) чотириполюсників важливими для практики підгрупами є чотириполюсники *симетричної* та *антисиметричної* структури.

Чотириполюсник зветься симетричним, якщо при зміні напрямку передачі енергії цим чотириполюсником, напруги та струми на його полюсах не змінюються.

При виконанні умов симетричності, як випливає з (1.45), виконуються умови

$$|a| = 1, a_{11} = a_{22}, \quad (1.47)$$

$$z_{12} = -z_{21}, z_{11} = -z_{22}, \quad (1.48)$$

$$y_{12} = -y_{21}, y_{11} = -y_{22}. \quad (1.49)$$

Зі співвідношень (1.47) – (1.49) видно, що симетричним може бути виключно зворотний чотириполюсник.

Електрична симетрія чотириполюсника визначається співвідношеннями (1.47) – (1.49) і, звичайно, не потребує його геометричної симетрії. Однак, якщо у зворотному чотириполюснику існує геометрична симетрія, то одночасно існує й електрична симетрія.

Зі співвідношень (1.38), (1.39) та (1.40) випливає також, що у симетричного чотириполюсника опори холостого ходу (короткого замикання) при прямій та зворотній передачі енергії співпадають.

Антисиметричним зветься чотириполюсник, у якого опори холостого ходу при прямій (зворотній) передачі зворотно пропорційні опору короткого замикання при зворотній (прямій) передачі.

З наведеного визначення антисиметричного чотириполюсника без посередньо впливають його основні властивості, а саме

$$z_{xx} z'_{kz} = z'_{xx} z_{kz} = \frac{a_{12}}{a_{21}} = R^2, \quad (1.50)$$

$$\frac{z_{xx}}{R} \frac{z'_{kz}}{R} = \frac{z'_{xx}}{R} \frac{z_{kz}}{R} = 1,$$

де

R - постійна, що має розмірність опору.

На підставі введеного визначення антисиметричного чотириполюсника (1.50) та матричних рівнянь (1.16) – (1.18) легко встановити відповідні зв'язки між елементами його матриць, а саме

$$\frac{a_{12}}{R} = a_{21}R, \quad (1.51)$$

$$\frac{z_{11}}{R} = y_{22}R, \quad (1.52)$$

$$\frac{z_{22}}{R} = y_{11}R. \quad (1.53)$$

Співвідношення (1.52), (1.53) дозволяють також визначити антисиметричний чотириполіусник як такий чотириполіусник, у якого нормований (по R) опір холостого ходу (короткого замикання) при прямій передачі дорівнює таким же чином нормованій провідності короткого замикання (холостого ходу) пр. зворотній передачі.

Переконаємось в тому, що будь-який зворотний несиметричний (в тому числі анти метричний) чотириполіусник можна подати у вигляді каскадного з'єднання симетричного й несиметричного чотириполіусників.

Дійсно, припустимо, що деякий несиметричний чотириполіусник описується матрицею

$$[a] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \text{ причому } a_{11} \neq a_{22}.$$

Ту саму матрицю $[a]$ чотириполіусника можна записати як

$$[a] = \begin{bmatrix} \sqrt{a_{11}a_{22}} & a_{12}\sqrt{\frac{a_{11}}{a_{22}}} \\ a_{21}\sqrt{\frac{a_{22}}{a_{11}}} & \sqrt{a_{11}a_{22}} \end{bmatrix} \times \begin{bmatrix} \sqrt{\frac{a_{11}}{a_{22}}} & 0 \\ 0 & \sqrt{\frac{a_{22}}{a_{11}}} \end{bmatrix}. \quad (1.54)$$

Співвідношення (1.54) можна довести безпосереднім перемноженням матриць.

Перша матриця у правій частині співвідношення (1.54) відповідає симетричному чотириполіуснику (адже елементи головної діагоналі цієї матриці дорівнюють один одному), друга – несиметричному.

Зазначимо, що структура останньої матриці схожа зі структурою матриці ідеального трансформатора (про який буде сказано пізніше), але, на відміну від неї, замість дійсного коефіцієнта трансформації N до складу цієї матриці входить комплексний (в загальному випадку) коефіцієнт асиметрії $\sqrt{a_{11}/a_{22}}$. Чотириполіусник такого типу називають *ідеальним перетворювачем*, його можна розглядати як каскадне з'єднання деякого трансформатора (з

коефіцієнтом трансформації, що в загальному випадку залежатиме від частоти) та фазового контуру (пристрою, який вносить в електричне коло фазовий зсув, не вносячи при цьому активних втрат). Маємо

$$\begin{bmatrix} \sqrt{\frac{a_{11}}{a_{22}}} & 0 \\ 0 & \sqrt{\frac{a_{22}}{a_{11}}} \end{bmatrix} = \begin{bmatrix} \frac{1}{N} & 0 \\ 0 & N \end{bmatrix} \times \begin{bmatrix} e^{-j\varphi} & 0 \\ 0 & e^{j\varphi} \end{bmatrix},$$

де коефіцієнт трансформації та фазовий кут фазового контуру, відповідно, визначаються як

$$N = \sqrt{\frac{a_{22}}{a_{11}}}, \quad \varphi = \arg \sqrt{\frac{a_{22}}{a_{11}}}.$$

1.3.7. Реактивні чотириполюсники

В понятті реактивного чотириполюсника ідеалізовані властивості, які використовуються на практиці під час математичного моделювання пристроїв з малими активними втратами енергії. Така ідеалізація зазвичай суттєво спрощує теоретичні викладки без помітної втрати точності результатів.

В реактивному чотириполюснику поглинання потужності відсутнє і, крім того, таким чотириполюсником не передається потужність в режимах короткого замикання та холостого ходу. Це означає, що всі опори чотириполюсника мають бути чисто реактивними. Інакше кажучи, мають бути реактивними всі елементи матриці $[z]$ такого чотириполюсника, а саме

$$z_{11} = jx_{11}, \quad z_{22} = jx_{22}, \quad z_{12} = jx_{12}, \quad z_{21} = jx_{21}. \quad (1.55)$$

Враховуючи співвідношення (1.55) в формулах (1.16) – (1.18), легко встановити властивості елементів матриці $[a]$ реактивного чотириполюсника, а саме:

a_{11} , a_{22} та $|a|$ - дійсні числа;

a_{12} , a_{21} - уявні числа.

1.3.8. Матриці деяких найпростіших чотириполюсників

В практичній діяльності часто доводиться стикатися з необхідністю визначення матриць, якими описуються властивості найпростіших електричних кіл, що утворюють чотириполюсники та складаються з одного або декількох елементів. Нижче наводяться схеми деяких з таких чотириполюсників та матриці, якими описуються їх властивості.

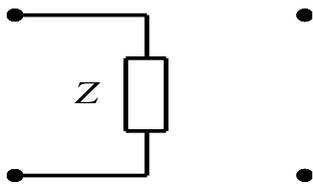


Рис. 1.11. Двополюсник, підключений до лівих полюсів

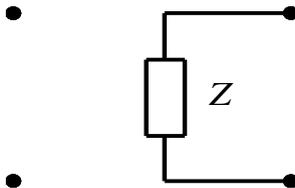


Рис. 1.12. Двополюсник, підключений до правих полюсів

На рис. 1.13 наведено схему чотириполюсника, що складається з одного опору, включеного послідовно відносно вхідних та вихідних полюсів.

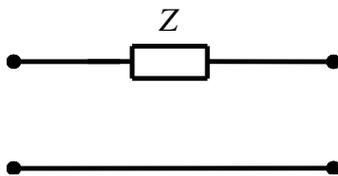


Рис. 1.13. Послідовний опір

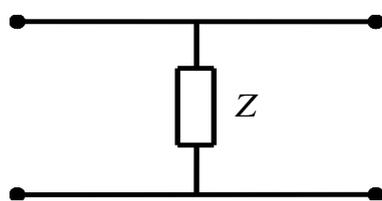


Рис. 1.14. Паралельний опір

На рис. 1.11 наведено схему чотириполюсника, який складається з двополюсника, підключеного до лівих полюсів цього чотириполюсника.

Властивості такого чотириполюсника можна описувати виключно матрицею провідностей, а саме:

$$[y] = \begin{bmatrix} y & 0 \\ 0 & 0 \end{bmatrix}, \text{ де } y = \frac{1}{z}.$$

На рис. 1.12 наведено схему чотириполюсника, який складається з двополюсника, підключеного до правих полюсів цього чотириполюсника.

Властивості такого чотириполюсника можна описувати лише матрицею провідностей, а саме

$$[y] = \begin{bmatrix} 0 & 0 \\ 0 & -y \end{bmatrix}, \text{ де } y = \frac{1}{z}.$$

Властивості такого чотириполюсника можна описувати матрицями передачі та провідностей, а саме

$$[a] = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix}, [y] = \begin{bmatrix} y & -y \\ y & -y \end{bmatrix}, \text{ де } y = \frac{1}{z}.$$

На рис. 1.14 наведено схему чотириполюсника, що складається з одного опору включеного паралельно відносно вхідних та вихідних полюсів.

Властивості такого чотириполюсника можна описувати матрицями передачі та опорів, а саме

$$[a] = \begin{bmatrix} 1 & 0 \\ y & 1 \end{bmatrix}, [z] = \begin{bmatrix} z & -z \\ z & -z \end{bmatrix}, \text{ де } y = \frac{1}{z}.$$

На рис. 1.15 наведено схему чотириполюсника, який складається з двох опорів, з'єднаних Г-подібно (Г-подібне з'єднання опорів).

Такий чотириполусник можна описувати в термінах матриць передачі, опорів та провідностей.

Отже, маємо

$$[a] = \begin{bmatrix} 1 & z_2 \\ y_1 & 1 + z_2 y_1 \end{bmatrix},$$

$$[z] = \begin{bmatrix} z_1 & -z_1 \\ z_1 & -(z_1 + z_2) \end{bmatrix},$$

$$[y] = \begin{bmatrix} y_1 + y_2 & -y_2 \\ y_2 & -y_2 \end{bmatrix},$$

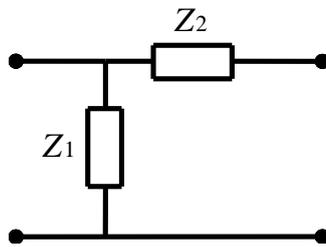


Рис. 1.15. Г-подібна схема з'єднання опорів

де

$$y_1 = \frac{1}{z_1}, \quad y_2 = \frac{1}{z_2}.$$

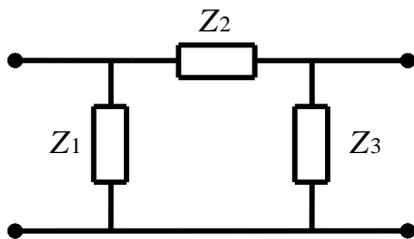


Рис. 1.16. П-подібна схема з'єднання опорів

На рис. 1.16 наведено схему чотириполусника, який складається з трьох опорів, що з'єднані за П-подібною схемою (П-подібне з'єднання опорів).

Такий чотириполусник можна описувати в термінах матриць передачі, опорів та провідностей.

Отже, маємо

$$[a] = \begin{bmatrix} 1 & z_2 \\ y_1 + y_3 + y_1 z_2 y_3 & 1 + z_2 y_1 \end{bmatrix},$$

$$[y] = \begin{bmatrix} y_1 + y_2 & -y_2 \\ y_2 & -(y_1 + y_2) \end{bmatrix},$$

$$[z] = \frac{1}{z_1 + z_2 + z_3} \begin{bmatrix} z_1(z_2 + z_3) & -z_1 z_3 \\ z_1 z_3 & -z_3(z_1 + z_2) \end{bmatrix},$$

де

$$y_1 = \frac{1}{z_1}, \quad y_2 = \frac{1}{z_2}, \quad y_3 = \frac{1}{z_3}.$$

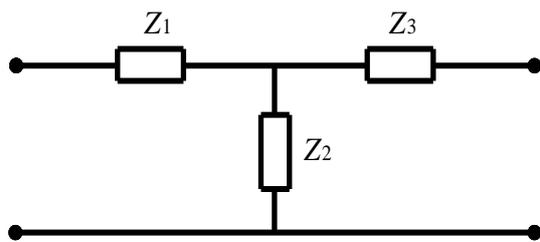


Рис. 1.17. Т-подібна схема з'єднання опорів

На рис. 1.17 наведено схему чотириполосника, який складається з трьох опорів, з'єднаних за Т-подібною схемою (Т-подібне з'єднання опорів).

Такий чотириполосник можна описувати в термінах матриць передачі, опорів та провідностей.

Отже, маємо

$$[a] = \begin{bmatrix} 1 + z_1 y_2 & z_1 + z_3 + z_1 y_2 z_3 \\ y_2 & 1 + y_2 z_3 \end{bmatrix},$$

$$[z] = \begin{bmatrix} z_1 + z_2 & -z_2 \\ z_2 & -(z_2 + z_3) \end{bmatrix},$$

$$[y] = \frac{1}{y_1 + y_2 + y_3} \begin{bmatrix} y_1(y_2 + y_3) & -y_1 y_3 \\ y_1 y_3 & -y_3(y_1 + y_2) \end{bmatrix},$$

де

$$y_1 = \frac{1}{z_1}, \quad y_2 = \frac{1}{z_2}, \quad y_3 = \frac{1}{z_3}.$$

На рис. 1.18 наведено схему чотириполосника, який складається з ідеального трансформатора.

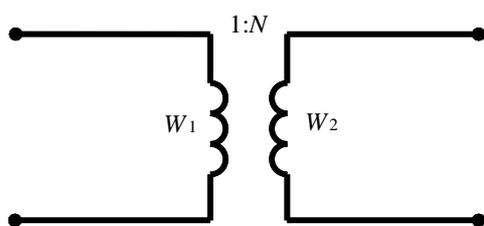


Рис. 1.18. Схема ідеального трансформатора

Такий чотириполосник подається в термінах матриці передачі.

Маємо

$$[a] = \begin{bmatrix} \pm \frac{1}{N} & 0 \\ 0 & \pm N \end{bmatrix},$$

де

$$N = \frac{w_1}{w_2} - \text{коєфіцієнт трансформації}$$

трансформатора,

w_1, w_2 - число витків первинної та вторинної обмоток трансформатора відповідно.

В наведеній матриці передачі ідеального трансформатора знак «плюс» відповідає зустрічному приєднанню обмоток до полюсів чотириполосника, а знак «мінус» – узгодженому.

1.3.9. Узгодження чотириполюсника

На практиці чотириполюсники зазвичай являють собою елементи кризь які відбувається передача електричної енергії. При цьому постає задача встановлення таких параметрів чотириполюсника, за яких ефект передачі енергії був би максимальним. Це приводить до розв'язання задачі узгодження чотириполюсника з зовнішніми електричними колами на його вході та виході. На сьогодні існує декілька варіантів розв'язання такої задачі, що обумовлене конкретними умовами використання чотириполюсника.

Узгодження чотириполюсника на максимум активної потужності на його виході

На практиці чотириполюсники часто використовують для передачі активної потужності. Ефективність такої передачі визначається не тільки власними параметрами чотириполюсника, тобто елементами його матриць передачі, опорів або провідностей, але й властивостями кінцевих пристроїв – опору навантаження та внутрішнього опору джерела електроживлення.

Для ефективної передачі активної потужності необхідно, щоб параметри вхідного та вихідного кола були визначеним чином пов'язані з внутрішніми параметрами чотириполюсника (тобто з елементами матриць, що описують властивості цього чотириполюсника).

Включення чотириполюсника між такими спеціально підібраними кінцевими пристроями зветься *узгодженням*.

Вважатимемо, що будь-яке перетворення опору навантаження та внутрішнього опору генератора здійснюється перетворювачами, які не містять активних втрат, і максимальна потужність генератора залишається незмінною при трансформації його внутрішнього опору.

Узгодження чотириполюсника на максимум активної потужності на виході потребує такої залежності між кінцевими опорами та внутрішніми параметрами цього чотириполюсника, при якій рівень активної потужності на його виході виявляється максимально можливим по відношенню до максимальної потужності джерела живлення (генератора), яка дорівнює

$$P_{\max} = \frac{E^2}{4R_{\Gamma}},$$

де

$R_{\Gamma} = \operatorname{Re}(z_{\Gamma})$ - активна складова опору генератора;

E - електрорушійна сила на виході генератора.

Для аналізу умов такого узгодження скористаємося відомим правилом, згідно якого генератор віддає максимальну потужність у зовнішнє коло лише за умови, коли опір навантаження та внутрішній опір генератора є комплексно спряженими величинами.

Застосувавши наведене правило до вхідного кола чотириполюсника (див. рис. 1.19) знаходимо, що максимальна активна потужність поступатиме у цей чотириполюсник за умови

$$z_{\Gamma}^* = z_{\text{ВХ}}, \quad (1.56)$$

де знаком $(^*)$ позначена комплексно спряжена величина.

Вихідне коло чотириполюсника (див. рис. 1.19), згідно теоремі Тевеніна, можна розглядати як еквівалентний генератор з деяким внутрішнім опором $z_{\text{ВІХ}}$. Цей внутрішній опір є вхідним опором чотириполюсника при зворотному напрямку передачі й незмінних опорах на полюсах. Отже, максимальна потужність, яка надходить з чотириполюсника до навантаження, забезпечуватиметься за умови, що

$$z_{\text{ВІХ}} = z_{\text{Н}}^*. \quad (1.57)$$

За умови одночасного виконання співвідношень (1.56) та (1.57) активна потужність в навантаженні буде максимальною.

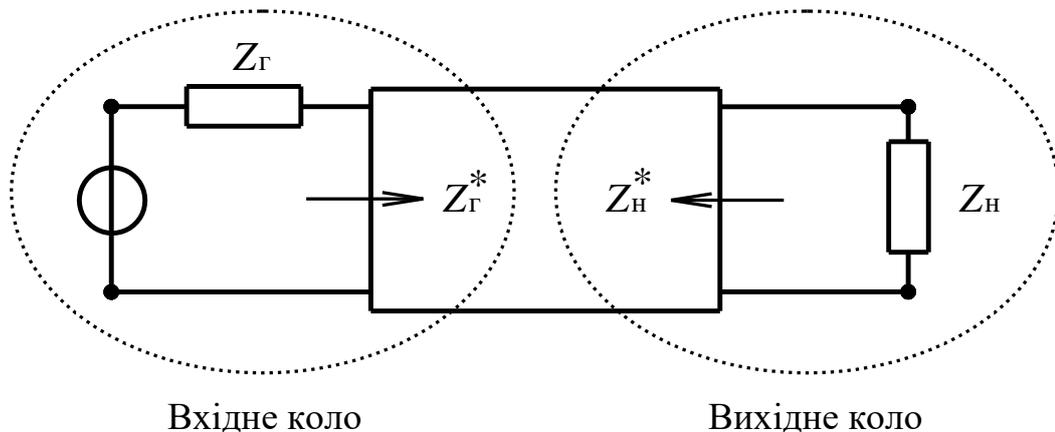


Рис. 1.19. До визначення умов узгодження чотириполюсника на максимум активної потужності на виході

Таким чином, для визначення умов максимальної передачі потужності чотириполюсником необхідно сумісно розв'язати рівняння (1.56) та (1.57) відносно опорів $z_{\text{Н}}$ та z_{Γ} .

Визначаючи $z_{\text{ВХ}}$ та $z_{\text{ВІХ}}$ через елементи матриці $[a]$, як це встановлено (1.36) та (1.37), а також врахувавши (1.56) і (1.57), отримуємо

$$\frac{a_{11}z_{\text{H}} + a_{12}}{a_{21}z_{\text{H}} + a_{22}} = z_{\Gamma}^*; \quad \frac{a_{22}z_{\Gamma} + a_{12}}{a_{21}z_{\Gamma} + a_{11}} = z_{\text{H}}^*. \quad (1.58)$$

Розв'язання системи рівнянь (1.58) визначає оптимальні значення опорів навантаження (z_{HO}) та генератора ($z_{\text{ГО}}$), а саме

$$z_{\text{HO}} = \sqrt{\frac{|a_{22}a_{12}| \cos(\varphi_{22} - \varphi_{12})}{|a_{11}a_{21}| \cos(\varphi_{11} - \varphi_{21})}} e^{j\varphi_{\text{HO}}}, \quad (1.59)$$

$$z_{\text{ГО}} = \sqrt{\frac{|a_{11}a_{12}| \cos(\varphi_{11} - \varphi_{12})}{|a_{22}a_{21}| \cos(\varphi_{22} - \varphi_{21})}} e^{j\varphi_{\text{ГО}}}, \quad (1.60)$$

де

$$\sin \varphi_{\text{HO}} = \frac{F_1}{|z_{\text{HO}}|}; \quad \varphi_{\text{ГО}} = \frac{F_2}{|z_{\text{ГО}}|};$$

$$F_1 = \frac{1}{2} \frac{|a_{11}a_{12}| \sin(\varphi_{11} - \varphi_{22}) + |a_{12}a_{21}| \sin(\varphi_{12} - \varphi_{21})}{|a_{11}a_{12}| \cos(\varphi_{11} - \varphi_{21})};$$

$$F_2 = \frac{1}{2} \frac{|a_{11}a_{22}| \sin(\varphi_{11} - \varphi_{22}) + |a_{12}a_{21}| \sin(\varphi_{12} - \varphi_{21})}{|a_{22}a_{21}| \cos(\varphi_{22} - \varphi_{21})}.$$

Узгодження чотириполюсника на відсутність відбиття від навантаження та генератора

Метод узгодження чотириполюсника з зовнішніми колами розв'язує задачу оптимального енергетичного режиму чотириполюсника. Але побудована відповідно до отриманих результатів система характеризується вузьким діапазоном частот, в якому цей оптимальний енергетичний режим реалізується. Тому суттєве поширення отримав інший метод узгодження чотириполюсника, заснований на попарній рівності опорів у його вхідному та вихідному колах (див. рис. 1.20).

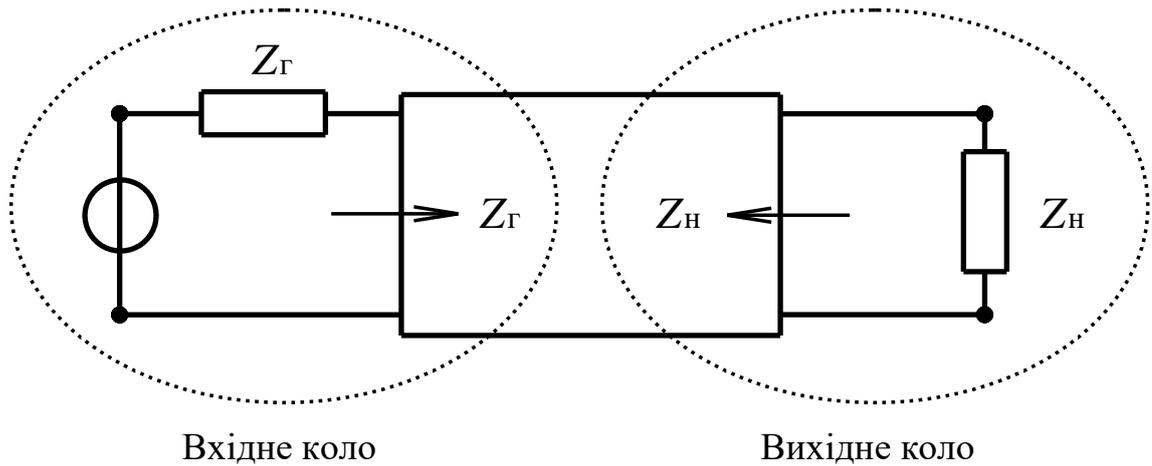


Рис. 1.20. До визначення умов узгодження чотириполюсника на відсутність відбиття

Цей метод вільний від вказаних недоліків, однак не забезпечує повного відбору потужності в системі генератор – чотириполюсник – навантаження. Фізичні основи методу полягають у тому, що чотириполюсник розглядається як деяка однорідна направляюча система. На границях такої системи можливі відбиття. Ці відбиття відсутні, якщо виконуються умови

$$z_{\Gamma} = z_{\text{ВХ}}, \quad (1.61)$$

$$z_{\text{Н}} = z_{\text{ВИХ}}. \quad (1.62)$$

Умовами відсутності відбиття є нульове значення на вході та виході чотириполюсника коефіцієнтів відбиття, а саме

$$\Gamma_1 = \frac{z_{\Gamma} - z_{\text{ВХ}}}{z_{\Gamma} + z_{\text{ВХ}}} = 0; \quad \Gamma_2 = \frac{z_{\text{ВИХ}} - z_{\text{Н}}}{z_{\text{ВИХ}} + z_{\text{Н}}} = 0. \quad (1.63)$$

За рахунок деякого зменшення потужності, що передається від генератора до навантаження на виході чотириполюсника, при такому методі узгодження досягається більша полоса частот узгодженості та підвищується сталість роботи всієї електричної системи.

Враховуючи співвідношення (1.61), (1.62) та визначаючи опори генератора і навантаження в термінах матриці $[a]$, отримуємо

$$\frac{a_{11}z_{\text{Н}} + a_{12}}{a_{21}z_{\text{Н}} + a_{22}} = z_{\Gamma}; \quad \frac{a_{22}z_{\Gamma} + a_{12}}{a_{21}z_{\Gamma} + a_{11}} = z_{\text{Н}}. \quad (1.64)$$

Розв'язок рівнянь (1.64) дає оптимальні (з точки зору відсутності відбиття на вході та виході чотириполюсника) значення опорів навантаження та

генератора, а саме

$$z_H = z_{c2} = \sqrt{\frac{a_{22}a_{12}}{a_{11}a_{21}}}; \quad z_T = z_{c1} = \sqrt{\frac{a_{11}a_{12}}{a_{22}a_{21}}}. \quad (1.65)$$

Оптимальні значення опорів, які визначаються співвідношеннями (1.65) зветься *характеристичними* опорами. Ці опори мають фізичний зміст вхідних опорів нескінченного ланцюга однакових чотириполіусників, включених «назустріч» один одному, адже саме за такий спосіб включення можна досягти виконання умов (1.61) – (1.63).

1.3.10. Робочі параметри чотириполіусника

Робочі (або експлуатаційні) параметри чотириполіусника враховують не лише його внутрішні властивості, але й властивості вхідного та вихідного кола, що підключені до входу та виходу чотириполіусника відповідно.

Введення всіх робочих (експлуатаційних) параметрів базується на порівнянні ефекту передачі в реальній схемі включення чотириполіусника з ефектом передачі крізь деяку стандартну систему, включену між тим самим генератором та навантаженням. Цей метод зазвичай називають методом заміщення.

Робоча постійна передачі

Робоча постійна передачі визначається як половина натурального логарифму (або десятикратне значення десяткового логарифму) відношення потужності, яку видає генератор в узгоджене (в сенсі рівності комплексних опорів) навантаження, до потужності, яка поступає від цього генератора через чотириполіусник в задане навантаження (див. рис. 1.21).

При визначенні робочої постійної передачі визначатимемо потужність як добуток комплексів напруги та струму, що надає можливість оцінювати і амплітудні, і фазові співвідношення.

Таким чином, за умови прямої передачі (відповідні змінні позначатимемо індексом «1») робоча постійна передачі визначатиметься як

$$g_{p1} = b_{p1} + ja_{p1} = \frac{1}{2} \ln \frac{E^2 / (4z_T)}{U_2 I_2},$$

де

b_{p1} - робоче затухання;

a_{p1} - фаза робочої постійної передачі.

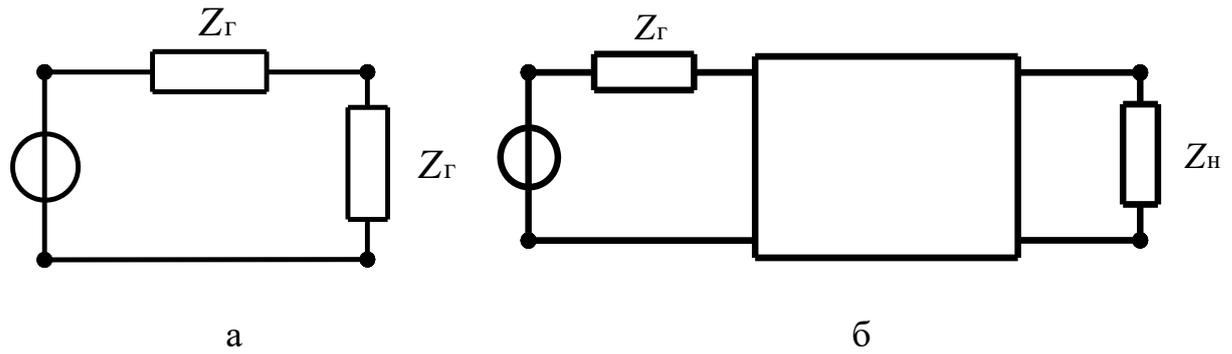


Рис. 1.21. До визначення змісту робочої постійної передачі

Величина робочого затухання визначається відношенням уявних потужностей на вході й виході чотирьохполосника, а саме

$$b_{\text{pl}} = \frac{1}{2} \ln \frac{|E^2 / (4z_{\text{Г}})|}{|U_2 I_2|} \text{ (неп)},$$

або

$$b_{\text{pl}} = 10 \lg \frac{|E^2 / (4z_{\text{Г}})|}{|U_2 I_2|} \text{ (дБ)}.$$

Величина фази робочої постійної передачі характеризує змінення фазових співвідношень у структурі чотирьохполосника і визначається як

$$a_{\text{pl}} = \frac{1}{2} \left\{ \arg \frac{E^2 / (4z_{\text{Г}})}{U_2 I_2} + 2k\pi \right\}.$$

Врахувавши співвідношення

$$U_1 = E - I_1 z_{\text{Г}}, \quad U_2 = I_2 z_{\text{Н}},$$

$$U_1 = a_{11} U_2 + a_{12} I_2, \quad I_1 = a_{21} U_2 + a_{22} I_2,$$

формулу для визначення робочої постійної передачі можна привести до вигляду

$$g_{\text{pl}} = \frac{1}{2} \ln \left\{ \frac{1}{2} \left(a_{11} \sqrt{\frac{z_{\text{Н}}}{z_{\text{Г}}}} + \frac{a_{12}}{\sqrt{z_{\text{Н}} z_{\text{Г}}}} + a_{21} \sqrt{z_{\text{Н}} z_{\text{Г}}} + a_{22} \sqrt{\frac{z_{\text{Г}}}{z_{\text{Н}}}} \right) \right\}^2.$$

Робоче затухання, виражене в децибелах, записується як

$$b_{p1} = 10 \lg L_{p1} =$$

$$= 10 \lg \left| \frac{1}{2} \left(a_{11} \sqrt{\frac{z_H}{z_\Gamma}} + \frac{a_{12}}{\sqrt{z_H z_\Gamma}} + a_{21} \sqrt{z_H z_\Gamma} + a_{22} \sqrt{\frac{z_\Gamma}{z_H}} \right) \right|^2.$$

Функція

$$L_{p1} = \left| \frac{1}{2} \left(a_{11} \sqrt{\frac{z_H}{z_\Gamma}} + \frac{a_{12}}{\sqrt{z_H z_\Gamma}} + a_{21} \sqrt{z_H z_\Gamma} + a_{22} \sqrt{\frac{z_\Gamma}{z_H}} \right) \right|^2 \quad (1.66)$$

зветься *функцією робочого затухання*.

Зазначимо, що під час синтезу електричних кіл за робочими параметрами, функція робочого затухання є заданою величиною.

При зворотному напрямку передачі й незмінних опорах на вході та виході чотириполюсника робоча постійна передачі визначатиметься як

$$g_{p2} = \frac{1}{2} \ln \left\{ \frac{1}{|a|} \frac{1}{2} \left(a_{22} \sqrt{\frac{z_\Gamma}{z_H}} + \frac{a_{12}}{\sqrt{z_H z_\Gamma}} + a_{21} \sqrt{z_H z_\Gamma} + a_{11} \sqrt{\frac{z_H}{z_\Gamma}} \right) \right\}^2.$$

Порівнюючи співвідношення для визначення постійної робочої передачі чотириполюсника в прямому та зворотному напрямку, приходимо до висновку

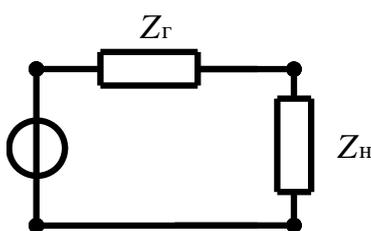
$$g_{p1} - g_{p2} = \ln |a|.$$

Для зворотних чотириполюсників (для яких $|a| = 1$), матимемо

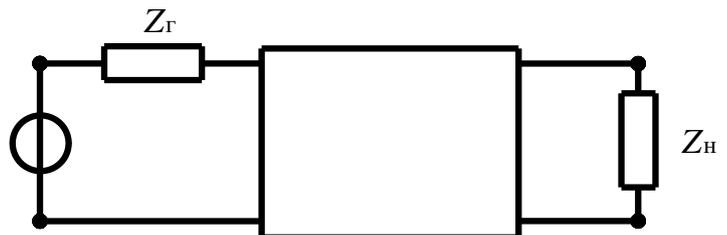
$$g_{p1} = g_{p2}.$$

Внесена постійна передачі

Внесена постійна передачі визначається як половина натурального логарифму (або десятикратне значення десятикового логарифму) відношення потужності, яку віддає генератор в задане навантаження за умови безпосереднього з'єднання цих генератора і навантаження (див. рис. 1.22а), до потужності яка надходить від цього генератора в те саме навантаження крізь даний чотириполюсник (див. рис. 1.22б).



а



б

Рис. 1.22. До визначення внесеної постійної передачі

Як і раніше потужність визначатимемо добутком комплексних значень напруги та струму.

За умови прямої передачі відповідно до визначення внесеної постійної передачі матимемо

$$g_{\text{вн1}} = b_{\text{вн1}} + ja_{\text{вн1}} = \frac{1}{2} \ln \frac{E^2}{U_2 I_2} \frac{z_{\text{н}}}{(z_{\text{г}} + z_{\text{н}})^2},$$

де

$b_{\text{вн1}}$ - внесене затухання;

$a_{\text{вн1}}$ - внесена фаза.

Зазначимо, що за прийнятих умов дослідження процесу передачі крізь чотириполюсник внесене затухання визначає відношення уявних потужностей.

Згідно з визначеннями, робоча та внесена постійні передачі співпадають одна з одною за умови однаковості опорів навантаження та генератора.

Можна показати, що

$$g_{\text{вн1}} = \frac{1}{2} \ln \left\{ a_{11} \frac{z_{\text{н}}}{z_{\text{н}} + z_{\text{г}}} + a_{12} \frac{1}{z_{\text{н}} + z_{\text{г}}} + a_{21} \frac{z_{\text{н}} z_{\text{г}}}{z_{\text{н}} + z_{\text{г}}} + a_{22} \frac{z_{\text{г}}}{z_{\text{н}} + z_{\text{г}}} \right\}^2.$$

За умови змінення напрямку передачі при незмінних опорах на вході та виході чотириполюсника індекси «н» та «г» міняються містами, а всередині матриці передачі відбуваються відповідні перестановки. Як результат матимемо

$$g_{\text{вн1}} = g_{\text{вн2}} = \ln |a|.$$

Зазначимо, що енергетичні оцінки передачі на підставі розглянутих параметрів (власне затухання, робоче затухання, внесене затухання) є неповноцінними за визначенням, адже ці параметри введені на підставі уявних потужностей, тоді як корисний ефект пов'язаний з активною потужністю.

Діюче затухання

Діюче затухання чотириполюсника визначається як половина натурального логарифму (або десятикратне значення десяткового логарифму) відношення максимально можливої активної потужності, яку може віддавати генератор, до тої активної потужності, яка надходить від цього генератора

крізь чотиріполюсник до заданого навантаження.

Тобто

$$b_d = \frac{1}{2} \ln \frac{E^2 / (4R_\Gamma)}{\operatorname{Re}(U_2 I_2^*)}.$$

Діюче затування не є аналітичною функцією комплексної частоти (параметра $j\omega$), що суттєво ускладнює використання цієї характеристики під час синтезу електричного кола на підставі заданої частотної характеристики діючого затування.

Зазначимо, що за умови

$$z_\Gamma = R_\Gamma, \quad z_H = R_H$$

діюче затування співпадає з робочим затуванням.

Опір передачі

Опір передачі або приведений опір чотиріполюсника визначається відношенням ЕРС, що діє на вході чотиріполюсника, до струму в його навантаженні

$$z_{\text{пр}} = \frac{E}{I_2}.$$

Зв'язок цього параметра з параметрами чотиріполюсника та опорам z_Γ , z_H визначається співвідношенням

$$z_{\text{пр}} = a_{11}z_H + a_{12} + a_{21}z_H z_\Gamma + a_{22}z_\Gamma. \quad (1.67)$$

Порівняння співвідношень (1.66) та (1.67) дозволяє встановити зв'язок між опором передачі та функцією робочого затування для зворотного чотиріполюсника, а саме

$$L_p = \left| \frac{z_{\text{пр}}}{2\sqrt{z_H z_\Gamma}} \right|^2. \quad (1.68)$$

Якщо опір навантаження та внутрішній опір генератора є дійсними числами, то (як вже було вказано) діюче та робоче затування співпадають, а отже співпадатимуть і умови найбільшої ефективності передачі та найбільшої зручності реалізації процесу синтезу електричного кола. Саме тому при сучасних методах синтезу електричних кіл з заданими частотними властивостями намагаються включати чотиріполюсник між активними (а якщо це можливе то і однаковими) опорам.

1.3.11. Принципи нормування матриць чотириполюсника

Нормування матриць (рівнянь), якими описуються параметри чотириполюсника, пов'язане з потребою за визначених обставин відобразити ці параметри, спираючись не на категорії струму та напруги (як це було зроблено вище) а на категорію потужності. Така потреба виникає, зокрема, за умови аналізу електричних кіл у надвисокочастотному діапазоні, де поняття струму та напруги, як правило, взагалі не визначені, тоді як потужність легко вимірюється.

Нормування параметрів чотириполюсника полягає в тому, що з основних рівнянь, якими ці параметри описуються, виключаються величини, зміст яких з тих чи інших обставин не є однозначним або складним для вимірювання. При цьому до основних рівнянь чотириполюсника вводиться нова змінна – потужність на полюсах цього чотириполюсника. Для того, щоб врахувати не лише амплітудні, але й фазові властивості чотириполюсника категорію потужності слід розуміти як добуток комплексних значень напруги та струму. Якщо опори зі входу/виходу чотириполюсника – дійсні величини, що дорівнюють ρ , то потужність (в указаному вище сенсі) дорівнюватиме U^2/ρ або $I^2\rho$. Саме ці величини й необхідно ввести до рівнянь, якими описуються параметри чотириполюсника.

Наприклад для рівнянь, які описують параметри чотириполюсника в термінах матриці передачі, процес нормування відображається як

$$\frac{U_1}{\sqrt{\rho_1}} = \left(a_{11} \sqrt{\frac{\rho_2}{\rho_1}} \right) \frac{U_2}{\sqrt{\rho_2}} + \left(\frac{a_{12}}{\sqrt{\rho_1 \rho_2}} \right) I_2 \sqrt{\rho_2} \quad (1.69)$$

$$I_1 \sqrt{\rho_1} = \left(a_{21} \sqrt{\rho_1 \rho_2} \right) \frac{U_2}{\sqrt{\rho_2}} + \left(a_{22} \sqrt{\frac{\rho_1}{\rho_2}} \right) I_2 \sqrt{\rho_2}.$$

Визначимо нормовані напруги та струми як

$$U_1^H = \frac{U_1}{\sqrt{\rho_1}}; \quad U_2^H = \frac{U_2}{\sqrt{\rho_2}}; \quad I_1^H = I_1 \sqrt{\rho_1}; \quad I_2^H = I_2 \sqrt{\rho_2}. \quad (1.70)$$

Коефіцієнти системи рівнянь (1.69) утворюють нормовану класичну матрицю передачі, а саме

$$[A] = \begin{bmatrix} a_{11} \sqrt{\frac{\rho_2}{\rho_1}} & \frac{a_{12}}{\sqrt{\rho_1 \rho_2}} \\ a_{21} \sqrt{\rho_1 \rho_2} & a_{22} \sqrt{\frac{\rho_1}{\rho_2}} \end{bmatrix}. \quad (1.71)$$

В матричній формі рівняння (1.71), якими зв'язані нормовані величини струму та напруги, набувають вигляду

$$\begin{bmatrix} U_1^H \\ I_1^H \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \times \begin{bmatrix} U_2^H \\ I_2^H \end{bmatrix}. \quad (1.72)$$

Зазначимо, що на підставі співвідношень (1.70) – (1.72) функцію робочого затухання чотириполюсника можна подати як

$$L_p = \left| \frac{1}{2} (A_{11} + A_{12} + A_{21} + A_{22}) \right|^2.$$

Аналогічним чином нормуються рівняння, якими визначаються матриці опорів та провідностей, а саме

$$[Z] = \begin{bmatrix} \frac{z_{11}}{\rho_1} & \frac{z_{12}}{\sqrt{\rho_1 \rho_2}} \\ \frac{z_{21}}{\sqrt{\rho_1 \rho_2}} & \frac{z_{22}}{\rho_2} \end{bmatrix}, \quad (1.73)$$

$$[Y] = \begin{bmatrix} y_{11} \rho_1 & y_{12} \sqrt{\rho_1 \rho_2} \\ y_{21} \sqrt{\rho_1 \rho_2} & y_{22} \rho_2 \end{bmatrix}. \quad (1.74)$$

Зі співвідношень (1.71), (1.73), (1.74) видно, що всі елементи нормованих матриць безрозмірні.

Таким чином, нормування рівнянь, що описують чотириполюсники, забезпечує

- перетворення величин напруги і струму на вході/виході чотириполюсника в величину, яка дорівнює кореню квадратному з потужності;

- безрозмірний характер елементів матриць.

Зв'язок між елементами різних типів нормованих матриць чотириполюсника за формою ідентичний зв'язку між елементами ненормованих матриць, але слід пам'ятати, що ці матриці встановлюють

зв'язок між різними характеристиками (категоріями), які діють на входах/виходах чотириполюсника.

КОНТРОЛЬНІ ПИТАННЯ

1. Що таке життєвий цикл матеріального об'єкту?
2. Назвіть основні стадії життєвого циклу матеріального об'єкту.
3. Які етапи входять до стадії проектування матеріального об'єкту?
4. Що таке споживча якість матеріального об'єкту?
5. Що таке споживчі властивості матеріального об'єкту?
6. Надайте сучасне визначення категорії якості матеріального об'єкту?
7. В чому полягає системність поняття якості матеріального об'єкту?
8. Які види моделей МО використовуються на практиці?
9. Які види шкал використовуються на практиці?
10. За якою шкалою можна здійснювати ранжування об'єктів?
11. За якою шкалою можна отримати об'єктивну інформацію щодо параметрів об'єкту?
12. Охарактеризуйте основні етапи моделювання матеріального об'єкту.
13. В чому полягає циклічність моделювання матеріального об'єкту?
14. Чим фізична модель об'єкту відрізняється від самого об'єкту?
15. Що таке узагальнений критерій якості?
16. В чому полягає схема Фішберна визначення узагальненого критерію якості?
17. В чому полягає принцип об'єктивності визначення вагових коефіцієнтів узагальненого критерію якості?
18. Чи можна до критерію якості включати експертні оцінки параметрів об'єкту?
19. Чи можна одночасно використовувати для визначення властивостей об'єкту дані вимірювання та експертного оцінювання?
20. Які параметри матеріального об'єкту вважаються об'єктивними?
21. Які параметри матеріального об'єкту вважаються суб'єктивними?
22. Які параметри матеріального об'єкту відносяться до категорії кількісних?
23. Які параметри матеріального об'єкту відносяться до категорії якісних?
24. Поясніть суть системи інформаційної підтримки життєвого циклу матеріального об'єкту?
25. Яку мету переслідує побудова математичних моделей?
26. Яку мету переслідує побудова фізичних моделей?
27. За яких умов доцільно формувати фізичну модель матеріального об'єкту?
28. Чи може один об'єкт характеризуватися декількома математичними

моделями?

29. Чим розрізняються детерміновані й імовірнісні математичні моделі?

30. Що таке визначальний параметр матеріального об'єкту?

31. Сформулюйте принцип Парето.

32. Що таке нечітке число?

33. В чому полягає сутність методу Дельфи?

34. Наведіть базові засади організації експертного дослідження матеріального об'єкту.

35. В чому полягає метод «мозкового штурму»?

36. Що таке нечіткий інтервал?

37. Що собою являє чотириполіусник?

38. Які види параметрів чотириполіусника найуживаніші на сьогодні?

39. Як узгодити чотириполіусник на максимум передачі активної потужності?

40. Як узгодити чотириполіусник на мінімум коефіцієнту відбиття?

Яке узгодження чотириполіусника більш широкосмугове?

41. Як визначити матрицю передачі каскадного з'єднання чотириполіусників?

42. Як визначається матриця провідності паралельного з'єднання чотириполіусників?

43. Як визначається матриця опорів послідовного з'єднання чотириполіусників?

44. Яку мету переслідує процес нормування матриць чотириполіусника?

45. Які параметри пов'язують нормовані матриці чотириполіусника?

РОЗДІЛ 2

АНАЛІТИЧНІ МЕТОДИ МАТЕМАТИЧНОГО МОДЕЛЮВАННЯ ЕЛЕКТРОМАГНІТНИХ ПРИСТРОЇВ ТА ЕЛЕКТРОМЕХАНІЧНИХ ПЕРЕТВОРЮВАЧІВ ЕНЕРГІЇ

Формула може залишатися незмінною, але її інтерпретації можуть бути різними

А. Ейнштейн

Електромеханічним перетворювачем (ЕМП) називатимемо пристрій, що забезпечує перетворення електричної енергії в механічну або механічної в електричну. Як правило, такі пристрої проявляють властивість зворотності, тобто забезпечують перетворення енергії в обох напрямках. Зазвичай вважається, що електрична енергія підводиться до ЕМП шляхом підключення його до електричного кола, який містить відповідні джерела електроживлення. Тобто електрична енергія підводиться до ЕМП у вигляді електричного струму або електричної напруги.

Електромеханічні перетворювачі, в яких електрична енергія перетворюється на механічну, зводяться *двигунами* або *приводами* (від. англ. *actuator* – актюатор).

Електромеханічні перетворювачі, в яких механічна енергія перетворюється на електричну, зводяться *генераторами* або *сенсорами* (від англ. *sensor*).

Перетворення електричної енергії в механічну (і навпаки) відбувається, як правило, не безпосередньо, а поетапно.

Перший етап полягає в перетворенні електричної енергії в енергію електричного або магнітного поля.

Другий етап – у перетворенні енергії вказаного поля у механічну енергію.

Зворотне перетворення механічної енергії в електричну також зазвичай відбувається через енергію відповідного (електричного або магнітного) поля.

Принциповим фактором, який визначає конструктивні відмінності й області використання конкретного ЕМП, є фізичний принцип взаємозв'язку електричної і механічної енергії. Тому за базову класифікаційну ознаку всіх існуючих ЕМП можна прийняти відповідний фізичний закон, що відображає механізм перетворення одного виду енергії у інший вид.

На сьогодні найчастіше використовуються наступні види перетворювачів механічних коливань, зсувів, обертань тощо в коливання електричного струму (та навпаки):

- електродинамічні;
- електростатичні (конденсаторні);
- п'єзоелектричні;
- магніострикційні.

Принцип дії кожного з цих видів ЕМП базується на відповідному фізичному законі, яким і визначаються конструктивні особливості ЕМП та області його використання. В рамках даного навчального посібника розглядатимемо виключно електромеханічні перетворювачі електродинамічного виду, які на сьогодні знайшли широке застосування у різних галузях науки та техніки. Для первинного ознайомлення з іншими видами перетворювачів можна рекомендувати роботи [19, 20].

До класу електромеханічних перетворювачів енергії електродинамічного виду відносяться, зокрема, електричні машини постійного та змінного струму, значна група електровимірювальних приладів, пристрої телемеханічних систем та систем автоматики (датчики, реле тощо), електроакустичні пристрої (гучномовці, мікрофони тощо), пристрої реєстрації та відтворення інформації та багато іншого.

2.1. БАЗОВІ ФІЗИЧНІ ЗАКОНИ ФУНКЦІОНУВАННЯ ЕЛЕКТРОМЕХАНІЧНИХ ПЕРЕТВОРЮВАЧІВ ЕНЕРГІЇ

Вірно визначте слова та встановіть правила, і звільніть, тим самим, світ від половини непорозумінь

Р. Декарт²

Електромеханічний перетворювач реалізується в найрізноманітніших конструктивних формах, але базові засади його устрою зводяться до переміщення в постійному однорідному магнітному полі відповідної системи провідників. Ці явища, пов'язані зі взаємодією електричних зарядів та магнітних полів, на сьогодні достатньо глибоко вивчені та викладені в теорії електромагнітного поля. Для повноти та цільності викладення подальшого матеріалу, а також з метою введення відповідних визначень та позначень, якими надалі неодноразово користуватимемося, розглянемо базові засади окремих (найсуттєвіших для розуміння подальшого матеріалу) явищ та приклади зовнішнього проявлення вказаної взаємодії магнітних полів з електричними зарядами.

1. Відомо, що в кожному провіднику, який переміщується в магнітному полі, наводиться електрорушійна сила (ЕРС), величина якої пропорційна довжині провідника, швидкості його руху та інтенсивності магнітного поля, причому напрямок цієї індукованої ЕРС залежить від напрямку переміщення провідника відносно магнітного поля (закон Фарадея).

Припустимо, що рівномірне магнітне поле представлено магнітними силовими лініями, як це зроблено на рис. 2.1, а його інтенсивність характеризується вектором магнітної індукції B .

² Древньогрецький філософ

Припустимо, далі, що провідник ab довжиною l , який розміщений перпендикулярно силовим лініям магнітного поля, переміщується зі швидкістю v під кутом α до магнітних силових ліній (див. рис. 2.1).

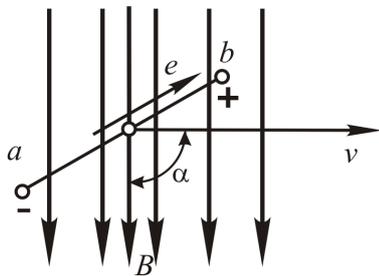


Рис. 2.1. Схема виникнення ЕРС при переміщенні провідника в магнітному полі

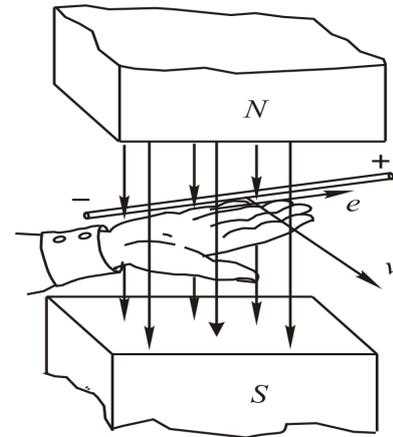


Рис. 2.2. Ілюстрація правила долоні правої руки

За вказаних умов величину ЕРС e , яка наводиться у провіднику ab , можна визначити за формулою,

$$e = Blv \sin \alpha .$$

За умови, що вектор швидкості перпендикулярний провіднику (тобто $\alpha = 90^\circ$), матимемо

$$e = Blv .$$

При цьому одиниці виміру фізичних величин, що входять до наведених співвідношень, визначатимуться як

e - ЕРС, що наводиться у провіднику – В;

B - індукція магнітного поля – Тл;

l - довжина провідника – м;

v - швидкість переміщення провідника – м/с;

α - кут між напрямком вектора швидкості та напрямком магнітних силових ліній (вектором індукції) магнітного поля – радіани або градуси.

Полярність індукованої ЕРС прийнято визначати за *правилом долоні правої руки* (див. рис. 2.2), яке формулюється як

якщо долонь правої руки розмістити так, щоб лінії магнітного поля входили в долонь, а відігнутий великий палець вказував напрямок переміщення провідника, то чотири витягнуті пальці вказуватимуть напрямком ЕРС (кінець провідника додатної полярності).

2. Теорія електромагнітного поля вказує на наявність силової взаємодії розміщеного в магнітному полі провідника зі струмом, і аналітично визначає її кількісну величину. А саме: якщо провідник, яким протікає струм,

знаходиться в магнітному полі, то на нього діє механічна сила, перпендикулярна силовим лініям (вектору індукції) цього поля та провіднику (див. рис. 2.3).

Величина цієї сили пропорційна інтенсивності (індукції) магнітного поля B , довжині провідника l та струму i , що протікає вказаним провідником (закон Біо-Савара-Лапласа).

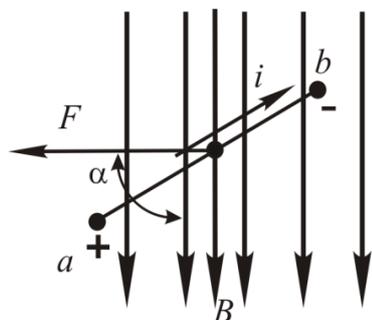


Рис. 2.3. Схема виникнення сили, яка діє на провідник зі струмом у магнітному полі

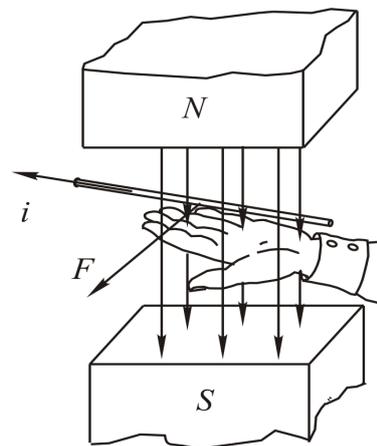


Рис. 2.4. Ілюстрація правила долоні лівої руки

Отже, на розміщений в рівномірному магнітному полі з індукцією B провідник ab , довжиною l , яким протікає струм i , діє механічна сила F , величина якої визначається співвідношенням

$$F = Bli$$

за умови, що вектор магнітної індукції та провідник взаємно перпендикулярні.

Зазначимо, що вектор сили завжди буде перпендикулярний до провідника та вектора індукції.

Напрямок механічної сили, що діє на розміщений в магнітному полі провідник зі струмом, прийнято визначати за *правилом долоні лівої руки* (див. рис. 2.4):

якщо долонь лівої руки розмістити так, щоб силові лінії магнітного поля входили у долонь, а чотири пальці вказували напрямку струму у провіднику, то відігнутий великий палець вказуватиме напрямку сили, яка діятиме на вказаний провідник.

3. Найчастіше описані в пп. 1, 2 явища проявляється одночасно, супроводжують одне одного.

Розглянемо переміщення в магнітному полі під дією зовнішніх сил деякого провідника, яким тече визначений струм (див. рис. 2.5).

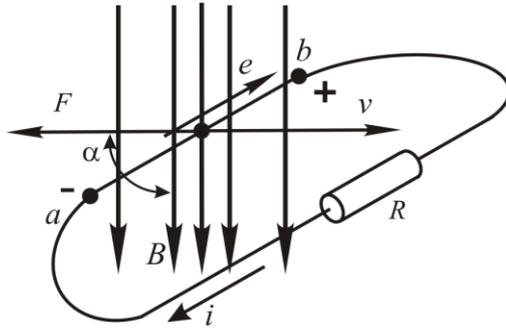


Рис. 2.5. Провідник зі струмом, що рухається в магнітному полі

Відповідно до явища, описаного у п. 1 вплив магнітного поля на провідник, що рухається в цьому полі, приводить до появи ЕРС, величина якої за умови перпендикулярності векторів швидкості провідника та індукції магнітного поля становитиме

$$e = Blv.$$

Якщо зовнішніми провідниками підключити до вказаного провідника резистор, опір якого дорівнює R , як це показано на рис. 2.5, то цим провідником потече струм, величина якого визначатиметься як

$$i = \frac{e}{R}.$$

За таких умов провідник, який рухається в магнітному полі, слід одночасно розглядати як провідник зі струмом, що знаходиться в магнітному полі. При цьому на цей провідник діятиме механічна сила F , яка спрямована протилежно напрямку переміщення провідника та є силою опору зовнішньому впливу, який саме й переміщує провідник зі швидкістю v .

Отже сила F є силою механічної протидії (механічного опору) руху провідника. Саме така ситуація притаманна всім електричним машинам, які перетворюють електричну енергію в механічну та механічну енергію в електричну.

4. Нагадаємо про відому з курсу фізики властивість впорядкованого руху електричних зарядів, а саме

навкруги будь-якого провідника зі струмом існує магнітне поле (див. рис. 2.6а), напруженість якого пропорційна величині цього струму.

Для визначення напрямку вектора напруженості магнітного поля H , що утворюється навколо провідника зі струмом, використовують правило правоходового («правого») гвинта (див. рис. 2.6а):

якщо «правий» гвинт обертати так, щоб напрямок його переміщення співпадав з напрямком струму, то напрямок обертання вказуватиме напрямок силових ліній магнітного поля

За умови згорнутого у контур провідника (рис. 2.6б) струм, який протікає цим провідником, утворює магнітне поле, магнітний потік якого пропорційний величині вказаного струму.

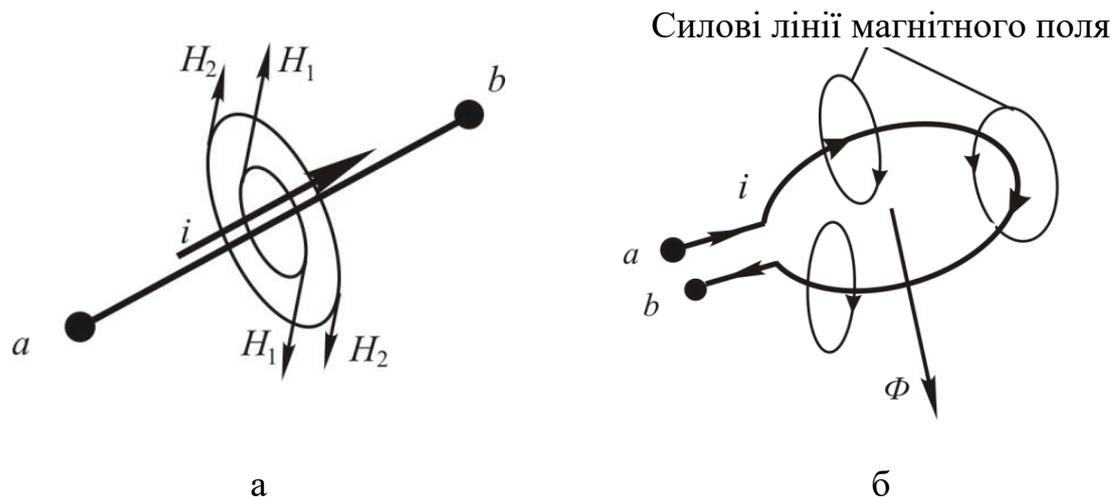


Рис. 2.6. Магнітне поле навкруги прямолінійного (а) та згорнутого у контур (б) провідника зі струмом

Якщо магнітний потік індукції магнітного поля крізь обмежену контуром поверхню зобразити відповідним вектором Φ , то напрямок цього вектора буде таким, як показано на рис. 2.6б. Для визначення напрямку цього вектора зазвичай використовують те саме правило «правого» гвинта.

Взаємодією магнітного поля та струму, що протікає розміщеним у цьому полі провідником, пояснюється й наступне явище, яке дозволяє зрозуміти принцип дії електричних машин. Мова йде про взаємодію зовнішнього магнітного поля зі струмом провідника, виконаного в формі прямокутної рамки або витка.

Якщо виток провідника зчеплений з магнітним полем, магнітний потік якого проходить крізь площу витка і змінюється у часі, то в цьому витку наводиться ЕРС, величина якої пропорційна швидкості змінення вказаного магнітного потоку, а саме

$$e = -\frac{d\Phi}{dt}.$$

При цьому ЕРС має такий напрямок, що при замиканні контуру витка в ньому протікатиме струм, власне магнітне поле якого буде *перешкоджати* зміненню магнітного потоку. Відповідно до вибраного напрямку магнітного потоку при його збільшенні потенціал клеми *a* витка буде вищий за потенціал клеми *b*. При зменшенні потоку полярність ЕРС змінюється на протилежну.

Незважаючи на загальні базові фізичні принципи роботи всіх типів електромеханічних перетворювачів, математичне моделювання кожного з них має свої характерні особливості. В рамках даної роботи переважно розглядатимемо математичне моделювання електромеханічних перетворювачів, що за конструктивним виконанням відповідають

електричним машинам постійного та змінного струму. До складу цих машин входять рухома (ротор) та нерухома (статор) частини, які в процесі роботи обертаються одна по відношенню до другої навколо загальної осі.

2.2. БАЗОВІ ПРИНЦИПИ ФУНКЦІОНУВАННЯ ЕЛЕКТРОМЕХАНІЧНИХ ПЕРЕТВОРЮВАЧІВ ЕНЕРГІЇ

У будь-якій теорії значущих принципів не багато, і лише окремі з них дозволяють зануритися у суть речей та явищ

NM

Крім розглянутих основних законів електромеханіки для розуміння процесів, що відбуваються в електромеханічних перетворювачах енергії, слід враховувати *три* базові принципи (положення), а саме:

1. Електромеханічне перетворення енергії не може здійснюватися без втрат. Коефіцієнт корисної дії (ККД) такого перетворення за будь-яких умов залишається меншим 100 %;
2. Всі електричні машини зворотні, одна й та сама машина може працювати як в режимі двигуна, так і в режимі генератора;
3. Електромеханічне перетворення енергії здійснюється нерухомими один до одного полями.

Перший принцип електромеханіки відображає загальний закон збереження енергії.

В електромеханічних перетворювачах енергія одного виду (електрична енергія, що надходить від зовнішнього електричного кола з відповідним джерелом електроживлення) перетворюється в енергію двох (рідше трьох чи більше) видів. В цьому виді перетворювачів основне (корисне) перетворення вхідної енергії супроводжується деяким додатковим (непродуктивним, супутнім) перетворенням цієї енергії (зазвичай на теплову).

Отже, в електромеханічних перетворювачах перетворення електричної енергії $W_{\text{ел}}$ в механічну енергію $W_{\text{мех}}$ (рівно як і механічної енергії на електричну) завжди супроводжується виділенням теплової енергії $W_{\text{теп}}$. З цієї точки зору узагальнену схему перетворення енергії в електромеханічному перетворювачі можна подати як на рис. 2.7.

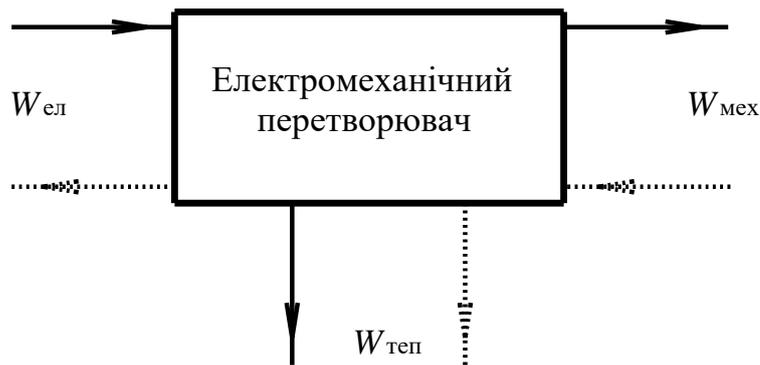


Рис. 2.7. Схема перетворення енергії в ЕМП

Згідно наведеного рис. 2.7, електромеханічний перетворювач можна подати як деяку систему, пристрій з шістьма виводами (шестиполіусник), в якому кожна пара виводів імітує відповідні впливи або реакції.

На практиці під час проектування та експлуатації ЕМП зазвичай вважається, що витoki теплової енергії пов'язані з деяким контуром, об'єм якого нескінченний, а температура незмінна.

За таких припущень ЕМП можна розглядати як чотиріполіусник (див. рис. 2.8). Як правило така модель ЕМП приймається при розв'язанні задач, коли процеси перетворення енергії всередині ЕМП не мають значення.

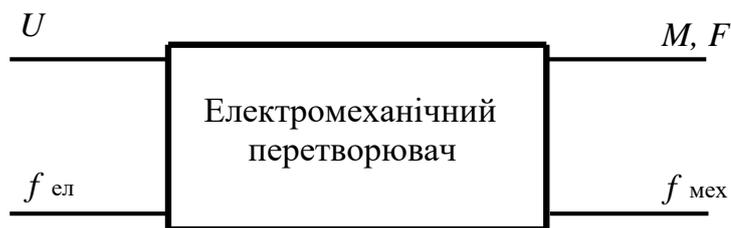


Рис. 2.8. ЕМП як чотиріполіусник

Згідно першому принципу електромеханіки в ЕМП відбувається (за будь-яких умов) перетворення частини електричної або частини механічної енергії на теплову енергію. Це режим холостого ходу ЕМП.

Потік теплової енергії зазвичай не змінює свого напрямку. Отже втрати в ЕМП незворотні. Однак на сьогодні існують перетворювачі, які перетворюють теплову енергію в електричну та механічну.

Другий принцип електромеханіки відображає зворотний характер процесів, що відбуваються в електромеханічному перетворювачі.

Робота будь-якого ЕМП в двох режимах є найважливішою перевагою таких перетворювачів перед перетворювачами енергії інших видів.

За умови, що до ЕМП підводиться електрична енергія з електромережі, він працюватиме в режимі *актюатора* (двигуна), якщо ж до ЕМП підводиться механічна енергія – встановлюється режим *сенсора* (генератора, генерації

електричної енергії).

Третій принцип електромеханіки характеризує просторову взаємодію магнітних полів, якими супроводжується процес перетворення енергії в електромеханічному перетворювачі.

Нерухомі відносно одне одного магнітні поля створюють в структурі електромеханічного перетворювача результуюче поле та електромагнітний момент, а поля, які переміщуються в повітряному зазорі (який обов'язково присутній в ЕМП всіх видів), створюють потік теплової енергії, що приводить до втрат та зменшенню ККД.

2.3. ЕНЕРГІЯ, МЕХАНІЧНІ ТА ЕЛЕКТРОРУШІЙНІ СИЛИ В ЕЛЕКТРОМЕХАНІЧНИХ СИСТЕМАХ

При переміщенні рухомої частини ЕМП відбувається обмін енергією між електричним колом, механічною системою та електромагнітним полем. За будь-яких умов введені в процесі математичного моделювання електромеханічної системи спрощення та припущення щодо конструктивного виконання та врахування тих чи інших явищ, що супроводжують процес функціонування такої системи, не можуть порушувати основні закони фізики та наведені вище принципи електромеханіки.

2.3.1. Баланс видів енергії

Зосередившись на корисному перетворенні енергії в ЕМП, нехтуватимемо тепловими втратами, а активний (резистивний) опір будемо вважати частиною опору зовнішнього електричного кола.

Крім теплових втрат в реальних ЕМП завжди мають місце втрати, що обумовлені

- змінням енергії електричного поля,
- випромінюванням енергії, яке викликане змінням у часі електричного та магнітного полів,
- втратами енергії в магнітному полі.

Випромінюванням енергії на промислових частотах можна нехтувати. Енергія електричного поля в електромагнітних пристроях також мала, внаслідок малих значень ємностей, які зазвичай існують між окремими елементами електричних кіл та притаманні визначеним конструктивним виконанням окремих елементів (наприклад, ємність між витками котушки).

Втрати на гістерезис та вихрові струми в магнітних колах більшості пристроїв, що розглядатимуться, також невеликі. Цими втратами на даному етапі аналізу базових принципів моделювання ЕМП також нехтуватимемо.

З урахуванням прийнятих припущень рівняння балансу відображатиме зміння основних видів енергії в електромеханічній системі і може бути подане як

$$\Delta W_{\text{ел}} \approx \Delta W_{\text{маг}} + \Delta W_{\text{мех}}, \quad (2.1)$$

де

$\Delta W_{\text{ел}}$ - змінення електричної енергії, що подається до ЕМП;

$\Delta W_{\text{маг}}$ - змінення енергії магнітного поля, яке саме й спричиняє відповідні перетворення електричної енергії у механічну;

$\Delta W_{\text{мех}}$ - змінення механічної енергії, яке спрямовується на виконання визначеної роботи відповідними конструктивними елементами ЕМП.

Знак приблизної рівності у (2.1) відображає наявність прийнятих припущень.

З (2.1) видно, що за умови, коли електрична енергія $\Delta W_{\text{ел}}$, яка підводиться до ЕМП, більша за енергію $\Delta W_{\text{маг}}$ магнітного поля, її частина переходить у механічну енергію $\Delta W_{\text{мех}}$. Тобто магнітна енергія (енергія магнітного поля) є, так би мовити, проміжним елементом в процесі перетворення електричної енергії в механічну.

Таким чином, механічна енергія (точніше величина змінення механічної енергії) визначається як функція струму i , що протікає у відповідних елементах ЕМП, та координати x його рухомої частини. Замість струму можна використати й категорію потокозчеплення Ψ , яке однозначно пов'язане зі струмом у відповідному елементі ЕМП.

Таким чином рівняння балансу енергії можна подати в термінах струмів у вигляді

$$\Delta W_{\text{мех}}(i, x) = \Delta W_{\text{ел}}(i, x) - \Delta W_{\text{маг}}(i, x), \quad (2.2)$$

або в термінах потокозчеплення як

$$\Delta W_{\text{мех}}(\Psi, x) = \Delta W_{\text{ел}}(\Psi, x) - \Delta W_{\text{маг}}(\Psi, x). \quad (2.3)$$

Середня сила, яка реалізується в процесі перетворення (2.2) або (2.3), визначається із співвідношення

$$\Delta W_{\text{мех}} = f_{\text{сер}} \Delta x, \quad (2.4)$$

де

Δx - змінення координати x відповідного конструктивного елемента ЕМП в процесі його функціонування.

З урахуванням (2.2) зі співвідношення (2.4) отримуємо величину середньої сили, яка розвивається в ЕМП. Величину цієї сили можна подати в термінах струму, що протікає відповідними елементами ЕМП, як

$$f_{\text{сер}}(i, x) = \frac{\Delta W_{\text{ел}}(i, x)}{\Delta x} - \frac{\Delta W_{\text{маг}}(i, x)}{\Delta x}. \quad (2.5)$$

Величину тої самої середньої сили можна подати і в термінах потокозчеплення, використовуючи при цьому співвідношення (2.3) та (2.4), а саме

$$f_{\text{сер}}(\Psi, x) = \frac{\Delta W_{\text{ел}}(\Psi, x)}{\Delta x} - \frac{\Delta W_{\text{маг}}(\Psi, x)}{\Delta x}. \quad (2.6)$$

Таким чином, маючи співвідношення для визначення електричної та магнітної енергії, що притаманні даній електромеханічній системі, можна отримати співвідношення для визначення відповідної сили (а отже й відповідного моменту), а також співвідношення, які описують поведінку даної системи в динаміці.

2.3.2. Енергія системи нерухомих провідників зі струмом

Визначимо енергію, яка надходить до систем, що складається з n нерухомих провідників (контурів). Кожним з цих провідників в процесі їх встановлення у нерухомий стан протікає визначений струм i_k , яким визначається відповідне потокозчеплення Ψ_k . При цьому розглядатимемо інтегроване значення енергії, яка надається відповідним конструктивним складовим ЕМП всіма струмами i_k в процесі зростання цих струмів від нуля до деяких кінцевих значень. Це рівнозначно відповідному змінненню потокозчеплення, яке викликається вказаним змінненням цих струмів.

Припустимо, що до виводів k -го контуру, резистивний опір якого становить r_k , приєднане зовнішнє електричне коло, яке забезпечує на виводах цього контуру напругу u_k . Зміннення струму в контурі викликає змінненні потокозчеплення та появу на виводах цього контуру ЕРС e_k .

Електричний баланс в такому контурі визначатиметься рівнянням, яке можна подати в термінах струмів

$$u_k + e_k = r_k i_k \quad (2.7)$$

або в термінах потокозчеплення

$$u_k = r_k i_k + \frac{d\Psi_k}{dt}. \quad (2.8)$$

Помноживши ліву та праву частини рівняння (2.8) на i_k , визначаємо потужність, що витрачається джерелом живлення, а саме

$$u_k i_k = r_k i_k^2 + i_k \frac{d\Psi_k}{dt}. \quad (2.9)$$

За проміжок часу dt джерелом живлення в k -му контурі ЕМП виконуватиметься робота, котра визначається як добуток потужності на відповідний інтервал часу (протягом якого ця потужність витрачається). Отже маємо

$$u_k i_k dt = dW_{\text{ел}k} = r_k i_k dt + i_k d\Psi_k. \quad (2.10)$$

Зазначимо, що частина вказаної роботи, яка пов'язана з енергією, що витрачається у вигляді тепла (ця частина дорівнює $r_k i_k^2 dt$) у рівнянні (2.10) не врахована.

Частина роботи, що визначається складовою $i_k d\Psi_k$, пов'язана з енергією, яка запасається в магнітному полі k -го контуру. Ця енергія може бути подана як

$$dW_{\text{маг}k} = i_k d\Psi_k.$$

Таким чином, магнітна енергія n контурів пов'язана зі збільшенням їх потокозчеплення співвідношенням

$$W_{\text{маг}} = \sum_{k=1}^n \int_0^{\Psi_k} i_k d\Psi_k \quad (2.11)$$

Для лінійної системи рівняння (2.11) зручно подати в термінах власних та взаємних індуктивностей відповідних контурів. При цьому вважатимемо, що потокозчеплення k -го контуру є лінійною функцією струмів всіх контурів (включаючи, звичайно, й k -й контур), тобто

$$\Psi_k = \sum_{s=1}^n L_{ks} i_s, \quad (2.12)$$

де

L_{ks} - взаємна індуктивність k -го та s -го контурів;

$L_{kk} = L_k$ - власна індуктивність k -го контуру.

Виділяючи у співвідношенні (2.12) власні індуктивності як окремі складові, отримуємо

$$\Psi_k = L_k i_k + \sum_{\substack{s=1 \\ s \neq k}}^n L_{ks} i_s. \quad (2.13)$$

Співвідношення (2.13) використаємо для визначення величини змінення потокозчеплення k -го контуру при відповідному змінненні величини струму, тобто для запису (2.13) в термінах кінцевих прирощень.

Маємо

$$\Delta\Psi_k = L_k \cdot \Delta i_k + \sum_{\substack{s=1 \\ s \neq k}}^n L_{ks} \cdot \Delta i_s. \quad (2.14)$$

Отже підінтегральна частина рівняння (2.11), записана в термінах кінцевих прирощень, виглядатиме як

$$\Delta W_{\text{маг}k} = i_k \cdot \Delta\Psi_k = i_k \left(L_k \cdot \Delta i_k + \sum_{s=1}^n L_{ks} \cdot \Delta i_s \right). \quad (2.15)$$

Співвідношення (2.15) з урахуванням підсумовування по всіх n контурах, але без інтегрування по поточкозчепленням цих контурів, зводиться до співвідношення

$$\Delta W_{\text{маг}} = \sum_{k=1}^n i_k \left(L_k \cdot \Delta i_k + \sum_{s=1}^n L_{ks} \cdot \Delta i_s \right). \quad (2.16)$$

Перш ніж перейти до подальшого перетворення співвідношення (2.16) розглянемо окремий випадок електромеханічної системи, що утворена двома контурами ($n = 2$).

За умови, що система складається з двох контурів, співвідношення (2.16) приймає вигляд

$$\Delta W_{\text{маг}} = L_1 \cdot i_1 \cdot \Delta i_1 + L_{12} \cdot i_1 \cdot \Delta i_2 + L_2 \cdot i_2 \cdot \Delta i_2 + L_{21} \cdot i_2 \cdot \Delta i_1$$

Приймаючи до уваги, що $L_{12} = L_{21}$ (тобто взаємна індуктивність першого контуру по відношенню до другого контуру дорівнює взаємній індуктивності другого контуру по відношенню до першого контуру), отримуємо

$$\frac{\Delta W_{\text{маг}}}{\Delta t} = \frac{\Delta \left(\frac{1}{2} L_1 i_1^2 \right)}{\Delta t} + \frac{\Delta \left(\frac{1}{2} L_2 i_2^2 \right)}{\Delta t} + \frac{\Delta (L_{12} i_1 i_2)}{\Delta t}.$$

Переходячи від кінцевих прирощень до нескінченно малих прирощень (тобто до диференціалів), матимемо

$$dW_{\text{маг}} = d \left(\frac{1}{2} L_1 i_1^2 \right) + d \left(\frac{1}{2} L_2 i_2^2 \right) + d (L_{12} i_1 i_2).$$

Отже повна магнітна енергія, яка запасатиметься в системі двох контурів, визначається як

$$W_{\text{маг}} = \frac{1}{2} L_1 i_1^2 + \frac{1}{2} L_2 i_2^2 + L_{12} i_1 i_2. \quad (2.17)$$

Аналогічно, розглядаючи систему з n контурів, можна записати

$$W_{\text{маг}} = \frac{1}{2} \sum_{k=1}^n L_k i_k^2 + \frac{1}{2} \sum_{k=1}^n \sum_{s=1}^n L_{ks} i_k i_s. \quad (2.18)$$

Множник $\left(\frac{1}{2}\right)$ перед другою складовою правої частини (2.18) враховує,

що $L_{ks} = L_{sk}$, а отже ця величина при подвійному підсумовуванні зустрічатиметься двічі.

Провівши нескладні перетворення, з (2.18) отримуємо

$$W_{\text{маг}} = \frac{1}{2} \sum_{k=1}^n i_k \left(L_k i_k + \sum_{s=1}^n L_{ks} i_s \right) \quad (2.19)$$

або

$$W_{\text{маг}} = \frac{1}{2} \sum_{k=1}^n \left(L_k i_k^2 + \sum_{s=1}^n L_{ks} i_s i_k \right). \quad (2.20)$$

З урахуванням (2.13) матимемо

$$W_{\text{маг}} = \frac{1}{2} \sum_{k=1}^n i_k \Psi_k. \quad (2.21)$$

Отже повна енергія, яка надходить в систему контурів ЕМП протягом усього процесу встановлення струмів, визначатиметься як

$$W_{\text{маг}} = \sum_{k=1}^n \int_0^{\Psi_k} i_k d\Psi_k = \frac{1}{2} \sum_{k=1}^n i_k \Psi_k. \quad (2.22)$$

Співвідношенням (2.22) визначається загальна кількість енергії, яка запасється в магнітному полі струмів утвореної k контурами електромеханічної системи.

Вплив нелінійності

Як правило, котушки ЕМП містять осердя з якогось магнітного матеріалу (по-більшості з електротехнічної сталі). Індукція насичення осердя – величина хоча й велика, але кінцева. Отже в процесі змінення струму в таких котушках проявлятимуться нелінійні ефекти, пов'язані з нелінійністю характеристик

намагнічування матеріалу осердя. В цілому, наведений вище аналіз процесів в системі контурів залишається справедливим. Однак кінцевий результат розв'язання інтегрального рівняння (2.11), яким визначається енергія, що запасастся, звичайно буде іншим.

За умови, що початкове потокозчеплення котушки дорівнює нулю ($\Psi = 0$), тобто крива намагнічування проходить через початок відповідних координат, рівняння (2.11) справедливе й енергія, що запасастся нелінійною системою, визначатиметься так само як і для лінійної системи. В той же час функціональна залежність потокозчеплення від струму, що протікає контурами електромеханічної системи, суттєво ускладнюється, а це в загальному випадку не дозволяє отримати відповідні аналітичні вирази для визначення величини енергії.

Як правило, вказані нелінійні фактори обмежують можливість збільшення енергії, що в ЕМП спрямовується на виконання корисної роботи. Відповідні математичні моделі стають надто складними й мають обмежене використання в процесі практичного проектування та експлуатації ЕМП.

В той же час, звичайно, вивчення таких ефектів та прогнозування роботи ЕМП у відповідних режимах безумовно підвищує рівень знань, дозволяє глибше пізнати суть відповідних процесів та явищ, сформулювати практично корисні поради щодо розроблення оптимальних конструкцій ЕМП та технології їх виготовлення, а також визначати оптимальні умови та параметри експлуатації електромеханічних систем.

2.3.3. Сили та моменти, що виникають в електромеханічних системах

Розглянуті вище енергетичні співвідношення можна використати для встановлення формул, які надають можливість обчислити сили (моменти), що виникають в електромеханічній системі.

Для спрощення розглядатимемо систему з однією обмоткою збудження. Нехтуючи резистивним опором обмотки збудження, запишемо умову її електричної рівноваги у вигляді

$$u \approx e = L \frac{di}{dt} = \frac{d\Psi}{dt}. \quad (2.23)$$

Запишемо рівняння енергетичного балансу (аналогічно (2.1)) в диференціальній формі. Маємо

$$dW_{\text{ел}} = dW_{\text{маг}} + dW_{\text{мех}}.$$

При цьому

$$dW_{\text{мех}} = F \cdot dx.$$

Переходячи до категорії потужності, яка, як відомо, характеризує швидкість змінення енергії у часі, отримуємо

$$F \frac{dx}{dt} + \frac{dW_{\text{маг}}}{dt} = \frac{dW_{\text{ел}}}{dt} = Ui = ei = i \frac{d\Psi}{dt}. \quad (2.24)$$

Зважаючи на те, що

$$W_{\text{маг}} = \frac{1}{2} \Psi i,$$

похідна магнітної енергії за часом в загальному випадку визначатиметься як

$$\frac{dW_{\text{маг}}}{dt} = \frac{1}{2} \left(\Psi \frac{di}{dt} + i \frac{d\Psi}{dt} \right). \quad (2.25)$$

Таким чином, на підставі співвідношень (2.23), (2.24) та (2.25) отримуємо

$$F \frac{dx}{dt} + \frac{1}{2} i \frac{d\Psi}{dt} = i \frac{d\Psi}{dt}. \quad (2.26)$$

Тобто

$$2F \frac{dx}{dt} = i \frac{d\Psi}{dt}. \quad (2.27)$$

З (2.27) безпосередньо видно, що за умови підтримання в контурі електромеханічної системи незмінного струму, джерело електричної енергії має забезпечувати потужність, яка достатня для виконання *подвійної* роботи. Фізично це означає, що здійснення ЕМП визначеної роботи потребує від джерела живлення енергії, яка достатня не лише для безпосереднього виконання цієї роботи, але й для забезпечення поповнення запасу енергії контурів цього ЕМП. Це положення зазвичай називають *правилом «50-50»*.

З рівняння (2.25) неважко отримати співвідношення для розрахунку сили, що розвивається в даному ЕМП відповідним потокозчепленням, а саме

$$F = \frac{1}{2} i \frac{d\Psi}{dx} \Big|_{i=\text{const}} = \frac{dW_{\text{маг}}}{dx} \Big|_{i=\text{const}}. \quad (2.28)$$

За умови незмінності потокозчеплення рівняння (2.25) зводиться до вигляду

$$F \frac{dx}{dt} + \frac{1}{2} \Psi \frac{di}{dt} = 0. \quad (2.29)$$

З (2.29) неважко отримати

$$F = -\frac{1}{2} \Psi \frac{di}{dx} \Big|_{\Psi=const} = -\frac{dW_{\text{маг}}}{dx} \Big|_{\Psi=const} . \quad (2.30)$$

З порівняння (2.28) та (2.29) видно, що розрахована за цими формулами сила, як і очікувалося, має однакову величину, але протилежні напрямлення.

Сили та моменти, що діють між контурами ЕМП в термінах взаємної індуктивності

Застосуємо отримані результати до практично важливого випадку, коли ЕМП має два контури з власними індуктивностями L_1, L_2 та взаємною індуктивністю $L_{12} = L_{21} = L_{\text{вз}}$.

Можливі варіанти конструктивного виконання такого ЕМП схематично подано на рис. 2.9. При цьому рухомий елемент ЕМП, наведеного на рис. 2.9а, здійснює поступальний рух, а рухомий елемент ЕМП, наведеного на рис. 2.9б обертається навколо деякої осі.

Вважатимемо, що в контурах протікають струми i_1, i_2 . При цьому в магнітному полі всього ЕМП запасатиметься енергія

$$W_{\text{маг}} = \frac{1}{2} L_1 i_1^2 + \frac{1}{2} L_2 i_2^2 + L_{\text{вз}} i_1 i_2 .$$

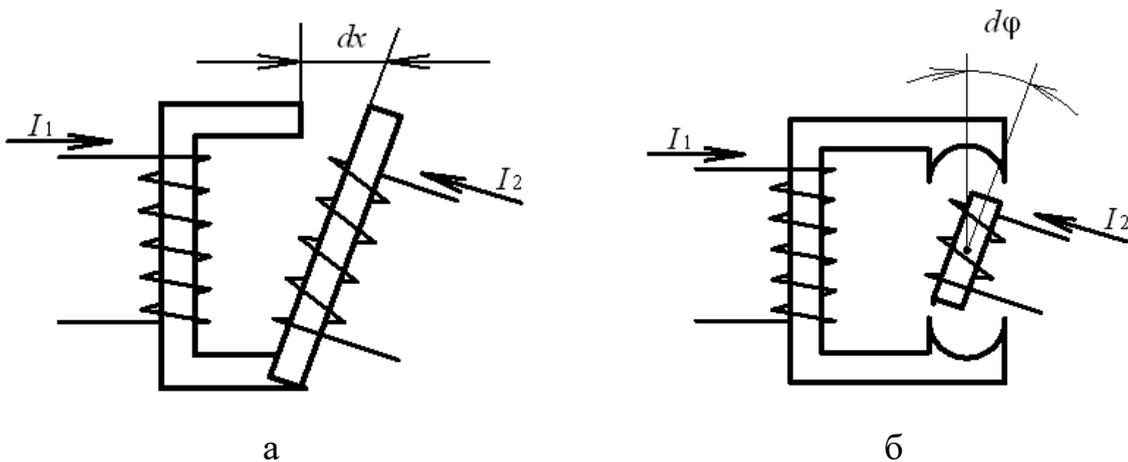


Рис. 2.9. Варіанти конструктивного виконання ЕМП з контурами,
що зв'язані магнітним зв'язком

Припустимо, що контуру 1 (див. рис. 2.9а) надане мале переміщення dx вздовж деякої координатної осі X . Це переміщення в даному ЕМП обумовлюється силою, що діє вздовж координати X та відповідно до (2.28) за умови конкретного та постійного значення струмів i_1, i_2 визначатиметься як

$$F_{1X} = \frac{dW_{\text{маг}}}{dx} = \frac{1}{2} i_1^2 \frac{dL_1}{dx} + \frac{1}{2} i_2^2 \frac{dL_2}{dx} + i_1 i_2 \frac{dL_{\text{вз}}}{dx}. \quad (2.31)$$

Аналогічно для варіанту ЕМП поданому на рис. 2.9б можна знайти рівняння, яким визначатиметься момент контуру 1 по відношенню до контуру 2. При цьому елементарне лінійне переміщення dx слід визначити через елементарне кутове переміщення $d\varphi$ (адже контур 2 згідно рис. 2.9б обертається відносно контуру 1). Маємо

$$dx = d\left(\pi D \frac{\varphi}{2\pi}\right) = \frac{D}{2} d\varphi.$$

Отже миттєве значення обертового моменту контуру 1 відносно контуру 2 визначатиметься як

$$m_{1\varphi} = \frac{1}{2} i_1^2 \frac{dL_1}{d\varphi} + \frac{1}{2} i_2^2 \frac{dL_2}{d\varphi} + i_1 i_2 \frac{dL_{\text{вз}}}{d\varphi}. \quad (2.32)$$

В більшості ЕМП власні індуктивності контурів практично не залежать від переміщення останніх. За припущень незмінності власних індуктивностей з (2.31), (2.32) матимемо

$$F_{1X} = i_1 i_2 \frac{dL_{\text{вз}}}{dx} \quad \text{та} \quad m_{1\varphi} = i_1 i_2 \frac{dL_{\text{вз}}}{d\varphi} \quad (2.33)$$

Зважаючи на те, що потокозчеплення контурів пов'язане зі струмами, які в цих контурах протікають, співвідношеннями

$$\Psi_1 = L_1 i_1 + L_{\text{вз}} i_2 \quad \text{та} \quad \Psi_2 = L_2 i_2 + L_{\text{вз}} i_1, \quad (2.34)$$

рівняння (2.33) можна переписати в дещо іншому вигляді, а саме

$$F_{1X} = i_1 \frac{d\Psi_1}{dx} \quad \text{та} \quad m_{1\varphi} = i_1 \frac{d\Psi_1}{d\varphi}. \quad (2.35)$$

Нагадаємо, що при виводі співвідношень (2.35) прийнято незалежність власних індуктивностей контурів від їх взаємного переміщення.

Рівнянням (2.35) може бути надане широке тлумачення. Ці рівняння показують, що сила або момент виникають на контурі з постійною власною індуктивністю незалежно від того, чим саме викликається змінення потокозчеплення – струмами в обох контурах, що розглядаються, або в деяких інших контурах, які зв'язані з даними контурами відповідним магнітним зв'язком.

2.3.4. Електрорушійна сила в електромеханічній системі

Отримані та проаналізовані вище рівняння для визначення сили та моменту надають реальне уявлення про впливи електричного кола на механічну частину електромеханічної системи, тобто описують один з видів зв'язку її електричної та магнітної складових.

В той же час, в електромеханічних системах існує ще один вид зв'язку, так би мовити зв'язок зворотної дії. Цей зв'язок вказує як механічна складова системи впливає на електричну. Кількісно характер такого впливу визначається на підставі закону Фарадея.

Складені на підставі закону Фарадея рівняння для процесів, що відбуваються в електромеханічній системі, за формою не відрізняються від відповідних рівнянь, складених для звичайних електричних кіл. Однак наявність та змінення відповідних механічних параметрів (змінних) в таких рівняннях приводить до появи в останніх додаткових складових, що відсутні у рівняннях, складених виключно для електричних кіл.

Для аналізу фізичної природи ЕРС, яка виникає (генерується) в електромеханічній системі, запишемо закон Фарадея в загальному вигляді, наприклад для першого контуру деякої двоконтурної електромеханічної системи. Маємо

$$e = \frac{d\Psi(i_1, i_2, x)}{dt} = \frac{d\Psi}{di_1} \frac{di_1}{dt} + \frac{d\Psi}{di_2} \frac{di_2}{dt} + \frac{d\Psi}{dx} \frac{dx}{dt}. \quad (2.36)$$

Кожна складова правої частини співвідношення (2.36) відображає відповідну складову загальної ЕРС, причому фізична природа цих складових різна, а саме

- перша складова відображає ЕРС самоіндукції, що наводиться в даному контурі за рахунок змінення власного струму (точніше магнітного потоку, що у даному контурі цим струмом обумовлюється);

- друга складова відображає ЕРС взаємної індукції, що наводиться в даному контурі за рахунок змінення струму в другому контурі (точніше магнітного потоку, що у даному контурі обумовлюється струмом, який протікає у другому контурі);

- третя складова характеризує ЕРС, яка обумовлюється взаємним переміщенням складових (контурів) електромеханічної системи.

Часто перші дві складові співвідношення (2.36) розглядають сумісно, називаючи відповідну ЕРС *трансформаторною*. Це обумовлене тим, що такі самі ЕРС діють в системах, що не мають рухомих складових, наприклад у трансформаторах.

Третя складова співвідношення (2.36) характеризує складову загальної ЕРС притаманну виключно електромеханічним системам. ЕРС, що характеризується цією складовою, є безпосереднім *результатом механічного*

переміщення й внаслідок залежності її величини від швидкості руху зазвичай зветься *ЕРС руху*.

Само по собі наведене розділення загальної ЕРС на окремі складові не є суттєвим. Важливіше чітко уявляти, що в будь-якій електромеханічній системі існує *взаємний вплив* між її електричною та механічною частинами.

Таким чином, структурно будь-яку електромеханічну систему можна подати як на рис. 2.10.

Як видно з рис. 2.10, будь-яка електромеханічна система (будь-який ЕМП) складається з

- електричної структури (електричне коло, електромагнітний пристрій) – складова, позначена як (1);
- механічної структури – складова, позначена як (2).

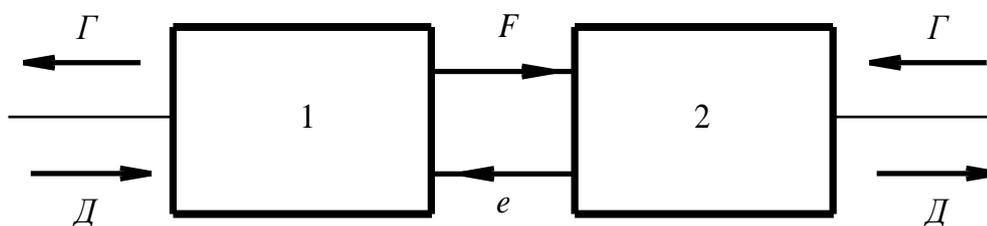


Рис. 2.10 Узагальнена структурна схема електромеханічної системи

Електричне коло, яке входить до складу електромеханічної системи, можна описати рівняннями Кірхгофа, причому, звичайно, слід обов'язково враховувати й ЕРС переміщення (руху).

Механічна структура (складова) електромеханічної системи описується рівняннями динамічної рівноваги (рівняннями Д'Аламбера). При цьому серед сил, що діють в електромеханічній системі слід враховувати й сили електромагнітного походження.

Отже, електромагнітні сили та ЕРС руху являють собою два види зв'язку між електричною та механічною структурами електромеханічної системи. Наявність цих зв'язків визначає, зокрема, принцип оберненості будь-якого ЕМП – можливість його використання як в режимі генератора (Г), за умови, що вихідною величиною (категорією) цього ЕМП є ЕРС (e), так і в режимі двигуна (Д), за умови, що вихідною величиною (категорією) ЕМП є сила (F)

2.4. БАЗОВІ ЗАСАДИ МОДЕЛЮВАННЯ ІДЕАЛІЗОВАНОГО ЕЛЕКТРОМЕХАНІЧНОГО ПЕРЕТВОРЮВАЧА

Зазвичай надто важко розібратися в тому, що відбувається насправді, а

часто це й не потрібно: необхідно лише знати – працює Ваша ідея або ні

NM

Електромеханічні перетворювачі енергії, які розглядаються в даній роботі, є електричними машинами – двигунами або генераторами. Ці машини можуть мати найрізноманітніше конструкційне виконання. Однак можна виділити загальні характерні риси, що притаманні всім електричним машинам незалежно від їх конструкції та призначення і на підставі такого узагальнення розробити відповідні моделі конструктивного виконання ЕМП, які в подальшому використовуватимемо для розроблення відповідних математичних моделей цих ЕМП.

2.4.1. Моделі конструкційного виконання електромеханічних перетворювачів

Переважна більшість електричних машин є пристроями «магнітного» типу, тобто в них превалює магнітне поле. У зв'язку з цим основу фізичної структури складають провідники, що розміщені в магнітних матеріалах.

Для реалізації обертового руху електричні машини виконуються у вигляді концентричних магнітних циліндрів з провідниками, розміщеними по сторонах загального повітряного зазору, що утворюється між вказаними циліндрами. Рухома складова такої електричної машини зветься *ротором*, а нерухома – *статором*. Магнітна поверхня з будь-якої сторони вказаного повітряного зазору може мати достатньо довільну форму. Однак по-більшості реальні конструкції можна звести до двох моделей, які наведені на рис. 2.11.

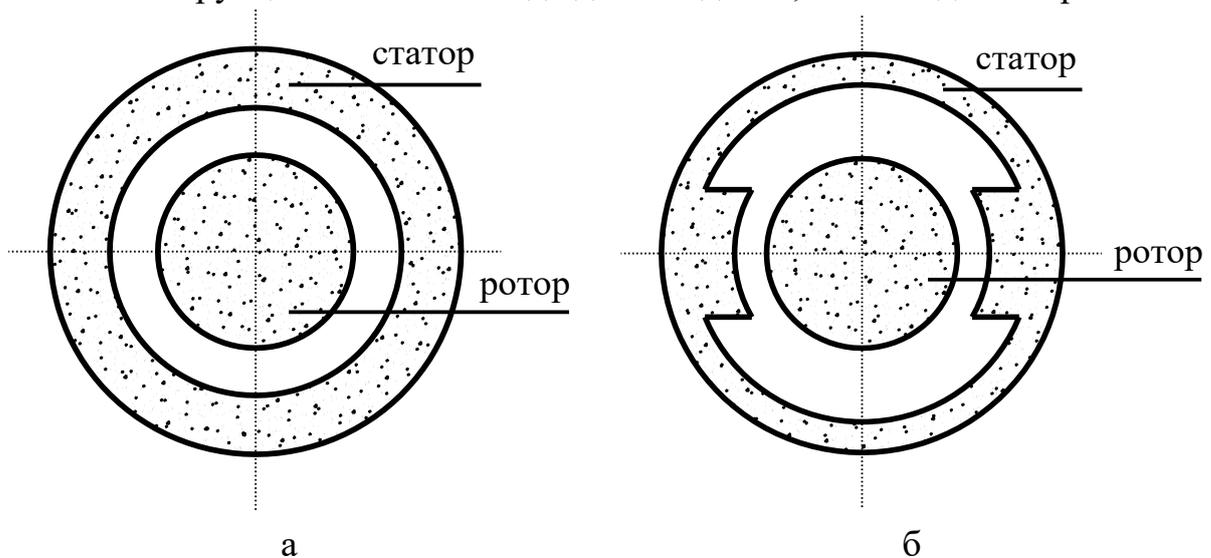


Рис. 2.11. Моделі конструкційного виконання обертових електромеханічних перетворювачів з концентричними магнітними поверхнями з обох (а) та з одної (б) сторони повітряного зазору

Першій моделі (типу) конструкційного виконання (рис. 2.11а) притаманні концентричні магнітні поверхні, які характеризуються симетрією з обох сторін повітряного зазору.

Другій моделі (рис. 2.11б) – притаманні концентричні магнітні поверхні, одна з яких гладка, а на другій відповідним чином сформовані спеціальні виступи.

Для спрощення подальшого аналізу вважатимемо, що

- магнітна проникність матеріалу циліндрів у порівнянні з магнітною проникністю повітря нескінченно велика;

- електромагнітна система містить обмотку, виконану з котушок проводу, які розміщені на поверхні ротора (статора) або з обох сторін повітряного зазору так, що результуючий аксіальний струм на кожній поверхні дорівнює нулю;

- сумарне магнітне поле при наявності декількох обмоток (струмів) може бути отримане методом простого підсумовування окремих складових (утворених кожною з обмоток окремо); тобто вважатимемо, що в структурі діє принцип суперпозиції.

2.4.2. Магнітне поле в повітряному зазорі. Обмоткова функція

Розглянемо варіант (модель) конструкційного виконання електричної машини, наведений на рис. 2.12.

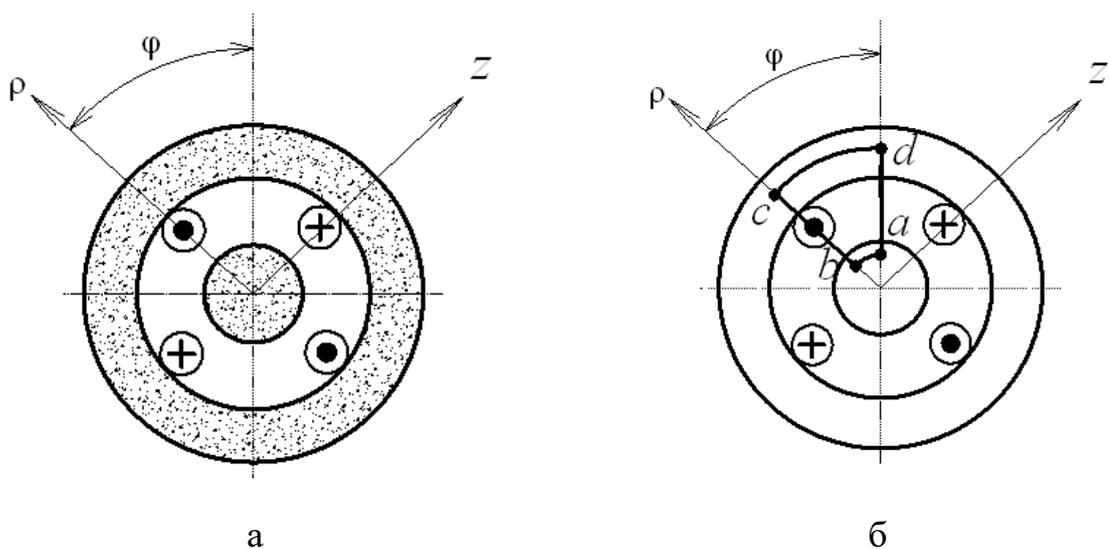


Рис. 2.12. Модель конструкційного виконання електричної машини
(до розрахунку магнітного поля в повітряному зазорі)

Для аналізу наведеної моделі конструкційного виконання електричної машини скористаємось циліндричною системою координат з осями ρ , φ , Z . При цьому орієнтуємо ось Z вздовж осі моделі. За таких умов полярні координати ρ , φ характеризуватимуть положення точки на площині, що перпендикулярна до осі Z .

В загальному випадку магнітне поле в повітряному зазорі тривимірне, тобто має складові в напрямку всіх трьох осей.

Зважаючи на те, що для математичного моделювання структури найбільший інтерес становить потокозчеплення обмоток, які знаходяться на поверхні ротора або статора, найсуттєвішою слід вважати компоненту магнітного поля, яка нормальна до поверхні, тобто радіальна (спрямована вздовж осі ρ) компонента. Ця компонента, в свою чергу, є функцією трьох координат: ρ , φ , Z . Однак задачу можна суттєво спростити, якщо припустити, що обмотка має довжину, яка дорівнює довжині циліндрів прийнятої до розгляду конструкційної моделі. За таких умов магнітне поле стає незалежним від координати Z . Така структура поля зазвичай зветься плоскою.

З другої сторони, якщо радіус ротора r значний, суттєво перевершує величину δ повітряного зазору (а зазвичай ця умова виконується на практиці), радіальна складова магнітного поля (спрямована вздовж координати ρ) в повітряному зазорі буде змінюватися несуттєво і цим змінням можна нехтувати. Отже задача розрахунку магнітного поля в зазорі нами, врешті-решт, зведена до розрахунку змінення магнітного поля виключно вздовж координати φ .

Скориставшись законом повного струму для будь-якого замкненого контуру $abcd$ (див. рис. 2.12б), можна записати

$$\int_{abcd} \vec{H} d\vec{l} = \text{струм всередині контуру } abcd. \quad (2.37)$$

Для складових контуру $abcd$, які проходять вздовж осердя (bc , ad), інтеграли дорівнюють нулю, адже за припущеннями магнітна проникність сталі вважається нескінченно великою.

Вважаючи додатною ту складову поля H_δ в зазорі, яка спрямована всередину машини, матимемо

$$\delta [H_\delta(\varphi) - H_\delta(0)] = \sum i. \quad (2.38)$$

Співвідношення (2.38) справедливе для будь-якої точки повітряного

зазору з координатою φ . Зважаючи на те, що всі контури, які проходять крізь повітряний зазор, замикаються в точці $\varphi = 0$, величина $H_{\delta}(0)$ буде присутня в кожному з рівнянь (яким описується відповідний контур), тобто є константою.

Таким чином, за умови, що струм всередині контуру визначений та відома константа $H_{\delta}(0)$, можна розв'язати рівняння (2.38) відносно $H_{\delta}(\varphi)$, визначивши, тим самим, магнітне поле в повітряному зазорі електричної машини.

Найпростіша обмотка

Розглянемо модель конструктивного виконання електричної машини, в якій обмотка складається з чотирьох провідників, з'єднаних так, що струми в сусідніх провідниках протікають в протилежних напрямках (див. рис. 2.13а).

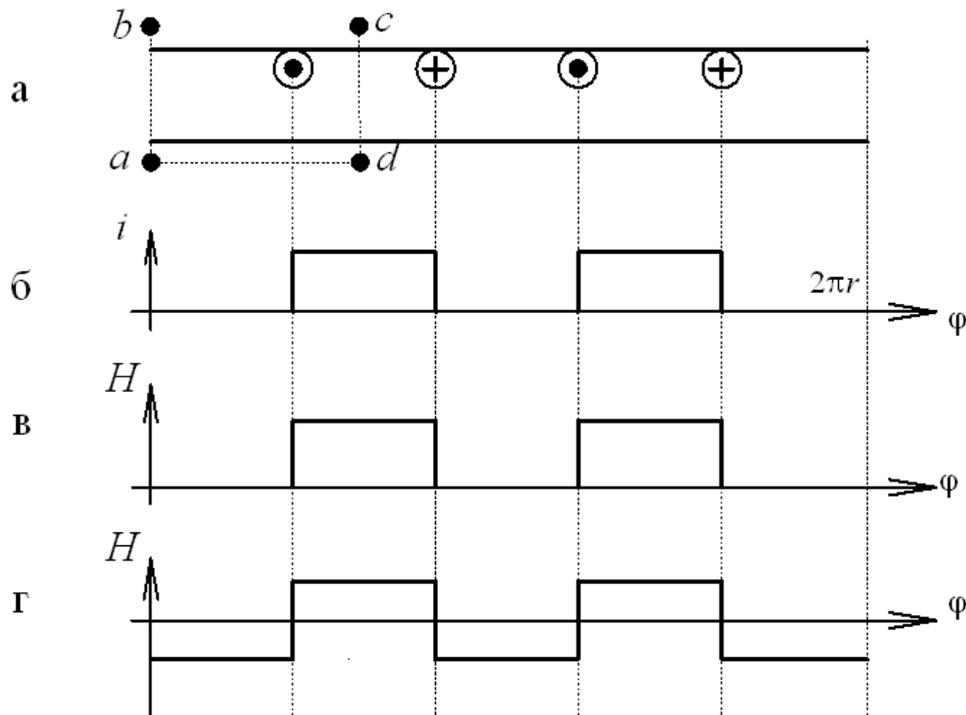


Рис. 2.13. Модель найпростішої обмотки (а) та розподілення струму (б) і магнітного поля (в, г) в повітряному зазорі

Згідно з законом неперервності потоку (законом Гауса) повний магнітний потік Φ , що перетинає замкнену поверхню, має дорівнювати нулю.

Отже

$$\Phi = \int \vec{B} d\vec{s} = \mu_0 \int \vec{H} d\vec{s} = 0. \quad (2.39)$$

Для моделі, яка розглядається, співвідношення (2.39) означатиме, що повний магнітний потік, який перетинає повітряний зазор (утворений між

обмотками статора та ротора), дорівнює нулю. Отже, середня величина $H_{\delta}(\varphi)$ має дорівнювати нулю.

Таким чином, для отримання нульового значення середньої величини $H_{\delta}(\varphi)$, необхідно досягти симетрії відповідного розподілення магнітного поля по відношенню до координати φ (див. рис. 2.13в). Це означатиме, що постійна складова $H_{\delta}(0)$ дорівнюватиме

$$H_{\delta}(0) = -\frac{i}{2\delta}. \quad (2.40)$$

При цьому напруженість магнітного поля в повітряному зазорі електричної машини (точніше, в повітряному зазорі прийнятої конструктивної моделі) матиме вигляд як на рис. 2.13г.

Довільна обмотка

Розглянемо тепер задачу визначення розподілення $H_{\delta}(\varphi)$ для довільної обмотки.

Струм всередині будь-якого замкненого контуру можна знайти, підрахувавши кількість провідників всередині цього контуру з урахуванням напрямку струмів, що протікають цими провідниками. При цьому рівняння (2.38) для загального випадку, який саме й розглядається, можна записати як

$$\delta \cdot [H_{\delta}(\varphi) - H_{\delta}(0)] = [n_{\varphi}(\varphi)] \cdot i, \quad (2.41)$$

де

$n_{\varphi}(\varphi)$ - результуюча кількість провідників з додатним для прийнятого напрямку обходу контуру струмом, які зосереджені між початком відліку та довільною віссю, положення якої визначається кутом φ .

Функція $n_{\varphi}(\varphi)$ характеризує обмотку електричної машини. Цю функцію завжди можна знайти, підрахувавши (як це описано вище) кількість провідників, з яких утворена відповідна обмотка.

Розв'язавши рівняння (2.41) відносно $H_{\delta}(\varphi)$, матимемо

$$H_{\delta}(\varphi) = \frac{i}{\delta} \cdot n_{\varphi}(\varphi) + H_{\delta}(0). \quad (2.42)$$

Скористаємось законом Гауса для поверхні ротора, яка примикає до повітряного зазору.

Площа елементарної частки цієї поверхні визначається як

$$ds = r \cdot d\varphi \cdot dz.$$

Зважаючи на те, що $H_{\delta}(\varphi)$ нормальна до поверхні ротора, скалярний добуток двох векторів можна подати як звичайний добуток двох величин. В результаті отримуємо

$$\mu_0 \int \vec{H} d\vec{s} = \mu_0 r \int_0^l \int_0^{2\pi} H_{\delta}(\varphi) dz d\varphi = \mu_0 r l \int_0^{2\pi} H_{\delta}(\varphi) d\varphi = 0. \quad (2.43)$$

Підставивши у (2.43) величину $H_{\delta}(\varphi)$, визначену як (2.42), отримаємо

$$\mu_0 r l \int_0^{2\pi} \left[\frac{i}{\delta} n_{\varphi}(\varphi) + H_{\delta}(0) \right] d\varphi = 0$$

або

$$\int_0^{2\pi} \frac{i}{\delta} n_{\varphi}(\varphi) d\varphi = \int_0^{2\pi} H_{\delta}(0) d\varphi = -2\pi H_{\delta}(0).$$

Звідки

$$H_{\delta}(0) = \frac{i}{\delta} n_{\varphi}(\varphi) d\varphi = -\frac{1}{\delta} n_{\varphi\text{сер}}. \quad (2.44)$$

Таким чином, з (2.42) матимемо

$$H_{\delta}(\varphi) = \frac{i}{\delta} \cdot [n_{\varphi}(\varphi) - n_{\varphi\text{сер}}]. \quad (2.45)$$

Введемо позначення

$$N_{\varphi}(\varphi) = n_{\varphi}(\varphi) - n_{\varphi\text{сер}}. \quad (2.46)$$

Таким чином, для визначення напруженості магнітного поля в зазорі електричної машини з урахуванням (2.46) отримуємо рівняння

$$H_{\delta}(\varphi) = \frac{i}{\delta} \cdot N_{\varphi}(\varphi). \quad (2.47)$$

Функціональну залежність $N_{\varphi}(\varphi)$ зазвичай називають *обмотковою функцією*. Цю функцію можна отримати шляхом підрахунку провідників зі струмом з подальшим приведенням результату до нульового середнього значення на інтервалі $0 \leq \varphi \leq 2\pi$ згідно (2.46).

Таким чином, напруженість магнітного поля в повітряному зазорі можна вважати відомою, як тільки знайдена обмоткова функція, адже залежності напруженості вказаного магнітного поля та обмоткової функції від координати

φ ідентичні, а їх величина відрізняється у i/δ разів, причому величина i/δ відома та залишається незмінною за вибраної конструкції ЕМП та режиму роботи цього пристрою.

Синусна обмотка

В електричних машинах використовуються найрізноманітніші обмотки, кожна з яких має відповідну (власну) обмоткову функцію.

Гармонійний аналіз (тобто розкладання функції в ряд Фур'є) дає загальний метод аналізу будь-яких обмоток. За допомогою такого аналізу обмоткову функцію будь-якої обмотки можна подати як суму нескінченного числа гармонік, амплітуда та період яких зменшуються з підвищенням номера. Зважаючи на те, що типовими членами ряду Фур'є є синуси (косинуси), можна ввести поняття *синусної обмоткової функції* та *синусної обмотки*. За таких умов, використовуючи принцип суперпозиції, можна *будь-яку реальну обмотку подати у вигляді суми відповідних гармонійних обмоток*.

Як видно з наведеного вище, обмоткова функція електромеханічної системи періодична. Період цієї функції для двополюсного ЕМП дорівнює 2π рад. В багатополюсних обмотках, коли $2p \geq 4$ (p - число пар полюсів електромеханічної системи), період обмоткової функції дорівнює $2\pi/p$ рад.

Під час розкладення обмоткової функції в ряд Фур'є зручно ввести нову змінну, для якої будь-яка пара полюсів, незалежно від протяжності в просторі, буде займати кут 2π рад.

З цією метою вводиться нова категорія - *електричний кут*, який відповідно до наведеного пов'язаний з геометричним кутом співвідношенням

$$\varphi_e = p\varphi. \quad (2.48)$$

Таким чином, обмоткова функція, аргументом якої є електричний кут, завжди матиме період 2π рад. та записуватиметься у вигляді

$$N_\varphi(\varphi) = \sum_v N_v \sin(v\varphi_e) = \sum_v N_v \sin(vp\varphi). \quad (2.49)$$

Відповідно до наведеного вище щодо можливості подання будь-якої обмотки у вигляді суми відповідних гармонійних обмоток, можна вважати, що кожен член правої частини рівняння (2.49) відповідає обмотці, яка створює в повітряному зазорі електромеханічної системи магнітне поле, розподілене відповідно до синусної функції.

Та з таких часткових (елементарних) уявних обмоток, яка має найбільший просторовий період, зветься *основною*, інші – *гармонійними*.

2.4.3. Потокозчеплення та індуктивність обмотки

Загальний випадок

Відповідно до загального визначення, магнітний потік Φ можна знайти як

$$\Phi = \int_S \vec{B} d\vec{s}, \quad (2.50)$$

де

$d\vec{s}$ - елементарна складова деякої довільної поверхні S , яка, наприклад обмежена провідниками, з котрих саме й складається обмотка.

Зважаючи на те, що рівняння (2.50) передбачає відомим розподілення магнітного поля в повітряному зазорі, величину елемента ds поверхні S вибираємо в повітряному зазорі як

$$ds = r \cdot dz \cdot d\varphi.$$

За таких умов (2.50) можна подати як

$$\Phi = rl \int B_\delta(\varphi) d\varphi. \quad (2.51)$$

При цьому потокозчеплення визначатиметься співвідношенням

$$\Psi = rl \int_0^{2\pi} n_\varphi(\varphi) B_\delta(\varphi) d\varphi. \quad (2.52)$$

Нагадаємо, що функція $n_\varphi(\varphi)$ - це результуюча кількість провідників з додатним для прийнятого напрямку обходу контуру струмом, при цьому підрахунку піддаються ті провідники, що зосереджені між початком відліку та довільної віссю, положення якої визначається кутом φ . Зазначимо, що функція $n_\varphi(\varphi)$ не тільки надає уявлення про характер розподілення магнітного поля в повітряному зазорі, але й визначає границі інтегрування для розрахунку потокозчеплення.

З урахуванням рівняння (2.46), яким визначається обмоткова функція, отримуємо наступне співвідношення для розрахунку потокозчеплення

$$\Psi = rl \int_0^{2\pi} N_\varphi(\varphi) B_\delta(\varphi) d\varphi + n_{\varphi_{\text{сер}}} rl \int_0^{2\pi} B_\delta(\varphi) d\varphi. \quad (2.53)$$

Зважаючи на те, що відповідно до закону Гауса середнє значення індукції $B_\delta(\varphi)$ в інтервалі $0 \leq \varphi \leq 2\pi$ дорівнює нулю, другий інтеграл в (2.53) також дорівнюватиме нулю, а отже

$$\Psi = rl \int_0^{2\pi} N_{\varphi}(\varphi) B_{\delta}(\varphi) d\varphi. \quad (2.54)$$

При визначенні власної індуктивності деякої обмотки a величина індукції $B_{\delta}(\varphi)$ пов'язана з магнітним полем (обумовлюється магнітним полем), яке утворюється саме цією обмоткою. Тому можемо записати цю індукцію як

$$B_{\delta a}(\varphi) = \mu_0 H_{\delta a}(\varphi) = \mu_0 \frac{i_a}{\delta} N_{\varphi a}(\varphi). \quad (2.55)$$

Отже, відповідно до визначення власної індуктивності та рівняннями (2.54) і (2.55) співвідношення для розрахунку власної індуктивності обмотки a можна подати як

$$L_a = \frac{\Psi_a}{i_a} = \mu_0 \frac{rl}{\delta} \int_0^{2\pi} N_{\varphi a}^2(\varphi) d\varphi. \quad (2.56)$$

Для визначення взаємної індуктивності двох обмоток a -ї та b -ї у співвідношення (2.54) слід додати індукцію магнітного поля, що утворюється b -ю обмоткою, а саме

$$B_{\delta b}(\varphi) = \mu_0 \frac{i_b}{\delta} N_{\varphi b}(\varphi).$$

Отже взаємна індуктивність a -ї обмотки по відношенню до b -ї обмотки з урахуванням наведеного визначатиметься як

$$L_{ab} = \frac{\Psi_{ab}}{i_b} = \mu_0 \frac{rl}{\delta} \int_0^{2\pi} N_{\varphi a}(\varphi) N_{\varphi b}(\varphi) d\varphi. \quad (2.57)$$

За аналогією можна записати взаємну індуктивність b -ї обмотки по відношенню до a -ї обмотки, а саме

$$L_{ba} = \frac{\Psi_{ba}}{i_a} = \mu_0 \frac{rl}{\delta} \int_0^{2\pi} N_{\varphi a}(\varphi) N_{\varphi b}(\varphi) d\varphi. \quad (2.58)$$

Із співвідношень (2.57) та (2.58) видно, що

$$L_{ab} = L_{ba}$$

Таким чином, ідеалізовану електромеханічну систему (як, звичайно, і ідеалізовану електричну машину) можна характеризувати, використовуючи лише поняття обмоткової функції.

Універсальність співвідношення (2.58) полягає ще й в тому, що воно показує незалежність характеру взаємного індуктивного зв'язку двох обмоток від того розподілені ці обмотки або зосереджені. Цей висновок важливий, адже дозволяє для цілей математичного моделювання замінити будь-яку розподілену обмотку еквівалентною зосередженою обмоткою, що часто суттєво спрощує задачу.

Синусна обмотка

Для будь-якої обмотки ротора або статора електричної машини v -та гармоніка обмоткової функції може бути подана як

$$N_{\varphi v}(\varphi) = N_v \cos(v\varphi_e + \gamma_v). \quad (2.59)$$

Вибір функції синусу або косинусу, звичайно, може бути довільним, адже при відповідному виборі просторового кута γ_v обидві форми виявляються ідентичними.

Для подальшого аналізу зручніше буде використовувати саме функцію косинусу, бо при цьому просторовий кут γ_v вказуватиме на розміщення осі гармонійної функції.

Власну індуктивність будь-якої v -ї гармонійної обмотки за умови використання співвідношення (2.59) можна отримати після підстановки цієї гармонійної обмоткової функції у рівняння (2.56) та врахування (2.59). А саме

$$\begin{aligned} L_v &= \mu_0 \frac{rl}{\delta} \int_0^{2\pi} N_v^2 \cos^2(v\varphi_e + \gamma_v) d\varphi = \\ &= \mu_0 \frac{rl}{\delta} N_v^2 \int_0^{2\pi} \cos^2(vp\varphi + \gamma_v) d\varphi = \mu_0 \frac{\pi rl}{\delta} N_v^2 \end{aligned} \quad (2.60)$$

Зі співвідношення (2.60) видно, що власна індуктивність не залежить від орієнтації осі обмотки у просторі, яка визначається величиною кута γ_v .

Таким чином, власна індуктивність синусної обмотки постійна - визначається геометричними розмірами обмотки та повітряного зазору, а також кількістю провідників (витків) обмотки.

Припустимо, що крім розглянутої v -ї обмотки є ще одна (k -та) синусна обмотка, яка характеризується обмотковою функцією

$$N_{\varphi k}(\varphi) = N_k \cos(k\varphi_e + \gamma_k). \quad (2.61)$$

В загальному випадку ця обмотка може мати іншу кількість полюсів та інше розподілення у просторі. Обидві обмотки можуть бути гармонійними складовими одної або двох різних обмоток. В будь-якому випадку взаємну

індуктивність цих двох обмоток (v -ї та k -ї) можна отримати на підставі рівняння (2.58), а саме

$$\begin{aligned} L_{vk} &= \mu_0 \frac{rl}{\delta} \int_0^{2\pi} N_v \cos(v\varphi_e + \gamma_v) N_k \cos(k\varphi_e + \gamma_k) d\varphi = \\ &= \mu_0 \frac{rl}{\delta} N_v N_k \int_0^{2\pi} \cos(v\rho\varphi + \gamma_v) \cos(k\rho\varphi + \gamma_k) d\varphi = 0 \end{aligned} \quad (2.62)$$

Нуль у правій частині співвідношення (2.62) має місце при будь-яких параметрах обмоток за умови $v \neq k$. Це випливає з властивості ортогональної системи гармонійних функцій, до яких належать і використані нами функції косинусу.

З рівняння (2.62) безпосередньо випливає

- між синусними обмотками, що мають різну кількість полюсів немає магнітного зв'язку;

- взаємна індуктивність відсутня між будь-якими гармонійними обмотками, які є складовими однієї реальної обмотки;

- власна індуктивність будь-якої обмотки дорівнює сумі власних індуктивностей синусних обмоток, які є її складовими.

Визначимо взаємну індуктивність двох синусних обмоток, які мають однакову кількість полюсів ($v = k$). Зважаючи на те, що взаємна індуктивність характеризує різні обмотки, введемо для таких обмоток позначення a та b . При цьому обмоткові функції таких обмоток запишуться як

$$\begin{aligned} N_{\varphi av}(\varphi) &= N_{av} \cos(v\varphi_e + \gamma_{av}). \\ N_{\varphi bv}(\varphi) &= N_{bv} \cos(v\varphi_e + \gamma_{bv}). \end{aligned} \quad (2.63)$$

В загальному випадку ці функції можуть мати різні амплітуди та розподілення у просторі.

Після підстановки (2.63) в (2.57) та відповідних перетворень отримуємо

$$L_{abv} = \mu_0 \frac{\pi rl}{\delta} N_{bv} N_{av} \cos(\gamma_{av} - \gamma_{bv}). \quad (2.64)$$

Зі співвідношення (2.64), зокрема, видно, що взаємна індуктивність двох синусних обмоток з однаковим числом полюсів визначається функцією косинусу збільшеного у v разів просторового кута між осями основних складових, які відповідають реальним обмоткам.

Дійсно, величину $(\gamma_{av} - \gamma_{bv})$ можна подати як

$$v = \left(\frac{\gamma_{av}}{v} - \frac{\gamma_{bv}}{v} \right) = v\gamma, \quad (2.65)$$

де

γ - електричний кут між осями основних синусних обмоток (тобто між максимумами обмоткових функцій).

Таким чином,

$$L_{abv} = \mu_0 \frac{\pi r l}{\delta} N_{bv} N_{av} \cos(v\gamma). \quad (2.66)$$

або

$$L_{abv} = L_{abv m} \cos(v\gamma). \quad (2.67)$$

де індексом (m) позначене максимальне значення взаємної індуктивності.

Співвідношення (2.67) має очевидний фізичний зміст: *взаємна індуктивність обмоток (магнітний зв'язок обмоток) матиме максимальну величину, за умови, що осі цих обмоток співпадають, і мінімальну, коли ці осі взаємноперпендикулярні.*

2.4.4. Момент ідеалізованої електричної машини

Неявнополюсна конструкція

Раніше нами було отримане співвідношення для визначення електромагнітного моменту, що існує при взаємодії двох обмоток, а саме

$$M = i_a i_x \frac{dL}{d\varphi}. \quad (2.68)$$

Із співвідношення видно, що момент M існуватиме (не дорівнюватиме нулю) лише за умови залежності індуктивності обмоток від кутової координати φ . Ця умова виконуватиметься коли одна з обмоток розміщена на статорі машини, а друга – на роторі.

Введене поняття синусної обмотки дозволяє провести необхідний аналіз в самому загальному вигляді. Оскільки між синусними обмотками, що мають різну кількість полюсів, не може бути індуктивного зв'язку, розглянемо випадок, коли на статорі і на роторі розміщені синусні обмотки v -ї гармоніки.

На рис. 2.14 зображені основні гармонічні складові обмоток статора (a) та ротора (x).

Припустимо, що кут φ_r характеризує положення ротора. Якщо кут (в електричних градусах) між осями гармонічних обмоток, що розглядаються, дорівнює γ , а $\varphi_{re} = p\varphi_r = 0$ в момент часу $t = 0$ (див. рис. 2.14а), то при

переміщенні ротора (в момент часу $t = t_1$) цей кут дорівнюватиме $(\gamma + \varphi_{re})$, як показано на рис. 2.14б.

Запишемо взаємну індуктивність a -ї та x -ї обмоток, скориставшись для цього співвідношенням (2.67)

$$L_{axv} = L_{axv m} \cos[v(\varphi_{re} + \gamma)], \quad (2.69)$$

де

$$L_{axv m} = \mu_0 \frac{\pi r l}{\delta} N_{av} N_{xv}.$$

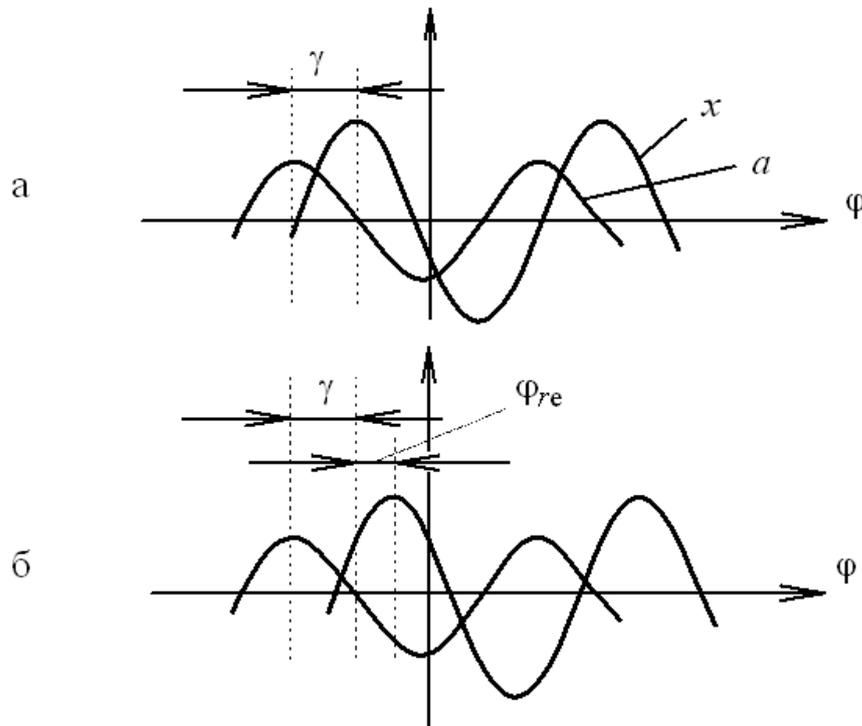


Рис. 2.14. Основні гармонічні складові обмотки статора (а) та ротора (б)

За таких умов відповідно до (2.68) момент дорівнюватиме

$$\begin{aligned} M_v &= i_a i_x \frac{dL_{axv}}{d\varphi_r} = i_a i_x L_{axv m} \frac{d \cos[v(\varphi_{re} + \gamma)]}{d\varphi_r} = \\ &= -v p i_a i_x L_{axv m} \sin[v(p\varphi_r + \gamma)]. \end{aligned} \quad (2.70)$$

Ввівши позначення

$$\theta = \varphi_{re} + \gamma,$$

отримуємо

$$(2.71)$$

Зазначимо, що кут θ є поточним (миттєвим) електричним кутом між осями двох гармонійних обмоток. Він різний для кожної гармоніки, бо неоднакові просторові кути γ_{av} та γ_{xv} відповідних членів рядів Фур'є для обмоток. Цей кут зазвичай називають *кутом моменту*.

Результуючий магнітний момент, який діє між обмотками електричної машини, може бути знайдений шляхом простого підсумовування гармонійних моментів, визначених відповідно до (2.71).

Співвідношення (2.71) показує, що при взаємодії синусних обмоток утворюється магнітний момент, який залежить від

- числа пар полюсів електромеханічної системи (що визначається множителем (vp)),
- струмів в обмотках,
- максимальної величини взаємної індуктивності обмоток,
- синуса кута між осями обмоток.

Знак $(-)$ у співвідношенні (2.71) показує, що утворений момент спрямовується на зменшення кута θ , тобто намагається привести обмотки, які взаємодіють, у положення, де їх осі співпадають.

Можливі й інші форми запису співвідношення (2.71).

Наприклад, якщо підставити (2.69) в (2.71), то після нескладних перетворень отримуємо

$$\begin{aligned}
 M_v &= -vp\pi r l \delta \frac{N_{av} i_a}{\delta} \frac{\mu_0 N_{xv} i_x}{\delta} \sin(v\theta) = \\
 &= -vp\pi r l \delta H_{avm} B_{xvm} \sin(v\theta),
 \end{aligned}
 \tag{2.72}$$

де

H_{avm} , B_{xvm} - характеристики (напруженість та індукція відповідно) магнітних полів, які утворені обмотками статора та ротора, а саме: максимальні значення v -х гармонік напруженості магнітного поля та магнітної індукції відповідно.

Порівняння двох наведених форм запису моменту взаємодії двох обмоток дозволяє зробити деякі висновки, а саме

- перша форма запису (співвідношення (2.71)) визначає момент як функцію параметрів та змінних електричного кола (момент визначається в термінах струмів та індуктивностей);
- друга форма запису (співвідношення (2.72)) визначає магнітний момент через параметри, які характеризують магнітне поле в зазорі машини (момент визначається в термінах напруженості магнітного поля та магнітної індукції).

Таким чином, замість моменту, який спрямований на приведення синусних обмоток в положення, за якого вісі цих обмоток співпадають, можна характеризувати момент як категорію, дія якої спрямована на приведення у

відповідне положення осей двох магнітних полів. Отже, співвідношення (2.72) є кількісним описом сил взаємодії між магнітними полюсами на поверхні статора й ротора. Момент, який утворюється цими силами, пропорційний напруженості обох магнітних полів (точніше, магнітних полів обох обмоток, що взаємодіють) та збільшеного у ν разів синусу просторового кута між осями цих полів.

Доки два магнітних поля синусоїдально розподілені у повітряному зазорі між статором та ротором, співвідношення (2.72) надаватиме дійсну величину моменту незалежно від того, як ці поля утворюються. Але при цьому слід правильно вибирати кут моменту. Щоб підкреслити цю відносну незалежність моменту від походження вказаних величин, що характеризують магнітне поле, перепишемо співвідношення (2.72) в дещо іншому вигляді, а саме

$$M_{\nu} = -\nu p \pi r l \delta H_{\nu m} B_{\nu m} \sin(\nu \theta). \quad (2.73)$$

Зазначимо, що за будь-яких умов (при всіх перетвореннях співвідношення для визначення моменту) величина θ характеризує просторовий кут між осями двох магнітних полів.

Співвідношення (2.73) найбільш придатне для якісного аналізу, тоді як запис моменту в термінах теорії кіл, тобто співвідношення (2.72), найзручніший при дослідженні характеристик електромеханічної системи.

Третю форму запису співвідношення, яким визначається момент, можна отримати якщо у (2.73) ввести пов'язаний з полюсом загальний магнітний потік на полюс машини (магнітний потік, що припадає на один полюс електричної машини).

Величина магнітного потоку на полюс для ν -ї гармоніки магнітного поля ротора або статора залежить від розподілення провідників у відповідній обмотці й визначається як

$$\begin{aligned} \Phi_{\nu} &= \int_s B_{\nu} ds = \iint_Z B_{\nu} \cdot r \cdot dz \cdot d\varphi = rl \int_0^{\frac{\pi}{p\nu}} B_{\delta} N_{\nu} \sin(\nu \varphi_e) d\varphi = \\ &= rl \int_0^{\frac{\pi}{p\nu}} B_{\nu m} \sin(\nu p \varphi) d\varphi = \frac{2}{\nu p} rl B_{\nu m}. \end{aligned} \quad (2.74)$$

Врахувавши (2.74), співвідношення (2.73) для магнітного моменту можна подати як

$$M_{\nu} = -\frac{\pi}{2} (\nu p)^2 \Phi_{\nu} F_{\nu m} \sin(\nu \theta), \quad (2.75)$$

де

$$F_{vm} = \delta H_{vm} = \delta \left(\frac{iN_v}{\delta} \right) - \text{максимальна магніторушійна сила, що діє в}$$

повітряному зазорі.

Співвідношення (2.75) особливо зручно використовувати за умови, що Φ_v є загальним утвореним всіма обмотками магнітним потоком полюса електромеханічної системи.

Явнополюсна конструкція

Модель конструктивного виконання електричної машини з виступаючими полюсами, але «оберненої» конструкції – з полюсами на роторі (що є типовим для синхронних електричних машин), представлена на рис. 2.15.

Як було показано вище, загальне співвідношення для визначення магнітного моменту має вигляд

$$M = \frac{dW_{\text{маг}}}{d\varphi},$$

де магнітна енергія двоконтурної електромеханічної системи, модель якої розглядається, визначається як

$$W_{\text{маг}} = \frac{1}{2} L_a i_a^2 + \frac{1}{2} L_x i_x^2 + L i_a i_x$$

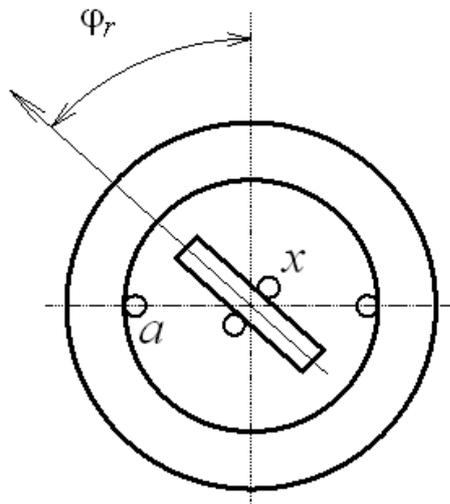


Рис. 2.15. Модель конструктивного виконання електричної машини з полюсами на роторі

Для даної моделі лише потокозчеплення ротора, обумовлене його власним магнітним полем, не залежить від положення ротора в просторі. За таких умов момент дорівнюватиме

$$M = \frac{dW_{\text{маг}}}{d\varphi} = \frac{1}{2} i_a^2 \cdot \frac{dL_a}{d\varphi} + i_a i_x \frac{dL}{d\varphi} = M'' + M'. \quad (2.76)$$

У співвідношенні (2.76) власна індуктивність обмотки a є функцією кута повороту ротора, адже її власне потокозчеплення залежить від величини повітряного зазору. Цей повітряний зазор, як видно з рис. 2.15, мінімальний, коли вісь ротора співпадає з віссю магнітного поля, яке утворюється обмоткою a , і максимальний при горизонтальному положенні ротора, коли вісь магнітного поля, утвореного обмоткою a , ортогональна осі ротора.

З метою спрощення задачі та надання результату більшої наочності весь подальший аналіз виконуватимемо, враховуючи лише основні гармонічні обмоткові функції.

Відповідно до рис. 2.15 просторовий кут φ_r , яким визначається положення ротора, в загальному випадку може бути представлений (в електричних градусах) у вигляді суми $(\varphi_{re} + \gamma)$, де γ - початковий кут (при $t = 0$) між осями основних гармонійних обмоткових функцій, а φ_{re} - приращення цього кута у часі.

На підставі (2.66) та 2.67) взаємна індуктивність обмоток, що розглядаються, визначається як

$$L_{ax} = L_{axm} \cos(p\varphi_r + \gamma), \quad (2.77)$$

де

$$L_{axm} = \mu_0 \frac{\pi r l}{\delta} N_a N_x. \quad (2.77a)$$

При повороті ротора величина повітряного зазору в явнополісній моделі електричної машини змінюється, а отже змінюється і величина опору (магнітного опору) на шляху магнітного потоку обумовленого струмом статора (тобто обмоткою a).

Характер змінення магнітного опору має вигляд як на рис. 2.16.

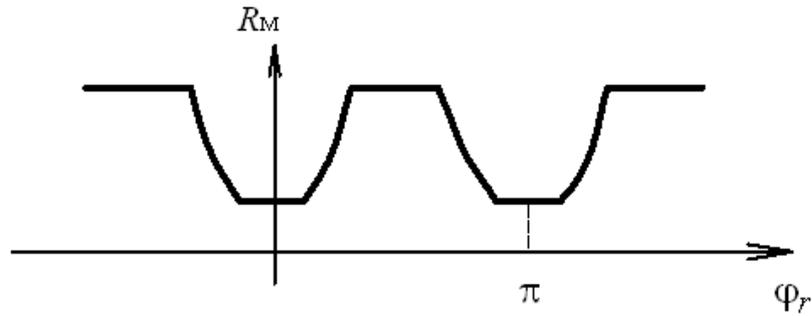


Рис. 2.16. Характер змінення магнітного опору повітряного зазору явнополюсної моделі

Величина магнітного опору за аналогією з резистивним опором електричного кола визначається співвідношенням

$$R_M(\varphi_r) = \frac{\delta(\varphi_r)}{\mu_0 S}. \quad (2.78)$$

На підставі (2.78) провідність повітряного зазору визначатиметься як

$$\Lambda_M(\varphi_r) = \frac{1}{R_M(\varphi_r)} = \mu_0 \frac{S}{\delta(\varphi_r)}. \quad (2.79)$$

Стосовно електричних машин, де магнітний потік зосереджений в магнітопроводі, можна говорити про магнітну провідність для магнітного потоку на полюсному діленні, площа якого S дорівнює

$$S = \frac{\pi r l}{p}.$$

Отже співвідношення (2.79) можна подати як

$$\Lambda_M(\varphi_r) = \mu_0 \frac{r l}{\delta(\varphi_r)} \frac{\pi}{p}. \quad (2.80)$$

Розклавши функцію $\Lambda_M(\varphi_r)$ в ряд Фур'є та обмежившись лише двома першими членами цього ряду (що зазвичай забезпечує достатню для практичних цілей точність), можна записати

$$\Lambda_M(\varphi_r) \approx \Lambda_0 + \Lambda \cos(2p\varphi_r), \quad (2.81)$$

де

$$\Lambda_0 = \frac{\Lambda_{\max} + \Lambda_{\min}}{2}; \quad \Lambda = \frac{\Lambda_{\max} - \Lambda_{\min}}{2}.$$

Графічна інтерпретація співвідношення (2.82) наведена на рис. 2.17.

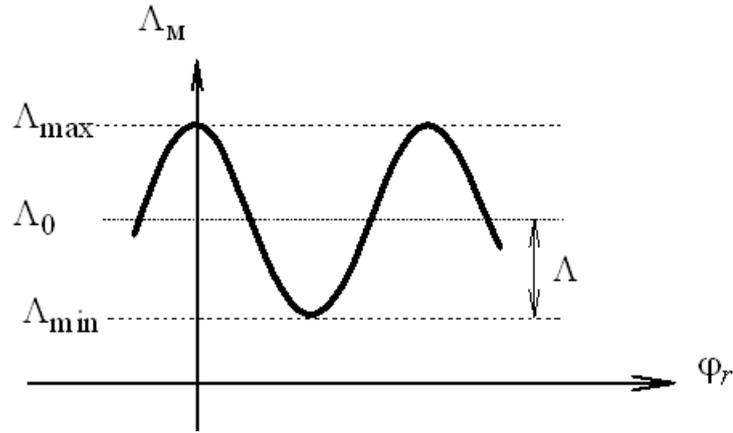


Рис. 2.17. Залежність від кутової координати магнітної провідності повітряного зазору явнополюсної моделі

Власна індуктивність обмотки a відповідно до (2.56) з урахуванням (2.81) може бути подана як

$$\begin{aligned}
 L_a &= \mu_0 r l \int_0^{2\pi} \frac{[N_a \cos(p\varphi_r + \gamma)]^2}{\delta(\varphi_r)} d\varphi_r = \\
 &= \int_0^{2\pi} \frac{p}{\pi} \Lambda_M(\varphi) [N_a \cos(p\varphi_r + \gamma)]^2 d\varphi_r = \\
 &= p\Lambda_0 N_a^2 + p \frac{\Lambda}{2} N_a^2 \cos 2\gamma = pN_a^2 \left(\Lambda_0 + \frac{\Lambda}{2} \cos 2\gamma \right).
 \end{aligned}$$

На підставі останнього співвідношення з урахуванням (2.80), (2.81) для моменту часу $t = 0$, коли осі основних гармонік синусних обмоток зміщені на кут γ , отримуємо співвідношення для визначення моменту, що розвивається в електромеханічній системі, а саме

$$\begin{aligned}
 M &= i_a i_x \frac{dL}{d\gamma} + \frac{1}{2} i_a^2 \cdot \frac{dL_a}{d\gamma} = \\
 &= - \left[p\mu_0 \frac{\pi r l}{\delta} i_a i_x N_a N_x \sin \gamma + p i_a^2 \frac{N_a}{4} (\Lambda_{\max} - \Lambda_{\min}) \sin 2\gamma \right] \quad (2.82)
 \end{aligned}$$

Таким чином, момент має дві складові, при цьому
 - одна складова залежить від струмів обох контурів та за структурою не відрізняється від моменту неявно полюсної моделі (див. (2.71));
 - друга складова присутня навіть за умови відсутності збудження з боку

ротора, тобто залежить лише від одного струму, а саме від струму статора, а також пропорційна різниці провідностей для магнітного потоку по двом взаємно перпендикулярним осям явнополюсного ротора і, відповідно, тим більше, чим більша ця різниця.

2.4.5. Магнітні поля обмоток електромеханічної системи

Однофазна обмотка

Миттєве значення індукції магнітного поля, яке утворюється основною синусною обмоткою, можна представити як

$$b(\varphi) = \mu_0 \frac{i}{\delta} N \cos(\varphi + \gamma)$$

За умови, що обмоткою тече струм $i = I_m \cos \omega t$, наведене співвідношення набирає вигляду

$$b(\varphi) = \mu_0 \frac{I_m}{\delta} N \cos \omega t \cdot \cos(\varphi + \gamma) = B_m \cos \omega t \cdot \cos(\varphi + \gamma), \quad (2.83)$$

де

I_m - амплітуда струму, що тече обмоткою;

B_m - амплітуда індукції магнітного поля, яке утворюється основною синусною обмоткою.

Таким чином, магнітне поле є функцією, як часу, так і просторової координати. Індукція цього магнітного поля, як гармонічна функція просторового кута, може бути представлена у вигляді деякого просторового вектору. Таке представлення цілком аналогічне використанню часових векторів при аналізі електричних кіл синусоїдального струму.

Магнітне поле, що описується співвідношенням (2.83), нерухоме у просторі, але змінюється (пульсує) у часі, внаслідок чого зазвичай зветься *стаціонарним пульсуючим* магнітним полем.

Магнітному полю, що збуджується однією синусною обмоткою при протіканні нею змінного струму, можна надати цікаву та практично важливу інтерпретацію, якщо скористатися відомим тригонометричним співвідношенням

$$\cos \alpha \cdot \cos \beta = \frac{1}{2} [\cos(\alpha - \beta) + \cos(\alpha + \beta)].$$

З урахуванням цього співвідношення (2.83) можна подати як

$$b(\varphi, t) = \frac{B_m}{2} \{ \cos[\omega t - (\varphi + \gamma)] + \cos[\omega t + (\varphi + \gamma)] \}. \quad (2.84)$$

З (2.84) видно, що загальне магнітне поле в зазорі можна подати як суму двох складових, які обертаються в різні сторони. Назвавши одну з цих складових «прямою», а другу - «зворотною» введемо відповідні позначення, а саме

$$\begin{aligned} b_{\text{пр}}(\varphi, t) &= \frac{B_m}{2} \cos[\omega t - (\varphi + \gamma)], \\ b_{\text{зв}}(\varphi, t) &= \frac{B_m}{2} \cos[\omega t + (\varphi + \gamma)]. \end{aligned} \quad (2.85)$$

Уявимо, що ми спостерігаємо за якимись точками цих двох хвиль, в яких значення магнітної індукції однакові. За таких умов мають виконуватися співвідношення

$$\cos[\omega t - (\varphi + \gamma)] = \text{const}, \quad \cos[\omega t + (\varphi + \gamma)] = \text{const}.$$

З цих співвідношень випливає

$$\omega t - (\varphi + \gamma) = \text{const}, \quad \omega t + (\varphi + \gamma) = \text{const}. \quad (2.86)$$

Диференціювання співвідношень (2.86) по t , відповідно, приводить до рівнянь

$$\frac{d\varphi}{dt} = \omega; \quad \frac{d\varphi}{dt} = -\omega \quad (2.87)$$

Нагадаємо, що похідні від кута за часом є кутовими швидкостями обертання хвилі у повітряному зазорі.

Відповідно до (2.87) перший член правої частини співвідношень (2.84) дійсно характеризує пряму хвилю $\left(\frac{d\varphi}{dt} > 0\right)$, тоді як другий – зворотну хвилю $\left(\frac{d\varphi}{dt} < 0\right)$.

Залежність (2.86) показує, що за один період змінення струму (тобто за час, коли ωt змінюється на кут 2π), магнітне поле в зазорі повертається на той самий кут, тобто на кут 2π електричних радіан. Відповідно до цього кутова швидкість Ω магнітного поля дорівнюватиме циклічній частоті ω мережі живлення, тобто

$$\Omega = \omega = 2\pi f, \quad (2.88)$$

де

f - коливальна частота струму в електромережі живлення.

З іншої сторони, кутова швидкість обертання поля в механічних одиницях кута дорівнює

$$\Omega_{\text{мех}} = \frac{\omega}{p} = \frac{2\pi f}{p}. \quad (2.89)$$

Зважаючи на те, що $\Omega_{\text{мех}} = 2\pi n$ (n - кількість обертів ротора за одиницю часу) то

$$n = \frac{f}{p}. \quad (2.90)$$

Для реалізації розрахунків інтерес становить також ν -та гармонічна складова магнітного поля, яка визначається як

$$b_{\nu}(\varphi, t) = B_{\nu m} \cos(\omega t) \cdot \cos[\nu(\varphi + \gamma)] \quad (2.91)$$

або після нескладних перетворень як

$$b_{\nu}(\varphi, t) = \frac{B_{\nu m}}{2} \left\{ \cos[\omega t - \nu(\varphi + \gamma)] + \cos[\omega t + \nu(\varphi + \gamma)] \right\}. \quad (2.92)$$

Отже, ν -та гармонічна складова пульсуючого магнітного поля також може бути представлена у вигляді двох компонент, що обертаються. Однак за час одного періоду збуджуючого струму ν -та гармонічна складова переміщується лише на $2\pi/\nu$ електричних радіан.

Іншими словами, швидкість обертання цієї складової в ν разів менша частоти обертання основної складової, тобто

$$\Omega_{\nu} = \frac{\omega}{\nu}; \quad \Omega_{\text{мех}\nu} = \frac{\omega}{\nu p}. \quad (2.93)$$

Багатофазна обмотка

Задача створення одного обертового магнітного поля постійної величини може бути розв'язана шляхом відповідного розміщення в просторі двох або більше однофазних обмоток так, щоб магнітне поле, яке обертається в протилежному напрямку, компенсувалося.

Найпростішим засобом досягнення шуканого результату є симетрична двофазна обмотка. Така обмотка складається з двох обмоток, магнітні осі яких орієнтовані під прямим кутом одна до другої, а струми зсунуті за фазою на кут $\pi/2$. Таким чином, додаткова обмотка створює магнітне поле, яке можна описати співвідношенням

$$\begin{aligned}
 b'(\varphi, t) &= B'_m \cos\left(\omega t - \frac{\pi}{2}\right) \cos\left(\varphi + \gamma - \frac{\pi}{2}\right) = \\
 &= B'_m \sin(\omega t) \sin(\varphi + \gamma).
 \end{aligned}
 \tag{2.94}$$

Тригонометрична тотожність

$$\sin \alpha \cdot \sin \beta = \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)]$$

дозволяє привести (2.94) до вигляду

$$b'(\varphi, t) = \frac{B'_m}{2} \left\{ \cos[\omega t - (\varphi + \gamma)] - \cos[\omega t + (\varphi + \gamma)] \right\}. \tag{2.95}$$

Індукція результуючого магнітного поля двох таких однофазних обмоток визначається як

$$b_{\text{рез}}(\varphi, t) = b(\varphi, t) + b'(\varphi, t).$$

На підставі співвідношень (2.84) та (2.95) неважко встановити, що при $N' = N$ магнітні поля обмоток, які обертаються у зворотному напрямку, дорівнюють одне одному за амплітудою, відрізняючись за фазою. Тому індукція результуючого магнітного поля приводиться до вигляду

$$b_{\text{рез}}(\varphi, t) = B_m \cos[\omega t - (\varphi - \gamma)].$$

Окремий інтерес становить використання трьох фазних обмоток.

Фази симетричної трифазної обмотки утворюють магнітне поле, складові індукції якого визначаються як (для спрощення вважатимемо $\gamma = 0$)

$$\begin{aligned}
 b_A(\varphi, t) &= B_m \cos(\omega t) \cos \varphi = \\
 &= \frac{B_m}{2} [\cos(\omega t - \varphi) + \cos(\omega t + \varphi)]
 \end{aligned}
 \tag{2.96}$$

$$\begin{aligned}
 b_B(\varphi, t) &= B_m \cos\left(\omega t - \frac{2\pi}{3}\right) \cos\left(\varphi - \frac{2\pi}{3}\right) = \\
 &= \frac{B_m}{2} \left[\cos(\omega t - \varphi) + \cos\left(\omega t + \varphi - \frac{4\pi}{3}\right) \right]
 \end{aligned}
 \tag{2.97}$$

$$\begin{aligned}
 b_C(\varphi, t) &= B_m \cos\left(\omega t + \frac{2\pi}{3}\right) \cos\left(\varphi + \frac{2\pi}{3}\right) = \\
 &= \frac{B_m}{2} \left[\cos(\omega t - \varphi) + \cos\left(\omega t + \varphi + \frac{4\pi}{3}\right) \right]
 \end{aligned} \tag{2.98}$$

З урахуванням (2.96) - (2.98) індукцію результуючого магнітного поля трифазної обмотки можна подати як

$$b(\varphi, t) = b_A(\varphi, t) + b_B(\varphi, t) + b_C(\varphi, t) = \frac{3B_m}{2} \cos(\omega t - \varphi) \tag{2.99}$$

За умови використання симетричної m -фазної обмотки матимемо

$$b(\varphi, t) = \frac{mB_m}{2} \cos(\omega t - \varphi). \tag{2.100}$$

Таким чином, орієнтовані синусні обмотки утворюють обертове магнітне поле постійної величини. Однак реальні обмотки можуть лише наближатися до синусних обмоток, бо завжди одночасно утворюють магнітні поля гармонічних складових.

Як вже було зазначено, вищі гармонічні складові магнітного поля обертаються повільніше основного коливання. Лише один цей факт робить наявність гармонічних складових індукції магнітного поля небажаними. Однак крім цього, виявляється, що поля деяких вищих гармонік обертаються в сторону, яка протилежна обертанню основного поля, що неприпустиме.

Наприклад, при наявності третьої гармоніки в магнітному полі, яке утворюється двофазною симетричною обмоткою, матимемо

$$b_A(\varphi, t) = B_m \cos(\omega t) \left[N_1 \cos \varphi + N_3 \cos(3\varphi) \right] \tag{2.101}$$

$$\begin{aligned}
 b_B(\varphi, t) &= \\
 &= B_m \cos\left(\omega t - \frac{\pi}{2}\right) \left[N_1 \cos\left(\varphi - \frac{\pi}{2}\right) + N_3 \cos\left(3\varphi - \frac{3\pi}{2}\right) \right] = \\
 &= B_m \sin(\omega t) \left[N_1 \sin \varphi - N_3 \sin \varphi \right].
 \end{aligned} \tag{2.102}$$

Індукція результуючого магнітного поля такої обмотки становитиме

$$b(\varphi, t) = B_m \left[N_1 \cos(\omega t - \varphi) + N_3 \cos(\omega t + 3\varphi) \right]. \tag{2.103}$$

Із (2.103) наочно видно, що поле третьої гармоніки обертається в протилежний бік зі швидкістю у три рази меншою за швидкість обертання основної складової.

2.4.6. Електрорушійна сила обмотки

Рівняння електричної рівноваги для будь-якої обмотки визначаються відповідно до другого закону Кірхгофа. Наприклад, для деякої обмотки a ці рівняння матимуть вигляд

$$u_a + e_a = r_a i_a,$$

де e_a згідно закону Фарадея визначається повним потокозчепленням обмотки, яке в загальному випадку можна подати як

$$\begin{aligned} \Psi_a = & L_a i_a + L_{ab} i_b + L_{ac} i_c + \dots + \\ & + L_{ax}(\varphi_r) i_x + L_{ay}(\varphi_r) i_y + \dots \end{aligned} \quad (2.104)$$

До співвідношення (2.104) входять, зокрема, деякі взаємні індуктивності, які залежать від положення ротора, що відмічено залежністю відповідних величин від кута φ_r повороту ротора. Це відноситься до взаємних індуктивностей обмоток, які розміщені по різні сторони повітряного зазору.

Для тих складових потокозчеплення, які не залежать від φ_r , складові результуючої ЕРС будуть аналогічні відомим з теорії кіл ЕРС само- та взаємоіндукції, тобто визначатимуться як

$$L_a \frac{di_a}{dt} + L_{ab} \frac{di_b}{dt} + \dots \quad (2.105)$$

Раніше було зазначено, що ці складові не вносять вкладу в процес перетворення енергії – енергія, пов'язана з ними, іде виключно на збільшення або зменшення енергії, що запасена в магнітному полі повітряного зазору.

Визначаючи ЕРС, які пов'язані з залежними від кута φ_r індуктивностями, необхідно мати на увазі, що цей кут сам може бути функцією часу. Така ситуація може спричинятися, наприклад, переміщенням ротора. Складові ЕРС такого типу визначаються як

$$L_{ax}(\varphi_r) \frac{di_x}{dt} + i_x \frac{dL_{ax}(\varphi_r)}{d\varphi_r} \frac{d\varphi_r}{dt} + \dots \quad (2.106)$$

У співвідношенні (2.106) перша складова являє собою звичайну ЕРС взаємоіндукції, яка може бути об'єднана з ЕРС, що розглядалися вище.

Друга складова співвідношення (2.106) характерна лише для електромеханічних перетворювачів енергії та являє собою ЕРС, що створена виключно завдяки руху ротора, а отже безпосередньо пов'язана з процесом перетворення енергії.

Для дослідження ЕРС, визначеної другою складовою співвідношення (2.106), скористаємося поняттям синусної обмотки. З метою спрощення

обмежимося розгляданням взаємодії однієї пари гармонічних обмоток.

Для ν -х гармонічних синусних обмоток, які належать реальним обмоткам a та x , ЕРС обертання дорівнюватиме

$$\begin{aligned} e_a &= i_x \frac{dL_{ax\nu}(\varphi_r)}{dt} = i_x \frac{d}{dt} \left[L_{ax\nu m} \cdot \cos(\nu(p\varphi_r + \gamma)) \right] = \\ &= -i_x \nu p L_{ax\nu m} \sin(\nu(\varphi_{re} + \gamma)) \cdot \frac{d\varphi_r}{dt}. \end{aligned} \quad (2.107)$$

За умови постійної швидкості обертання ротора матимемо

$$\varphi_{re} = p\varphi_r = \omega_r t.$$

При цьому співвідношення (2.107) прийме вигляд

$$e_a = -i_x \nu \omega_r L_{ax\nu m} \sin(\nu(\omega_r t + \gamma)). \quad (2.108)$$

Якщо струм i_x постійний, тобто

$$i_x = I_x,$$

то ЕРС є синусною функцією часу і її діюче значення дорівнюватиме

$$E_a = \frac{1}{\sqrt{2}} I_x \nu \omega_r L_{ax\nu m} = \frac{2\pi}{\sqrt{2}} f_\nu I_x L_{ax\nu m}, \quad (2.109)$$

де

$$f_\nu = \nu f.$$

ЕРС можна подати і в іншій формі, підставивши в (2.109) значення $L_{ax\nu m}$ зі співвідношення (2.77а).

В результаті отримуємо

$$E_a = \frac{\pi^2}{\sqrt{2}} \nu p f_\nu N_{av m} \Phi_{\nu m}, \quad (2.110)$$

де

$\Phi_{\nu m}$ - амплітуда магнітного потоку ν -ї гармоніки на один полюс.

Магнітне поле, викликане постійним струмом в обмотці x , має незмінну величину та нерухоме по відношенню до ротора. Внаслідок обертання ротора, це поле переміщується відносно обмотки статора, примушуючи змінюватися його потокозчеплення.

Як правило, послідовність аналізу більшості типів реальних електричних машин відрізняється від наведеної тим, що на першому етапі визначається ЕРС, яка наводиться в обмотці рухомих синусоїдально розподіленим у

просторі магнітним полем.

Якщо, наприклад, зосереджена обмотка має w витків та p пар полюсів, то її обмоткова функція являє собою прямокутну хвилю з амплітудою $w/(2p)$.

Гармонічний аналіз дає наступне співвідношення для визначення амплітуди v -ї гармоніки прямокутної хвилі

$$N_{av} = \frac{4}{\pi} \frac{w}{2p} \frac{1}{v}. \quad (2.111)$$

При цьому співвідношення (2.110) приймає вигляд

$$E_{av} = \frac{\pi^2}{\sqrt{2}} v p f_v \frac{4w}{2\pi v p} \Phi_{vm} = \sqrt{2} \pi f_v w \Phi_{vm}. \quad (2.112)$$

2.4.7. Зв'язок магнітного поля в повітряному зазорі з напругою на обмотці

Розглядаючи вище обертове магнітне поле, ми заради зручності вважали струми в обмотках заданими функціями часу. При такому припущенні було визначено поле в повітряному зазорі та встановлено характер його просторового розподілення.

Однак у більшості електричних машин, крім статорної обмотки, є ще й обмотки на роторі, якими також протікають струми. Це робить задачу визначення магнітного поля в термінах струмів значно складнішою.

Саме тому, як правило, заданими функціями вважають не струми обмоток, а прикладені до них напруги. Встановивши зв'язок магнітного поля в повітряному зазорі з напругами на обмотках можна показати, що вони однозначно визначають величину магнітного поля.

Спершу розглянемо простий пристрій у вигляді тороїдального осердя з двома обмотками (див. рис. 2.18). В такому пристрої існуватиме пульсуюче магнітне поле.

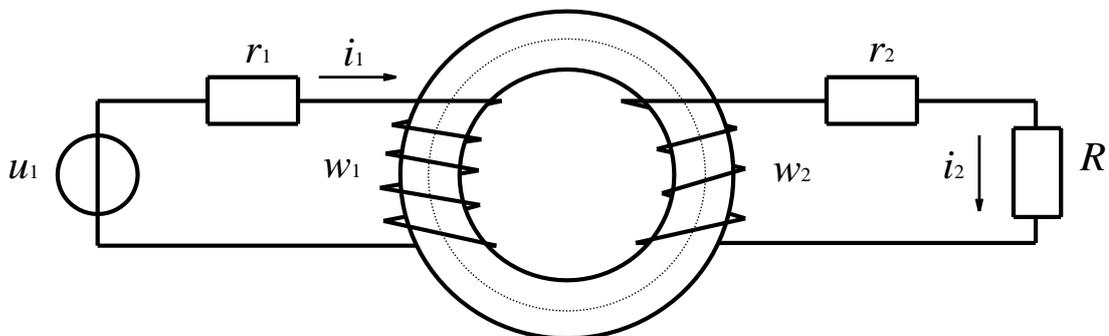


Рис. 2.18. Електромагнітний пристрій з тороїдальним осердям

Нехтуючи потоками розсіювання, співвідношення рівноваги (рівняння стану) для первинного та вторинного кіл можна записати, скориставшись другим законом Кірхгофа. Маємо

$$u_1 = r_1 i_1 + w_1 \frac{d\Phi}{dt}; \quad u_2 = (r_2 + R) i_2 + w_2 \frac{d\Phi}{dt}$$

Якщо задана напруга u_1 на виводах первинної обмотки, то магнітний потік можна знайти наближено, припустивши, що

$$r_1 i_1 \ll w_1 \frac{d\Phi}{dt}.$$

За таких умов для первинного електричного кола матимемо

$$u_1 \approx w_1 \frac{d\Phi}{dt}. \quad (2.113)$$

Отже, якщо нехтувати активним опором та магнітним потоком розсіювання, прикладена напруга повністю визначає магнітний потік в осерді. Фізично це означає, що в осерді має встановитися такий потік, який наводить ЕРС, здатну врівноважити прикладену напругу. Цей потік не залежить від вторинного струму.

За умови, що прикладена до первинної обмотки напруга визначається як

$$u = \sqrt{2}U \cos(\omega t),$$

розв'язок рівняння (2.113) відносно магнітного потоку матиме вигляд

$$\Phi \approx \frac{\sqrt{2}U}{\omega w} \sin(\omega t) = \Phi_m \sin(\omega t). \quad (2.114)$$

Таким чином, між діючим значенням синусоїдної прикладеної напруги та амплітудою магнітного потоку в осерді існуватиме наступне співвідношення

$$U \approx \frac{1}{\sqrt{2}} \omega w \Phi_m = \sqrt{2} \pi f w \Phi_m = 4,44 f w \Phi_m. \quad (2.115)$$

Встановимо тепер зв'язок між магнітним полем, що обертається в повітряному зазорі електромеханічної системи, та напругою, яка прикладена до виводів обмотки цієї системи.

Якщо можна нехтувати активним (резистивним) опором обмоток та потоками розсіювання, то для кожної з двох обмоток (a та b) можна записати

$$u_a \approx \frac{d\psi_a}{dt}; \quad u_b \approx \frac{d\psi_b}{dt}. \quad (2.116)$$

Припустимо, що в обмотках двофазної симетричної машини протікають струми

$$i_a = I_m \cos(\omega t); \quad i_b = I_m \sin(\omega t).$$

Відповідно до отриманих вище співвідношень, результуюча індукція основної гармонічної складової магнітного поля, яке утворюється в повітряному зазорі обома обмотками, дорівнюватиме

$$b(\varphi, t) = \mu_0 \frac{I_m}{\delta} N \cos[\omega t - (\varphi + \gamma)] = B_m \cos[\omega t - (\varphi + \gamma)]. \quad (2.117)$$

Потокозчеплення з будь-якою обмоткою, обумовлене цим полем, визначатиметься згідно співвідношення (2.149) як

$$\Psi = rl \int_0^{2\pi} N_\varphi(\varphi) b(\varphi) d\varphi.$$

Наприклад, для обмотки a матимемо

$$\begin{aligned} \Psi_a &= rl \int_0^{2\pi} N_a \cos(\varphi + \gamma) B_m \cos[\omega t - (\varphi + \gamma)] \cdot d\varphi = \\ &= \pi rl N_a B_m \cos(\omega t). \end{aligned} \quad (2.118)$$

ЕРС, що наводиться в такій обмотці визначатиметься як

$$e_a = -\frac{d\Psi_a}{dt} = -\pi rl N_a B_m \omega \sin(\omega t). \quad (2.119)$$

Цілком аналогічно для обмотки b маємо

$$e_b = -\frac{d\Psi_b}{dt} = -\pi rl N_b B_m \omega \sin(\omega t). \quad (2.120)$$

Якщо до обмоток прикладені напруги

$$u_a = -\sqrt{2}U \sin(\omega t), \quad u_b = -\sqrt{2}U \sin(\omega t), \quad (2.121)$$

то згідно зі співвідношеннями (2.116), (2.120) та (2.121) між діючим значенням напруги та амплітудою магнітної індукції магнітного поля, яке обертається, існує залежність

$$U \approx \frac{\pi rl N \omega}{\sqrt{2}} B_m. \quad (2.122)$$

Потік на один полюс, який утворюється основною синусною обмоткою, дорівнює

$$\begin{aligned}\Phi &= rl \int B_m \cos(\omega t - \varphi_e) d\varphi = \\ &= rl \int_0^{\pi/p} B_m \cos(\omega t - p\varphi) d\varphi = \frac{2}{p} rl B_m \sin(\omega t).\end{aligned}$$

Враховуючи це співвідношення, (2.115) можна переписати як

$$U \approx \frac{\pi}{2\sqrt{2}} p \omega N \Phi_m = \sqrt{2} \pi f \left(\frac{\pi}{2} p N \right) \Phi_m. \quad (2.123)$$

Як видно, співвідношення (2.123) відрізняється від (2.115) лише постійним множником, який залежить від розподілення обмотки.

2.5. УЗАГАЛЬНЕНА МОДЕЛЬ ЕЛЕКТРИЧНОЇ МАШИНИ

Незважаючи на різноманіття типів електричних машин, всі вони засновані на загальних фізичних явищах, що виникають під час руху провідників в магнітному полі. Це дозволяє побудувати узагальнену модель конструктивного виконання електричної машини, на підставі якої розробити відповідну узагальнену математичну модель. Така математична модель, звичайно, включатиме опис всіх характерних для електричних машин процесів, і може бути використана для дослідження різних типів їх реалізації.

2.5.1. Фізичні засади побудови узагальненої моделі

В основі теорії узагальненої електричної машини лежить можливість заміни реальної, в загальному випадку m -фазної та багатополусної машини, двофазною машиною з взаємно ортогональними обмотками на статорі й аналогічними обмотками на роторі.

Можливість такої заміни заснована на деяких базових принципах, якими враховуються фізичні засади реалізації будь-якої електричної машини, а саме:

- за допомогою двох ортогональних обмоток можна створити всі види магнітного поля, яке реалізується в електромеханічних перетворювачах енергії. При живленні двох ортогональних обмоток постійним струмом отримуємо результуюче стаціонарне магнітне поле. При живленні цих обмоток змінним струмом з однаковими фазами отримуємо результуюче пульсуюче магнітне поле. При живленні цих обмоток змінними струмами з фазовим зсувом у чверть періоду отримуємо бігуче або обертове магнітне поле;

- завжди можна вибрати ортогональні обмотки так, щоб вони утворювали той самий магнітний момент, що й реальні обмотки складної конфігурації.

В багатьох електричних машинах доводиться розглядати момент, що діє на рамочну (прямокутну) котушку зі струмом. Така котушка схематично

показана на рис. 2.19 в прямокутних координатах (d, q) .

Припустимо, що магнітне поле котушки спрямоване вздовж осі q , а сама котушка утворена з w витків, якими протікає струм i . Вісь котушки утворює з віссю d кут γ . З котушкою зчеплений магнітний потік Φ , який залежить від орієнтації цієї котушки у просторі.

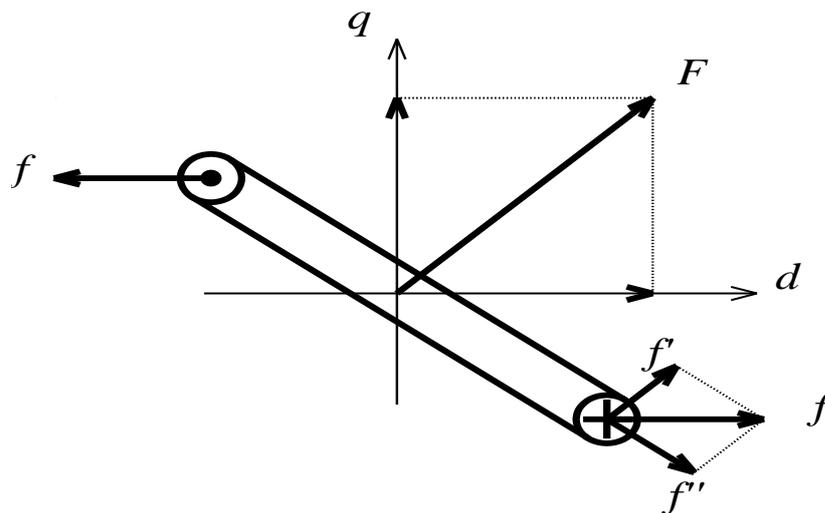


Рис. 2.19. Модель котушки зі струмом в прямокутній системі координат

У створеному зовнішньому магнітному полі на котушку діятиме магніторушійна сила

$$F = iw.$$

Вважатимемо, що на кожну сторону (довжиною l) котушки діє сила

$$f = B_q l i w,$$

завдяки якій в ЕМП утворюється електромагнітний момент

$$M_{\text{маг}} = 2r B_q i l = \Phi_q F_d, \quad (2.124)$$

де

Φ_q - максимальний магнітний момент вздовж осі q ;

F_d - складова магніторушійної сили вздовж осі d .

Можна показати, що при довільному напрямку магнітного поля утвореного складовими магнітної індукції B_q, B_d , отримуємо

$$M_{\text{маг}} = \Phi_q F_d - \Phi_d F_q. \quad (2.125)$$

Співвідношення (2.125) можна записати в дещо іншому вигляді, а саме як

$$M_{\text{маг}} = \Psi_q i_d - \Psi_d i_q, \quad (2.126)$$

де

$$i_d = i \cos \gamma; i_q = i \sin \gamma; \Psi_d = w \Phi_d; \Psi_q = w \Phi_q.$$

Співвідношення (2.126) можна безпосередньо застосовувати для електромеханічної системи, яка складається з двох однакових ортогональних та механічно жорстко з'єднаних котушок.

Таким чином, модель електричної машини з двома ортогональними обмотками на статорі та двома аналогічними обмотками на роторі дозволяє імітувати *будь-яку реальну* електричну машину як за значеннями електромагнітного моменту, який ця машина здатна розвивати, так і за структурою магнітного поля в повітряному зазорі між статором і ротором.

2.5.2. Алгоритм математичного моделювання узагальненої електричної машини

Як було показано вище, конструктивну модель, яка адекватно представляє клас обертових електричних машин, можна подати як двофазну двополюсну машину, принципова схема якої наведена на рис. 2.20.

Нагадаємо припущення, які слід прийняти з метою побудови узагальненої математичної моделі на підставі прийнятої узагальненої конструктивної моделі електричної машини:

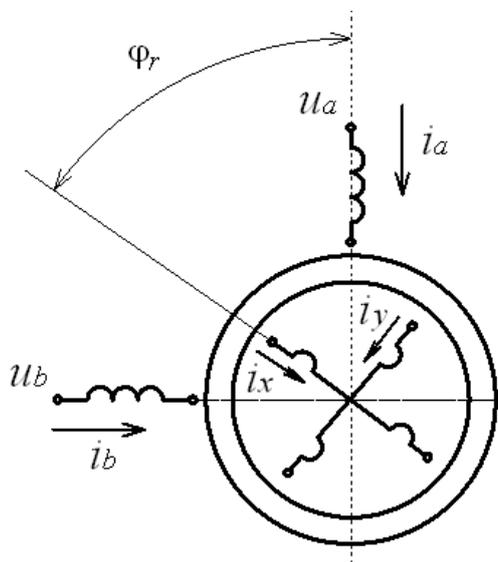


Рис. 2.20. Принципова схема двофазної двополюсної електричної машини

- повітряний зазор рівномірний;
- втрати на гістерезис та вихрові струми відсутні;
- магнітна проникність сталі суттєво більша за магнітну проникність повітря;
- двофазні обмотки статора та ротора симетричні та синусоїдально розподілені у просторі.

В загальному випадку модель електричної машини з урахуванням прийнятих припущень можна характеризувати за допомогою десяти змінних, а саме:

- чотири змінних, що відображають напруги на затискачах обмоток,
- чотири змінних, що відображають струми в обмотках,
- одна змінна, що відображає кут, яким визначається положення ротора,

- одна змінна, що відображає момент навантаження на валу.

При постановці задачі математичного моделювання ЕМП п'ять з наведених десяти параметрів, а саме: напруги та момент, зазвичай вважаються заданими. Інші п'ять параметрів мають бути знайдені за результатами моделювання.

Отже математичну модель узагальненої електричної машини можна подати як систему з п'яти рівнянь, які містять п'ять невідомих параметрів. Чотири з цих рівнянь описуватимуть чотири кола обмоток, а п'яте – умову механічної рівноваги ротора.

На підставі другого закону Кірхгофа можна записати

$$\begin{aligned} u_a &= r_1 i_a + \frac{d\Psi_a}{dt}; & u_b &= r_1 i_b + \frac{d\Psi_b}{dt}; \\ u_x &= r_2 i_x + \frac{d\Psi_x}{dt}; & u_y &= r_2 i_y + \frac{d\Psi_y}{dt}, \end{aligned} \quad (2.127)$$

де індекси 1 та 2 відносяться до кіл статора та ротора відповідно.

Рівняння механічної рівноваги ротора отримуємо, підсумовуючи всі моменти, які на цей ротор діють.

Маємо

$$J \frac{\partial^2 \varphi_r}{\partial t^2} + M_2 \left(\varphi_r, \frac{d\varphi_r}{dt} \right) = M_{\text{маг}}, \quad (2.128)$$

де

J - результуючий момент інерції обертових частин;

$M_2 \left(\varphi_r, \frac{d\varphi_r}{dt} \right)$ - зовнішній момент (момент навантаження електричної машини);

$M_{\text{маг}}$ - електромагнітний момент, що розвивається електричною машиною.

Загальне визначення індуктивностей, наведене раніше, дозволяє виразити потокозчеплення через струми в обмотках та індуктивності цих обмоток як

$$\begin{aligned} \Psi_a &= L_a i_a + L_{ab} i_b + L_{ax} i_x + L_{ay} i_y \\ \Psi_b &= L_{ba} i_a + L_b i_b + L_{bx} i_x + L_{by} i_y \\ \Psi_x &= L_{xa} i_a + L_{xb} i_b + L_x i_x + L_{xy} i_y \\ \Psi_y &= L_{ya} i_a + L_{yb} i_b + L_{yx} i_x + L_y i_y \end{aligned} \quad (2.129)$$

При відомих обмоткових функціях індуктивності обмоток можуть бути

знайдені за отриманими вище співвідношеннями. Наприклад, власну індуктивність обмотки a можна визначити як

$$L_a = \mu_0 \frac{\pi r l}{\delta} N_1^2.$$

Взаємна індуктивність тої самої обмотки, наприклад з обмоткою x , визначається як

$$L_{B3} = \mu_0 \frac{\pi r l}{\delta} N_1 N_2 \cos \varphi_r.$$

Таким чином, можна записати

$$\begin{aligned} L_a &= L_b = L_1 = \text{const}; \quad L_x = L_y = L_2 = \text{const}; \\ L_{ax} &= L_{xa} = L_{by} = L_{yb} = L_{B3} \cos \varphi_r; \\ L_{ay} &= L_{ya} = -L_{B3} \sin \varphi_r; \quad L_{bx} = L_{xb} = L_{B3} \sin \varphi_r; \\ L_{ab} &= L_{ba} = L_{xy} = L_{yx} = 0. \end{aligned} \tag{2.130}$$

Після підстановки (2.130) у (2.129), а отриманий результат у співвідношення (2.127) отримуємо рівняння рівноваги електричних кіл математичної моделі узагальненої електричної машини у вигляді

$$\begin{aligned} u_a &= r_1 i_a + L_1 \frac{di_a}{dt} + L_{B3} \frac{d}{dt} (i_x \cos \varphi_r) - L_{B3} \frac{d}{dt} (i_y \sin \varphi_r) \\ u_b &= r_1 i_b + L_1 \frac{di_b}{dt} + L_{B3} \frac{d}{dt} (i_x \sin \varphi_r) + L_{B3} \frac{d}{dt} (i_y \cos \varphi_r) \\ u_x &= L_{B3} \frac{d}{dt} (i_a \cos \varphi_r) + L_{B3} \frac{d}{dt} (i_b \sin \varphi_r) + L_2 \frac{di_x}{dt} + r_2 i_x \\ u_y &= -L_{B3} \frac{d}{dt} (i_a \sin \varphi_r) + L_{B3} \frac{d}{dt} (i_b \cos \varphi_r) + L_2 \frac{di_y}{dt} + r_2 i_y \end{aligned} \tag{2.131}$$

Зважаючи на те, що деякі взаємні індуктивності (див. співвідношення (2.130) залежать від кута φ_r , співвідношення для визначення моменту матимуть вигляд

$$\begin{aligned}
M_{\text{маг}} &= \frac{dW_{\text{маг}}}{d\varphi_r} = \\
&= i_a i_x \frac{dL_{ax}}{d\varphi_r} + i_a i_y \frac{dL_{ay}}{d\varphi_r} + i_b i_x \frac{dL_{bx}}{d\varphi_r} + i_b i_y \frac{dL_{by}}{d\varphi_r},
\end{aligned} \tag{2.132}$$

Відповідно до визначення зв'язок геометричних та електричних кутів визначається як

$$\varphi_e = p\varphi_r.$$

Якщо у співвідношення (2.132) підставити значення взаємних індуктивностей з (2.130) та виконати операції диференціювання, отримуємо

$$M_{\text{маг}} = pL_{\text{вз}} \left[(i_b i_x - i_a i_y) \cos \varphi_{re} - (i_a i_x - i_b i_y) \sin \varphi_{re} \right]. \tag{2.133}$$

З урахуванням наведеного рівняння механічної рівноваги матиме вигляд

$$\begin{aligned}
\frac{J}{p} \frac{\partial^2 \varphi_{re}}{\partial t^2} + M_2 &= \\
= L_{\text{вз}} \left[(i_b i_x - i_a i_y) \cos \varphi_{re} - (i_a i_x - i_b i_y) \sin \varphi_{re} \right] &= \frac{M_{\text{маг}}}{p}
\end{aligned} \tag{2.134}$$

Рівняння (2.131) та (2.134) повністю та цілком строго описують математичну модель електричної машини, що розглядається. Вони характеризують поведінку цієї моделі в будь-якому режимі роботи машини, включаючи її динамічні властивості.

На жаль, рівняння (2.131) та (2.134) нелінійні, адже містять добуток змінних. Це унеможливує їх безпосереднє розв'язання в загальному (аналітичному) вигляді.

2.5.3. Особливості математичних моделей електричних машин різних типів

Під час обертання ротора кут φ_{re} , який входить до рівняння (2.133) для визначення моменту за аналогією з рівнянням (2.69) може бути поданий як

$$\varphi_{re} = \omega_r t + \gamma \tag{2.135}$$

де

γ - кут між осями відповідних обмоток (в даному випадку між обмотками a та x) в момент часу $t = 0$.

Для реалізації процесу перетворення енергії усереднений за часом момент має відрізнятись від нуля, тобто

$$M_{\text{сер}} = \int_t M_{\text{маг}} dt \neq 0 \quad (2.136)$$

Як впливає з рівняння (2.133), співвідношення (2.136) означає, що інтеграли від складових

$$i_1 i_2 \cos(\omega_r t + \gamma) \text{ та } i_1 i_2 \sin(\omega_r t + \gamma),$$

де індекси відносяться до статорної та роторної обмотки відповідно, не мають давати нулеві значення.

Припустимо, що струм статора має циклічну частоту ω_1 , тобто

$$i_1 = I_{1m} \cos(\omega_1 t), \quad (2.137)$$

а струм ротора – частоту ω_2 , тобто

$$i_2 = I_{2m} \cos(\omega_2 t). \quad (2.138)$$

За таких умов з урахуванням відомої тригонометричної тотожності

$$\cos(\omega_1 t) \cos(\omega_2 t) \cos(\omega_r t) = \cos(\omega_1 \mp \omega_2 \mp \omega_r) t \quad (2.139)$$

можна зробити висновок, що значення інтегралу від косинуса з таким аргументом буде відрізнятись від нуля за умови, що

$$\omega_1 \pm \omega_2 \pm \omega_r = 0. \quad (2.140)$$

Отже,

співвідношення частот (2.140) є необхідною умовою для здійснення перетворення в узагальненій електричній машині електричної енергії в механічну.

Із вказаної умови випливають наступні *три* можливості.

1. Якщо

$$i_1 = I_{1m} \cos(\omega_r t), \text{ то } M_{\text{сер}} \neq 0 \text{ за умови } i_2 = \text{const},$$

тобто за умови, що

$$\omega_1 = \omega, \omega_2 = 0.$$

Таким умовам відповідає *синхронна* електрична машина, у якій частота струму в обмотці статора (ω_1) жорстко зв'язана з кутовою частотою обертання ротора (ω_r). При цьому в обмотці ротора тече постійний струм.

2. Якщо

$$i_1 = \text{const},$$

то для забезпечення умови $M_{\text{сер}} \neq 0$ необхідно, щоб

$$i_2 = I_{2m} \cos(\omega_2 t).$$

Звідси отримуємо умови

$$\omega_1 = 0, \omega_2 = \omega.$$

Така математична модель описуватиме

- *синхронну машину*, але «обернену» у порівнянні з визначеною у п. 1; це означає, що індуктор (система збудження) такої машини має розміщуватися на статорі;

- *машину постійного струму*.

3. $M_{\text{сер}} \neq 0$ також за умови, що

$$\omega_2 = \omega_1 - \omega_r,$$

тобто

$$i_2 = I_{2m} \cos[(\omega_1 - \omega_r)t], \quad i_1 = I_{1m} \cos(\omega_1 t).$$

Така математична модель описуватиме *асинхронну електричну машину*.

2.6. МАТЕМАТИЧНІ МОДЕЛІ ЕЛЕКТРИЧНИХ МАШИН РІЗНИХ ТИПІВ

2.6.1. Багатофазна синхронна машина

Найпростішою з точки зору принципу дії є синхронна машина.

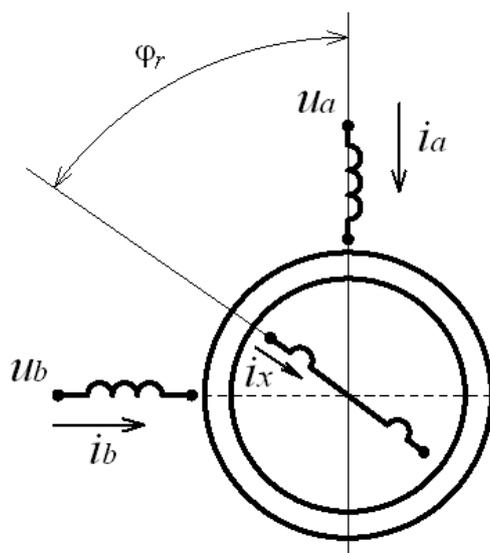


Рис. 2.21. Конструктивна модель двофазної синхронної машини

Елементарна конструктивна модель двофазної синхронної електричної машини наведена на рис. 2.21.

Така модель складається з двофазної симетричної обмотки статора, яка має синусоїдальне розподілення, та однієї синусоїдально розподіленої обмотки ротора.

Припустимо, що до статора такої машини прикладена симетрична двофазна система напруги, а саме

$$u_a = \sqrt{2}U_1 \cos(\omega_1 t);$$

$$u_b = \sqrt{2}U_1 \cos\left(\omega_1 t - \frac{\pi}{2}\right) = \sqrt{2}U_1 \sin(\omega_1 t).$$

Припустимо, далі, що в обмотці ротора синхронної машини протікає постійний струм

$$i_x = I_x.$$

Синхронний режим роботи електромеханічного перетворювача реалізується за умови, що поле статора обертається синхронно з полем ротора (тобто синхронно з ротором).

Електромагнітний момент в такому ЕМП утворюється як результат спрямування магнітних полів статора й ротора до положення, при якому їх просторові осі співпадатимуть. Згідно (2.73) величина цього моменту для основної гармоніки дорівнюватиме

$$M = -p\pi r l \delta H_m B_m \sin \theta. \quad (2.141)$$

Нагадаємо, що під час виведення цього співвідношення було зазначено: H_m , B_m можуть мати різне походження і їм відповідатиме цілком визначений кут θ .

Під час аналізу синхронних електричних машинах найзручніше вибрати прийняти, що H_m - це напруженість магнітного поля, утвореного обмоткою ротора, а B_m - індукція результуючого магнітного поля в повітряному зазорі. Кутом моменту за таких умов буде кут між осями магнітного поля ротора та результуючого магнітного поля в повітряному зазорі.

Прийнявши вказані умови, подамо H_m , B_m в термінах відповідних напруг та струмів.

Зважаючи на те, що обмотка ротора синусоїдально розподілена у просторі, H_m являтиме собою добуток

$$H_m = \frac{I_x}{\delta} N_x, \quad (2.142)$$

де

N_x - максимум обмоткової функції.

Величину індукції B_m можна знайти як функцію прикладеної напруги, якщо нехтувати активним опором та потоками розсіювання обмоток. За таких умов на підставі співвідношення (2.122) матимемо

$$B_m \approx \frac{\sqrt{2}U_1}{\pi r l \omega_1 N_a} \dots \quad (2.143)$$

Підставивши (2.142) та (2.143) у (2.141), матимемо

$$M = -\frac{P}{\omega_1} \frac{N_x}{N_a} \sqrt{2} U_1 I_x \sin \theta$$

або

$$M = -\left(\frac{\sqrt{2} N_x}{m_1 N_a} \right) \frac{m_1 U_1 I_x}{\Omega_{\text{мех}}} \sin \theta.$$

Таким чином, момент синхронної електричної машини пропорційний напрузі на затискачах, струму збудження (тобто струму ротора) і синусу просторового кута між осями магнітного поля в повітряному зазорі машини та магнітного поля ротора.

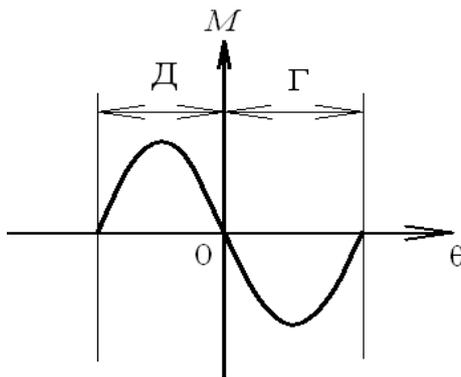


Рис. 2.22. Кутова характеристика моменту синхронної машини

Зважаючи на те, що U_1 , I_x можна регулювати зовні, змінення навантаження супроводжується зміненням кута моменту θ .

Залежність магнітного моменту електричної машини від кута моменту, зветься *кутовою характеристикою*.

Загальний вигляд кутової характеристики подано на рис. 2.22, з якого видно, що на кутовій характеристиці можна виділити *дві зони Д та Г*, при цьому

зоні Д відповідає режим роботи двигуна, а зоні Г – режим роботи генератора.

Загальне співвідношення між полем повітряного зазору та полями ротора і статора можна встановити за допомогою відповідних векторних діаграм (див. рис. 2.23).

Коли машина працює в режимі двигуна (рис. 2.23а), результуючий момент додатний, тобто діє в напрямку обертання ротора, а тому, як це видно зі співвідношення (2.141) кут θ має бути від'ємним.

За таких умов магнітне поле в повітряному зазорі машини дорівнюватиме сумі магнітних полів статора і ротора. Якщо момент навантаження, прикладений до валу ротора, зменшується, то при незмінній напрузі статора та незмінному струмі збудження кут моменту зменшується. Як видно з рис. 2.23а, зменшуватиметься при цьому й величина поля статора.

Режиму холостого ходу відповідатиме діаграма, подана на рис. 2.23б.

Коли машина не розвиває моменту, кут моменту дорівнює нулю й осі магнітних полів повітряного зазору та ротора співпадатимуть – теоретично їх значення мають бути однакові.

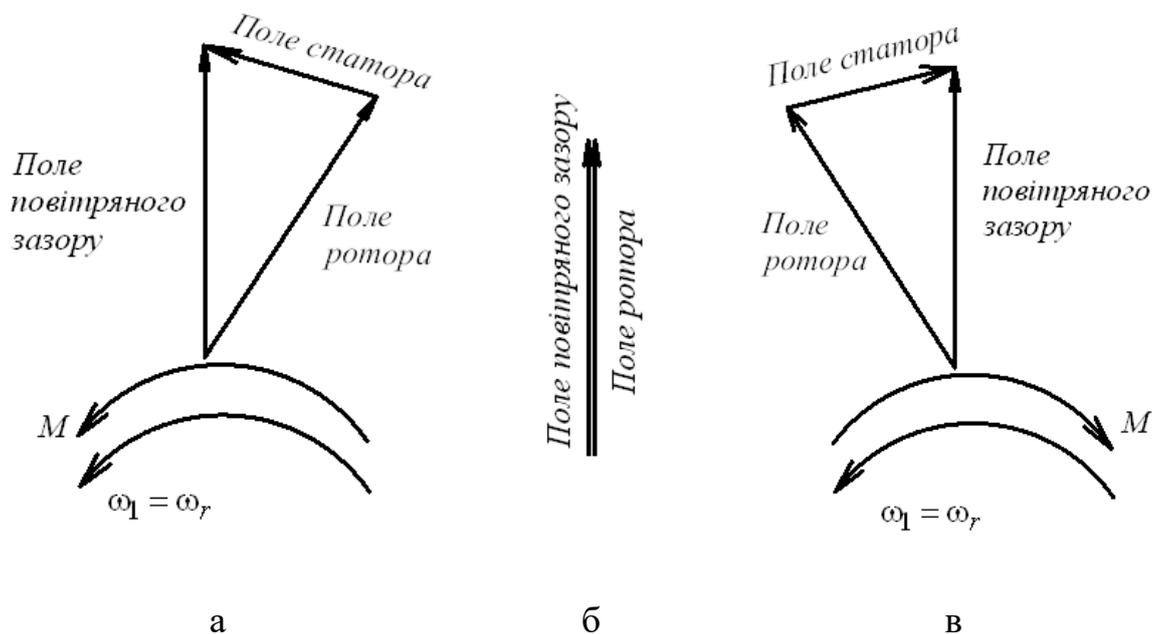


Рис. 2.23. Векторні діаграми магнітних полів синхронної машини в режимі двигуна (а), холостого ходу (б) та генератора (в)

Зменшуючи навантаження далі, ми маємо прикласти до валу машини момент, спрямований у протилежний бік. Тоді машина буде розвивати від’ємний момент, тобто момент, що діє у напрямку, протилежному напрямку обертання (кут θ має додатне значення). Іншими словами машина працюватиме в режимі генератора (рис. 2.23в). Поле статора знову збільшується відповідно до збільшення струму статора.

Той факт, що поле статора змінюється як за величиною, так і за напрямком, означає, що струм статора змінює свою величину і фазу у часі.

В реальній синхронній електричній машині неможливо повністю знищити гармонічні складові магнітного поля. Поля повітряного зазору та ротора містять різні гармонічні складові, які можуть взаємодіяти, створюючи відповідні гармонічні складові моменту.

Будь-яка гармоніка не утворює середнього синхронного моменту, адже її поле обертається зі швидкістю більш низькою, ніж швидкість обертання основного магнітного поля. Однак якщо якісь гармоніки присутні одночасно в магнітних полях повітряного зазору та ротора, то при плавному відносному переміщенні цих полів з’являться *пульсації* моменту. Частота цих пульсацій залежатиме від порядку гармонік, що взаємодіють. Такі пульсації зазвичай суттєво не порушують роботу синхронних електричних машин.

2.6.2. Багатофазна асинхронна електрична машина

Обертювий момент в багатофазній асинхронній електричній машині

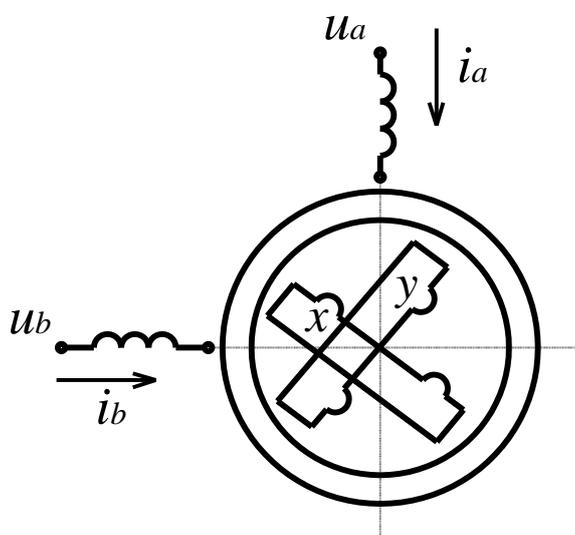


Рис. 2.24. Конструктивна модель асинхронної двофазної машини

утворюється за результатом взаємодії магнітного поля, що обертається в повітряному зазорі, з полем ротора, яке обумовлене струмами, наведеними в короткозамкненій обмотці ротора вказаним обертювим магнітним полем.

Найпростіша конструктивна модель такої машини (див. рис. 2.24) складається із статора з симетричною двофазною обмоткою та ротора, який також має симетричну двофазну обмотку, кожна фаза якої замкнена накоротко.

Припустимо, що обмотка статора машини живиться симетричною двофазною системою напруг, які можна подати як

$$u_a = \sqrt{2}U_1 \cos(\omega_1 t), \quad u_b = \sqrt{2}U_1 \sin(\omega_1 t).$$

Асинхронному режиму роботи електричної машини відповідає будь-яка швидкість обертання ротора, крім синхронної.

Момент, що розвивається асинхронною машиною можна знайти на підставі загального співвідношення (2.141) для визначення моменту.

Вважатимемо, що величина H_m відповідає амплітуді напруженості магнітного поля ротора, а B_m - амплітуді індукції магнітного поля в повітряному зазорі машини. Отже кут θ аналогічний куту моменту синхронної машини.

Процес утворення обертювого моменту в асинхронній машині не можна подати настільки наочно, як це було зроблено під час аналізу синхронної машини. Головне відрізнєння полягає в тому, що за будь-якої швидкості обертання ротора його обмотки переміщуються відносно поля повітряного зазору. Тому в цих обмотках наводиться ЕРС, величина та частота якої прямо пропорційна різниці швидкостей обертання магнітного поля в повітряному зазорі та ротора.

Наведена ЕРС викликає струм в короткозамкненій обмотці ротора. При цьому внаслідок симетрії двофазної обмотки ротора з'являється відповідне обертюве магнітне поле. Це поле обертається відносно ротора зі

швидкістю, яка пропорційна частоті струму в обмотках ротора. Однак ця частота, в свою чергу, пропорційна різниці швидкостей обертання ротора та поля в повітряному зазорі. Отже, сума швидкостей обертання ротора та його магнітного поля завжди дорівнюватиме синхронній швидкості, а тому поле ротора обертається синхронно з полем повітряного зазору при будь-якій швидкості обертання ротора.

Таким чином, в асинхронній машині момент утворюється за будь-якої швидкості обертання ротора, завдяки здатності поля ротора завжди залишатися нерухомим відносно магнітного поля повітряного зазору.

З метою отримання зручного для практичного використання співвідношення для розрахунку моменту перш за все встановимо відповідні співвідношення для обчислення величин H_m , B_m , θ .

Оскільки поле в повітряному зазорі визначається прикладеною напругою, то співвідношення для обчислення B_m матиме той самий вигляд, що і для синхронної машини (див. 2.143)), а саме

$$B_m \approx \frac{\sqrt{2}U_1}{\pi r l \omega_1 N_a}.$$

Амплітуда напруженості обертового магнітного поля ротора може бути подана як функція струму ротора, а саме

$$H_m = \frac{\sqrt{2}I_x}{\delta} N_x. \quad (2.144)$$

ЕРС ротора створюється обертовим полем повітряного зазору, а тому неважно записати співвідношення, яке пов'язуватиме цю ЕРС з напругою на обмотці статора. Ці дві величини розрізняються лише тому, що обмотки можуть мати різну кількість витків і поле повітряного зазору переміщується відносно цих обмоток з різною швидкістю.

За прийнятих раніше припущень ($r_1 = x_1 = 0$) напруга на обмотці статора дорівнюватиме її ЕРС, а ЕРС статора та ротора лінійно залежать від швидкості обертання поля та числа витків обмоток, тобто

$$E_x \approx U_1 \frac{N_x}{N_a} \frac{\omega_1 - \omega_r}{\omega_1}, \quad (2.145)$$

де величиною різниця $(\omega_1 - \omega_r)$ визначається швидкість обертання ротора по відношенню до поля повітряного зазору.

Ввівши величину

$$s = \frac{\omega_1 - \omega_r}{\omega_1}, \quad (2.146)$$

яка зветься ковзанням, співвідношення (2.145) можна подати у вигляді

$$E_x \approx s \frac{N_x}{N_a} U_1. \quad (2.147)$$

Діюче значення струму ротора визначається як

$$I_x = \frac{E_x}{Z_x} \approx s \frac{N_x}{N_a} \frac{U_1}{Z_x}, \quad (2.148)$$

де

Z_x - повний опір фази ротора.

Зазначимо, що нехтування активним опором та потоком розсіювання обмотки статора не приводить до помітного викривлення результатів аналізу. В той же час, нехтувати активним опором обмотки ротора не можна, адже за таких припущень $Z_x = 0$ і співвідношення (2.148) втрачає зміст.

За нормальних умов роботи машини ЕРС, що наводиться в обмотці статора, дуже близька до прикладеної напруги. Зважаючи на те, що в середньому ковзання асинхронного двигуна в нормальному режимі роботи мало відхиляється від величини 0,03, відповідно до (2.147) приблизно таку саму величину матиме ЕРС ротора. Ще раз підкреслимо, що, незважаючи на малу величину Z_x , нехтувати цією величиною не можна, адже пов'язана з нею ЕРС від потоку розсіювання складає значну частину всієї ЕРС ротора.

Частота струму ротора може бути записана як

$$f_x = f_1 \frac{\omega_1 - \omega_r}{\omega_1} = f_1 s, \quad (2.149)$$

де

f_1 - частота напруги, прикладеної до обмотки статора.

Повний опір фази короткозамкненої обмотки ротора

$$Z_x = \sqrt{r_x^2 + (2\pi f_x L_x)^2} = \sqrt{r_x^2 + (s\omega_1 L_x)^2}, \quad (2.150)$$

де

ω_1 - кутова частота електромережі.

За таких умов діюче значення струму ротора визначатиметься як

$$I_x = \frac{N_x}{N_a} \frac{sU_1}{\sqrt{r_x^2 + (s\omega_1 L_x)^2}}. \quad (2.151)$$

Для визначення величини кута θ розглянемо векторну діаграму, яка відповідає моменту часу, коли ЕРС в обмотці ротора дорівнюватиме

максимальній величині (див. рис. 2.25).

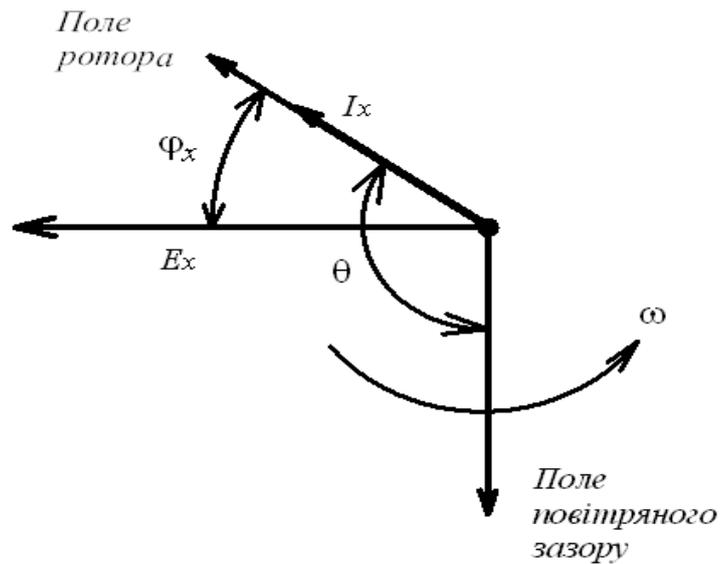


Рис. 2.25. Векторна діаграма магнітних полів асинхронної машини

ЕРС, яка наводиться обертовим магнітним полем в обмотці ротора, відстає за фазою від збуджуючого її потоку на кут $\pi/2$. В свою чергу, струм в реальній обмотці відстає за фазою у часі від наведеної ЕРС на кут ψ_x . Отже, вісь поля ротора має відставати від осі поля повітряного зазору на кут $(\pi/2 - \psi_x)$.

За таких умов, як видно з рис. 2.25, кут моменту визначатиметься як

$$\theta = -(\pi/2 + \psi_x). \quad (2.152)$$

З урахуванням усього наведеного момент, який розвивається асинхронною електричною машиною, можна подати як

$$M = \frac{2p}{\omega_1} \left(\frac{N_x}{N_a} \right)^2 \frac{sU_1^2}{\sqrt{r_x^2 + (s\omega_1 L_x)^2}} \cos \psi_x. \quad (2.153)$$

Коефіцієнт потужності кола ротора дорівнюватиме

$$\cos \psi_x = \frac{r_x}{Z_x} = \frac{r_x}{\sqrt{r_x^2 + (s\omega_1 L_x)^2}}. \quad (2.154)$$

Таким чином, остаточно співвідношення для визначення моменту приймає вигляд

$$\begin{aligned}
 M &= \frac{2p}{\omega_1} \left(\frac{N_x}{N_a} \right)^2 \frac{r_x s U_1^2}{r_x^2 + (s\omega_1 L_x)^2} = \\
 &= \left[\frac{1}{m_1} \left(\sqrt{2} \frac{N_x}{N_a} \right)^2 \right] \frac{m_1 r_x s U_1^2}{\Omega_{\text{мех}} \left[r_x^2 + (s\omega_1 L_x)^2 \right]}.
 \end{aligned}
 \tag{2.155}$$

Зазначимо, що нехтування величиною активного опору та величиною індуктивності розсіювання обмотки статора впливатиме виключно на величину отриманого значення моменту, тоді як форма залежності $M = f(s)$ суттєво не змінюватиметься.

Найчастіше асинхронна машина приєднується до електромережі стабільної напруги постійної частоти. За таких умов у співвідношенні (2.155) залишається лише дві змінні величини s та M .

Взаємозв'язок вказаних змінних зветься *механічною характеристикою* асинхронної електричної машини.

Механічну характеристику прийнято зображати графічно. Конкретна форма цієї кривої залежить від параметрів машини. Найбільш типовий вигляд механічної характеристики наведено на рис. 2.26.

Як видно з рис. 2.26 шкалу швидкості (ковзання) можна розділити на *три* області, в яких поведінка машини суттєво різниться.

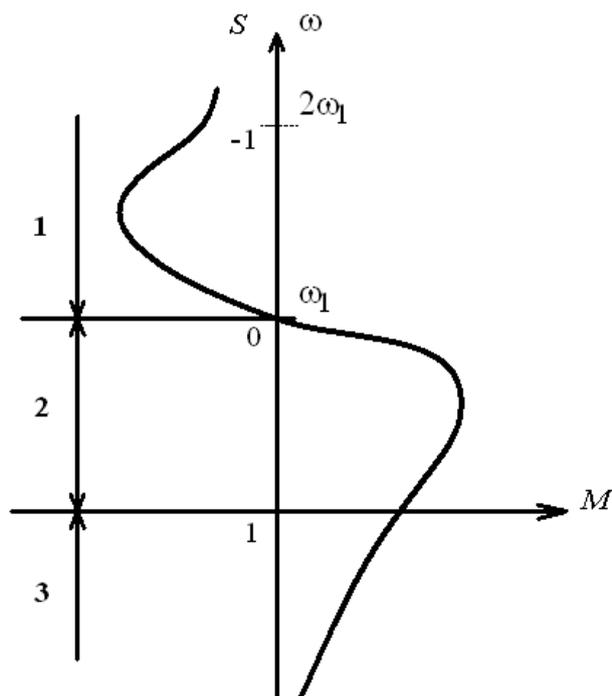


Рис. 2.26. Типова форма механічної характеристики асинхронної машини

За умови ковзання більше 1 (тобто за умови від'ємних швидкостей) асинхронна машина розвиватиме додатний момент. Але оскільки швидкість від'ємна, момент діє назустріч обертанню ротора (машина працює в режимі гальмування), а тому споживає як електричну, так і механічну енергію (область 3).

Область між величинами ковзання від 1 до 0 (тобто між нульовою та синхронною швидкостями обертання) відповідає роботі машини в режимі двигуна, адже момент, що розвивається, та швидкість мають додатні значення (область 2). Саме ця область має найбільше практичне значення.

При від'ємних значеннях ковзання момент стає від'ємним і машина працюватиме в режимі генератора. (область 1).

Зважаючи на те, що більшість асинхронних машин працюють в режимі двигуна, обмежимося розгляданням лише цього режиму.

Робота асинхронної машини в режимі двигуна

Швидкість обертання ротора асинхронного двигуна визначається моментом навантаження на його валу. Номінальний режим роботи, як вже було зазначено, має місце при швидкостях обертання дуже близьких до синхронної. Механічна характеристика в цій області мало відрізняється від лінійної, в чому легко переконатися, адже за умови $s \rightarrow 0$ виконується нерівність

$$r_x \gg s\omega_1 L_x.$$

За таких умов співвідношення (2.155), яким визначається момент, суттєво спрощується, приймаючи вигляд

$$M \approx \frac{2p}{\omega_1} \left(\frac{N_x}{N_a} \right)^2 \frac{U_1^2}{r_x} s. \quad (2.156)$$

Для оцінювання впливу на величину моменту різних параметрів, знайдемо ковзання, за якого момент досягатиме максимального значення. З цією метою диференціюємо співвідношення (2.156) по s , а отриману похідну прирівнюємо до нуля. Як результат отримуємо

$$s_m = \frac{r_x}{\omega_1 L_x}. \quad (2.157)$$

Підставивши (2.157) в (2.156) знаходимо максимальне значення моменту, а саме

$$M_m = \frac{p}{\omega_1} \left(\frac{N_x}{N_a} \right)^2 \frac{U_1^2}{\omega_1 L_x}. \quad (2.158)$$

Отже, при будь-якому змінній r_x змінюється положення точки максимального моменту, а його величина залишається незмінною. Це надає можливість управляти характером залежності $M = f(s)$ шляхом регулювання відношення $\frac{r_x}{\omega_1 L_x}$.

Так же як і в синхронній машині, в реальному асинхронному двигуні неможливо забезпечити ідеальне синусоїдальне розподілення магнітного поля в повітряному зазорі. Внаслідок того, що число провідників обмотки завжди кінцеве в полі повітряного зазору завжди присутні гармонічні складові.

Якщо обмотки ротора та статора містять гармонічні складові з однаковою кількістю полюсів, то ці полюси можуть взаємодіяти між собою, утворюючи відповідні моменти (вищого порядку), як і основні обмотки.

Загальний ефект від дії моментів вищого порядку буде таким, нібито з валом даної машини скріплені вали інших машин, які мають більшу кількість полюсів. Зважаючи на те, що деякі гармонічні складові обертаються назустріч основному полю, вони утворюють від'ємні моменти. За таких умов можливе зниження моменту, що розвивається машиною, та поява провалів на результуючій механічній характеристиці.

2.6.3. Електрична машина постійного струму

Як було зазначено вище, в машині постійного струму магнітні поля

ротора й статора нерухомі у просторі, тобто можуть взаємодіяти та створювати відповідний момент.

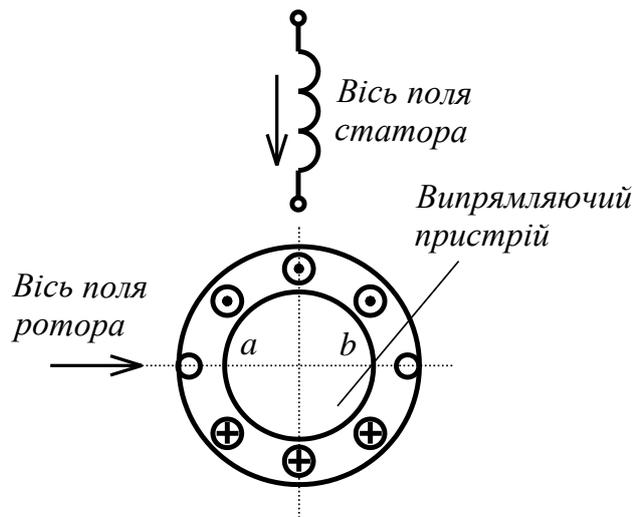


Рис. 2.27. Схема устрою машини постійного струму

Поле ротора залишається нерухомим під час обертання ротора завдяки спеціальному пристрою (*колектору*), що знаходиться між зовнішньою електричною мережею та обмоткою ротора. Хоча провідники ротора переміщуються (див. рис. 2.27), струм в кожному з них перемикається (за допомогою колектора) так, щоб поле в повітряному зазорі машини залишалось нерухомим у просторі. Це поле буде еквівалентне полю, що утворюється розподіленою нерухомою обмоткою.

Момент, що розвивається машиною постійного струму, як і раніше, визначимо на підставі основного співвідношення (2.141), яке для зручності наведемо ще раз

$$M = -p\pi r l \delta H_m B_m \sin \theta. \quad (2.159)$$

Припустимо, що поле статора орієнтоване у просторі так, як показано на рис. 2.27. Зрозуміло, що вибравши положення поля ротора відносно поля статора, ми визначаємо кут моменту. З метою досягнення найбільшої ефективності цей кут в машинах постійного струму складає $\pi/2$.

Припустимо далі, що B_m є максимальною індукцією поля статора, тоді H_m має відповідати амплітуді напруженості магнітного поля ротора.

Зважаючи на те, що рівномірно розподілена обмотка створює у просторі поле, яке має форму трикутника (див. рис. 2.28), напруженість магнітного поля ротора можна подати рядом Фур'є, а саме у вигляді

$$h_x(\varphi) = \frac{8}{\pi^2} \left(\frac{1}{8p} \frac{n_x I_x}{\delta} \right) \left(\sin \varphi - \frac{\sin(3\varphi)}{3^2} + \frac{\sin(5\varphi)}{5^2} - \dots \right) \quad (2.160)$$

де

n_x - загальна кількість провідників на поверхні ротора.

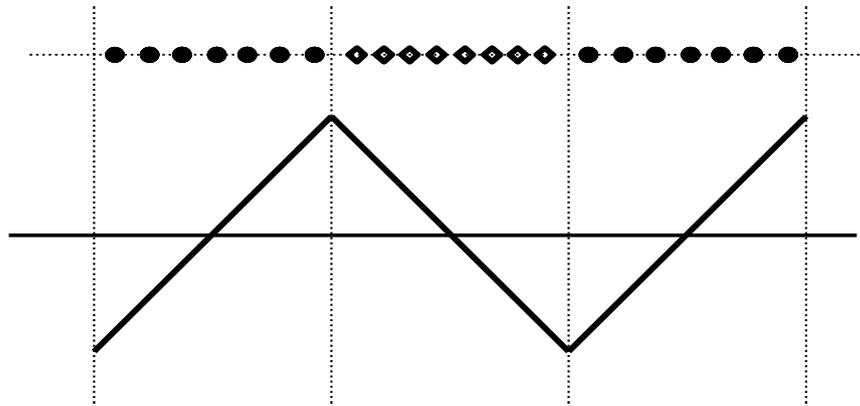


Рис. 2.28. Форма магнітного поля, яке утворюється рівномірно розподіленою обмоткою

Більшість машин постійного струму мають зосереджену обмотку статора, яка зазвичай розміщується на явно виражених полюсах. Розподілення магнітна індукції при цьому матиме вигляд прямокутної хвилі розкладання якої в ряд Фур'є можна подати як

$$b_a(\varphi) = \frac{4}{\pi} B_a \left(\cos \varphi - \frac{\cos(3\varphi)}{3} + \frac{\cos(5\varphi)}{5} - \dots \right). \quad (2.161)$$

Отже, магнітні поля статора та ротора мають вигляд, як на рис. 2.29.

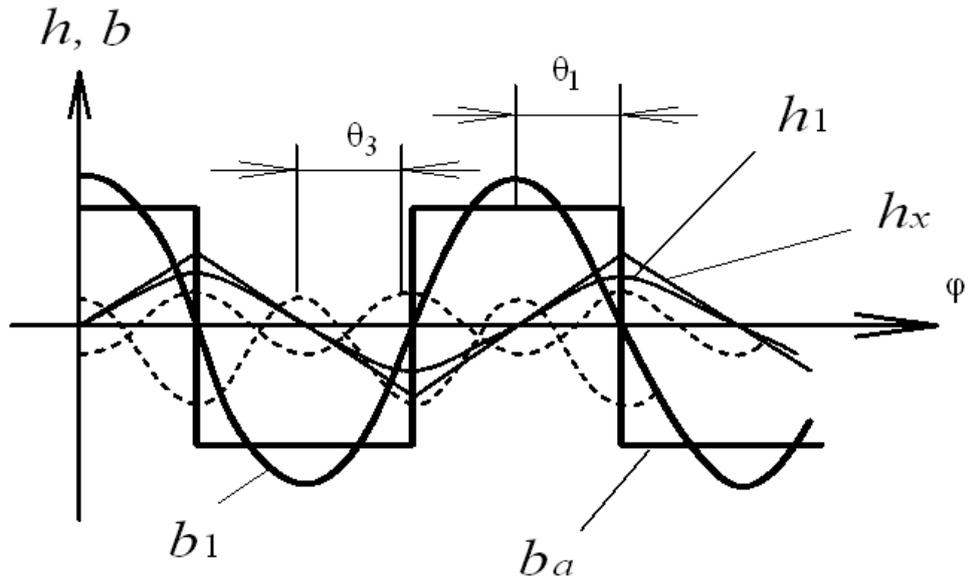


Рис. 2.29. Магнітні поля ротора та статора в повітряному зазорі машини постійного струму

Як видно з рис. 2.29, момент ν -ї гармоніки матиме той самий знак, що й основний момент, тобто діє в ту саму сторону. Кути зсуву обох гармонік складають оптимальну величину $\pi/2$. Величини цих моментів можна отримати за допомогою основного співвідношення, записаного для відповідної гармонічної складової.

Маємо

$$M_{\nu} = -(\nu p) \pi r l \delta H_{\nu m} B_{\nu m} \sin(\nu \theta_{\nu}). \quad (2.162)$$

У співвідношенні (2.162) максимальні значення (амплітуди) гармонік магнітних полів статора та ротора згідно співвідношень (2.160) та (2.161) відповідно, визначатимуться як

$$B_{\nu m} = \frac{B_{1m}}{\nu} = \frac{B_m}{\nu}, \quad H_{\nu m} = \frac{H_{1m}}{\nu^2} = \frac{H_m}{\nu^2}. \quad (2.163)$$

Зважаючи на те, що кут моменту $(\nu \theta)$ для всіх гармонік дорівнює $\pi/2$, співвідношення (2.162) можна подати як

$$M_{\nu} = -p \pi r l \delta H_{1m} B_{1m} \frac{1}{\nu^2}. \quad (2.164)$$

На підставі (2.164) загальна величина моменту може бути подана у вигляді

$$M = M_1 \sum_{v=1}^{\infty} \frac{1}{v^2}. \quad (2.165)$$

При непарному значенні v матимемо

$$\sum_{v=1}^{\infty} \frac{1}{v^2} = \frac{\pi^2}{8}.$$

і співвідношення для визначення моменту набуває вигляду

$$M = \frac{\pi^2}{8} M_1. \quad (2.166)$$

Таким чином, на відміну від машин змінного струму в машині постійного струму гармоніки збільшують величину моменту в $\pi^2/8$ разів, що відповідає величині, яка приблизно на 23 % перевищує момент основної гармоніки.

Використовуючи загальне співвідношення для основної гармоніки моменту, подамо момент, що розвивається машиною постійного струму, в дещо іншій формі запису.

При взаємодії основних гармонік магнітних полів маємо

$$M_1 = p\pi r l \delta H_{1m} B_{1m} = \left(\frac{\pi r l}{p} B_{a1m} \right) p^2 \delta H_{x1m}. \quad (2.167)$$

З урахуванням співвідношень (2.160) та (2.161) отримуємо

$$M_1 = \left(\frac{\pi r l}{p} B_a \right) \frac{pn_x}{2\pi} \frac{8}{\pi^2} I_x. \quad (2.168)$$

Врахувавши (2.166), отримуємо

$$M = k \Phi_a I_x, \quad (2.169)$$

де

$$\Phi_a = \frac{\pi r l}{p} B_a \text{ - магнітний потік на полюс;}$$

$$k = \frac{pn_x}{2\pi} \text{ - коефіцієнт, який визначається параметрами обмотки ротора,}$$

тобто є константою для кожної машини.

На завершення аналізу машини постійного струму дослідимо два її електричних кола, а саме: коло статора та коло ротора

В обмотці статора (збудження) в усталеному режимі роботи машини ЕРС не наводиться, оскільки поле повітряного зазору нерухоме у просторі та має постійну амплітуду.

Обмотка ротора, внаслідок її руху, має змінне потокозчеплення, а тому в

цій обмотці наводиться ЕРС. Не зважаючи на те, що в кожній котушці обмотки ротора наводиться змінна ЕРС, дія колектора така, що напруга у зовнішньому колі (між точками a , b на рис. 2.27) є напругою постійного струму. Всі котушки обмотки ротора в кожній з паралельних віток між вказаними точками з'єднані послідовно. Оскільки всі котушки однієї вітки знаходяться в односпрямованому магнітному полі й рухаються в одному напрямку, вони мають однакову полярність. Отже, напруга між точками a , b завжди матиме однакову полярність.

Для визначення наведеної ЕРС скористаємось методом, заснованим на рівнянні балансу потужностей.

Механічна потужність, що розвивається машиною постійного струму, при швидкості $\Omega_{\text{мех}}$ та моменті M визначаються як

$$P_{\text{мех}} = M\Omega_{\text{мех}} = k\Phi_a I_x \Omega_{\text{мех}}. \quad (2.170)$$

Оскільки в обмотці статора не наводиться ЕРС, вся ця потужність має виникати, завдяки наявності ЕРС в роторі. Таким чином, електромагнітна потужність в колі ротора, яка може бути перетворена в механічну, визначається як

$$P_{\text{ем}} = E_x I_x. \quad (2.171)$$

Якщо прирівняти праві частини співвідношень (2.170) та (2.171), отримуємо

$$E_x = k\Phi_a \Omega_{\text{мех}}. \quad (2.172)$$

2.7. ЕКВІВАЛЕНТНІ ЕЛЕКТРИЧНІ СХЕМИ ЕЛЕКТРОМЕХАНІЧНИХ ПЕРЕТВОРЮВАЧІВ ЕНЕРГІЇ

*Математик цінує красоту формул,
а інженера, навіть за умови його
високої чутливості до елегантності,
перш за все турбують задачі
практичного призначення*

NM

Еквівалентні електричні схеми (еквівалентні схеми заміщення – ЕСЗ), широко використовуються при аналізі електромеханічних перетворювачів енергії, потужних електроенергетичних систем, ліній передавання електроенергії тощо.

Як правило, ЕСЗ складаються з пасивних елементів (резисторів, індуктивностей, ємностей), з'єднаних так, щоб отримана електрична схема описувалася тими ж диференціальними рівняннями, що і схема, яка досліджується.

Якщо елементами ЕСЗ є лише лінійні елементи (тобто елементи, параметри яких не залежать від струмів та напруг, що діють в даному електричному колі), схему можна описати лінійними диференціальними рівняннями. Такі схеми заміщення зазвичай називають *лінійними* схемами.

Лінійними схемами заміщення можна представляти виключно клас лінійних систем, що, звичайно, обмежує їх використання. Однак, завдяки своїй простоті лінійні ЕСЗ широко використовуються й для аналізу нелінійних систем після введення відповідних обмежень, які дозволяють лінеаризувати систему в визначеному діапазоні зміни параметрів.

Найширше ЕСЗ використовуються для аналізу систем, що збуджуються синусоїдальними сигналами незмінної частоти. Ефективним засобом розв'язання задач, які виникають при аналізі таких систем, є комбінація ЕСЗ та відповідних векторних діаграм.

Розглянемо, як приклад, побудову схеми заміщення синхронної електричної машини, на підставі загальних співвідношень, що описують її роботу та надані в попередніх розділах.

Нагадаємо умови, за яких узагальнена модель електромеханічного перетворювача енергії буде відображати синхронну машину:

- симетрична двофазна напруга живлення обмотки статора

$$u_a = \sqrt{2}U \sin(\omega t); \quad u_b = -\sqrt{2}U \cos(\omega t);$$

- постійний струм збудження в одній обмотці ротора та відсутність струму в другій

$$i_x = -I; \quad i_y = 0;$$

- постійна швидкість обертання ротора

$$\Omega_r = \Omega_{\text{мех}} = \frac{\omega}{p}.$$

Нагадаємо, що кутове положення ротора машини характеризується кутом

$$\varphi_r = \omega t + \theta,$$

де

θ - кут початкового положення ротора в момент часу $t = 0$.

За наведених умов загальні співвідношення, якими визначається двофазна напруга живлення обмоток статора, набувають вигляду

$$u_a = \sqrt{2}U \sin(\omega t) = r_1 i_a + L_1 \frac{di_a}{dt} + \omega L I \sin(\omega t + \theta)$$

$$u_b = -\sqrt{2}U \cos(\omega t) = r_1 i_b + L_1 \frac{di_b}{dt} + \omega L I \cos(\omega t + \theta).$$

$$i_x = -I; i_y = 0$$

$$m = pLI \left[i_a \sin(\omega t + \theta) - i_b \cos(\omega t + \theta) \right].$$

Перші два з чотирьох наведених співвідношень дозволяють зразу отримати схему заміщення синхронної машини, як це наведено на рис. 2.30.

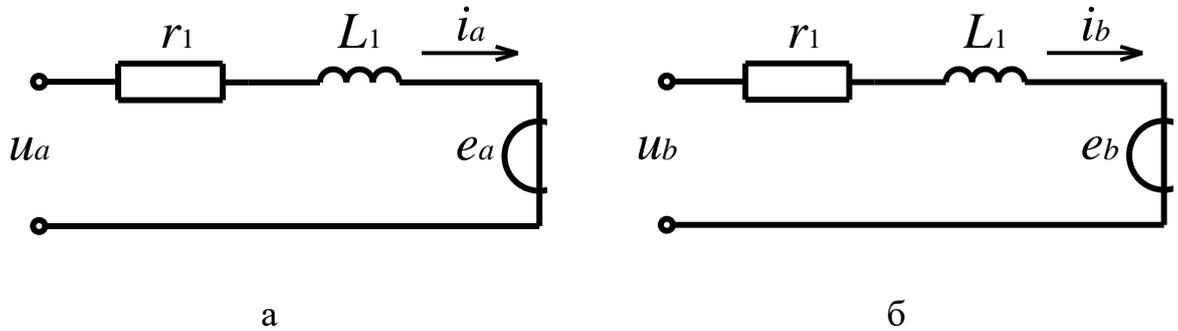


Рис. 2.30. Еквівалентні схеми заміщення синхронної електричної машини

Вочевидь, кола, наведені на рис. 2.30, точно відповідають вказаним співвідношенням.

Для практичних цілей, звичайно, достатньо однієї з поданих на рис. 2.30 схем електричних кіл, адже вони ідентичні: різниця полягає лише в тому, що струми та напруги цих кіл зсунуті за фазою на кут $\pi/2$. Тому прийнято зображати лише одну схему, розглядаючи її як ЕСЗ відповідної фази машини.

Зважаючи на те, що всі напруги та струми синусоїдальні, на ЕСЗ зазвичай вказують комплексні значення параметрів та змінних. ЕРС зветься зовнішньою або наведеною, а опір $x_1 = \omega L_1$ - синхронним індуктивним опором.

Момент (миттєве значення), що розвивається машиною, зручно також зв'язати зі схемою заміщення.

З цією метою перепишемо співвідношення для визначення миттєвого значення моменту у вигляді

$$\begin{aligned} m &= \frac{P}{\omega} \left[i_a \omega LI \sin(\omega t + \theta) - i_b \omega LI \cos(\omega t + \theta) \right] = \\ &= \frac{P}{\omega} (e_a i_a - e_b i_b). \end{aligned} \tag{2.173}$$

Із співвідношення (2.173) видно, що момент пропорційний загальній миттєвій електромагнітній потужності p .

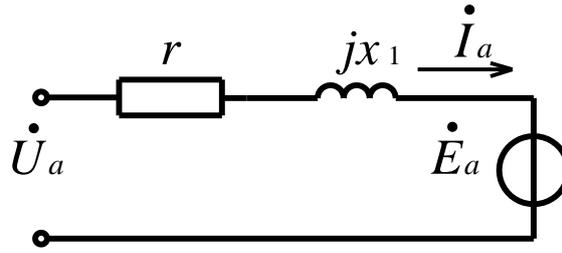


Рис. 2.31. Спрощена схема заміщення фази обмотки

Щоб отримати просту та зручну до практичного використання формулу для визначення моменту, нехтуватимемо активним опором обмотки статора. Тоді струм в електричному колі, поданому на рис. 2.31 дорівнюватиме

$$\dot{i}_a = \frac{\dot{U}_a - \dot{E}_a}{jx_1} = -j \frac{\dot{U}_a}{x_1} + j \frac{\dot{E}_a}{x_1}. \quad (2.174)$$

На підставі співвідношення (2.174) можна записати миттєве значення струму, що протікає обмоткою, у вигляді

$$i_a = -\frac{\sqrt{2}U}{x_1} \cos(\omega t) + \frac{\omega LI}{x_1} \cos(\omega t + \theta). \quad (2.175)$$

Струм i_b матиме по відношенню до струму i_a фазовий зсув $\pi/2$, тобто

$$i_b = -\frac{\sqrt{2}U}{x_1} \sin(\omega t) + \frac{\omega LI}{x_1} \sin(\omega t + \theta). \quad (2.176)$$

Підставляючи (2.175) та (2.176) у співвідношення (2.173), після незначних спрощень отримуємо

$$M = -pLI \frac{\sqrt{2}U}{x_1} \sin \theta. \quad (2.177)$$

Введемо у співвідношення (2.177) діюче значення наведеної ЕРС

$$E_a = \frac{\omega LI}{\sqrt{2}}.$$

Як результат отримуємо співвідношення для визначення моменту у вигляді

$$M = -\frac{2p}{\omega} \frac{UE_a}{x_1} \sin \theta = -\frac{2}{\Omega_{\text{мех}}} \frac{UE_a}{x_1} \sin \theta. \quad (2.178)$$

Отримане співвідношення дуже схоже на те, яке було отримане нами за

допомогою аналізу магнітних полів, що обертаються. Ці співвідношення будуть цілком однакові за умови нехтування потоком розсіювання статора. В цьому легко переконатися, записавши співвідношення, отримане в розділі 2.6.1. у вигляді

$$M = -p \frac{\sqrt{2}L_{3B}IU}{\omega L_1} \sin \theta.$$

За умови відсутності потоку розсіювання статора матимемо

$$\frac{L_{3B}}{L_1} = \frac{N_2}{N_1}.$$

Як результат для визначення моменту синхронної машини остаточно матимемо формулу

$$M = -\frac{2p}{\omega} \frac{N_2}{N_1} \sqrt{2}UI \sin \theta.$$

2.8. ЕЛЕКТРОМАГНІТНІ ПРИСТРОЇ ТРАНСФОРМАЦІЇ ЕНЕРГІЇ

За умови, що провідники та магнітне поле не мають відносно один одного переміщення, не може звичайно мати місце взаємне перетворення електричної та механічної енергії. В цьому окремому випадку можливий лише процес перетворення електричної енергії в електричну енергію, часто зі зміною повного опору системи.

Електромагнітні пристрої, що виконують вказану функцію звуться трансформаторами енергії або просто – *трансформаторами*.

2.8.1. Фізичні засади дії трансформатора

На рис. 2.32 наведена одна з можливих схем конструктивного виконання (конструктивна модель) трансформатора, а саме схема однофазного двох обмоткового трансформатора.

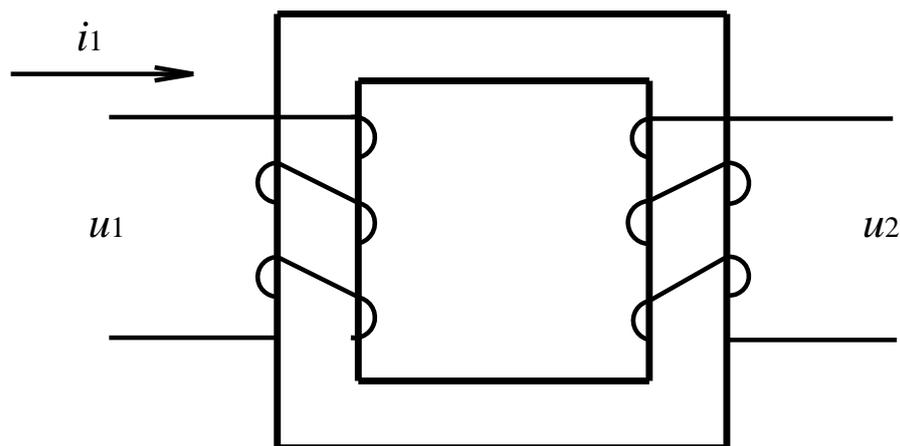


Рис. 2.32. Схема однофазного двохобмоткового трансформатора

Робочий процес у трансформаторі пояснюється наведенням ЕРС в провіднику при зміні магнітного потоку у часі.

Одна з обмоток трансформатора зазвичай зветься *первинною*, а друга – *вторинною*. Вторинною обмоткою вважається та, до якої приєднане навантаження. Отже потужність передається із первинної обмотки у вторинну.

Зазначимо, що трансформатор може встановлюватися і в таких мережах, де напрямок потоку потужності час від часу змінюється. За таких умов наведене уявлення про обмотки, звичайно, втрачає зміст, а тому доцільніше обмотки трансформатора просто нумерувати.

При використанні трансформаторів в електромережах можуть реалізуватися гальванічні зв'язки між різними його обмотками. При цьому важливо знати залежність між полярністю ЕРС (магнітних потоків) та напрямком струмів в обмотках.

Домовимось напрямком струмів, наприклад, в двох індуктивно зв'язаних котушках вважати узгодженими (додатними), якщо напрямки магнітних потоків, що цими струмами утворюються, співпадають.

Напрямок струм та утвореного ним магнітного потоку узгоджуються відповідно до правила правоходового гвинта (штопора). Відповідно до цього правила, напрямок магнітного потоку співпадає з напрямком поступового руху гвинта за умови, що він обертається згідно того як циркулює струм витками котушки.

На рис. 2.33 наведені дві індуктивно зв'язані котушки на загальному осерді.

Напрямок переміщення
правоходового гвинта

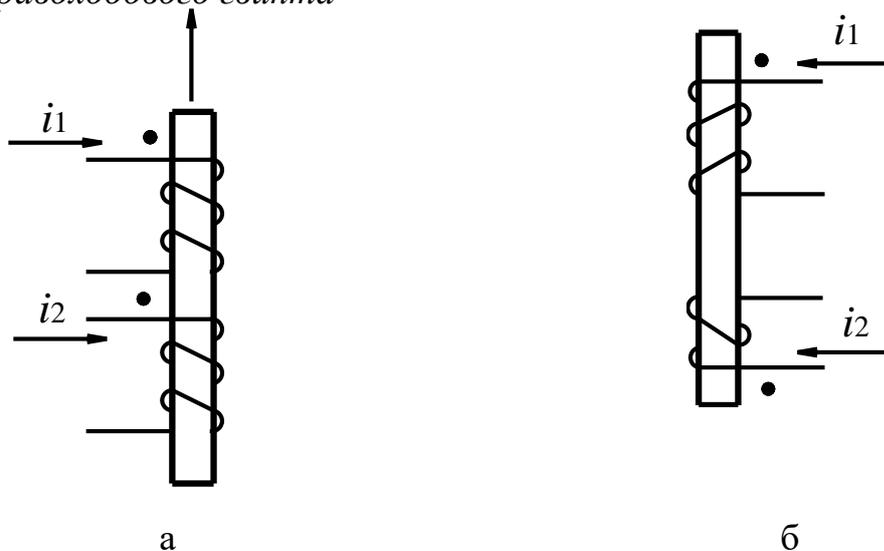


Рис. 2.33. Ілюстрація правила штопора та варіантів узгодженого включення обмоток

В залежності від напрямку намотки котушки вибрані такі напрямки струмів, при яких їх магнітні потоки співпадають. Отже рис. 2.33 ілюструє узгоджений напрямок струмів. При узгодженому напрямку струмів в двох індуктивно зв'язаних котушках ті виводи котушок, відносно яких струми спрямовані однаково, зветься *однойменними* або *однополярними*. Такі виводи зазвичай на схемах помічають зірочками або зачерненими кругами, як це показано і на рис. 2.33.

Таким чином, однополярні виводи індуктивно зв'язаних котушок характерні тим, що при однаковому напрямку струмів відносно цих виводів, магнітні потоки (утворені цими струмами) співпадатимуть, а отже додаватимуться. Завдяки таким поміткам на практиці немає необхідності вказувати напрямок намотки котушок, а достатньо лише помітити їх однополярні виводи.

Схематичне зображення індуктивно зв'язаних котушок з позначеннями однополярних виводів та вибраних напрямків струмів наведено на рис. 2.34, причому рис. 2.34а відповідає узгодженому напрямку, а рис. 2.34б - зустрічному.



Рис. 2.34. Зображення узгодженої (а) та зустрічної (б) намотки індуктивно зв'язаних котушок

2.8.2. Основні співвідношення та еквівалентна схема заміщення однофазного двохобмоткового трансформатора

Електрична схема однофазного трансформатора з двома обмотками наведена на рис. 2.35.

Для визначеності до виводів другої обмотки трансформатора, наведеного на рис. 2.35, підключене навантаження

$$Z_H = R_H + jX_H.$$

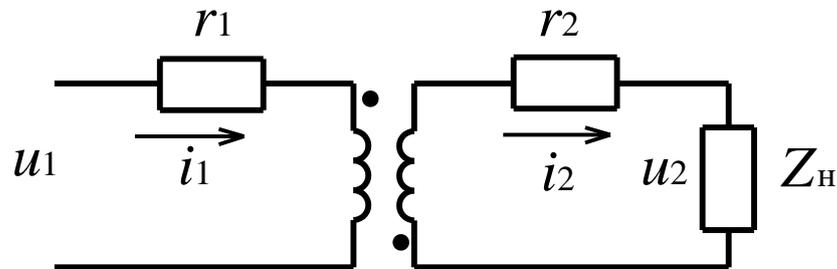


Рис. 2.35. Електрична схема однофазного трансформатора з двома обмотками

Для наведеної схеми на підставі другого закону Кірхгофа маємо

$$\begin{aligned} u_1 + e_{L1} + e_L &= r_1 i_1 \\ e_{L2} + e_L &= r_2 i_2 + Z_H i_2 \end{aligned} \quad (2.179)$$

Розкриваючи у співвідношенні (2.179) зміст електрорушійних сил самоіндукції та взаємної індукції, а також враховуючи, що

$$u_2 = Z_H i_2,$$

отримуємо

$$\begin{aligned} u_1 &= r_1 i_1 + L_1 \frac{di_1}{dt} + L_{12} \frac{di_2}{dt} \\ 0 &= r_2 i_2 + L_2 \frac{di_2}{dt} + L_{21} \frac{di_1}{dt} + u_2. \end{aligned} \quad (2.180)$$

За умови, що напруга u_1 змінюється за гармонічним законом, та врахувавши співвідношення

$$L_{12} = L_{21} = L_{3B},$$

рівняння (2.180) можна записати в комплексній формі, а саме

$$\begin{aligned} \dot{U}_1 &= r_1 \dot{I}_1 + j\omega L_1 \dot{I}_1 + j\omega L_{3B} \dot{I}_2 \\ 0 &= r_2 \dot{I}_2 + j\omega L_2 \dot{I}_2 + j\omega L_{3B} \dot{I}_1 + \dot{U}_2. \end{aligned} \quad (2.181)$$

Введемо позначення

$$\omega L_1 = x_1, \quad r_2 + R_H = r, \quad \omega L_2 + X_H = x. \quad (2.182)$$

З урахуванням (2.182) співвідношення (2.181) можна подати як

$$\begin{aligned} \dot{U}_1 &= (r_1 + jx_1) \dot{I}_1 + j\omega L_{3B} \dot{I}_2 \\ 0 &= (r + jx) \dot{I}_2 + j\omega L_{3B} \dot{I}_1. \end{aligned} \quad (2.183)$$

Розв'язавши систему рівнянь (2.183) наприклад методом підстановки, отримуємо

$$\dot{I}_1 = \frac{\dot{U}_1}{\left[r_1 + \frac{(\omega L_{3B})^2}{r^2 + x^2} r \right] + j \left[x_1 - \frac{(\omega L_{3B})^2}{r^2 + x^2} x \right]} = \frac{\dot{U}_1}{Z_{BX}}, \quad (2.184)$$

де

$Z_{BX} = r_{BX} + jx_{BX}$ - комплексний опір вхідного (еквівалентного) опору всього електричного кола, утвореного трансформатором та опором навантаження.

Зі співвідношення (2.184) видно, що еквівалентний активний опір більше за опір r_1 первинної обмотки трансформатора. Збільшення еквівалентного активного опору пов'язане з тим, що незворотне перетворення енергії в контурі вторинної обмотки відбувається за рахунок енергії, яка передається до контуру вторинної обмотки від контуру первинної обмотки, де саме й підключене джерело електричної енергії. Оскільки для заданого струму активна потужність, що визначає незворотні перетворення енергії прямо пропорційна активному опору, то поглинання енергії у вторинному контурі приводить до збільшення опору всього кола.

Еквівалентний активний опір може бути

- більше x_1 , за умови ($x < 0$),
- менше x_1 , за умови ($x > 0$).

Фізичну картину процесів, що відбуваються в електричному колі з трансформатором при вказаних варіантах, можна проілюструвати за допомогою векторних діаграм.

На рис. 2.36 представлені граничні випадки, коли контур вторинної обмотки має чисто реактивний (індуктивний – рис. 2.36а або ємнісний – рис. 2.36б) характер.

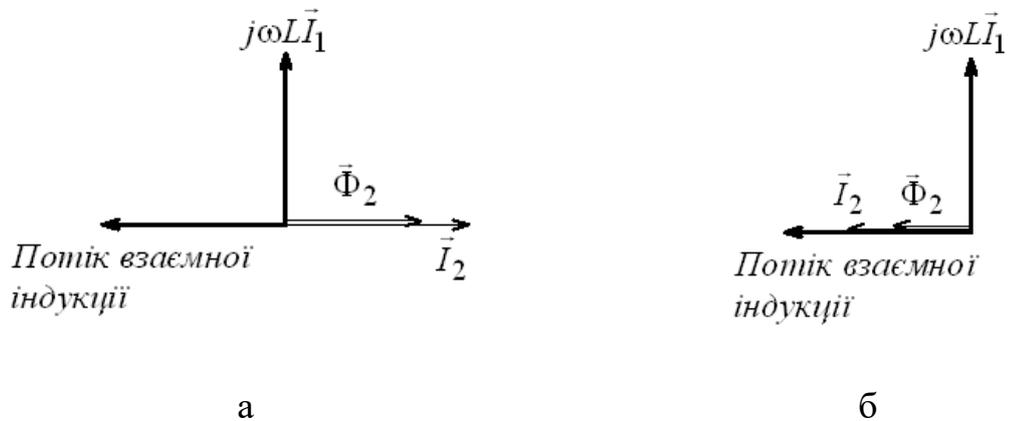


Рис. 2.36. Векторні діаграми за умови індуктивного (а) та ємнісного (б) характеру опору контуру вторинної обмотки

При індуктивному характері опору контуру вторинної обмотки (див. рис. 2.36а) магнітний потік, який утворюється струмом цього контуру, буде зменшувати потік взаємної індукції, що веде до зменшення реактивного опору контуру первинної обмотки.

При ємнісному характері опору контуру вторинної обмотки (див. рис. 2.36б) навпаки, магнітний потік, який утворюється струмом цього контуру, сприятиме збільшенню потоку взаємної індукції, а отже приводить до збільшення реактивного опору контуру вторинної обмотки трансформатора.

Повертаючись до умов електричної рівноваги в контурах первинної та вторинної обмоток, які відображаються співвідношеннями (2.181), неважко помітити, що вони рівносильні наступним рівнянням

$$\begin{aligned} \dot{U}_1 &= r_1 \dot{I}_1 + j\omega(L_1 - L_{зв}) \dot{I}_1 + j\omega L (\dot{I}_1 + \dot{I}_2) \\ 0 &= r_2 \dot{I}_2 + j\omega(L_2 - L_{зв}) \dot{I}_2 + j\omega L_{зв} (\dot{I}_1 + \dot{I}_2) + Z_H \dot{I}_2. \end{aligned} \quad (2.185)$$

Схема електричного кола, яке описуватиметься системою рівнянь (2.185), має вигляд як на рис. 2.37

В наведеній на рис. 2.37 схемі струми I_1, I_2 та напруги U_1, U_2 дорівнюють відповідним струмам та напругам притаманним реальному трансформатору, а отже наведена схема є еквівалентною електричною схемою трансформатора або ЕСЗ.

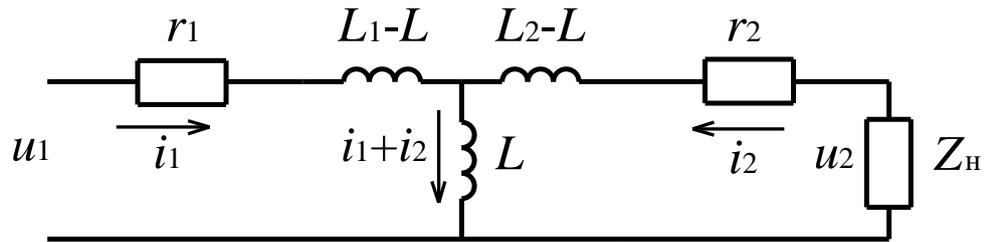


Рис. 2.37. Еквівалентна електрична схема трансформатора

Як відомо, ступінь магнітного зв'язку контурів прийнято характеризувати величиною

$$c = \frac{L_{3В}}{\sqrt{L_1 L_2}}.$$

Величина c зазвичай зветься коефіцієнтом зв'язку котушок, індуктивності яких становлять L_1, L_2 при величині їх взаємної індуктивності $L_{3В}$.

Оскільки у природі не існує магнітних ізоляторів, то завжди має місце розсіювання магнітного потоку, тобто в реальних випадках завжди $c < 1$.

Таким чином, за умови, що

$$L_1 = L_2,$$

мають виконуватися співвідношення

$$L_1 - L_{3В} > 0, L_2 - L_{3В} > 0.$$

Видно, що за умови неоднакових значень L_1, L_2 одна з наведених різностей може виявитися від'ємною.

Отже, різності

$$(L_1 - L_{3В}), (L_2 - L_{3В})$$

у співвідношенні (2.185) матимуть фізичний зміст лише за умови однакової кількості витків первинної та вторинної обмоток трансформатора, представляючи індуктивності розсіювання L_{1s}, L_{2s} відповідних обмоток трансформатора.

При неоднаковій кількості витків обмоток на практиці зазвичай використовують так звану *приведену схему заміщення* трансформатора (див. рис. 2.38).

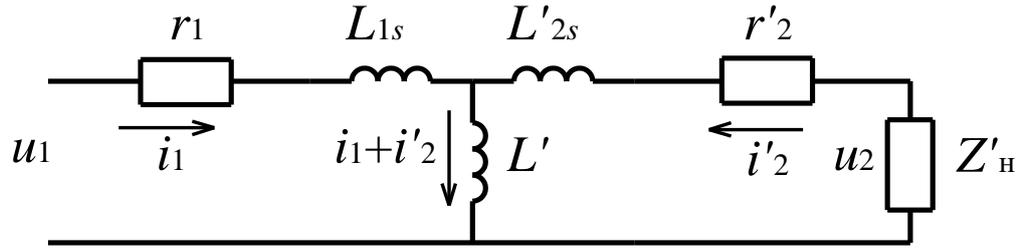


Рис. 2.38. Приведена схема заміщення трансформатора

Процедура переходу до приведеної схеми заміщення трансформатора полягає в тому, що напруга U_2 та струм I_2 у контурі вторинної обмотки трансформатора замінюються фіктивними їх значеннями, які визначаються так щоб збереглися енергетичні співвідношення в ЕСЗ та реальному трансформаторі.

Дотримуючись вказаного правила, матимемо

$$i'_2 = \frac{i_2}{k}, \quad U'_2 = kU_2, \quad (2.186)$$

де

$k = \frac{w_1}{w_2}$ - константа, яка зветься коефіцієнтом трансформації.

Таким чином, співвідношення (2.185) для приведеної ЕСЗ матимуть вигляд

$$\begin{aligned} \dot{U}_1 &= r_1 \dot{I}_1 + j\omega L_{1s} \dot{I}_1 + j\omega L'_{3B} (\dot{I}_1 + \dot{I}'_2) \\ 0 &= r'_2 \dot{I}'_2 + j\omega L'_{2s} \dot{I}'_2 + j\omega L'_{3B} (\dot{I}_1 + \dot{I}'_2) + \dot{U}'_2, \end{aligned} \quad (2.187)$$

де

$$\begin{aligned} L_{1s} &= L_1 - L'_{3B} = L_1 - kL_{3B} \\ L'_{2s} &= L_2 - L'_{3B} = k^2 \left(L_2 - \frac{L_{3B}}{k} \right) = k_2 L_{2s} \\ Z'_H &= k_2 Z_H. \end{aligned} \quad (2.188)$$

Індуктивні опори ωL_{1s} , ωL_{2s} являють собою опори розсіювання відповідних обмоток, а індуктивний опір $\omega L_{3B} k$ - опір вітки намагнічування.

Відповідно струм, що тече в цій вітці й дорівнює $\left(I_1 + \frac{I_2}{k} \right)$, зветься струмом намагнічування (намагнічуючим струмом) трансформатора.

2.8.3. Досконалий та ідеальний трансформатори

Розглянемо деякі властивості трансформатора в граничних (ідеалізованих) випадках.

Припустимо, що активні втрати в обмотках трансформатора відсутні, а коефіцієнт взаємної індукції дорівнює одиниці, тобто

$$r_1 = r_2, \quad c = 1.$$

За таких умов рівняння трансформатора (2.181) можна подати у вигляді

$$\begin{aligned} \dot{U}_1 &= j\omega L_1 \dot{I}_1 + j\omega L_{3B} \dot{I}_2 \\ 0 &= j\omega L_2 \dot{I}_2 + j\omega L_{3B} \dot{I}_1 + \dot{U}_2. \end{aligned} \quad (2.189)$$

Перетворюючи співвідношення (2.189) так, щоб \dot{U}_1 , \dot{I}_1 були виражені через \dot{U}_2 , \dot{I}_2 , отримуємо

$$\begin{aligned} \dot{U}_1 &= -\frac{L_1}{L} \dot{U}_2 + j\omega \left(L_{3B} - \frac{L_1 L_2}{L_{3B}} \right) \dot{I}_2 \\ \dot{I}_1 &= -\frac{\dot{U}_2}{j\omega L_{3B}} - \frac{L_2}{L_{3B}} \dot{I}_2. \end{aligned} \quad (2.190)$$

З (2.190) видно, що за умови $c = 1$, матимемо

$$L_{3B} - \frac{L_1 L_2}{L_{3B}} = 0.$$

При цьому, ввівши позначення

$$\frac{L_1}{L_{3B}} = a,$$

з (2.190) отримуємо

$$\dot{U}_1 = -a \dot{U}_2, \quad \dot{I}_1 = \frac{\dot{U}_1}{j\omega L_{3B}} - \frac{\dot{I}_2}{a}. \quad (2.191)$$

Трансформатор, для якого при будь-якому навантаженні виконується умова

$$\frac{\dot{U}_1}{\dot{U}_2} = a,$$

зветься *досконалим* трансформатором.

Якщо додатково до введених умов прийняти умову

$$L_1 = \infty,$$

то у співвідношенні (2.191) струмом $\frac{\dot{U}_1}{\omega L_1}$ можна нехтувати у порівнянні зі

струмом $\frac{\dot{I}_2}{a}$, а отже матимемо

$$\dot{U}_1 = -a\dot{U}_2, \quad \dot{I}_1 = -\frac{\dot{I}_2}{a}. \quad (2.192)$$

Трансформатор, для якого виконуються співвідношення (2.192), зветься *ідеальним*. Такий трансформатор дійсно має властивість перетворювати струми та напруги (незалежно від величини опору, включеного у контур вторинної обмотки) у визначене число разів (це число, як видно з (2.192) дорівнює a).

Для ідеального трансформатора на підставі співвідношень (2.192) отримуємо

$$\frac{\dot{U}_1}{\dot{I}_1} = Z_{\text{вх}} = \frac{-a\dot{U}_2}{-\frac{\dot{I}_2}{a}} = a^2 Z_{\text{н}}. \quad (2.193)$$

З (2.193) видно, що за допомогою ідеального трансформатора можна також перетворювати у визначене число разів комплексні (в загальному випадку) опори. Ця обставина особливо важлива для раціонального конструювання окремих елементів електричних кіл, наприклад елементів узгодження окремих ділянок електричного кола за їх опором.

Досконалий трансформатор можна подати як ідеальний трансформатор, до виводів якого приєднана індуктивність, як показано на рис. 2.39.

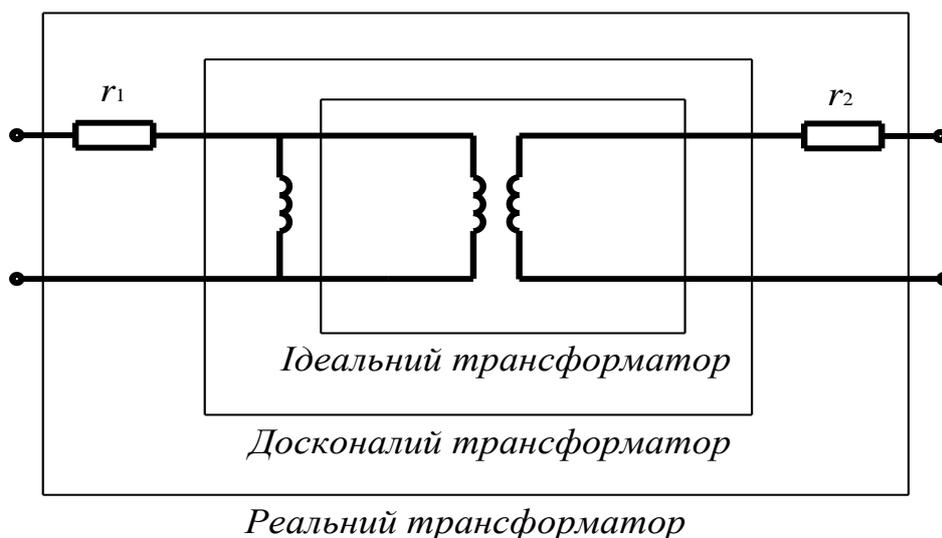


Рис. 2.39. До визначення ідеального, досконалого та реального трансформаторів

Реальний трансформатор можна подати як ідеальний трансформатор, до виводів якого відповідним чином приєднані додаткові індуктивності та активні опори r_1, r_2 обмоток, за умови що виконується умова $s < 1$ (див. рис. 2.39).

Властивості, близькі до ідеального та досконалого трансформатора, притаманні трансформатору з феромагнітним осердям при достатньо великій кількості витків та великій магнітній проникності феромагнітного матеріалу осердя.

2.9. КОНТРОЛЬНІ ПИТАННЯ І ПРИКЛАДИ РОЗВ'ЯЗАННЯ ЗАДАЧ

2.9.1. Контрольні питання

1. Що являє собою обмоткова функція обмотки електричної машини?
2. Розкрийте поняття «синусна обмотка».
3. Що таке електричні кути та як вони пов'язані з геометричними кутами?
4. Яка обмотка зветься основною і чому?
5. Поясніть фізичний зміст понять «власна індуктивність» та «взаємна індуктивність» обмотки.
6. Чим визначається взаємна індуктивність двох обмоток?
7. Чому магнітне поле, утворене однофазною обмоткою, зветься стаціонарним?
8. Поясніть поняття «пульсуюче магнітне поле».

-
9. За яких умов, магнітне поле, утворене двома ортогональними обмотками, буде стаціонарним?
 10. Сформулюйте необхідні умови утворення магнітного поля, яке обертається.
 11. Яке обертове магнітне поле зветься круговим?
 12. Як спрямована компонента магнітного поля, що є найсуттєвіша для утворення обертового моменту?
 13. Від яких величин залежить магнітний момент обертового ЕМП енергії?
 14. Яка мінімальна кількість пар полюсів необхідна для утворення кругового обертового магнітного поля?
 15. Сформулюйте та проаналізуйте фізичні основи реалізації будь-якої електричної машини.
 16. Як пов'язаний момент електричної машини з взаємною індуктивністю обмоток її ротора і статора?
 17. Чи залежить (якщо залежить, то як) магнітний момент від числа пар полюсів електромеханічної системи?
 18. Чи залежить (якщо залежить, то як) магнітний момент від сили струму, що протікає обмотками ротора та статора електромеханічної системи?
 19. Чи залежить (якщо залежить, то як) магнітний момент електромеханічної системи від кута між осями обмоток ротора та статора?
 20. Якими параметрами визначається власна індуктивність синусної обмотки?
 21. Яке співвідношення частот обертання ротора та струмів в обмотках ротора та статора притаманні синхронній електричній машині?
 22. Яке співвідношення частот обертання ротора та струмів в обмотках ротора та статора притаманні асинхронній електричній машині?
 23. Яке співвідношення частот обертання ротора та струмів в обмотках ротора та статора притаманні електричній машині постійного струму?
 24. Що таке еквівалентна схема заміщення електричної машини і яке її призначення?
 25. За яких умов еквівалентна схема заміщення вважається лінійною?

2.9.2. Приклади розв'язання задач

Приклад 1

Схематичне зображення конструкції електромеханічного перетворювача – генератора з двома обмотками наведено на рис. 2.13.

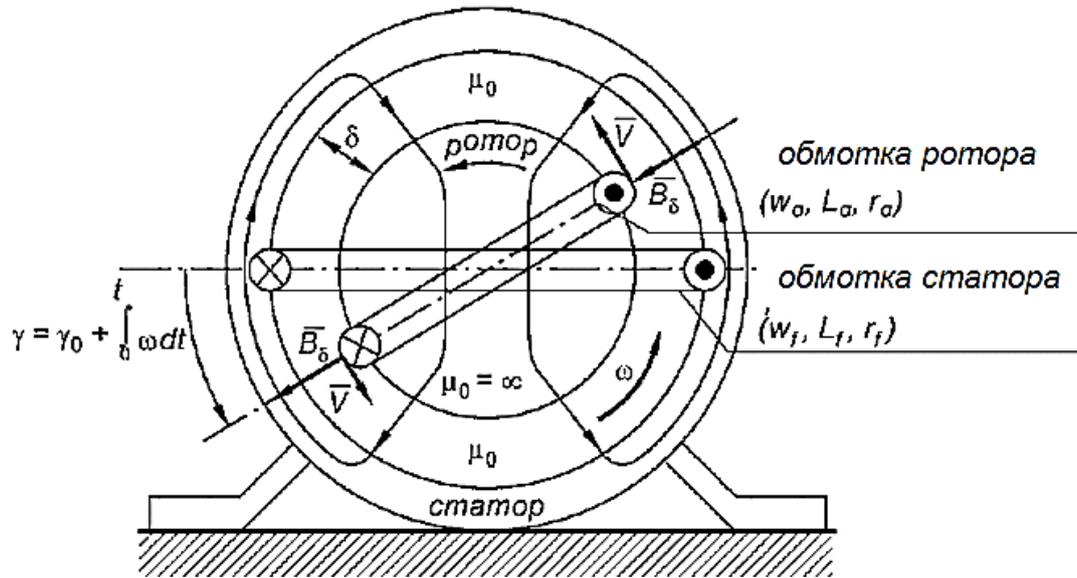


Рис. 2.13. Двохобмотковий перетворювач механічної енергії на електричну

Відомо

- діаметр ротора $D_a = 1$ м;
- довжина ротора $l_\delta = 1$ м;
- магнітна індукція в повітряному зазорі $B_\delta = 0,8$ Тл;
- число витків обмотки ротора $w_a = 100$;
- частота обертання ротора $n = 3000$ об/хв.;
- активна складова опору обмотки ротора $R_a = 0,6$ Ом;
- опір навантаження $R_{\text{нав}} = 39,4$ Ом.

Визначити

- ЕРС обмотки ротора;
- струм навантаження;
- напругу на затискачах генератора U_a ;
- втрати в обмотці ротора;
- електромагнітну потужність;
- електромагнітний момент;
- потужність на виході генератора P_2 .

Розв'язок задачі

Визначаємо лінійну швидкість на поверхні ротора

$$v = \frac{\pi D_a n}{60} = \frac{\pi \cdot 1 \cdot 3000}{60} = 157 \text{ м/с.}$$

Розраховуємо ЕРС, що наводиться в обмотці якоря

$$E = 2B_\delta l_\delta w_a v = 2 \cdot 1 \cdot 1 \cdot 100 \cdot 157 = 31400 \text{ В.}$$

Струм навантаження

$$I_a = \frac{E}{R_a + R_{\text{Нав}}} = \frac{31400}{39,4 + 0,6} = 785 \text{ А.}$$

Втрати в обмотці ротора

$$\Delta P_a = I_a^2 \cdot R_a = 785^2 \cdot 0,6 = 369,7 \text{ кВт.}$$

Напряга на затискачах генератора

$$U_a = E - I_a R_a = 31400 - 785 \cdot 0,6 = 30929 \text{ В.}$$

Електромагнітна потужність – потужність, що відляється на роторі

$$P_{\text{ем}} = EI_a = 31400 \cdot 785 = 24649 \text{ кВт.}$$

Електромагнітний момент

$$M_{\text{ем}} = \frac{P_{\text{ем}}}{\omega} = \frac{24649 \cdot 10^3 \cdot 60}{2\pi \cdot 3000} = 78500 \text{ Нм}$$

Електромагнітний момент можна також визначити як

$$\begin{aligned} M_{\text{ем}} &= F_{\text{ем}} \frac{D_a}{2} = 2B_\delta l_\delta I_a w_a \frac{D_a}{2} = \\ &= 2 \cdot 1 \cdot 1 \cdot 785 \cdot 100 \cdot 0,5 = 78500 \text{ Н} \cdot \text{м} \end{aligned}$$

Потужність, що віддається генератором

$$P_2 = P_{\text{ем}} - \Delta P_a = 24649 - 369,7 = 24279,3 \text{ кВт.}$$

ГЛАВА 4

МЕТОДИ ТЕОРІЇ ОПТИМІЗАЦІЇ В ЗАДАЧАХ МАТЕМАТИЧНОГО МОДЕЛЮВАННЯ МАТЕРІАЛЬНИХ ОБ'ЄКТІВ

Людина, яка оволоділа базовими принципами математики, має на один орган почуттів більше ніж інші люди

Чарльз Дарвін

Узагальнено теорію оптимізації можна визначити як сукупністю математичних методів (фундаментального та прикладного характеру), які орієнтовані на відшукування та ідентифікацію найкращих варіантів з чисельних альтернатив. Всі ці методи спрямовані на розроблення та реалізацію процедур, які дозволяють уникнути повного перебору та оцінювання всіх (або, як мінімум, надто значної кількості) можливих варіантів.

Процес оптимізації на сьогодні є основою практичної діяльності багатьох галузей техніки, технології, економіки, фінансової діяльності тощо.

На підставі фундаментальних досліджень в області теорії оптимізації, на сьогодні розроблено величезну кількість числових методів (а також відповідних алгоритмів та процедур), які призначені для широкого використання під час розв'язання найрізноманітніших задач, в тому числі й пов'язаних з пошуком екстремальних значень функціональних залежностей, що утворює методичну основу для використання вказаних методів в процесі дослідження математичних моделей МО [28 – 32].

Зазначимо, що від правильного вибору числового методу на тому чи іншому етапі використання математичної моделі МО, як правило, суттєво (напрямую) залежить успішність та ефективність розв'язання поставлених задач на будь-якому етапі його життєвого циклу.

4.1. ЗАГАЛЬНІ ПОНЯТТЯ ТЕОРІЇ ОПТИМІЗАЦІЇ

Постановка кожної задачі оптимізації включає два математичні об'єкти:

- множину розв'язків, які допустимі в контексті даної задачі;
- цільову функцію (або функціонал) котру слід на вказаній множині допустимих розв'язків.

При цьому вказана цільова функція і є математичною моделлю МО.

Саме з такої узагальненої точки зору розглядаються в теорії оптимізації різні класи екстремальних задач.

Для того, щоб використати фундаментальні математичні результати та числові методи теорії оптимізації для розв'язання конкретних задач, необхідно:

- встановити границі об'єкту; за умови оптимізації МО це означає необхідність встановлення границь можливого змінення його первинних (в тому числі вхідних) та вихідних параметрів, або інших показників (критеріїв) якості; за умови мінімізації функції це чає необхідність визначення області її

можливих значень та можливих значень її аргументів;

- побудувати математичну модель досліджуваного об'єкту; тобто встановити правила, функціональні залежності, алгоритми що зв'язують первинні параметри об'єкту з його вихідними параметрами;

- скласти цільову функцію (або функціонал, за умови, що аргументами є функції);

- визначити критерій оптимальності – зазвичай актуалізувати вимогу щодо пошуку екстремуму цільової функції при наявності відповідних обмежень щодо змінення параметрів;

- вибрати алгоритм для розв'язання екстремальної задачі.

В загальному вигляді задачу оптимізації можна визначати як задачу пошуку екстремального значення функції за наявності деяких обмежень на можливі змінення аргументу. Математично таку задачу в n -вимірному просторі дійсних чисел (\mathcal{R}^n) можна подати як

$$\begin{aligned} \min f(\mathbf{x}), g_k(\mathbf{x}) = 0, \quad k = \overline{1, K}; \quad l_j(\mathbf{x}) \geq 0, \\ j = \overline{1, J}, \quad \mathbf{x} \in D \subset \mathcal{R}^n \end{aligned} \quad (4.1)$$

Вектор $\mathbf{x} \in D$ має компоненти x_1, x_2, \dots, x_n , котрі є невідомими задачі, сформульованої як (4.1).

Функція $f(\mathbf{x})$ зветься цільовою функцією (критерієм оптимальності, функцією якості і т.п.), а множина умов

$$g_k(\mathbf{x}), l_j(\mathbf{x}), \mathbf{x} \in D$$

визначає обмеження, які мають враховуватися під час розв'язання задачі.

Розв'язком задачі (4.1) є вектор \mathbf{x} , який задовольняє заданим обмеженням. Оптимальним розв'язком задачі (4.1) є вектор \mathbf{x}^* , який мінімізує функцію $f(\mathbf{x})$ на множині всіх можливих розв'язків, тобто

$$\forall \mathbf{x} \in D \quad f(\mathbf{x}^*) \leq f(\mathbf{x}).$$

Зазначимо, що задача максимізації функції зводиться до задачі пошуку мінімуму (мінімізації) функції реалізацією операції

$$F = -f(\mathbf{x}).$$

Характеристикою точності отриманого розв'язку задачі може слугувати будь-яка міра відстані між відповідними математичними об'єктами (див. главу 3)

Наприклад, як міру точності можна використати величину абсолютного відхилення значення функції, що мінімізується, яке вона досягла в точці $\mathbf{x}^{(n)} \in D$ від точного значення її мінімуму на множині D , тобто

$$\delta(\mathbf{x}) = \left| f(\mathbf{x}^{(n)}) - \min f(\mathbf{x}) \right|, \quad \mathbf{x} \in D, \quad \delta(\mathbf{x}) \geq 0 \quad (4.2)$$

З (4.2) видно, що чим менша невід’ємна величина $\delta(\mathbf{x})$, яка відображає абсолютне значення похибки мінімізації, тим точніше отриманий розв’язок поставленої задачі.

Недоліком використання абсолютного значення похибки є те, що вона змінюється за умови множення функції на деяку додатну константу α , тобто за умови реалізації переходу $f(\mathbf{x}) \rightarrow \alpha f(\mathbf{x})$.

В зв’язку з наведеним на практиці доцільніше використовувати оцінку точності у вигляді відношення

$$c = \frac{\min f(\mathbf{x}) - f(\mathbf{x}^*)}{f(\mathbf{x}^*)} \cdot 100 \% . \quad (4.3)$$

де

$f(\mathbf{x}^*)$ - значення мінімуму функції точне або отримане за «точним» (точнішим за той, що використовувався) алгоритмом.

4.2. ХАРАКТЕРИСТИКИ СХОДЖЕННЯ АЛГОРИТМІВ ПОШУКУ ЕКСТРЕМУМУ

Визначення характеристик алгоритмів пошуку екстремуму необхідне для можливості їх порівняння з метою вибору найефективнішого в рамках розв’язання конкретної задачі. Зазвичай такі алгоритми характеризуються необхідним об’ємом обчислювальних потужностей, швидкістю сходження, стабільністю до помилок, що можуть виникати в процесі реалізації алгоритму, чутливістю до значень параметрів тощо.

Однією з найважливіших характеристик, яка до того ж має більш-менш чітке математичне формулювання, є характеристика *сходження*. Більшість методів розв’язання екстремальних задач мають ітераційну природу, тобто алгоритм A реалізує процес

$$x_i = A(x_{i-1}) \quad (4.4)$$

де

x_i, x_{i-1} - значення деякого аргументу функції на поточній та попередній ітерації відповідно.

Алгоритми такого типу, реалізують процес пошуку екстремуму функції з деякої початкової точки x_0 , породжуючи при цьому послідовність

$\{x_0, x_1, \dots, x_i, \dots\}$, яка сходиться до точки екстремуму x^* . При цьому характер сходження, звичайно, може бути різним.

Глобальне сходження

Алгоритму A притаманна властивість глобального сходження, за умови, що для *будь-якої* початкової точки x_0 , послідовність $\{x_i\}$, яка визначається співвідношенням (4.4), сходиться до точки, що задовольняє необхідним умовам екстремуму.

Дана властивість відображає надійність роботи алгоритму.

Асимптотичне сходження та швидкість сходження

З практичної точки зору ефективність алгоритму пошуку екстремуму залежить від кількості ітерацій, які необхідні для отримання наближеного розв'язку x^* відповідної екстремальної задачі з заданою точністю.

Критерій ефективності алгоритму зазвичай визначають на підставі аналізу його *асимптотичного сходження*, тобто на підставі аналізу поведінки генерованої алгоритмом послідовності точок $\{x_i\}$ в околі деякої граничної точки x^* .

Прийнявши евклідову міру відстані між точками генерованої алгоритмом послідовності, розрізняють *лінійне*, *суперлінійне* та *сходження n -го порядку*.

Лінійне сходження характеризується виконанням нерівності

$$\limsup_{i \rightarrow \infty} \frac{\|x_{i+1} - x^*\|}{\|x_i - x^*\|} = \alpha < 1, \quad (4.5)$$

де

α - коефіцієнт лінійного сходження алгоритму.

Співвідношення (4.5) засвідчує той факт, що алгоритм пошуку екстремуму генерує послідовність значень x_i таких, що при достатньо великих значеннях i , межа верхньої границі відношення відстаней двох послідовних значень аргументу x_i, x_{i+1} до точки екстремуму x^* буде дорівнювати деякому числу, що менше за одиницю.

Суперлінійне сходження відповідає умові

$$\limsup_{i \rightarrow \infty} \frac{\|x_{i+1} - x^*\|}{\|x_i - x^*\|} \rightarrow 0.$$

Сходження n -го порядку реалізується алгоритмом, для якого виконується умова

$$\limsup_{i \rightarrow \infty} \frac{\|x_{i+1} - x^*\|}{\|x_i - x^*\|^p} < c < 1,$$

де

c - константа,

$p > 1$ - число, яким визначається порядок (а отже й швидкість) сходження: $p = 2$ - квадратичне сходження; $p = 3$ - кубічне сходження; $p = n$ - сходження n -го порядку.

Для оцінювання ефективності вибраних методів розв'язку екстремальної задачі зазвичай рекомендуються три характеристики:

- час, який витрачається на отримання результату,
- точність отриманого результату,
- чутливість до змінення параметрів сходження.

Визначення вказаних характеристик нерідко само по собі становить непросту задачу, для розв'язку якої використовуються і відповідні тестові функції, і евристичні правила, і власний досвід дослідника.

Зазначимо, що вибір ефективного алгоритму розв'язання екстремальної задачі, як правило, тим успішніший, чим більше інформації відомо про функцію, екстремум якої має бути знайдено.

До такої інформації, перш за все, слід віднести інформацію щодо наявності шуканого екстремуму у відповідному інтервалі. Для цього використовуються найрізноманітніші методи (включаючи, звичайно й евристичні), серед яких, мабуть найуживанішими є

- графічне представлення функції;
- аналітичний аналіз функції, що апроксимує досліджувану;
- відомості про об'єкт (процес), який відображається наданою функцією;
- відомості про базові засади розроблення математичної моделі досліджуваного об'єкту.

4.3. ДЕТЕРМІНОВАНІ МЕТОДИ ПОШУКУ ЕКСТРЕМУМУ ОДНОВИМІРНИХ ФУНКЦІЙ

Реалізація алгоритмів пошуку екстремуму багатовимірних функцій, як правило, супроводжується пошуком (нерідко багаторазовим) нулів або екстремумів деякої одновимірної функції. Пошук нуля одновимірної функції легко звести до розв'язання задачі пошуку екстремуму цієї функції.

На сьогодні розроблено чимало методів розв'язання вказаної задачі, вибір яких серед іншого визначається видом функції в області, де її значення наближаються до нуля.

По більшості методи одновимірної оптимізації можна розділити на дві групи. До першої групи віднесемо методи, засновані на звужуванні інтервалу невизначеності, а до другої – методи з використанням похідних.

4.3.1. Методи звужування інтервалу невизначеності

Розглянемо задачу пошуку мінімуму функції $f(x)$ на деякому інтервалі $[a, b]$. При цьому спиратимемося виключно на значення функції, залишаючи поза увагою інформацію щодо її похідних. Група таких методів отримала назву *методів звужування інтервалу невизначеності або методів нульового порядку*.

Задача наближеного відшукування мінімуму в методах звуження інтервалу невизначеності полягає в тому, щоб знайти таку множину значень аргументу x_1, x_2, \dots, x_k , що значення x^* аргументу, при якому досягається мінімальне значення $f^*(x^*)$ функції, потрапляє до інтервалу $x_{i-1} \leq x^* \leq x_i$, який зветься інтервалом невизначеності D .

Вочевидь, що на початку реалізації відповідного алгоритму інтервал невизначеності співпадає з відрізком $[a, b]$.

Існує декілька способів послідовного звуження інтервалу невизначеності, вибір серед яких базується на інформації щодо поведінки функції на інтервалі невизначеності.

Метод загального пошуку

Знаходитимемо мінімум функції $f(x)$ на деякому інтервалі $[a, b]$ при відсутності будь якої додаткової інформації щодо характеру змінення вказаної функції на даному інтервалі. За таких умов екстремум функції доречно шукати, використовуючи метод загального пошуку (метод перебору).

Реалізація методу загального пошуку полягає в розділенні інтервалу $[a, b]$ на декілька однакових частин (відрізків) з наступним розрахунком значень функції у вузлах отриманої сітки (див. рис. 4.1). Вважається, що абсциса, в якій за результатами розрахунків отримано мінімальне значення функції і є мінімумом цієї функції.

За результатом реалізації методу загального пошуку, як видно з рис. 4.1, інтервал невизначеності звужується до двох кроків сітки.

Зазвичай говорять про *дроблення інтервалу невизначеності*, котре характеризується коефіцієнтом дроблення α .

Розділивши інтервал невизначеності $[a, b]$ на n однакових частин, отримаємо $(n + 1)$ вузол. При цьому

$$\alpha = \frac{2}{n},$$

а кількість розрахунків значення функції становитиме

$$N = n + 1.$$

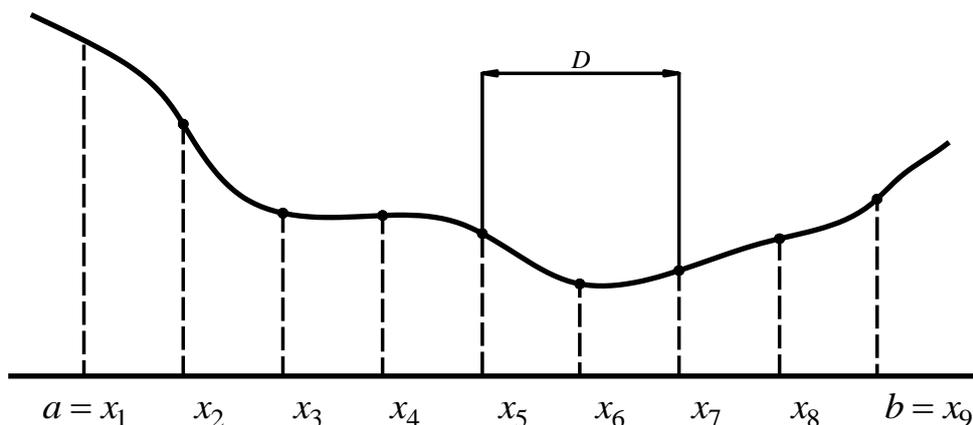


Рис. 4.1. Схема реалізації методу загального пошуку

Таким чином коефіцієнт дроблення інтервалу невизначеності дорівнюватиме

$$\alpha = \frac{2}{N-1}, \quad N = 3, 4, 5, \dots \quad (4.6)$$

З (4.6) видно, що для отримання, наприклад, $\alpha = 0,01$, необхідно буде розрахувати значення функції у 201 точці, а при $\alpha = 0,001$ - у 2001 точці.

Таким чином приходимо до висновку, що ефективність методу загального пошуку зі зменшенням інтервалу невизначеності швидко падає.

Метод ділення інтервалу пополам (метод половинного ділення)

За умови, що функція $f(x)$ на деякому інтервалі $[a, b]$ - унімодальна (тобто існує відповідна додаткова інформація щодо поведінки функції на інтервалі дослідження) можна використати метод, який передбачає ділення вказаного інтервалу на дві однакові частини, надалі кожен отриманий за такий спосіб інтервал знову на дві однакові частини (див. рис. 4.2). Це перший етап пошуку мінімуму. На цьому етапі після п'яти розрахунків значення функції (два на краях інтервалу та три - всередині) інтервал невизначеності звужується вдвічі, тобто на першому етапі $\alpha = 0,5$.

Надалі ділимо інтервал $[x_4, x_5]$ пополам, а в подальшому кожен отриманий половину цього інтервалу знову пополам. Зважаючи на те, що на цьому етапі значення функції на краях інтервалу та в його середині вже відомі (за даними першого етапу), розраховувати слід лише два значення функції.

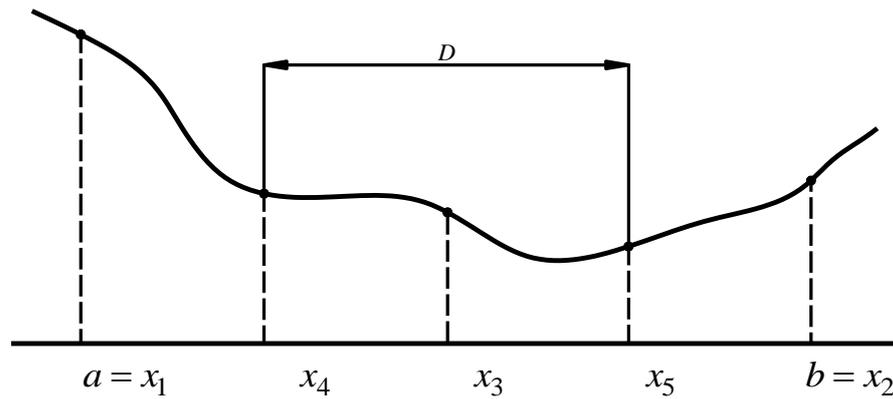


Рис. 4.2. Схема реалізації методу ділення інтервалу пополам

За результатами реалізації другого етапу інтервал невизначеності знову ж зменшується вдвічі.

Така схема реалізації алгоритму зберігається до кінця процесу.

За результатами N обчислень функції коефіцієнт дроблення початкового інтервалу складає

$$\alpha = (0,5)^{\frac{N-3}{2}}, \quad N = 5, 7, 9, \dots \quad (4.7)$$

Звернемо увагу на те, що у співвідношенні (4.7) $N = 5, 7, 9, \dots$. Це обумовлено тим, що інтервал невизначеності, починаючи з другого етапу, зменшується лише після обчислення двох значень функції.

Метод золотого перерізу

В цьому методі інтервал невизначеності на кожному кроці ділиться на *нерівні* частини, що зазвичай підвищує ефективність алгоритму у порівнянні з методом половинного ділення.

Розрахуємо значення функції $f(x)$ на кінцях відрізка невизначеності $[a, b]$ та позначимо $a = x_1$, $b = x_2$.

Розрахуємо також значення функції $f(x)$ у двох точках x_3 , x_4 , розміщених всередині відрізка $[a, b]$.

Порівняємо чотири розраховані значення функції та виберемо серед них найменше (див. рис. 4.3).

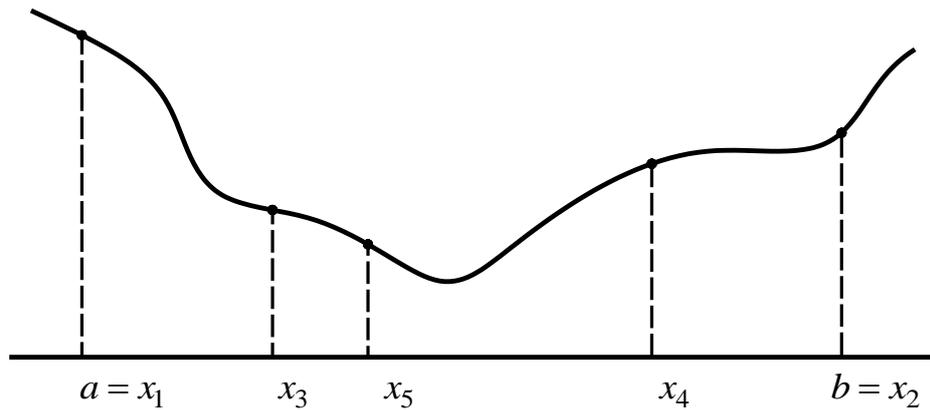


Рис. 4.3. Схема реалізації методу золотого перерізу

Припустимо, наприклад, що найменше значення функції - $f(x_3)$. Вочевидь, мінімум знаходиться в одному з відрізків, прилеглих до цього значення функції.

Отже відрізок $[x_4, b]$ можна виключити з подальшого розгляду, зосередивши пошук виключно на відрізку $[a, x_4]$.

На цьому перший етап реалізації методу золотого перерізу вважається закінченим.

Другий етап реалізується аналогічно першому.

На відрізку $[a, x_4]$ слід вибрати дві внутрішні точки, розрахувати в цих точках та на кінцях інтервалу значення функції, і зробити наступний крок. Під час реалізації цього етапу враховуємо, що на попередньому кроці вже були розраховані значення функції на кінцях відрізка $[a, x_4]$ та в одній його внутрішній точці x_3 .

Отже на цьому етапі достатньо вибрати на відрізку $[a, x_4]$ лише одну внутрішню точку x_5 , розрахувати значення функції в цій точці та провести необхідні порівняння.

Наведена процедура дозволяє вчетверо зменшити об'єм розрахунків на кожному (починаючи з другого) кроці процесу, який полягає в тому, щоб *розділити відібраний відрізок на три частини*, та виключити з подальшого розгляду (відкинути) один з крайніх відрізків за результатами відповідного аналізу.

Розглянемо детальніше процес ділення відрізка на три частини та виберемо формальний метод реалізації такої операції.

Схема процесу зменшення інтервалу невизначеності D під час реалізації наведеного алгоритму показана на рис. 4.4

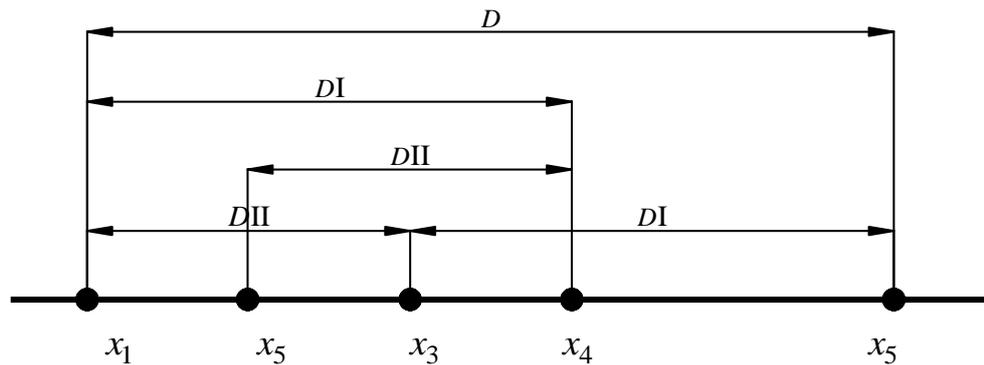


Рис. 4.4. Схема процесу зменшення інтервалу невизначеності під час реалізації методу золотого перерізу

Зважаючи на те, що в загальному випадку може бути «відкинутий» будь-який з відрізків $[x_1, x_3]$, $[x_4, x_2]$, виберемо точки x_3 , x_4 так, щоб значення довжини цих відрізків були однаковими.

Тобто вважатимемо, що в процесі реалізації

$$x_3 - x_1 = x_2 - x_4.$$

Після «відкидання» одного з відрізків (а саме $[x_4, x_5]$) отримуємо новий інтервал невизначеності D_I .

Позначимо

$$\varphi = \frac{D}{D_I}.$$

Аналогічно до наведеного продовжимо процес зменшення інтервалу невизначеності. Для цього інтервал D_I розділимо так, як це було зроблено для інтервалу D , тобто відповідно до співвідношень

$$\varphi = \frac{D}{D_I} = \frac{D_I}{D_{II}},$$

де

D_{II} - інтервал невизначеності, отриманий на даному (другому) кроці реалізації алгоритму.

Такий спосіб ділення інтервалу невизначеності називають діленням згідно золотого перерізу, або діленням в крайньому та середньому відношенні: весь інтервал відноситься до більшої частини, як більша частина до меншої.

Зважаючи на те, що D_{II} за довжиною дорівнює відрізку, який був відкинутий на попередньому етапі, тобто відрізку

$$D_{II} = D - D_I,$$

отримуємо

$$\frac{D}{DI} = \frac{DI}{D - DI} \Rightarrow \frac{DI}{D} = \frac{D}{DI} - 1,$$

або, з урахуванням введеного позначення

$$\frac{1}{\varphi} = \varphi - 1.$$

Після відповідного перетворення останнього співвідношення матимемо

$$\varphi^2 - \varphi - 1 = 0.$$

Додатний корінь такого рівняння становить

$$\varphi = \frac{\sqrt{5} + 1}{2} \approx 1,6180.$$

Отримане число в математиці відоме як *золоте відношення*, а наведений метод ділення відрізка, як *золотий переріз*. Саме тому розглянутий метод пошуку мінімуму й називають методом золотого перерізу.

Відношення

$$\frac{D}{DI} = \varphi \approx 1,6180$$

вказує у скільки разів скорочується інтервал невизначеності при одному додатковому обчисленні значення функції.

Враховуючи, що перші три розрахунки значення функції не скорочують інтервал невизначеності, можемо записати співвідношення для визначення коефіцієнту дроблення α після N розрахунків значення функції

$$\alpha = \left(\frac{1}{\varphi}\right)^{N-3} \approx (0,6180)^{N-3}.$$

За умови $N \rightarrow \infty$ довжина інтервалу невизначеності спрямовується до нуля як геометрична прогресія зі знаменником $\frac{1}{\varphi}$, тобто *метод золотого перерізу завжди сходиться*.

Зазначимо, що за умови наявності на інтервалі $[a, b]$ декількох мінімумів функції $f(x)$ процес реалізації методу золотого перерізу сходиться до одного з них, але не обов'язково до найменшого.

Встановлення початкового інтервалу невизначеності

Розглянуті вище, як і інші методи пошуку мінімуму функції шляхом

зменшення інтервалу невизначеності, передбачають наявність початкового значення вказаного інтервалу. Як видно з наведених алгоритмів, вибір такого початкового інтервалу невизначеності може суттєво вплинути на ефективність та успішність всього алгоритму.

З метою оптимального вибору початкового інтервалу на сьогодні зазвичай використовуються евристичні методи, серед котрих виділимо для розгляду *метод Свенна*.

Для пояснення сутності методу Свенна вважатимемо, що розв'язується задача пошуку мінімуму унімодальної функції $f(x)$ на всій числові осі.

Виберемо деяке початкове наближення x_0 та зробимо крок довжиною h , переходячи тим самим до точки $x_1 = x_0 + h$ (див. рис. 4.5).

За умови

$$f(x_1) > f(x_0),$$

змінимо напрямок кроку і приймемо

$$x_1 = x_0 - h.$$

Припустимо, що тепер

$$f(x_1) < f(x_0).$$

Подвоюємо довжину кроку, тобто приймаємо

$$h' = 2h.$$

Визначаємо нове значення аргументу

$$x_2 = x_1 + h',$$

розраховуємо відповідні значення функції та порівнюємо їх.

Повторюємо наведену процедури доки не виконуватиметься умова

$$f(x_n) > f(x_{n-1}).$$

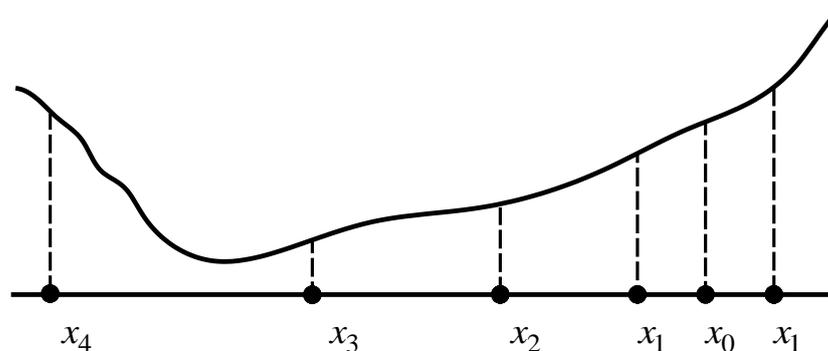


Рис. 4.5. Схема встановлення початкового інтервалу невизначеності за методом Свенна

Аналізуючи отримані результати неважко прийти до висновку, що

шуканий мінімум унімодальної функції розміщуватиметься на відріжку $[x_4, x_3]$ і може бути знайдений, наприклад, одним з розглянутих вище методів звужування інтервалу невизначеності.

Головна перевага розглянутих пошукових методів полягає в тому, що вони засновані виключно на значеннях функції, а отже не потребують виконання умов диференціювання та запису самої функції в аналітичному вигляді. Однак з такої переваги розглянутих методів випливає їх головний недолік – невелика швидкість сходження.

Для підвищення швидкості розв'язання екстремальних однопараметричних задач запропоновані методи, засновані на використанні похідних.

4.3.2. Ньютонівські методи

Вважатимемо, що унімодальна функція $f(x)$, мінімум якої розшукується, допускає подвійне диференціювання (двічі диференційована). Необхідна умова мінімуму такої функції записується як

$$f'(x^*) = 0.$$

Достатня умова мінімуму функції $f(x)$, як відомо, визначається нерівністю

$$f''(x^*) > 0.$$

Для знаходження мінімуму розв'язуватимемо рівняння

$$f'(x) = 0. \quad (4.8)$$

Задамо деяке початкове наближення x_k і розкладемо функцію $f(x)$ в цій точці в ряд Тейлора, обмежившись лише членами до другого порядку включно.

Будуємо квадратичну модель функції:

$$\tilde{f}(x) \approx f(x_k) + f'(x_k) \cdot (x - x_k) + \frac{1}{2} f''(x_k) \cdot (x - x_k)^2 \quad (4.9)$$

За умови

$$f''(x_k) \neq 0,$$

функція $\tilde{f}(x)$ матиме одну стаціонарну точку, для знаходження якої прирівнюємо нулю похідну $\tilde{f}'(x)$

$$\tilde{f}'(x) = f'(x_k) + f''(x_k) \cdot (x - x_k) = 0.$$

Розв'язавши це рівняння відносно x , прийнемо знайдений результат як чергове $(k + 1)$ -е наближення до мінімуму, а саме

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} \quad (4.10)$$

Формулу (4.10) можна отримати й інакше, якщо використати числовий метод розв'язання рівняння

$$g(x) = 0,$$

відомий як метод Ньютона (врахувавши, звичайно, що

$$g(x) = f'(x)).$$

За таких умов матимемо

$$x_{k+1} = x_k - \frac{g(x_k)}{g'(x_k)}.$$

Зазначимо, що алгоритму, поданому співвідношенням (4.10), притаманні два суттєві недоліки.

По-перше, розв'язок рівняння (4.8) може визначати не лише мінімум, але й максимум.

По-друге, функція (4.9) в окремих випадках може суттєво відрізнятися від досліджуваної функції $f(x)$ і крок $x_{k+1} - x_k$ може виявитися надто великим (див. рис. 4.6). Отже стратегію (4.10) слід уточнити.

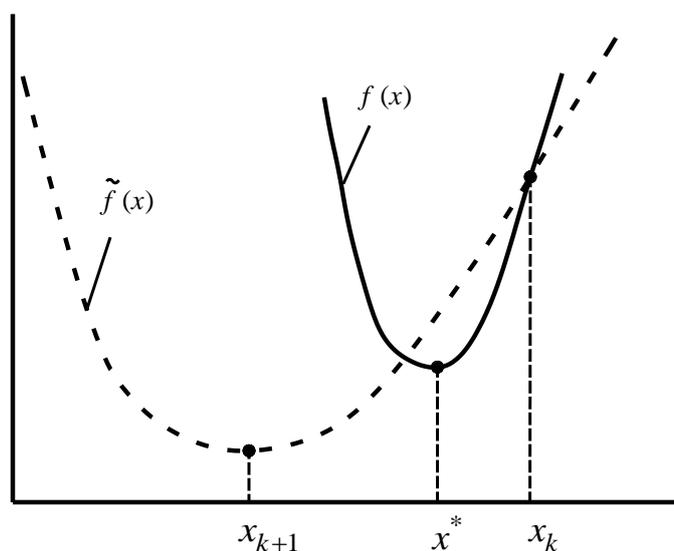


Рис. 4.6. Схема реалізація методу Ньютона

Щоб бути впевненими, що ми просуваємося до мінімуму, будемо на кожному кроці перевіряти виконання нерівності

$$f(x_{k+1}) < f(x_k).$$

Якщо ця нерівність виконується, то можна переходити до наступного кроку.

Якщо ж

$$f(x_{k+1}) > f(x_k), \text{ та } f'(x_k) \cdot (x_{k+1} - x_k) < 0$$

то значення функції $f(x)$ має спочатку зменшуватися в напрямку від x_k до x_{k+1} , а тому наступну прийнятну точку можна знайти роздроблюючи крок у зворотному напрямку, наприклад, прийнявши

$$x'_{k+1} = \frac{x_{k+1} + x_k}{2}.$$

Із співвідношення (4.10) видно, що добуток

$$f'(x_k) \cdot (x_{k+1} - x_k)$$

від'ємний тоді і тільки тоді, коли $f''(x_k)$ додатна. Це означає: якщо локальна модель, яка використовується для отримання Ньютонівського кроку, має мінімум, а не максимум, то гарантується існування прийнятного напрямку кроку.

З іншої сторони, якщо

$$f''(x_k) < 0 \text{ та } f'(x_k) \cdot (x - x_k) > 0,$$

то при переході від x_k до x_{k+1} , значення функції $f(x)$ спочатку збільшується, а тому слід зробити крок у протилежному напрямку.

Критерієм припинення ітераційного процесу можна, наприклад, вибрати умову

$$\left| \frac{f'(x_{k+1})}{f(x_{k+1})} \right| < \varepsilon, \quad (4.11)$$

де

ε - заздалегідь вибрана точність.

Описаний метод з основним кроком, заданим (4.10), та наведеними уточненнями зазвичай називають *методом Ньютона або Ньютона-Рафсона*.

Зазначимо, що при розв'язанні деяких задач, похідні функції $f(x)$ розрахувати неможливо з тих, чи інших причин. За таких умов може бути використана відповідна *модифікація методу Ньютона*, якою передбачена апроксимацію першої та другої похідних функції шляхом заміни диференціалів кінцевими прирощуваннями так, як це описано нижче.

Виберемо деяке наближення x_k екстремального значення аргументу та

мале значення h кроку.

Розглянемо три значення аргументу та відповідні їм значення функції

$$x_k - h, x_k, x_k + h, f(x_k - h), f(x_k), f(x_k + h).$$

На підставі цих даних апроксимуємо першу та другу похідні функції як

$$f'(x_k) = \frac{f(x_k + h) - f(x_k - h)}{2h},$$

$$f''(x_k) = \frac{f'(x_k + h) - f'(x_k - h)}{2h} = \frac{f(x_k + h) - 2f(x_k) + f(x_k - h)}{2h^2}.$$

Підстановка отриманих значень похідних у (4.10) приводить до співвідношення

$$x_{k+1} = x_k - h \frac{f(x_k + h) - f(x_k - h)}{f(x_k + h) - 2f(x_k) + f(x_k - h)} \quad (4.12)$$

Формула (4.12) визначає основний крок алгоритму, який зазвичай зветься *квазіньютонівським* методом або *модифікованим методом Ньютона*. При цьому, звичайно, всі міркування щодо величини кроку ($x_{k+1} - x_k$), наведені під час виведення формул метода Ньютона-Рафсона, залишаються в силі.

4.4. ДЕТЕРМІНОВАНІ МЕТОДИ ПОШУКУ ЕКСТРЕМУМУ БАГАТОВИМІРНИХ ФУНКЦІЙ

*Лише дурні йдуть весь час уперед.
Розумні ж ідуть туди, куди їм треба.*

NN

Пошук екстремуму багатовимірної функції є стержнем багатьох алгоритмів, що використовуються під час математичного моделювання МО.

Нижче розглянуті окремі алгоритми розв'язання екстремальних задач у багатовимірному просторі, які на сьогодні набули найбільшої популярності. При цьому алгоритми, які на сьогодні можна вважати традиційними (або класичними), розглядаються переважно в оглядовому плані, тоді як опис деяких найсучасніших алгоритмів, тісно пов'язаних на методологічному рівні з технологією ШНМ, виділені в окремі розділи.

4.4.1. Загальні поняття про багатовимірну функцію

Поняття багатовимірної функції тісно пов'язане з поняттям її *рельєфу*, який наочно демонструє характер змінення функції в тій чи іншій області аргументів. Знання або, навіть, приблизне уявлення рельєфу функції суттєво

допомагає при виборі та реалізації більшості відомих алгоритмів розв'язання екстремальної задачі.

Поняття «рельєф функції» розглядатимемо на прикладі функції двох змінних - $z = F(x, y)$, яка описує деяку поверхню в тривимірному просторі. з координатами x, y, z .

Задача пошуку мінімуму такої функції, тобто задача

$$F(x, y) \rightarrow \min$$

означає пошук нижчої точки вказаної поверхні.

Як в топографії, подамо рельєф поверхні, що описується функцією $F(x, y)$, лініями рівня (ізолініями).

Для побудови ізоліній проведемо плоскості $z = const$ та знайдемо лінії їх перетину з поверхнею $F(x, y)$. Проекції цих ліній на плоскість x, y і зветься лініями рівня. Напрямок спадання функції вказуватимемо штрихами поруч з лініями рівня.

Відповідно до виду ліній рівня умовно виділимо три типи рельєфу: котловина, яр та невпорядкований.

Котловинний рельєф (див. рис. 4.7) характеризується лініями рівня, що близькі до еліпсів, але в загальному випадку не співпадають з ними.

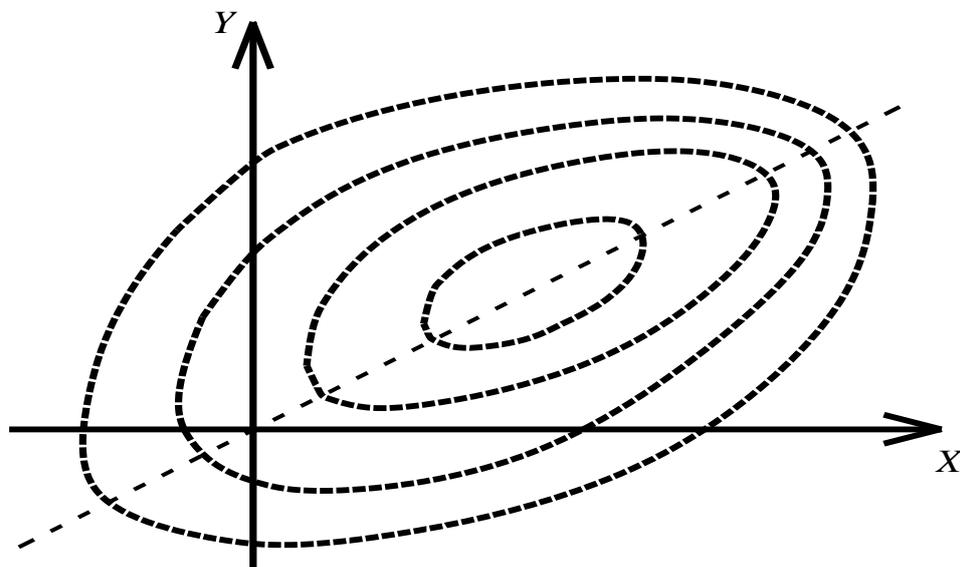


Рис. 4.7. Приклад рельєфу котловинного типу

Рельєф типу яр (див. рис. 4.8) характеризується лініями рівня, яким притаманний довгий витягнутий та скривлений вигляд.

На лініях рівня, притаманних рельєфу типу яр, часто можна спостерігати точки злому (або точки високої кривизни), де перша похідна розривна (або надто швидко змінює знак). Геометричне місце таких точок (на рис. 4.8 показане пунктирною кривою) зветься *істинним яром*, за умови, що кут злому (кут високої кривизни) спрямований в сторону зростання функції, або

гребенем, за умови, що кут злому спрямований в сторону спадання функції.

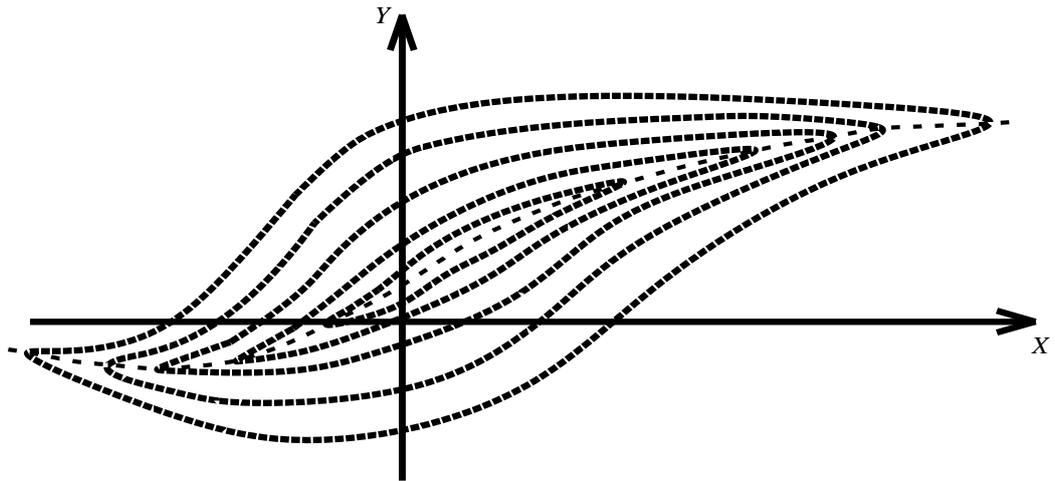


Рис. 4.8. Приклад рельєфу типу яр

Невпорядкований тип рельєфу характеризується наявністю багатьох максимумів та мінімумів.

На рис. 4.9, як приклад, наведено рельєф невпорядкованого типу, який відображає лінії рівня деякої періодичної функції, визначеної при додатних значеннях аргументів.

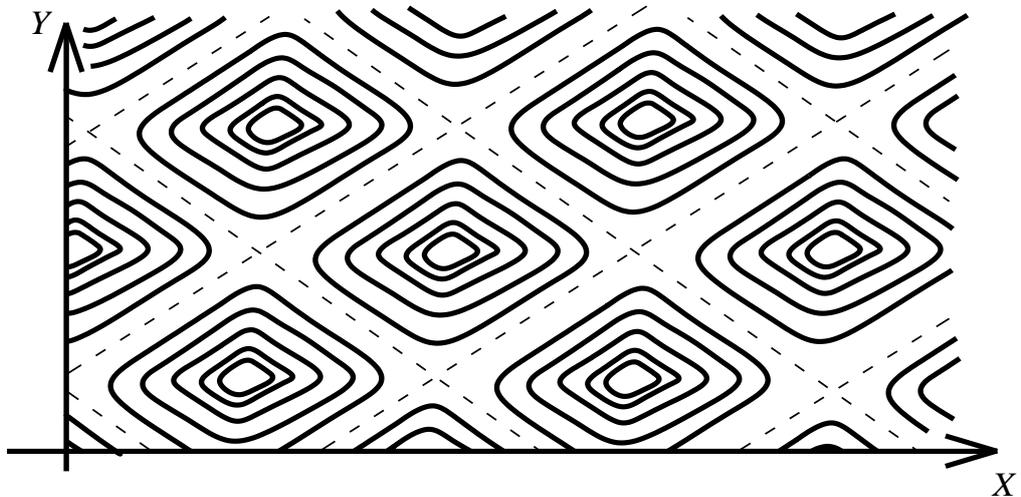


Рис. 4.9. Приклад рельєфу невпорядкованого типу

Всі методи пошуку мінімуму функції багатьох змінних зводяться до побудови траєкторії, вздовж якої функція спадає. Різні методи відрізняються лише способами побудови таких траєкторій. При цьому метод, що дає чудові результати при одному рельєфі нерідко виявляється неефективним на рельєфі іншого типу.

4.4.2. Метод координатного спуску (метод Гауса)

Сутність цього методу полягає в *поступовому просуванні до мінімуму*

кроками, що орієнтовані вздовж координатних осей.

Розглядатимемо функцію $F(x, y, z)$ трьох змінних.

Вибираємо як нульове наближення точку (x_0, y_0, z_0) .

Фіксуємо значення двох координат, наприклад, $y = y_0$, $z = z_0$.

За результатами такої процедури отримуємо функцію

$$f_1(x) = F(x, y_0, z_0),$$

що залежить лише від одної змінної x .

Використовуючи будь-який алгоритм розв'язання одновимірної екстремальної задачі, відшукуємо мінімум функції $f_1(x)$. Відповідне мінімуму цієї функції значення аргументу позначимо як x_1 .

Описана процедура відповідає реалізації кроку від точки (x_0, y_0, z_0) до точки (x_1, y_0, z_0) в напрямку, що паралельний осі x .

Припустимо, що за результатами реалізації вказаного кроку значення функції $F(x, y, z)$ зменшилося (якщо це не так, то слід повторити процедуру, зробивши крок в протилежний бік).

Надалі з нової точки (x_1, y_0, z_0) зробимо крок в напрямку паралельному осі y , тобто розглядатимемо функцію

$$f_2(y) = F(x_1, y, z_0).$$

Цей другий крок приводить у точку (x_1, y_1, z_0) , з якої робимо третій крок паралельно осі z та знаходимо мінімум функції

$$f_3(z) = F(x_1, y_1, z).$$

Таким чином переходимо в точку (x_1, y_1, z_1) .

На цьому завершується перший цикл спусків – завершується перша ітерація.

Надалі повторюватимемо цикли.

На кожному спуску досліджувана функція не зростатиме. Крім того, ця функція обмежена знизу її значенням в точці мінімуму, а саме, значенням

$$F^* = F(x^*, y^*, z^*)$$

Таким чином, описаний ітераційний процес сходиться до деякої границі $\tilde{F} \geq F^*$.

Чи перетвориться вказана нерівність на тотожність, тобто чи приведе реалізація наведених процедур координатних спусків до мінімуму функції і як швидко це відбудеться, залежатиме від функції (точніше від її рельєфу в

області реалізації процедури пошуку) та вибору початкового наближення.

Схема успішної реалізації алгоритму координатного спуску наведена на рис. 4.10.

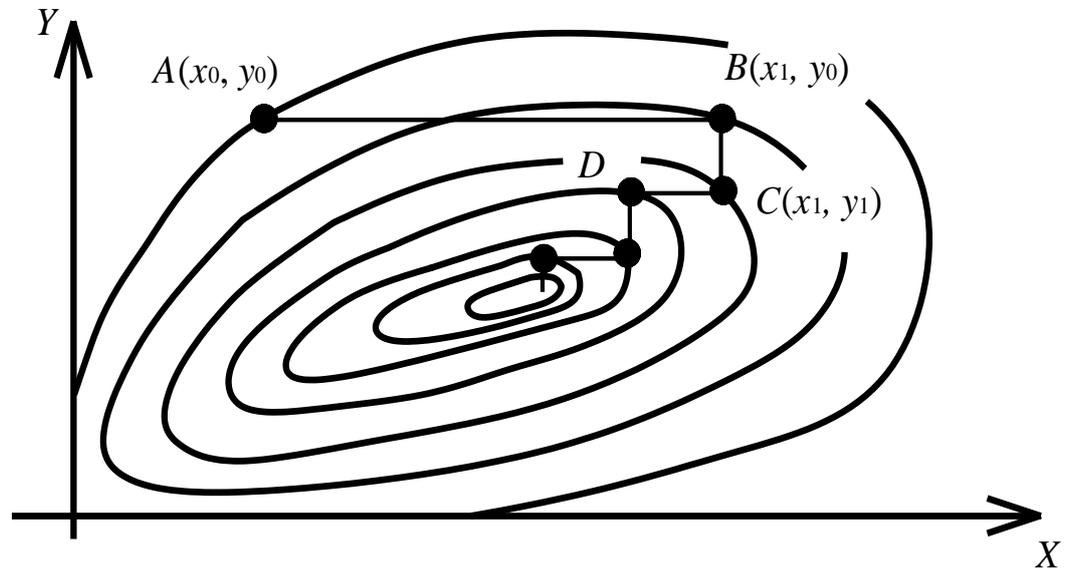


Рис. 4.10. Схема успішної реалізації алгоритму координатного спуску

Можливий варіант реалізації алгоритму координатного спуску, який приводить до зупинку в точці далекій від мінімуму показаний на рис. 4.11.

За умови, що досліджувана функція достатньо гладка в деякій околиці мінімуму, процес спуску за координатами зазвичай сходиться до цього мінімуму. Однак швидкість сходження надто сильно залежить від форми ліній рівня. Так, за наявності рельєфу типу яр, при попаданні траєкторії спуску на дно яру сходження зазвичай стає настільки повільним, що продовження процесу пошуку втрачає сенс.

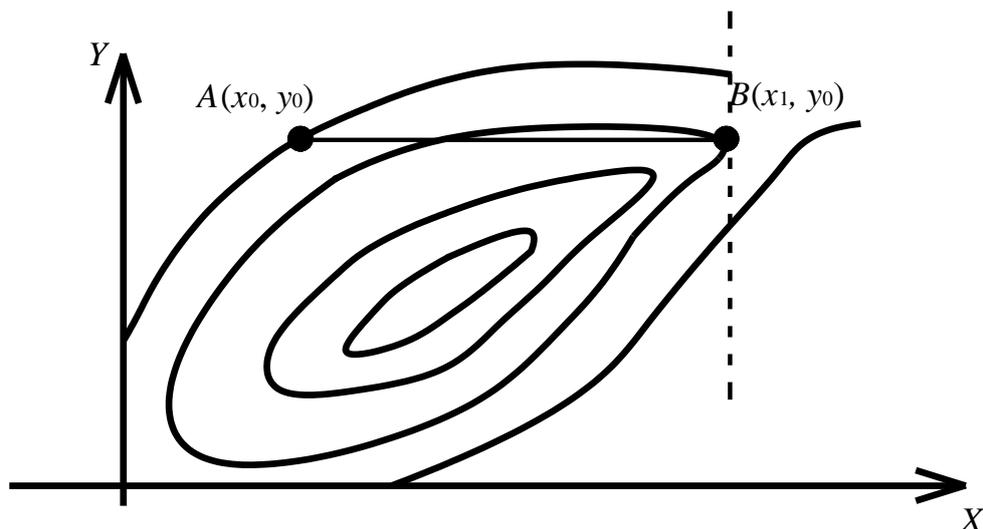


Рис. 4.11. Схема реалізації алгоритму координатного спуску, який приводить до зупинки в точці далекій від мінімуму

Як правило метод координатного спуску використовується як перша (тестова) спроба під час реалізації процесу відшукування мінімуму багатовимірної функції.

4.4.3. Метод ярів

Цей метод розроблявся спеціально для мінімізації багатовимірних функцій з рельєфом типу яр.

Для пояснення базових засад реалізації методу ярів розглядатимемо задачу пошуку мінімуму функції, відображаючи схему процесу на рис. 4.12.

Виберемо довільну точку A_0 та спустимося з неї (наприклад, використовуючи метод координатного спуску), роблячи не дуже багато кроків, тобто не вимагаючи від процесу цього спуску високої точності. Кінцеву точку такого спуску позначимо як r_0 .

За умови, що функція характеризується рельєфом типу яр, точка r_0 опиниться поблизу дна яру (див. рис. 4.12).

Надалі виберемо другу точку A_1 , що не надто віддалена від точки A_0 . З цієї точки також реалізуємо спуск і потрапляємо в деяку точку r_1 , яка також розмішуватиметься близько дна яру.

Проведемо крізь точки r_0, r_1 на дні яру пряму, що приблизно відображає лінію дна яру.

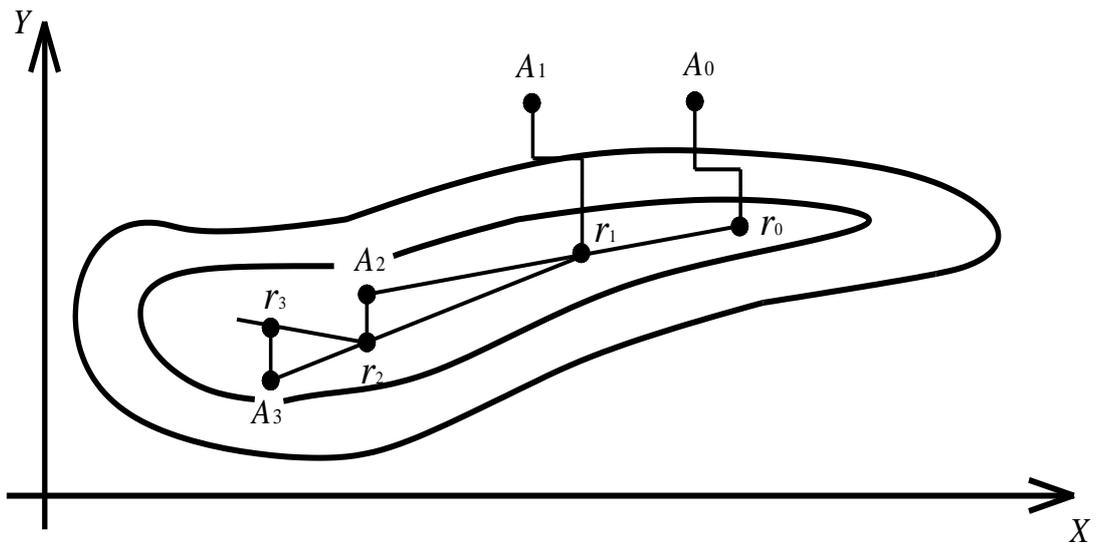


Рис. 4.12. Схема реалізації методу ярів

Зсунемося вздовж лінії дна яру в бік спадання функції та виберемо нову точку A_2 на цій прямій на відстані h від точки r_1 (див. рис. 4.12).

Величина h зветься «кроком яру» й для кожної функції підбирається індивідуально під час реалізації процесу пошуку її мінімуму (як правило евристичними методами).

Зазвичай дно яру не є відрізком прямої, тому точка A_2 реально лежатиме не на дні яру, а на його схилі. З цієї точки спускаємося на дно яру і потрапляємо в деяку точку r_2 . Надалі з'єднуємо точки r_1, r_2 прямою лінією, відмічаючи нову лінію дна яру, та робимо наступний крок вздовж дна яру.

Продовжуємо реалізацію наведеного процесу до тих пір, поки зменшуються значення функції на дні яру, тобто в точках r_0, r_1, \dots, r_n .

За умови

$$F(r_{n+1}) > F(r_n),$$

процес має бути припинений і точку r_{n+1} не слід використовувати.

Як бачимо, метод ярів розрахований на проходження вздовж яру з виходом в котловину поблизу мінімуму. В самій котловині зазвичай значення мінімуму краще уточнювати іншими методами.

4.4.4. Загальна характеристика методів, що використовують значення похідних функцій

Методи спуску та їх модифікації, такі, наприклад, як розглядалися вище, використовують виключно значення багатовимірної функції (методи нульового порядку). Зазвичай ця група методів має невелику швидкість сходження. Тому являє інтерес розглянути деякі з методів, що використовують

перші та другі похідні досліджуваної функції (відповідно, *методи першого та другого порядку*).

Розглядаючи методи пошуку мінімуму багатовимірної функції, вважатимемо, що функція $f(x): \mathfrak{R}^n \rightarrow \mathfrak{R}$, її градієнт $\nabla f(x)$ та матриця Гессе $\nabla^2 f(x)$ існують та безперервні в області визначення функції.

Більшість методів, заснованих на використанні похідних, реалізують ітераційний процес відповідно до співвідношення

$$x_{k+1} = x_k + \lambda_k s(x_k),$$

де

x_k, x_{k+1} - відповідно, значення аргументу функції на поточному (k -му) та наступному ($k+1$)-му кроці, яке розглядається як наближення до розв'язку x^* ;

λ_k - параметр, що характеризує довжину кроку;

$s(x_k) = s_k$ - параметр, що вказує напрямок пошуку на поточному кроці.

4.4.5. Градієнтні методи

Відомо, що градієнт функції в будь-якій точці x вказує напрямок найбільшого локального збільшення функції. Отже під час пошуку мінімуму можна спробувати *рухатися в напрямку, протилежному градієнту* функції в даній точці, тобто в напрямку найшвидшого спуску.

Такий підхід приводить до ітераційного процесу, який зветься *методом градієнтного спуску*, та подається формулою

$$x_{k+1} = x_k - \lambda_k \nabla f(x_k),$$

або

$$x_{k+1} = x_k - \lambda_k \frac{\nabla f(x_k)}{\|\nabla f(x_k)\|} = x_k - \lambda_k s_k,$$

де

$\|\nabla f(x_k)\|$ - норма (довжина) вектора градієнта функції.

Значення параметру s_k в даному випадку можна вважати одиницею, бо параметр λ , згідно наведеному співвідношенню, по суті, бере на себе функцію величини кроку, а s_k вказує лише напрямок, в якому цей крок має бути реалізований.

Звичайно, вибір параметра λ_k суттєво впливатиме на траєкторію спуску (див. рис. 4.13).

Як видно з рис. 4.13 при суттєвому значенні λ_k (крива 3) траєкторія

спуску (переміщення) являтиме коливальний процес, а надто велике значення λ_k як правило спричинятиме розходження ітераційного процесу.

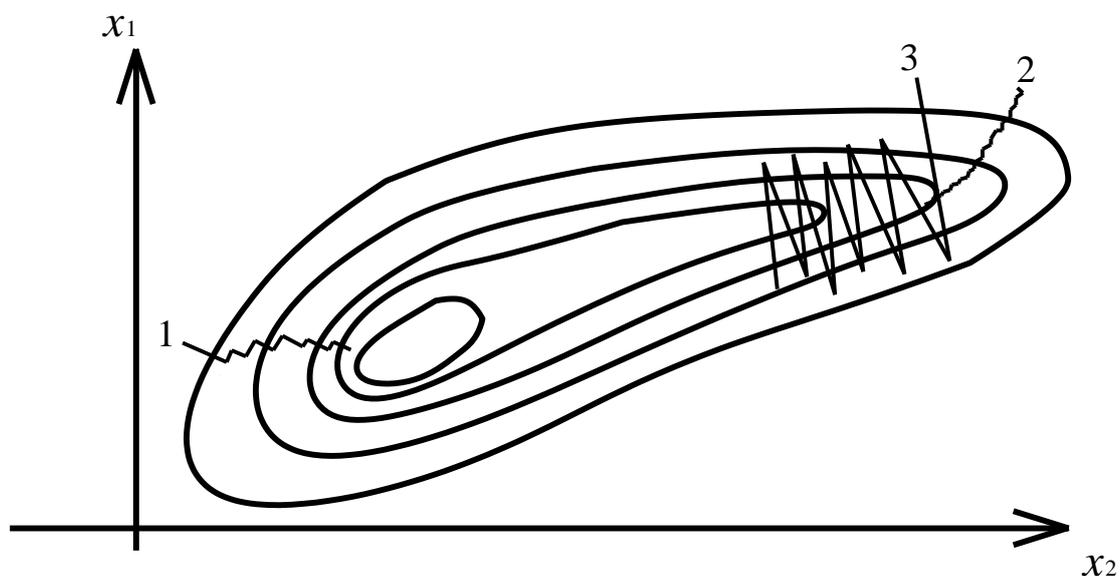


Рис. 4.13. Схема реалізації процесу градієнтного спуску
1 – крок вибрано правильно; 2 – крок надто малий; 3 – крок надто великий

При малих значеннях λ_k (крива 2 на рис. 4.13) траєкторія спуску буде плавною, але сходження процесу буде надто повільним.

Параметр λ_k можна приймати як постійну величину протягом усього процесу або вибирати різним на кожній ітерації.

Нерідко на кожному кроці процесу параметр λ_k вибирають на підставі процедури одновимірної мінімізації відповідної функції вздовж напрямку s_k , використовуючи для цього визначений метод одновимірної мінімізації. За таких умов відповідний ітераційний процес зазвичай зветься *методом найшвидшого спуску* або *методом Коші*.

За умови, що параметр λ_k визначається за результатами одновимірної мінімізації, тобто згідно співвідношення

$$\lambda_k = \arg \min_{\lambda} f(x_k + \lambda s_k),$$

градієнт в точці чергового наближення буде ортогональний до напрямку спуску на попередньому кроці (як це показано на рис. 4.13, крива 1).

Одновимірна мінімізація вздовж напрямку s_k покращує сходження методу, але одночасно приводить до зростання кількості обчислень значень функції $f(x)$ на кожній ітерації. Крім того, поблизу екстремуму норма градієнту $\|\nabla f(x_k)\|$ стає близькою до нуля, а отже сходження надто

уповільнюється.

Ефективність градієнтного методу суттєво залежить від виду функції, мінімум якої шукається. Наприклад, він мало придатний для функцій з рельєфом типу яр. В класичному (наведеному вище) варіанті цей метод використовується не часто. Однак, слід зазначити, що більшість відомих на сьогодні методів мінімізації багатовимірної функції на підставі значень функції та її похідних виникли як розвиток та модифікації саме градієнтного методу.

4.4.6. Метод Ньютона

Метод Ньютона пошуку екстремуму багатовимірної функції $f(x)$ заснований на квадратичній апроксимації (моделюванні) цієї функції, розкладанням її у ряд Тейлора в околі точки поточного наближення x_k з урахуванням членів до другого порядку включно.

На підставі сказаного, квадратичну модель функції $f(x)$ можна подати як

$$\tilde{f}(x_k + p) = f(x_k) + \nabla f^T(x_k) \cdot p + \frac{1}{2} p^T \cdot \nabla^2 f(x_k) \cdot p, \quad (4.13)$$

де

$(^T)$ – позначка операції транспонування матриці;

$\nabla^2 f(x_k)$ - матриця Гессе функції $f(x)$ в точці x_k (ця матриця має бути не вироджена в усій області пошуку).

Чергове значення аргументу

$$x_{k+1} = x_k + s_k,$$

знаходитимемо з умови мінімуму функції-моделі (4.13).

Необхідною умовою мінімуму є

$$\nabla \tilde{f}(x_{k+1}) = 0.$$

З урахуванням (4.13) матимемо

$$\nabla \tilde{f}(x_k + s_k) = 0 = \nabla f(x_k) + \nabla^2 f(x_k) \cdot s_k.$$

Наведене приводить до алгоритму, яким передбачається розв'язання відносно s_k на кожній k -й ітерації системи рівнянь:

$$\nabla^2 f(x_k) \cdot s_k = -\nabla f(x_k) \quad (4.14)$$

з подальшим визначенням точки наступного наближення до мінімуму за

формулою

$$x_{k+1} = x_k + s_k \quad (4.15)$$

Алгоритм (4.14), (4.15) і зветься методом Ньютона.

Позитивною стороною методу Ньютона є підвищена (у порівнянні з методом градієнтного спуску) швидкість сходження. В той же час, цьому методу притаманні визначені недоліки:

- не забезпечує глобального (до глобального мінімуму) сходження, а отже потребує відповідного («хорошого») початкового наближення;
- потребує аналітичного задавання першої та другої похідних;
- не містить інформації, яка б вказувала на те, що просування йде в бік максимуму або сідлової точки, де градієнт функції також дорівнює нулю.

Останній недолік пов'язаний з тим, що під час реалізації ітераційного процесу, кожен крок спрямовується в сторону стаціонарної точки локальної квадратичної моделі функції незалежно від «характеру стаціонарності» цієї точки, адже стаціонарність притаманна точці і мінімуму, і максимуму, і сідловій точці.

З огляду на поставлену задачу кожен крок буде виправданий лише за умови, що матриця Гессе у відповідній точці додатно визначена.

Взагалі кажучи, методу Ньютона не притаманна висока надійність, він потребує суттєвих витрат на апріорне дослідження функції в області мінімізації.

Зазначимо, що в різних модифікаціях метода Ньютона необхідно здійснювати велику кількість розрахунків, адже на кожній ітерації слід спочатку розрахувати елементи матриці розміром $n \times n$, а потому розв'язати систему лінійних рівнянь. Використання замість диференціалів кінцеві прирощення не покращує, а частіше погіршує ситуацію.

Саме тому на сьогодні побудовано чимало так званих *квазіньютонівських* алгоритмів. По більшості ці алгоритми спираються на можливість апроксимації кривизни нелінійної функції, мінімум якої розшукується, без явного формування її матриці Гессе. Зазвичай дані щодо кривизни функції накопичуються на підставі спостереження за зміною градієнта під час реалізації процедури спуску.

4.4.7. Метод Левенберга-Марквардта

Алгоритм Левенберга-Марквардта (*Levenberg-Marquardt Algorithm, LMA-алгоритм*) на сьогодні є одним з найпоширеніших алгоритмів теорії оптимізації, які використовуються в технології ШНМ. За продуктивністю цей алгоритм перевершує метод найскорішого спуску й інші градієнтні методи в найрізноманітніших задачах.

На методологічному рівні *LMA*-алгоритм є комбінацією методів

градієнтного спуску та Ньютона. При цьому він містить базові позитивні властивості обох вказаних методів. Просування в напрямку антиградієнту з деякої початкової точки x_0 , яка розміщується на значній відстані від точки мінімуму x^* , зазвичай, приводить до суттєвого зменшення функції, мінімум якої розшукується. З іншої сторони, напрямок ефективного пошуку в околі точки мінімуму визначається за методом Ньютона.

Відповідно до *LMA*-алгоритму напрямок пошуку s_k визначається на підставі розв'язання рівняння

$$(H_k + \lambda_k I) \cdot s_k = -\nabla f(x_k), \quad \lambda_k \geq 0, \quad (4.16)$$

де

I – одинична матриця;

$H_k = \nabla^2 f(x_k)$ - матриця Гессе.

Чергова точка ітераційного процесу розраховується як

$$x_{k+1} = x_k + s_k. \quad (4.17)$$

У формулі (4.17) коефіцієнт перед s_k дорівнює одиниці, адже введення параметрів λ_k до співвідношення (4.16) дозволяє змінювати довжину кроку та його напрямок.

На початковій стадії пошуку значення параметру λ_0 вибирається достатньо великим, наприклад 10^4 . За таких умов ліва частина рівняння (4.16) набуває вигляду

$$(H_0 + \lambda_0 I) \cdot s_k \approx (\lambda_0 I) \cdot s_k = \lambda_0 s_k. \quad (4.18)$$

Таким чином, за умови великого значення λ_0 , як видно з (4.16) та (4.18), напрямок пошуку визначається як

$$s_k = -\frac{\nabla f(x_k)}{\lambda_0},$$

тобто відповідно до градієнтного методу найшвидшого спуску.

З формули (4.16) видно, що при зменшенні λ_k до нуля напрямок вектора s_k змінюватиметься від напрямку, що притаманний градієнтному методу, до напрямку, який визначається за методом Ньютона.

Якщо після першого кроку отримана точка (аргумент) з меншим значенням функції, тобто

$$f(x_1) < f(x_0),$$

то параметр кроку λ слід вибрати за нерівністю

$$\lambda_1 < \lambda_0$$

і реалізувати ще один крок.

В протилежному випадку слід вибрати

$$\lambda_0 = \beta \lambda_0, \quad \beta > 1$$

і повторно реалізувати попередній крок.

Як вже вказувалося, одним з недоліків методу Ньютона є те, що за умови коли матриця Гессе у відповідній точці додатно не визначена, метод Ньютона не приводить до спадання функції $f(x)$. Саме тому модифікація матриці Гессе за співвідношенням

$$H_k + \lambda_k I$$

при відповідному виборі параметру λ_k необхідна для виправлення вказаної ситуації, адже матриця I завжди додатно визначена.

Загальний алгоритм Левенберга-Марквардта можна подати у вигляді:

1. задати:

- початкове наближення аргументу x_0 до шуканого значення x^* ,
- максимальну допустиму кількість ітерацій,
- параметр сходження ε ;

2. прийняти

$$k = 0, \quad \lambda_0 = 10^4;$$

3. розрахувати $\nabla f(x_k)$;

4. перевірити виконання умови

$$\|\nabla f(x^k)\| < \varepsilon;$$

якщо ця умова виконується то перейти до п. 11;

5. перевірити умову

$$k \geq M,$$

якщо ця умова виконується – перейти до п. 11;

6. розрахувати крок, розв'язавши систему рівнянь (4.16);

7. прийняти

$$x_{k+1} = x_k + s_k;$$

8. перевірити умову

$$f(x_{k+1}) < f(x_k);$$

якщо ця умова виконується перейти до п. 9, інакше – до п. 10;

9. прийняти

$$\lambda_{k+1} = \frac{1}{2}\lambda_k, \quad k = k + 1$$

та перейти до п. 3;

10. прийняти

$$\lambda_k = 2\lambda_k$$

та перейти до п. 7;

11. виведення результатів роботи алгоритму

$$x_k, f(x_k), \nabla f(x_k), \|\nabla f(x_k)\|, k.$$

Зазначимо, що коефіцієнт зменшення (збільшення) параметру λ в пп. 9, 10 має евристичний характер і, звичайно, може бути й іншим (в деяких випадках він може бути завбільшки 10).

Недоліком наведеного варіанту *LMA*-алгоритму є те, що при великих значеннях λ , обчислена матриця Гессе ніяк не використовується. Однак можна отримати деяку користь з другої похідної навіть в цьому випадку, змінюючи масштаб кожної компоненти градієнта відповідно до кривизни функції. Це приводить до збільшення кроку вздовж напрямку, де значення градієнта мале, а отже класична «проблема яру» не виникатиме.

Наведене положення можна врахувати, замінивши одиничну матрицю у формулі (4.16) на діагональну матрицю, елементами якої є елементи, що розміщені на головній діагоналі матриці Гессе.

Слід зазначити, що, незважаючи на евристичний (а отже не оптимальний) характер *LMA*-алгоритму, він добре працює під час розв'язання практичних задач.

Мабуть єдиний недолік *LMA*-алгоритму полягає в необхідності розрахунку зворотної матриці на кожному кроці. Навіть з огляду на те, що обчислення зворотної матриці на сьогодні зазвичай виконується з використанням прискорених методів обчислення псевдозворотної матриці (наприклад, таких методів як розкладання по сингулярним числам матриці), час однієї ітерації стає неприйнятним для навчання ШНМ, які призначені для роботи з об'єктами, що задаються декількома тисячами параметрів (така ситуація нерідко зустрічається під час використання технології ШНМ, особливо, для розв'язання задач управління МО або прогнозування). Для моделей же середніх розмірів (з декількома сотнями параметрів) *LMA*-алгоритм працює навіть швидше за найпростіший градієнтний метод.

4.5. СТАТИСТИЧНІ МЕТОДИ ПОШУКУ ЕКСТРЕМУМУ ФУНКЦІЇ

У кожній великій справі завжди доводиться якусь частку залишити на долю випадковості

Наполеон

На відміну від детермінованих методів пошук екстремуму функції статистичними (стохастичними, випадковими) методами передбачає обчислення значень функції на деякій множині *випадково вибраних значень аргументу*, проведення відповідного аналізу отриманих результатів та відбір області значень аргументів, що визначена як перспективна з точки зору наявності шуканого результату.

Статистичні методи (або методи випадкового пошуку) отримали достатньо широке поширення під час розв'язання найрізноманітніших прикладних задач.

Це пояснюється в першу чергу тим, що із зростанням розмірності задачі різко знижується ефективність детермінованих (регулярних) методів пошуку - так зване «прокляття розмірності».

По-друге, нерідко наявна інформація щодо функції, екстремум якої розшукується, надто мала, щоб впевнено використовувати детерміністські методи.

Найбільший ефект використання статистичних методів дає при розв'язанні задач великої розмірності або при пошуку глобального екстремуму.

Під статистичними методами пошуку зазвичай розуміють методи, які використовують елемент випадковості або під час збирання інформації щодо досліджуваної (цільової) функції при пробних кроках, або для покращення значень функції при робочому кроці. Випадковим чином може вибиратися напрямок спуску, довжина кроку тощо.

Статистичним алгоритмам притаманні визначені переваги, серед яких, перш за все, слід відзначити:

- простота реалізації та налаштування відповідних програм;
- надійність та завадостійкість;
- універсальність;
- можливість введення операцій навчання до алгоритму пошуку;
- можливість введення процедури прогнозування оптимальної точки (оптимального розв'язку задачі).

До основних недоліків статистичних методів зазвичай відносять необхідність реалізації великої кількості обчислень значення функції, екстремум якої розшукується, та повільне сходження процесу в околі екстремуму.

Як вже вказувалося, стабільно вважається, що переваги статистичних

методів проявляються із зростанням розмірності задачі, адже обчислювальні витрати при використанні детерміністських методів пошуку із зростанням розмірності зростають швидше, ніж при використанні статистичних методів.

4.5.1. Найпростіший випадковий пошук

Сутність алгоритму найпростішого випадкового пошуку розглядатимемо на прикладі реалізації алгоритму мінімізації функції $f(x)$ на відрізку $[A, B]$.

Відповідно до алгоритму на відрізку $[A, B]$ згідно рівномірного закону вибираємо випадкову точку x_1 та розраховуємо значення y_1 функції в цій точці $y_1 = f(x_1)$.

Надалі таким же чином (тобто випадково та з використанням рівномірного закону розподілення імовірності) вибираємо випадкову точку x_2 та розраховуємо $y_2 = f(x_2)$.

Запам'ятовуємо мінімальне з цих значень функції та точку (значення аргументу), в якій значення функції мінімальне.

Надалі генеруємо (звичайно, за тим же правилом) нову точку.

Робимо N чисельних експериментів за наведеною схемою.

Визнано кращою за результатами всіх експериментів точку приймаємо як розв'язок задачі. Тобто за такий спосіб в якості розв'язку задачі серед всіх «випадково» генерованих точок буде прийнята та, в якій значення функції мінімальне.

Оцінимо кількість експериментів, необхідних для визначення шуканої точки мінімуму з заданою точністю.

Прийнявши розмірність вектора змінних – n , об'єм відповідної n -вимірної області (n -вимірного паралелепіпеду), в якій ведеться пошук мінімуму, можна записати як

$$V = \prod_{i=1}^n (A_i - B_i).$$

За умови, що необхідно знайти розв'язок задачі з точністю ε_i ($i = \overline{1, n}$), по кожній змінній, точка-розв'язок за результатами реалізації алгоритму врешті-решт має опинитися в околі оптимальної точки, тобто в області (паралелепіпеді) об'ємом

$$V_\varepsilon = \prod_{i=1}^n \varepsilon_i.$$

Імовірність *попадання* в цей об'єм за результатами одного експерименту дорівнює

$$P_{\varepsilon} = \frac{V_{\varepsilon}}{V}.$$

Тоді як імовірність *не попадання* точки у згаданий об'єм становитиме

$$\bar{P}_{\varepsilon} = 1 - P_{\varepsilon}.$$

Як видно з процедури реалізації наведеного алгоритму найпростішого випадкового пошуку, всі експерименти відносно визначення кожного випадкового значення аргументу є незалежними. При цьому імовірність не попадання точки в околицю шуканої точки (точніше у відповідний об'єм) за N експериментів визначатиметься як

$$\bar{P} = (1 - P_{\varepsilon})^N.$$

Отже імовірність того, що розв'язок задачі буде знайдено за N експериментів, дорівнюватиме

$$P = 1 - (1 - P_{\varepsilon})^N.$$

Таким чином, оцінка необхідної для знаходження мінімуму функції з заданою точністю кількості експериментів N визначатиметься нерівністю

$$N \geq \frac{\ln(1 - P)}{\ln(1 - P_{\varepsilon})}.$$

Спираючись на задану точність ε_i , $i = \overline{1, n}$ та величину V , можна визначити P_{ε} , і, задавшись імовірністю P , подивитись, як змінюється необхідна кількість експериментів N в залежності від P_{ε} та P (див. табл. 4.1).

Таблиця 4.1

P_{ε}	P				
	0,8	0,9	0,95	0,99	0,999
0,1	16	22	29	44	66
0,025	64	91	119	182	273
0,01	161	230	299	459	688
0,005	322	460	598	919	1379
0,001	1609	2302	2995	4603	6905

З наведеної табл. 4.1 наочно видно, що зменшення заданої імовірності P_{ε} попадання у відповідну зону навколо шуканої точки, рівно як і збільшення

імовірності P (збільшення впевненості в тому, що розв'язок задачі може бути знайдений за результатами не більше N експериментів) швидко і суттєво збільшує кількість необхідних експериментів.

Зазначимо, що під час розв'язання екстремальних задач на областях зі складною геометрією, зазвичай вписують цю область в «правильну» геометричну фігуру – в загальному випадку найчастіше такою фігурою вибирають n -вимірний паралелепіпед (див. рис. 4.14).

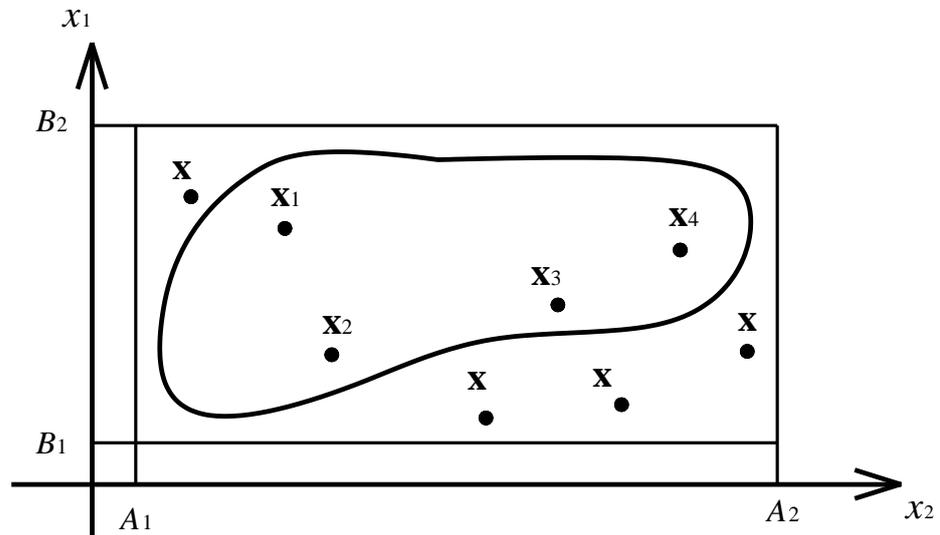


Рис. 4.14. Реалізація найпростішого випадкового пошуку на складній області

(в тестових процедурах задіяні лише значення X_1, X_2, X_3, X_4)

Це робиться з метою визначеності з інтервалами, в яких слід генерувати рівномірно розподілені тестові значення аргументів. Саме в межах цього паралелепіпеду і генерують випадкові точки (значення аргументів) за рівномірним законом, в подальшому залишаючи як тестові для реалізації експериментів (тестів) лише ті з них, котрі потрапляють до допустимої області (див. рис. 4.14).

Розрізняють *направлений* та *ненаправлений* випадковий пошук.

Ненаправлений випадковий пошук

При такому пошуку всі тестові експерименти (випробування) проводять повністю незалежно від результатів експериментів, що передували даному (тобто незалежно від усіх попередніх експериментів).

Сходження такого пошуку зазвичай надто мале, але такий підхід має важливу перевагу, пов'язану з можливістю розв'язання багато екстремальних задач (з пошуком глобального екстремуму).

Прикладом ненаправленого пошуку є розглянутий вище варіант реалізації найпростішого випадкового пошуку.

Направлений випадковий пошук

Під час реалізації процедури направленої випадкової пошуку окремі випробування пов'язані між собою. Результати вже проведених випробувань використовуються для формування наступних випробувань. Як правило, випадковість використовується при формуванні напрямку спуску.

Сходження таких методів, як правило, вище, але самі методи зазвичай приводять лише до локальних екстремумів.

4.5.2. Алгоритм парної проби

В даному алгоритмі чітко розділені пробний та робочий кроки.

Припустимо, що $\mathbf{x}^{(k)}$ - знайдене на k -му кроці найменше значення функції $f(\mathbf{x})$, мінімум якої розшукується. На підставі рівномірного закону генеруємо випадковий одиничний вектор ξ (напрямок майбутнього кроку) і по обидві сторони від первісної точки $\mathbf{x}^{(k)}$ реалізуємо дві проби (два випробування, два експерименти), тобто проводимо обчислення функції в точках

$$\mathbf{x}_1^{(k)} = \mathbf{x}^{(k)} + g \cdot \xi \quad \text{та} \quad \mathbf{x}_2^{(k)} = \mathbf{x}^{(k)} - g \cdot \xi$$

де

g - довжина пробного кроку.

Робочий крок (довжиною a) робиться в напрямку найменшого значення функції.

Тобто чергове наближення визначається співвідношенням

$$\begin{aligned} \mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} + \Delta \mathbf{x}^{(k)} = \\ &= \mathbf{x}^{(k)} + a \cdot \xi \cdot \text{sign}\left(f\left(\mathbf{x}^{(k)} - g \cdot \xi\right) - f\left(\mathbf{x}^{(k)} + g \cdot \xi\right)\right). \end{aligned}$$

Схема реалізації розглянутого алгоритму наведена на рис. 4.15.

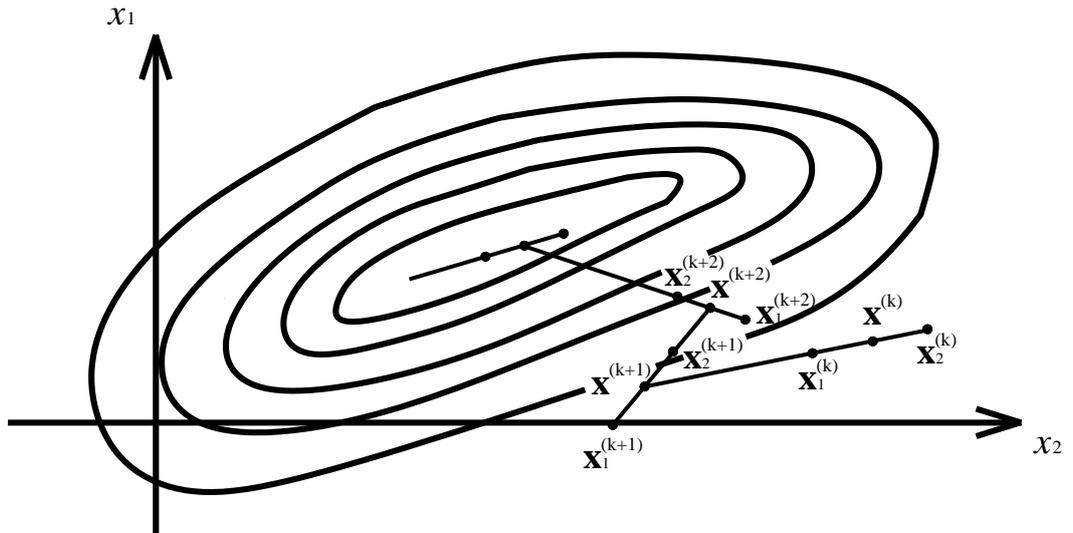


Рис. 4.15. Схема реалізації алгоритму парної проби

Особливістю наведеного алгоритму є його підвищена «схильність до блукання». Якщо вчасно не зупинити процес, то навіть знайшовши екстремум, алгоритм здатен в подальшому увести процес пошуку в ту чи іншу сторону.

4.5.3. Алгоритм найкращої проби

Для розгляду реалізації алгоритму вважатимемо, що на k -му кроці в багатовимірному просторі визначення функції $f(\mathbf{x})$ отримано точку $\mathbf{x}^{(k)}$.

Наступною процедурою алгоритму є генерування напрямків – генерування m випадкових одиничних векторів (тобто m випадково вибраних напрямків можливої реалізації чергового кроку) ξ_1, \dots, ξ_m .

Зазначимо, що кінці цих одиничних векторів утворюють коло одиничного радіусу з центром в точці $\mathbf{x}^{(k)}$.

Після генерування напрямків реалізуються пробні кроки в напрямках

$$g \cdot \xi_1, \dots, g \cdot \xi_m$$

та розраховуються значення функції в точках

$$\left(\mathbf{x}^{(k)} + g \cdot \xi_1\right), \dots, \left(\mathbf{x}^{(k)} + g \cdot \xi_m\right).$$

Надалі вибирається той напрямок ξ^* , котрий приводить до найбільшого зменшення значення функції, тобто вибирається напрямок

$$\xi^* = \arg \min_{i=1, m} f(\mathbf{x}^{(k)} + g \cdot \xi_i).$$

В отриманому за такий спосіб напрямку реалізується крок

$$\Delta \mathbf{x}^{(k)} = \lambda \cdot \xi^*,$$

де

λ - параметр, який можна визначити за будь-яким (в т.ч., звичайно, й евристичним) правилом, наприклад, як результат мінімізації функції в напрямку найкращої проби.

Зі збільшенням кількості експериментів вибраний напрямок наблизатиметься до напрямку $-\nabla f(\mathbf{x})$, тобто до напрямку, що протилежний градієнту функції у відповідній точці.

Схема реалізації алгоритму найкращої проби наведена на рис. 4.16.

Зі збільшенням кількості експериментів вибраний напрямок наблизатиметься до напрямку $-\nabla f(\mathbf{x})$, тобто до напрямку, що протилежний градієнту функції у відповідній точці.

Схема реалізації алгоритму найкращої проби наведена на рис. 4.16.

За умови, що функція $f(\mathbf{x})$ близька до лінійної, є можливість прискорити пошук, розглядаючи разом з найкращою й найгіршу пробу. При цьому робочий крок можна реалізувати або в напрямку найкращої проби, або в напрямку, протилежному найгіршій пробі.

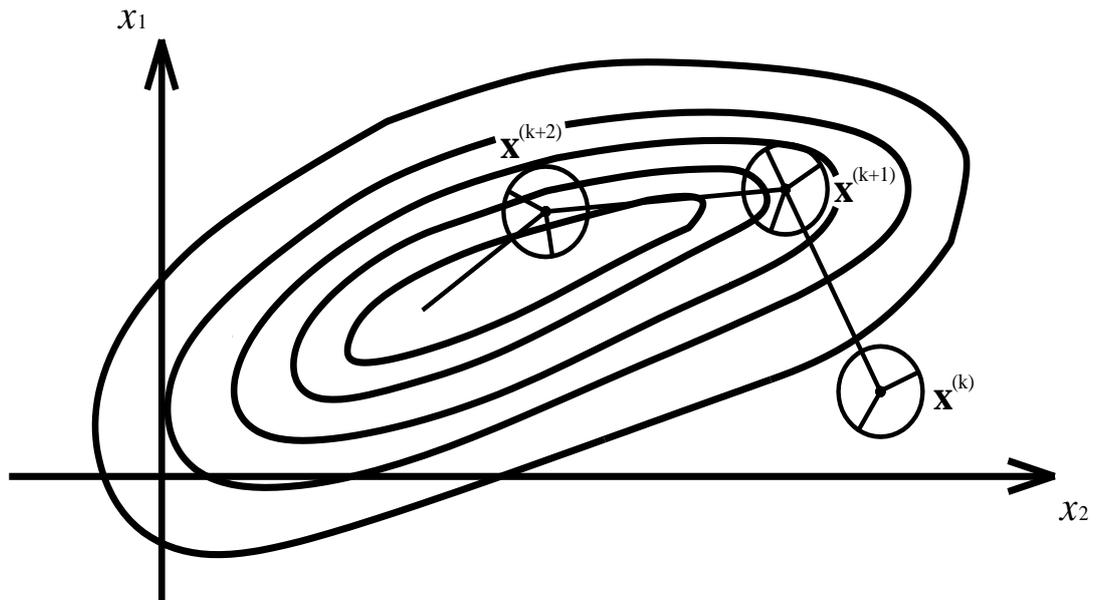


Рис. 4.16. Схема реалізації алгоритму найкращої проби (круги навколо точок $\mathbf{x}^{(k)}$, $\mathbf{x}^{(k+1)}$, $\mathbf{x}^{(k+2)}$ відображають геометричне місце кінцевих точок випадкових одиничних векторів)

4.5.4. Метод статистичного градієнта

Метод статистичного градієнту є розвиненням методу найкращої проби. Відповідно до цього методу з первісного стану $\mathbf{x}^{(k)}$ в m випадкових напрямках реалізується m незалежних проб $g \cdot \xi_1, \dots, g \cdot \xi_m$ (тут, як і в п. 4.5.3, ξ_1, \dots, ξ_m - одиничні вектори напрямків) та обчислюються значення функції у відповідних точках.

Значення прирощення функції при реалізації кожної j -ї згаданої проби запам'ятовуються, тобто запам'ятовуються величини

$$\Delta f_j(\mathbf{x}) = f(\mathbf{x}^{(k)}) + g \cdot \xi_j - f(\mathbf{x}^{(k)}).$$

Після цього формується векторна сума

$$\Delta \mathbf{f}(\mathbf{x}) = \sum_{j=1}^m \xi_j \cdot \Delta f_j(\mathbf{x})$$

В граничному випадку при $m \rightarrow \infty$ напрямком вектора $\Delta \mathbf{f}(\mathbf{x})$ співпадатиме з напрямком градієнта функції $f(\mathbf{x})$, екстремум якої розшукується.

При кінцевому значенні m вектор $\Delta \mathbf{f}(\mathbf{x})$ є статистичною оцінкою напрямку градієнта.

Згідно алгоритму в напрямку, протилежному градієнту (точніше його статистичній оцінці), реалізується черговий робочий крок. Як результат чергове наближення визначається співвідношенням

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \lambda \cdot \frac{\Delta \mathbf{f}(\mathbf{x})}{\|\Delta \mathbf{f}(\mathbf{x})\|}.$$

При виборі оптимального значення λ , котре мінімізує функцію $f(\mathbf{x})$ в заданому (точніше, вибраному) напрямку, отримуємо статистичний варіант методу найскорішого спуску.

Схема реалізації наведеного алгоритму показана на рис. 4.17.

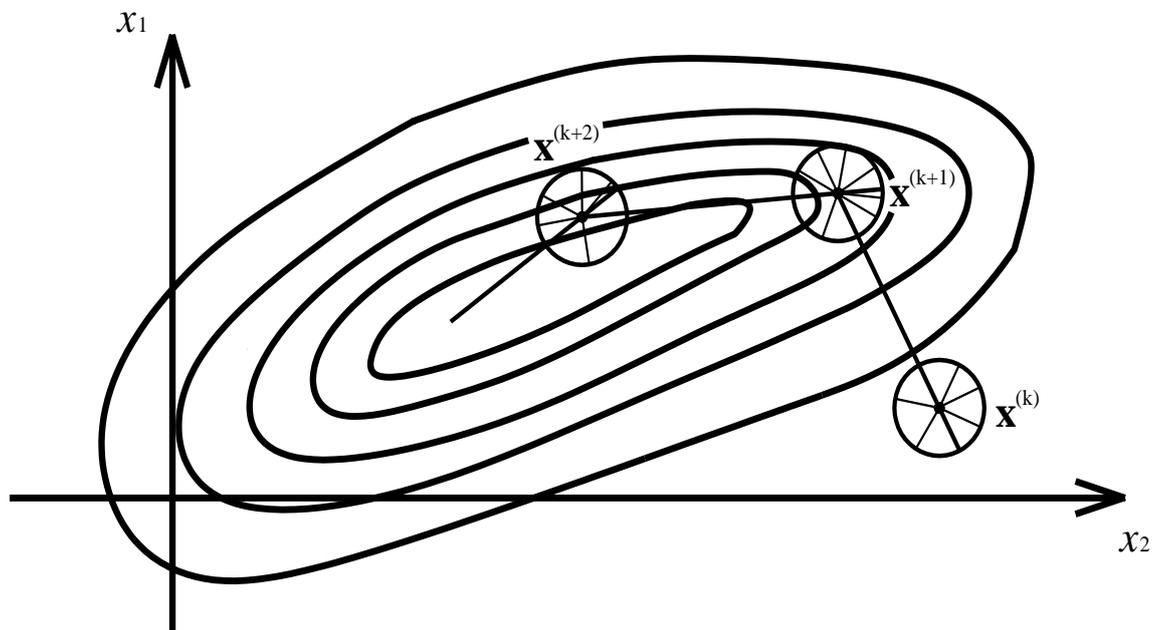


Рис. 4.17. Схема реалізації алгоритму статистичного градієнту (круги навколо точок $\mathbf{x}^{(k)}$, $\mathbf{x}^{(k+1)}$, $\mathbf{x}^{(k+2)}$ відображають геометричне місце кінцевих точок випадкових одиничних векторів)

Характерне відрізнення алгоритму статистичного градієнту від алгоритму найкращої проби полягає в суттєвому збільшенні кількості «пробних» кроків навколо точки поточного стану, що викликане необхідністю відшукування напрямку градієнта.

Суттєва перевага даного алгоритму перед відповідними детермінованими алгоритмами полягає у можливості прийняття рішення щодо напрямку робочого кроку при $m < n$.

Зазначимо, що за умови, коли $m = n$ та реалізуються не випадкові

ортогональні робочі кроки, направлені вздовж осей координат, наведений алгоритм вироджується в градієнтний метод.

4.5.5. Алгоритм найкращої проби з направляючим гіперквадратом

Згідно цього алгоритму на першому етапі його реалізації всередині області визначення (всередині допустимої області аргументів) функції $f(\mathbf{x})$ будується гіперквадрат (див. рис. 4.18).

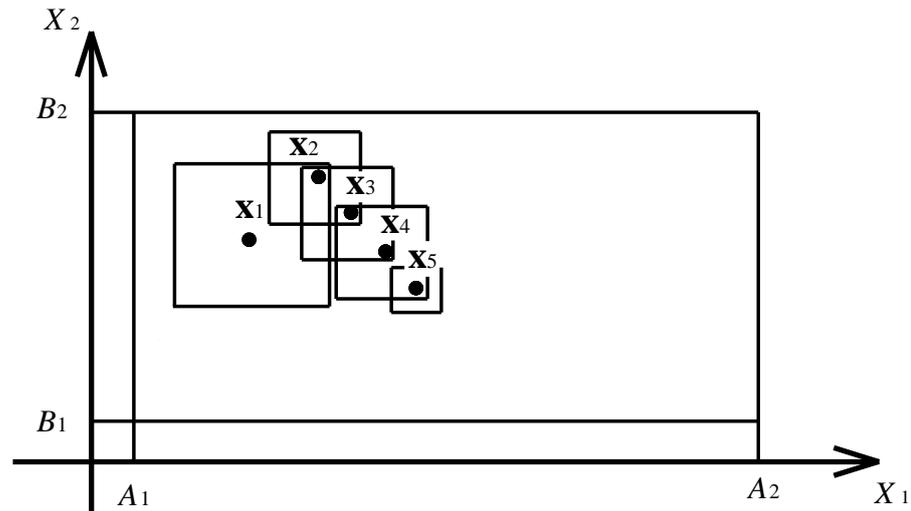


Рис. 4.18. Схема реалізації алгоритму найкращої проби з направляючим гіперквадратом

В побудованому гіперквадраті $A_1 A_2 B_2 B_1$, де

$$A_1 \leq x_1 \leq A_2; \quad B_1 \leq x_2 \leq B_2,$$

випадковим чином «розкидається» m точок $\mathbf{x}_1, \dots, \mathbf{x}_m$, в котрих обчислюються значення функції $f(\mathbf{x})$.

Серед побудованих за такий спосіб точок вибираємо найкращу (тобто точку, в якій значення функції мінімальне).

Спираючись на вибрану точку, будуємо новий гіперквадрат так, щоб його центр співпадав з вибраною точкою. Наведена процедура багаторазово повторюється.

Координати вершин гіперквадрата на $(k+1)$ -му кроці визначаються співвідношеннями

$$a_i^{(k+1)} = x_i^{(k+1)} - \frac{b_i^{(k)} - a_i^{(k)}}{2}; \quad b_i^{(k+1)} = x_i^{(k+1)} + \frac{b_i^{(k)} - a_i^{(k)}}{2},$$

де

$\mathbf{x}^{(k)}$ - найкраща точка в гіперквадраті, що побудований на k -му кроці.

Повторення наведеної послідовності дій приводить до направленої руху

гіперквадратна в сторону зменшення значення функції.

Можливі різні модифікації наведеного алгоритму.

Наприклад, сторони гіперквадрата можуть регулюватися згідно деякого правила, що неважко здійснити введенням відповідного параметра α . При цьому координати вершин гіперквадрата на k -му кроці визначатимуться як

$$a_i^{(k+1)} = x_i^{(k+1)} - \frac{b_i^{(k)} - a_i^{(k)}}{2\alpha}; \quad b_i^{(k+1)} = x_i^{(k+1)} + \frac{b_i^{(k)} - a_i^{(k)}}{2\alpha}.$$

Зазначимо, що вибір ефективного правила зазвичай приводить до суттєвого підвищення ефективності алгоритму.

В конкретних випадках замість направляючого гіперквадрату, звичайно, можна використовувати направляючі гіперсфери, гіперконуси і т.п.

4.6. КОМБІНОВАНІ АЛГОРИТМИ ГЛОБАЛЬНОГО ПОШУКУ

Комбінація алгоритмів випадкового та детермінованого пошуку набуває пріоритетного значення під час розв'язання багатоекстремальних задач. В загальному випадку розв'язання таких задач без елемента випадковості практично неможливе.

Розглянемо деякі загальні підходи до пошуку глобального екстремуму функції, аналізуючи чотири найуживаніші на сьогодні алгоритми.

Алгоритм I

Припустимо, що в допустимій області D значень аргументів функції випадковим чином вибрана точка $\mathbf{x}_1 \in D$ (див. рис. 4.19).

Прийнявши вказану точку за початкову та використовуючи деякий детерміністський метод або алгоритм направленої випадкового пошуку, здійснюється спуск в точку локального мінімуму $\mathbf{x}_1^* \in D$, в області тяжіння якого опинилася точка \mathbf{x}_1 .

Надалі вибирається нова випадкова точка $\mathbf{x}_2 \in D$. За тою ж схемою здійснюється спуск в точку локального мінімуму $\mathbf{x}_2^* \in D$ і т.д.

Пошук припиняється, коли деяке наперед задане число m раз не вдається знайти точку локального екстремуму зі значенням функції, меншим за попередні.

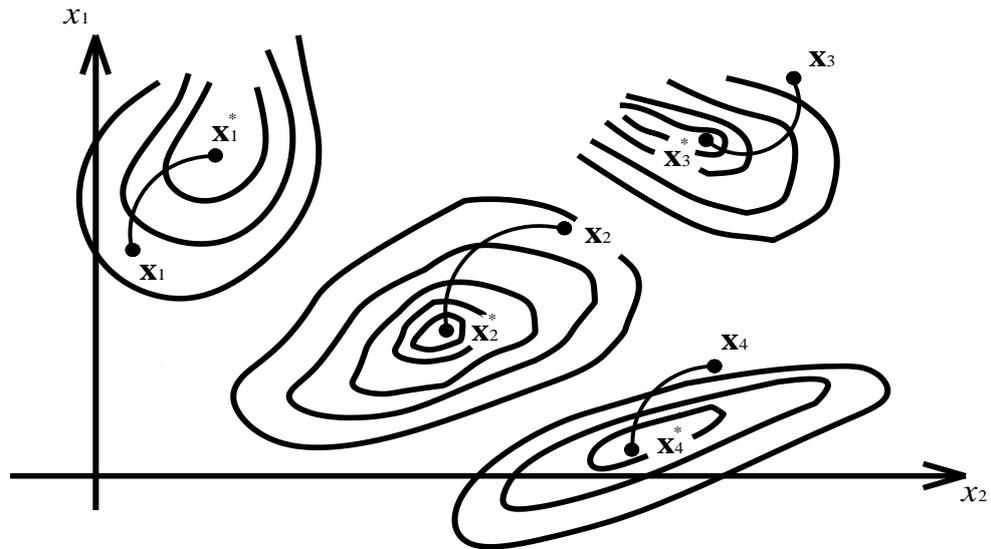


Рис. 4.19. Схема реалізації пошуку глобального екстремуму за Алгоритмом I

Алгоритм II

Припустимо, що на визначеному етапі мінімізації функції $f(\mathbf{x})$ в допустимій області D отримана деяка точка локального екстремуму $\mathbf{x}_1^* \in D$.

Переходимо до *ненаправленого випадкового* пошуку, який реалізується до отримання точки \mathbf{x}_2 такої, що

$$f(\mathbf{x}_2) < f(\mathbf{x}_1^*).$$

З точки \mathbf{x}_2 , використовуючи детерміністський алгоритм або направлений випадковий пошук, отримуємо точку локального екстремуму \mathbf{x}_2^* , в якій виконується нерівність

$$f(\mathbf{x}_2^*) < f(\mathbf{x}_1^*).$$

Далі за реалізацією випадкового процесу визначаємо нову точку \mathbf{x}_3 , для якої справедлива нерівність

$$f(\mathbf{x}_3) < f(\mathbf{x}_2^*),$$

та знову здійснюємо спуск в точку локального екстремуму \mathbf{x}_3^* , і т.д.

Пошук припиняється, якщо при генерації деякого граничного числа нових випадкових точок не вдається знайти кращої, за попередній локальний екстремум, який при цьому і приймається як розв'язок задачі.

Алгоритм III

Припустимо, що \mathbf{x}_1^0 - деяка початкова точка пошуку в допустимій області D аргументів функції (див. рис. 4.20). З цієї точки здійснюється спуск в точку локального екстремуму \mathbf{x}_1^* , де значення функції становить $f(\mathbf{x}_1^*)$.

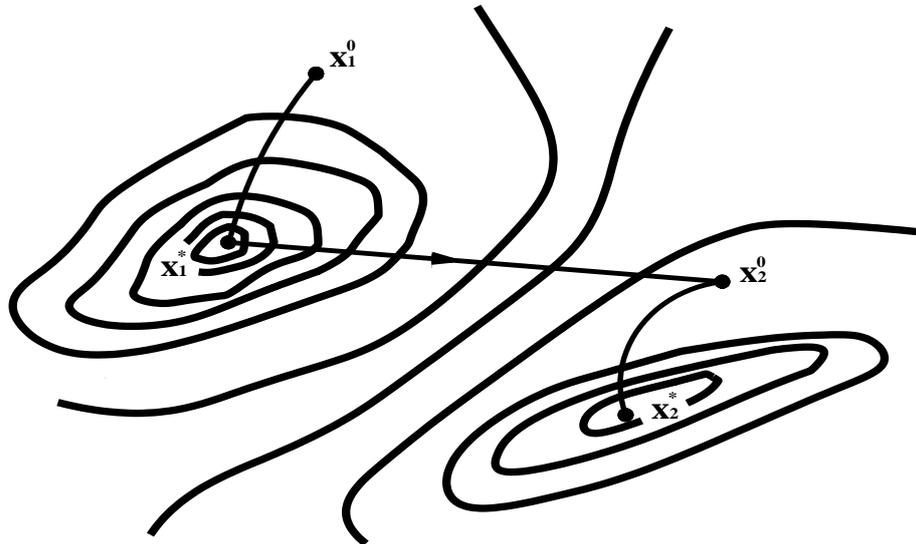


Рис. 4.20. Схема реалізації пошуку глобального екстремуму за Алгоритмом III

Далі з точки \mathbf{x}_1^* рухаємося або у випадково вибраному напрямку, або в напрямку $\mathbf{x}_1^* - \mathbf{x}_1^0$ до тих пір, поки функція знову не стане спадати (тобто поки не виходимо з області тяжіння \mathbf{x}_1^*).

Отримана за такий спосіб точка \mathbf{x}_2^0 приймається як початкова для наступного етапу спуску.

Як результат знаходимо новий локальний екстремум \mathbf{x}_2^* зі значенням функції $f(\mathbf{x}_2^*)$.

За умови

$$f(\mathbf{x}_2^*) < f(\mathbf{x}_1^*)$$

точка \mathbf{x}_1^* забувається й її місце займає точка \mathbf{x}_2^* .

За умови

$$f(\mathbf{x}_2^*) \geq f(\mathbf{x}_1^*)$$

повертаємося в точку \mathbf{x}_1^* та рухаємося з неї в новому випадковому напрямку.

Процес припиняється, якщо не вдається знайти кращий локальний мінімум після заданої кількості спроб або не вдається знайти “випадкового” напрямку, в якому функція знову починає спадати.

Зазначимо, що Алгоритм III дозволяє знайти глобальний екстремум у випадку багатозв’язної допустимої області.

Алгоритм IV

В допустимій області D розкидаємо m випадкових точок і вибираємо кращу з них, тобто ту, в якій значення функції мінімальне.

З вибраної за такий спосіб точки здійснюємо спуск до локального мінімуму. Надалі навкруги траєкторії спуску утворюємо *заборонену область* (див. рис. 4.21).

В цій забороненій області на наступних кроках реалізації алгоритму заборонена поява «нових» точок. Форма забороненої області, звичайно, може бути будь-якою, але зазвичай переваги віддаються простим геометричним фігурам (прямокутник, коло, еліпс і т.п.).

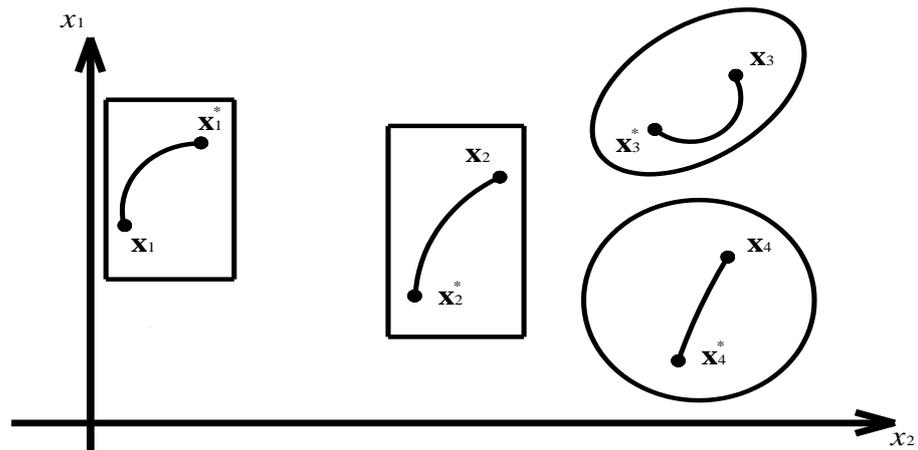


Рис. 4.21. Схема реалізації пошуку глобального екстремуму за Алгоритмом IV

В області, що залишилась «вільною» після реалізації вказаної процедури, випадковим чином розкидаємо нову сукупність випадкових точок. З кращої точки знову здійснюємо спуск в точку локального екстремуму. Навкруги нової траєкторії знову будуємо заборонену область і т.д.

Реалізація наведених процедур з виділенням заборонених областей поступово звужує допустиму область пошуку, а отже з кожним кроком ефективність пошуку (при незмінній кількості точок, які розкидаються на кожному кроці випадковим чином) підвищується.

Пошук припиняється за умови, що протягом наперед заданої кількості спроб не вдається знайти кращого локального екстремуму.

Комбінація випадкового пошуку з детермінованими методами використовується не лише для розв'язання багатоекстремальних задач. Часто таку комбінацію використовують в ситуаціях, коли використання виключно детермінованих методів стикається з тими чи іншими труднощами (зупинка алгоритму на дні вузького яру або в сідловій точці тощо). За таких умов нерідко крок у випадковому напрямку дозволяє подолати ситуації, які є тупиковими для детермінованих алгоритмів.

4.7. ГЕНЕТИЧНІ АЛГОРИТМИ

*Cujus est potentia, ejus est actum*³

Генетичні алгоритми (ГА) виникли за результатами спостереження й спроб копіювання природних процесів, що відбуваються у світі живих організмів, зокрема, еволюції й пов'язаної з нею селекцією (природним відбором) всередині популяції живих істот.

Звичайно, при співставленні технології ШНМ і ГА слід звернути увагу на принципово різну тривалість вказаних природних процесів, тобто на надто швидке оброблення інформації в нервовій системі і надто повільне протікання процесу природної еволюції. Однак при комп'ютерному моделюванні ці розрізнення виявляються несуттєвими.

4.7.1. Історія виникнення та основні визначення

Ідею ГА висловив Дж. Голланд (*J. Holland*) наприкінці шестидесятих – початку семидесятих років ХХ століття. Він зацікавився властивостями процесів природної еволюції (в тому числі тим фактом, що еволюціонують не самі живі істоти а їх *хромосоми* - структури, в яких зосереджена найбільша частина інформації, що унаслідкується). Дж. Голланд був впевнений у можливості скласти й реалізувати як комп'ютерну програму алгоритм, котрий буде розв'язувати складні задачі так, як це робить природа – шляхом еволюції. Саме тому він почав працювати над алгоритмами, що оперують послідовностями (обробляють послідовності) двійкових цифр (одиниць та нулів). Такі послідовності отримали назву *хромосом*.

Розроблені на підставі наведених засад алгоритми [31, 32] імітували еволюційні процеси в поколіннях хромосом. В них були реалізовані механізми *селекції* та *репродукції*, аналогічні тим, що використовуються при природній еволюції. Так же, як в природі, генетичні алгоритми здійснювали пошук "добрих" хромосом без використання якої б то ні було інформації про характер задачі, яка розв'язується. Необхідно було лише знайти оцінку кожної хромосоми, яка б відображала ступень *приспосованості* цієї хромосоми,

³ Діє той, хто сильніший (лат.)

тобто її привабливості з точки зору розв'язання поставленої задачі.

Механізм селекції полягає у виборі хромосом з найвищою оцінкою (тобто найбільш пристосованих), котрі репродукують частіше за особин з більш низькою оцінкою (тобто за особин, що гірше пристосовані).

Репродукція означає створення нових хромосом за результатом рекомбінації *генів* (структурна і функціональна одиниця спадкоємності) батьківських хромосом.

Рекомбінація – це процес, за результатами якого виникають нові комбінації генів.

Для реалізації процесу рекомбінації використовують дві операції:

- *схрещування*, яке дозволяє створити дві цілком нові хромосоми потомків шляхом комбінування генетичного матеріалу пари батьків;

- *мутації*, котра може викликати зміни в окремих хромосомах.

В теорії та практиці ГА, серед інших, використовуються терміни, що запозичені з генетики. Крім термінів «хромосома» та «гени», тлумачення яким надане вище, використовують також терміни:

- *популяція* - сукупність організмів одного виду, що тривалий час мешкають на одній території;

- *особина* - неподільна одиниця живої природи; найсуттєвіша ознака особини - строга взаємозалежність окремих частин: розділити особину на складові частини без втрати "індивідуальності" неможливо;

- *алелі* - різні форми одного й того ж гена, які розміщені на однакових ділянках, відрізках хромосом та визначають альтернативні варіанти розвитку однієї ознаки;

- *генотип* - сукупність всіх успадкованих властивостей особини;

- *фенотип* - сукупність всіх внутрішніх та зовнішніх ознак та властивостей особини, які сформувалися на базі генотипу в процесі її індивідуального розвитку; при відносно однакових генотипах у визначених межах можливі чисельні варіанти фенотипів;

- *локус* - фіксоване положення (локалізацію) на хромосомі, наприклад, положення гена;

- *гаплоїдні клітки* – живі клітки, які містять одинарний набір хромосом; кожна хромосома представлена в таких клітках однією копією, а кожен ген - одним алелем; у гаплоїдних клітках (гаплоїдах) проявляються всі мутації, що виникають;

- *нуклеотиди* – речовини (фосфорні ефіри нуклеозидів, нуклеозидфосфати), що відіграють важливу роль в енергетичних та інформаційних процесах всередині клітки.

Генетичні алгоритми відрізняються від традиційних методів оптимізації декількома базовими властивостями, а саме, ці алгоритми

1) обробляють не значення самих параметрів об'єктів, що є предметом даної задачі, а виключно відповідним чином закодовану форму цих параметрів;

2) здійснюють пошук розв'язку задачі на підставі не якоїсь окремої точки, а на підставі деякої сукупності цих точок (на підставі популяції точок);

3) використовують виключно значення цільової функції, і не використовують похідні цільової функції та іншу додаткову інформацію;

4) використовують імовірнісні, а не детерміновані правила вибору.

Перелічені чотири характерні властивості ГА, які стисло можна сформулювати як кодування параметрів, операції на популяціях, використання мінімуму інформації про задачу та рандомизація операцій приводять до стабільності ГА й до їх переваги над іншими широкоживаними технологіями.

Слід зазначити, що за допомогою ГА на сьогодні розв'язуються задачі, для котрих раніше використовувалися виключно ШНМ. За таких умов ГА розглядаються як незалежний від ШНМ альтернативний метод. В той же час, ГА широко використовуються сумісно зі ШНМ в рамках гібридних систем, призначених для розв'язання конкретної задачі.

4.7.2. Подання інформації про об'єкт в генетичних алгоритмах

Інформація щодо МО, який є предметом математичного дослідження, зазвичай подається у вигляді відповідних векторів, компоненти яких відображають значення тих чи інших його параметрів (характеристик). Вектори змінних, що відображають параметри МО в технології ГА записуються за аналогією з хромосомами як лінійна послідовність комбінацій нуклеотидів, а саме, у вигляді ланцюга символів двох, трьох або чотирьох літерного алфавіту. Привабливим є бінарне кодування (як найпростіше), яке найчастіше і використовується під час моделювання гаплоїдів.

За вказаних умов, кожна змінна x_i кодується фрагментом хромосоми, який складається з фіксованої кількості генів.

На рис. 4.22 показано відображення вектора змінних \mathbf{x} (що складається з п'яти компонентів) хромосомою \mathbf{X} . Кожна компонента вектора змінних відображається чотирма генами.

З рис. 4.22 видно, що всі локуси хромосом діалельні, тобто в будь-якій позиції фрагмента може стояти як нуль, так і одиниця.



Рис. 4.22. Приклад відображення вектора змінних \mathbf{x} хромосомою \mathbf{X}

Фрагменти, що у хромосомі розміщуються поруч, не відділяють один від одного якими би то ні було позначками. Однак при декодуванні хромосоми у

вектор змінних протягом усього процесу моделювання періоду еволюції використовується один і той же тип запису хромосоми – одна й та ж сама *маска картування*, яка визначає план розподілення інформації щодо спадкування за довжиною хромосоми.

Слід зазначити, що в технології ГА пряма операція кодування вектора змінних \mathbf{X} до хромосоми \mathbf{X} не використовується. Хромосоми генеруються випадковим чином, шляхом послідовного заповнення розрядів (генів), зразу в бінарному вигляді, і будь-які наступні змінення в популяції зачіпають спочатку генетичний рівень, а лише потому аналізуються фенотипові наслідки таких змін, але *ніколи не навпаки*.

Для декодування генетичної інформації з бінарної форми до десятинної, звичайно, можна використати будь-який двійково-десятковий код. Однак найчастіше використовується *код Грея*.

Характерна особливість коду Грея полягає в тому, що два сусідніх значення розрізняються лише в *одному розряді*.

Перетворення двійкового коду в код Грея здійснюється шляхом побітової операції "виключне АБО" з тим же числом, зсунутим вправо на один біт. Отже i -й біт коду Грея G_i подається через біти двійкового коду B_i та цього ж коду зсунутого на одиницю B_{i+1} як

$$G_i = B_i \oplus B_{i+1}.$$

В табл. 4.2, як приклад, наведена повна процедура декодування фрагменту хромосоми в проекцію x_i вектора змінних.

Процедура декодування виглядатиме наступним чином.

Від коду Грея переходимо до двійково-десятькового коду, а від нього – до натуральних цілих чисел. Відношення отриманого числа до максимального числа, доступного для кодування даною кількістю розрядів фрагмента (згідно табл. 4.2 – число 15) і дає шукане значення зсуву змінної відносно лівої границі a_i дозволеного діапазону її змінення, нормованого на ширину $b_i - a_i$ діапазону.

Таблиця 4.2

Код Грея	Двійково-десятковий код	Десятькове значення зсуву	Дійсне значення координати
0000	0000	0	a_i
0001	0001	1	$a_i + 1 \cdot (b_i - a_i)/15$
0011	0010	2	$a_i + 2 \cdot (b_i - a_i)/15$
0010	0011	3	$a_i + 3 \cdot (b_i - a_i)/15$

0110	0100	4	$a_i + 4 \cdot (b_i - a_i) / 15$
0111	0101	5	$a_i + 5 \cdot (b_i - a_i) / 15$
0101	0110	6	$a_i + 6 \cdot (b_i - a_i) / 15$
0100	0111	7	$a_i + 7 \cdot (b_i - a_i) / 15$
1100	1000	8	$a_i + 8 \cdot (b_i - a_i) / 15$
1101	1001	9	$a_i + 9 \cdot (b_i - a_i) / 15$
1111	1010	10	$a_i + 10 \cdot (b_i - a_i) / 15$
1110	1011	11	$a_i + 11 \cdot (b_i - a_i) / 15$
1010	1100	12	$a_i + 12 \cdot (b_i - a_i) / 15$
1011	1101	13	$a_i + 13 \cdot (b_i - a_i) / 15$
1001	1110	14	$a_i + 14 \cdot (b_i - a_i) / 15$
1000	1111	15	b_i

З табл. 4.2 видно, чому код Грея має явні переваги у порівнянні з двійково-десятковим кодом, адже останній породжує своєрідні тупики для пошукового процесу.

Як приклад, що підтверджує наведене, розглянемо будь-які три рядки табл. 4.2, які розміщені поруч один з одним. Вибираємо рядки, якими кодується зсув на 4, 5 та 6 одиниць.

Вважатимемо, що фрагменти хромосом, які розміщені в п'ятому рядку та число кодування 5, належать оптимальному вектору, який i є розв'язком поставленої задачі, а краща особина з поточної популяції містить фрагмент хромосоми з четвертого рядка. Така ситуація сприятлива для обох кодів. Достатньо виконати всього одну операцію – замінити в четвертому розряді фрагмента 0 на 1 – i розв'язок знайдено.

Цікавішим є випадок згідно якого краща особина містить фрагмент із шостого рядка. За умови використання коду Грея така ситуація не складніша за попередню – заміна 0 на 1 в третьому розряді знову веде до успіху. В той же час двійково-десятковий код примушує виконувати послідовно дві операції – замінити 1 на 0 в третьому розряді та 0 на 1 в четвертому. З якої б операції не розпочати, результат не наблизить нас до розв'язку задачі, адже перший варіант заміни перемістить нас в четвертий рядок, а другий – взагалі в сьомий.

З геометричної точки зору код Грея гарантує:

дві сусідні, що належать одному ребру, вершини гіперкубу α^L , на якому здійснюється пошук, завжди декодуються в дві найближчі точки простору дійсних чисел, які відстоять одна від одної на одну дискрету точності.

4.7.3. Генетичні оператори

Для того, щоб застосувати принципи успадковування та змінюваності до віртуальної популяції, необхідно ввести відповідні процедури - *генетичні оператори* [49, 53, 54] Ці оператори, серед іншого, характеризуються ймовірнісною властивістю. Тобто генетичні оператори не обов'язково мають застосовуватися до всіх особин, що схрещуються. Це, звичайно, вносить додатковий елемент невизначеності в процесі пошуку рішення. В даному контексті, невизначеність не трактується як негативний фактор, а є своєрідною «ступеню свободи» функціонування ГА.

До найпоширеніших та найнеобхідніших, для розв'язання практичних задач, генетичних операторів відносяться, перш за все, *кросинговер (кросовер)* та *мутація*. Серед інших, з точки зору частоти використання, можна виділити також оператор «*інверсія*». Інші оператори використовуються надто рідко, мають частковий, рецептурний характер.

Впливаючи з деякою імовірністю на генотипи батьківських особин, кожен генетичний оператор, з одної сторони, забезпечує передачу потомству життєво важливих ознак, а з другої – підтримує протягом еволюційно значимого періоду достатньо високий рівень його змінюваності. Відокремлення в потомстві нових, відрізнених від батьківських, фенотипічних ознак відкриває для популяції додаткові можливості адаптації, тобто сприяє збереженню її пошукової здатності.

Оператор кросинговеру

Оператор кросинговеру (кросоверу - crossover operator) є основним генетичним оператором, за рахунок якого відбувається обмін генетичним матеріалом між особинами. Тобто цей оператор моделює процес схрещування особин, а отже і появу потомків (нащадків).

Для ілюстрації функціонування цього оператора, розглянемо дві батьківські особини з хромосомами

$$\mathbf{X} = \{x_i, i = 1, 2, \dots, L\}, \quad \mathbf{Y} = \{y_i, i = 1, 2, \dots, L\}.$$

Випадковим чином визначимо точку всередині хромосоми, в якій обидві хромосоми діляться на дві частини та обмінюються цими частинами. Назвемо цю точку *точкою розриву*.

Зазначимо, що вказана точка в науковій літературі (особливо англомовній) іноді зветься *точкою кросинговеру (crossover point)*.

Описаний процес виділення точки розриву подано на рис.4. 23.

Реалізація операції кросинговеру приводить до того, що хромосома потомку включає два фрагмента, один з яких належав раніше одній хромосомі (наприклад, батьківській), а другий – іншій хромосомі (материнській).

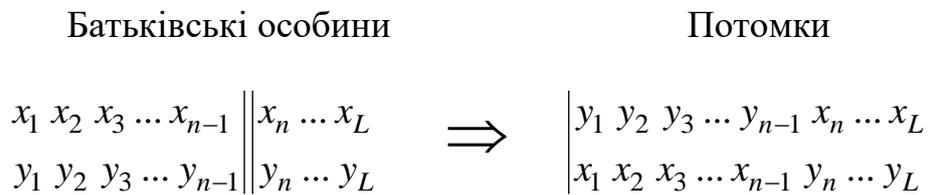


Рис. 4.23. Принцип функціонування оператора кросинговеру

Звернемо увагу, що дві батьківські особини можуть дати два потомки, хромосоми яких різняться послідовністю генів, набутих від батьківської та материнської хромосоми.

Саме завдяки операції кросинговеру (часто говорять - завдяки кросинговерному обміну) особини популяції обмінюються між собою генетичною інформацією.

Наведений тип кросинговеру зветься *одноточковим* (1-точковим, *single-point crossover*), бо при ньому батьківські хромосоми розрізаються лише в одній випадково визначеній точці.

Існують також *двохточковий* (2-точковий, *double-point crossover*) та *n-точковий* (*n-point crossover*) оператори кросинговеру.

В 2-х точковому кросинговері точок розриву дві, а *n*-точковий кросинговер є своєрідним узагальненням 1- та 2-точкового кросинговерів для ситуації, за якої $n > 2$.

Крім описаних типів кросинговеру використовують також *однорідний кросинговер* (*uniform crossover*), особливість якого полягає в тому, що значення кожного біта в хромосомі потомка визначається випадковим чином з відповідних бітів хромосом батьків.

Для реалізації такого оператора вводиться деяка величина p_0 так, щоб

$$0 < p_0 < 1,$$

і якщо генероване випадкове число більше за p_0 , то на *n*-у позицію першого потомка попадає *n*-й біт першого з батьків, а на *n*-у позицію другого - *n*-й біт другого з батьків. В протилежному випадку до першого потомка попадає біт другого з батьків, а до другого – першого з батьків. Така операція проводиться для всіх бітів хромосоми.

Саме завдяки наявності операцій кросинговеру особини популяції обмінюються між собою генетичною інформацією, тобто пошук набуває дійсно колективний характер.

Імовірність операції кросинговеру (схрещування) найвища серед генетичних операторів і зазвичай дорівнює 60 % і більше.

Оператор мутації

Оператор мутації (*mutation operator*) необхідний для "вибивання" популяції з локального екстремуму й сприяє захисту алгоритму від передчасного сходження.

Оператор мутації подібний точковим мутаціям у природі й інтерпретується як заміна існуючого окремого гена в хромосомі визначеної особини на протилежний (одиниці на нуль і навпаки).

Отже операція мутації реалізується інвертуванням випадково визначеного біту у хромосомі, як показано на рис. 4.24.

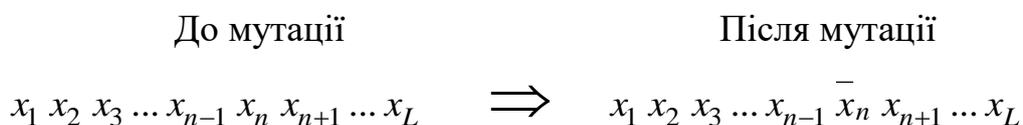


Рис. 4.24. Принцип реалізація операції мутації

Вочевидь, що в залежності від того в якому розряді фрагмента, яким кодується змінна, відбувається мутація, залежить величина відстані, яка відділяє потомка від батьків, при цьому, звичайно, йдеться не про відстань Хеммінга, а про простір дійсних чисел.

Як і кросинговер, мутація може проводитись не лише за одною випадковою точкою. Для інверсії можна вибирати декілька точок в хромосомі, причому кількість цих точок може бути випадковою. Крім того, можна інвертувати одразу деяку групу точок, що розміщені поруч.

Імовірність мутації значно менша за імовірність кросинговеру й рідко перевершує 1 %. Серед рекомендацій щодо вибору імовірності мутації пропонуються різні евристичні варіанти, наприклад, $1/L$ або $1/N$, де L - довжина хромосоми, N - розмір популяції.

Значна імовірність операції кросинговеру і мала імовірність операції мутації погоджується зі світом живих організмів, де схрещування поширене і набуває систематичного характеру, а мутації надто рідкі.

Оператор інверсії

Реалізація оператора інверсії (перестановки) веде до *порушення порядку* слідування фрагментів хромосом у потомка у порівнянні з батьківською хромосомою. Вплив цього оператора на хромосому ілюструється рис. 4.25.



Рис. 4.25. Реалізація операції інверсії
(гени, що піддавалися інверсії помічені темним кольором)

Зазначимо, що ні кросинговер, ні мутація не спираються під час генерування потомку на знання локального рельєфу поверхні цільової функції. В цьому сенсі ці операції можна вважати глобальними, на відміну від операції мутації, яку часто подають як засіб локального налаштування розв'язку екстремальної задачі.

4.7.4. Загальна схема реалізації генетичного алгоритму

Загальна схема реалізації ГА подана на рис. 4.26.

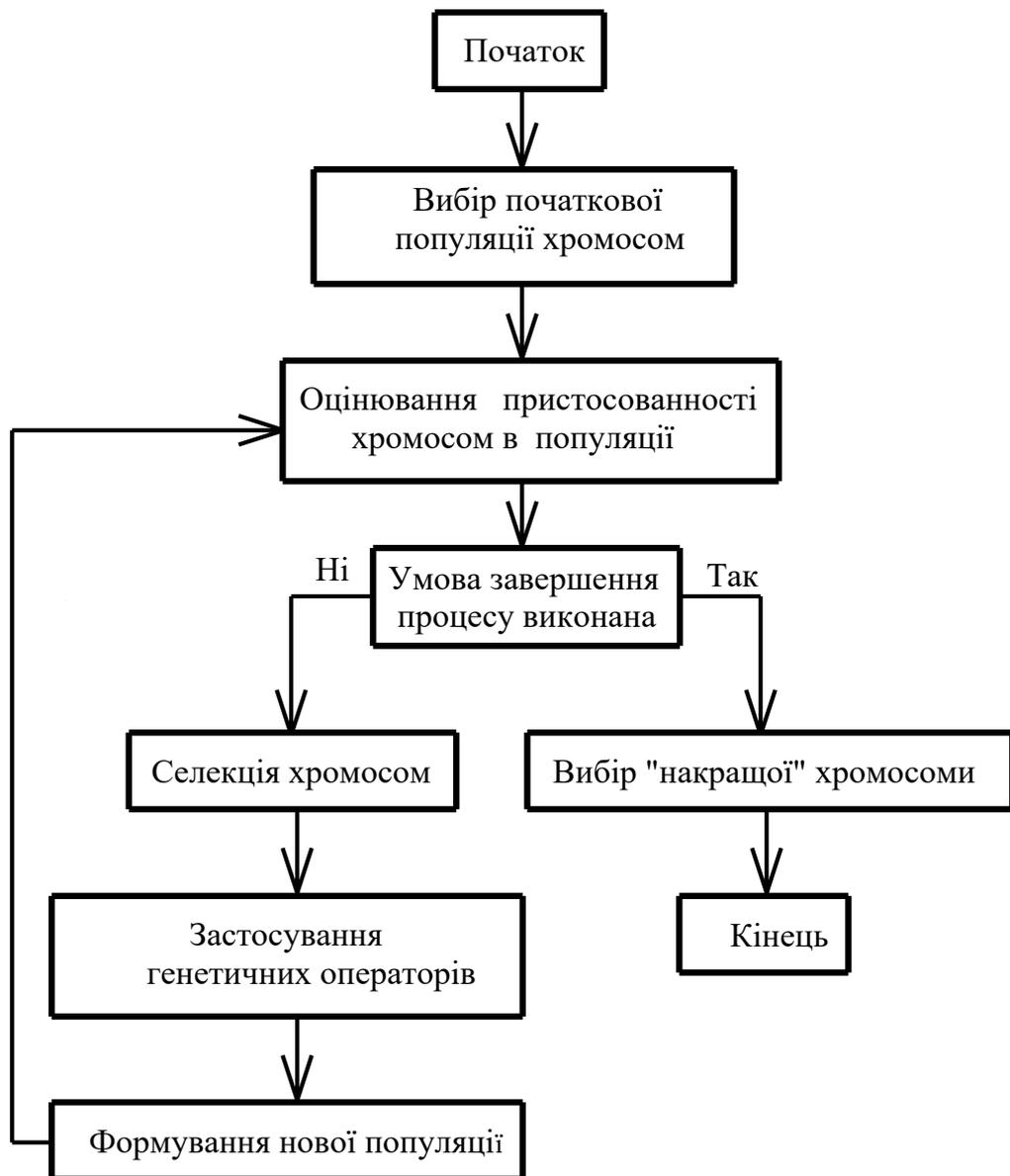


Рис. 4.26. Узагальнена схема реалізації генетичного алгоритму

Генетичний алгоритм в загальному вигляді складається з наступних процедур.

Вибір (формування) початкової популяції хромосом, та цільової функції (тобто функції пристосованості). Розрахунок значення функції пристосованості для всіх особин популяції – оцінка пристосованості хромосом в популяції.

Перевірка умов зупинки алгоритму.

Селекція хромосом – вибір тих хромосом (особин), які прийматимуть участь у створенні нащадків для наступної популяції.

Застосування генетичних операторів (кросинговеру, та за необхідністю мутації, інверсії тощо)

Розрахунок значення цільової функції для всіх утворених особин

Формування нового покоління особин (селекція)
Вибір «найкращої» хромосоми

Створення початкової популяції

Перед першим кроком реалізації ГА слід *випадковим чином* створити початкову популяцію; навіть якщо ця популяція виявиться повністю не конкурентоздатною, з великою імовірністю ГА все одно достатньо швидко переведе її у життєздатну популяцію.

Таким чином, на першому кроці можна не прикладати особливих зусиль, маючи на меті одразу зробити надто пристосованих особин, достатньо, щоб вони відповідали вибраному формату особин популяції, й надавали можливість розрахувати функцію пристосованості (*fitness function*).

Результатом реалізації першого кроку алгоритму є популяція **X**, що складається з N особин.

Розмноження

Для реалізації процедури розмноження в ГА зазвичай вибирається двоє батьків (це правило можна порушувати, але робиться таке надто рідко).

Розмноження в різних варіантах реалізації ГА визначається по різному, і, перш за все, залежить від форми подання даних. Головна вимога до процедури розмноження - щоб потомок або потомки мали можливість успадкувати риси обох батьків, «змішавши» ці риси якимось чином.

Зазначимо, що особини для розмноження зазвичай вибираються зі всієї популяції **X**, а не виключно з тих її особин, які вижили на попередньому кроці (хоча можливі й варіанти). Справа в тому, що один з головних недоліків ГА – «хронічна» нестача різноманіття (*diversity*) в особинах. Як правило, достатньо швидко виділяється один генотип, котрий являє собою локальний екстремуму, а тому всі особини популяції програють йому відбір, і вся популяція «забивається» копіями цієї особини. Існують різні способи боротьби з таким небажаним ефектом; один з них як раз і полягає в тому, щоб вибирати для розмноження не самих пристосованих, а взагалі всі особини.

Мутації

По відношенню до оператора мутації можна сказати все те ж саме, що й по відношенню до оператора розмноження.

В популяції є деяка доля мутантів m , яка є параметром генетичного алгоритму. На відповідному кроці алгоритму (на якому реалізується процедура мутації) слід вибрати mN особин, а потому змінити їх у відповідності до заздалегідь визначених операцій мутації. Зазначимо, що операція мутації не є надто поширеною у природі, а отже під час реалізації ГА слід встановлювати малу (зазвичай менше 1 %) імовірність мутації особин.

Відбір

На етапі відбору необхідно зі всієї популяції вибрати визначену частину особин, котрі залишаться «в живих» на цьому етапі еволюції. На сьогодні існують різні способи проведення процедури відбору.

Імовірність виживання особини, звичайно, має залежати від значення функції пристосованості (*Fitness function*). Частина s особин, що виживає, зазвичай є параметром ГА. Її просто задають заздалегідь. За результатами відбору з N особин популяції H мають залишитися sN особин, котрі увійдуть до підсумкової популяції H' . Інші особини гинуть.

Найпростіший пропорційний відбір зазвичай організовується за *принципом рулетки*.

Колесо рулетки розмічається так, що містить по одному сектору для кожної особини популяції. *Розмір кожного такого сектора робиться пропорційним відповідній величині функції пристосованості*. При такому відборі ті особини популяції, які характеризуються більшою пристосованістю (тобто ті, яким відповідає більше значення функції пристосованості), з більшою імовірністю вибиратимуться частіше, ніж особини з низькою пристосованістю.

Формування нової популяції

Особини, хромосоми яких отримані за результатами застосування генетичних операторів до хромосом тимчасової батьківської популяції, включаються до складу нової популяції. Ця нова популяція стає так званою поточною популяцією для даної ітерації ГА. На кожній черговій ітерації розраховується значення функції пристосованості для всіх хромосом (точніше, для всіх особин з наявними хромосомами) цієї популяції, після чого перевіряється умова зупинки алгоритму і або фіксується результат у вигляді хромосоми з найбільшим значенням функції пристосованості, або здійснюється перехід до наступного кроку ГА, тобто до селекції. В найпоширеніших варіантах реалізації ГА, як вже вказувалося, вся попередня популяція особин (з притаманними їм хромосомами) заміщується новою популяцією потомків (з притаманними їм хромосомами), яка має ту саму чисельність.

Розглянуті операції та процедури повторюються до тих пір, поки не почне задовольняти отриманий результат. Це відбувається, як правило, за умови досягнення заздалегідь вибраним критерієм, заздалегідь визначеного значення. Такою умовою може бути, наприклад досягнення кількості поколінь (циклів) заздалегідь вибраного максимуму.

4.7.5. Приклад використання генетичного алгоритму

Для повнішого уявлення базових засад реалізації ГА розглянемо приклад.

Припустимо, що задача полягає у визначенні аргументу, при якому досягає свого максимального значення функція

$$f(x) = 2x^2 + 1 \quad (4.19)$$

Вважатимемо, що змінна x може приймати значення виключно цілих чисел від 0 до 31.

Для застосування ГА необхідно, перш за все, подати значення змінної x у вигляді двійкових послідовностей, тобто закодувати значення цієї змінної.

Вочевидь, цілі числа з інтервалу від 0 до 31 можна подати послідовностями нулів і одиниць, використовуючи, наприклад, їх представлення у двійковій системі числення. Це не суперечить сказаному вище щодо переваг подання змінних у вигляді коду Грея. Адже форма запису генів у хромосомі особин може визначатися довільно, в тому числі, доречі, і в десятковій системі числення. *Головне дотримуватися вибраної форми запису хромосом протягом усього процесу реалізації алгоритму.*

За прийнятої форми подання змінної x , наприклад, число 0 записуватиметься як 00000, а число 31 – як 11111. Тобто в даному випадку хромосоми особин, загальною кількістю N , набувають вигляду двійкових послідовностей, що складаються з п'яти бітів, тобто ланцюжками, утвореними п'ятьма елементами.

Також очевидно, що в ролі функції пристосованості F виступатиме цільова функція, задана формулою (4.19).

За прийнятих умов пристосованість особини з хромосомою \mathbf{X}_i $i = \overline{1, N}$ визначатиметься значенням функції $f(x)$ для x , що дорівнюватиме фенотипу, який відповідає генотипу \mathbf{X}_i . Позначатимемо ці фенотипи як \mathbf{X}_i^* . В такому випадку значення функції пристосованості хромосоми \mathbf{X}_i , тобто $F(\mathbf{X}_i)$ дорівнюватиме $f(\mathbf{X}_i^*)$.

Приймаємо $N = 6$.

Вибираємо випадковим чином початкову популяцію, яка складатиметься (за припущенням) з 6 кодових послідовностей, тобто з 6 особин з відповідними хромосомами. Для цього можна, наприклад, 30 разів підкинути монету.

Таким чином отримуємо початкову популяцію особин з хромосомами:

$$\begin{aligned} \mathbf{X}_1 &= 10011, \quad \mathbf{X}_2 = 00011, \quad \mathbf{X}_3 = 00111, \\ \mathbf{X}_4 &= 10101, \quad \mathbf{X}_5 = 01000, \quad \mathbf{X}_6 = 11101. \end{aligned}$$

Наведеним хромосомам відповідають фенотипи (числа в десятковій системі, в якій саме і подана задача)

$$\mathbf{X}_1^* = 19, \quad \mathbf{X}_2^* = 3, \quad \mathbf{X}_3^* = 7,$$

$$\mathbf{X}_4^* = 21, \quad \mathbf{X}_5^* = 8, \quad \mathbf{X}_6^* = 29.$$

За формулою (4.19) розраховуємо значення функції пристосованості для кожної хромосоми (точніше для відповідної особини) в популяції:

$$F(\mathbf{X}_1) = 723, \quad F(\mathbf{X}_2) = 19, \quad F(\mathbf{X}_3) = 99,$$

$$F(\mathbf{X}_4) = 883, \quad F(\mathbf{X}_5) = 129, \quad F(\mathbf{X}_6) = 1683.$$

Селекцію хромосом проводитимемо за методом рулетки.

Колесо рулетки представлено на рис. 4.27.

Випадковим чином вибираємо 6 особин з відповідними хромосомами для репродукції.

Припустимо, що вибрані числа:

$$97, 26, 54, 13, 31, 88.$$

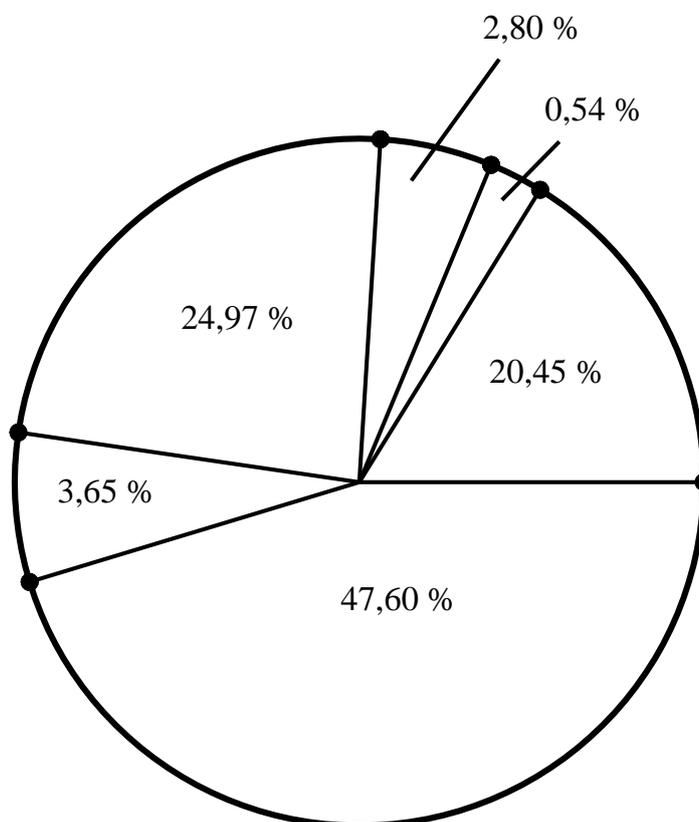


Рис. 4.27. Схема селекції хромосом за методом рулетки

Згідно з методом рулетки (див. рис. 4.27) це означатиме, що вибрані особини з хромосомами

$$\mathbf{X}_6, \mathbf{X}_4, \mathbf{X}_6, \mathbf{X}_1, \mathbf{X}_4, \mathbf{X}_6.$$

Припустимо, що схрещення відбувається з імовірністю $p_c = 1$.

Випадковим чином формуємо три пари особин для схрещення:

$$(\mathbf{X}_1 - \mathbf{X}_4), (\mathbf{X}_4 - \mathbf{X}_6), (\mathbf{X}_6 - \mathbf{X}_6).$$

Крім того, припустимо, що випадковим чином вибрані точки схрещення (точки реалізації оператора кросинговеру), а саме:

- для пари $(\mathbf{X}_1 - \mathbf{X}_4)$ - точка 3;
- для пар $(\mathbf{X}_4 - \mathbf{X}_6), (\mathbf{X}_6 - \mathbf{X}_6)$ - точка 2.

Схеми реалізації процедури схрещення (оператора кросинговеру) кожної з вказаних пар, відповідно, наведені на рис. 4.28, 4.29, 4.30.

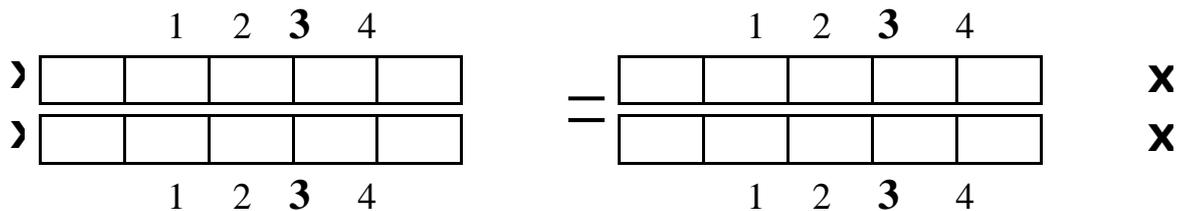


Рис. 4.28. Схеми реалізації схрещення пари $(\mathbf{X}_1 - \mathbf{X}_4)$

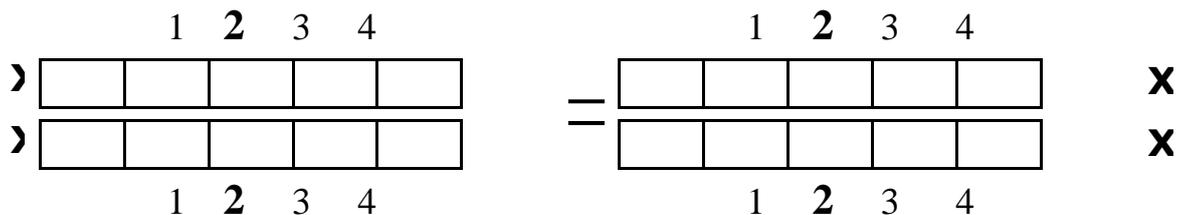


Рис. 4.29. Схеми реалізації схрещення пари $(\mathbf{X}_4 - \mathbf{X}_6)$

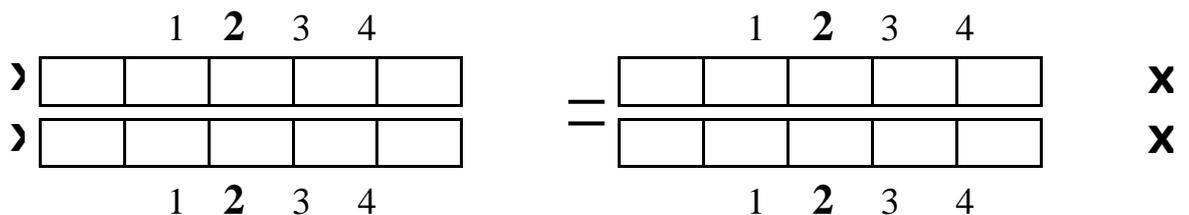


Рис. 4.30. Схеми реалізації схрещення пари $(\mathbf{X}_6 - \mathbf{X}_6)$

За умови, що імовірність мутації $p_m = 0$, в нову популяцію включаються особи з хромосомами:

$$\mathbf{X}_1^{(1)} = 10001, \quad \mathbf{X}_2^{(1)} = 10111, \quad \mathbf{X}_3^{(1)} = 10101,$$

$$\mathbf{X}_4^{(1)} = 11101, \quad \mathbf{X}_5^{(1)} = 11101, \quad \mathbf{X}_6^{(1)} = 11101.$$

Для розрахунку значень функції пристосованості особин з наведеними

хромосомами необхідно декодувати двійкові послідовності, якими ці хромосоми подані, та отримати фенотипи, що їм відповідають.

Позначатимемо ці фенотипи як $\mathbf{X}_i^{(1)*}$, $i = \overline{1, N}$.

За результатами декодування отримуємо:

$$\begin{aligned}\mathbf{X}_1^{(1)*} &= 17, \mathbf{X}_2^{(1)*} = 23, \mathbf{X}_3^{(1)*} = 21, \\ \mathbf{X}_4^{(1)*} &= 29, \mathbf{X}_5^{(1)*} = 29, \mathbf{X}_6^{(1)*} = 29.\end{aligned}$$

Відповідно, значення функції пристосованості особин нової популяції (з відповідними хромосомами), які розраховані за формулою (4.19), складатимуть:

$$\begin{aligned}F(\mathbf{X}_1^{(1)}) &= 579, F(\mathbf{X}_2^{(1)}) = 1059, F(\mathbf{X}_3^{(1)}) = 883, \\ F(\mathbf{X}_4^{(1)}) &= 1683, F(\mathbf{X}_5^{(1)}) = 1683, F(\mathbf{X}_6^{(1)}) = 1683.\end{aligned}$$

Як видно, середнє значення функції пристосованості особин отриманих за результатами схрещення зросло з 589 до 1262, що свідчить, про просування в напрямку екстремального значення.

Звернемо увагу, що за умови формування для схрещування на наступній ітерації пар

$$(\mathbf{X}_4^{(1)} - \mathbf{X}_2^{(1)}), (\mathbf{X}_2^{(1)} - \mathbf{X}_5^{(1)}), (\mathbf{X}_6^{(1)} - \mathbf{X}_2^{(1)})$$

з точкою схрещування 2 або 3, то серед інших буде отримана особина з хромосомою 11111, з фенотипом, який дорівнює 31, тобто саме тому числу при якому функція (4.19) досягає свого максимального значення. Значення функції пристосованості для особини з цією хромосомою виявляється найбільшим і складає 1923. Якщо таке сполучення пар в даній ітерації не відбудеться, то можна чекати утворення особини з хромосомою, яка має найбільше значення функції пристосованості, на наступних ітераціях.

Особина з хромосомою 11111 могла бути отримана й на поточній ітерації за умови формування для схрещування пари $(\mathbf{X}_1^{(1)} - \mathbf{X}_6^{(1)})$ з точкою схрещування 3.

Слід зазначити, що в малих популяціях часто зустрічаються ситуації, коли на початковому етапі декілька особин мають значно більші значення функції пристосованості, у порівнянні з іншими особинами даної популяції. Використання селекції на основі метода рулетки дозволяє в такому випадку надто швидко вибрати «найкращі» особини, іноді, навіть, на протязі «життя» одного покоління. Однак такий розвиток подій вважається небажаним, оскільки він стає головною причиною передчасного сходження ГА, як правило до неоптимального рішення. З цієї причини паралельно можуть

використовуватися й інші методи селекції, що відрізняються від методу рулетки, або здійснюється масштабування функції пристосованості.

4.7.6. Критичний погляд на генетичні алгоритми

Існує декілька приводів для критики щодо використання ГА у технології ШНМ (звичайно у порівнянні з іншими методами теорії оптимізації).

Багаторазова оцінка функції пристосованості для складних задач, часто є фактором, що обмежує використання ГА. Пошук оптимального розв'язку для складної задачі високої розмірності нерідко потребує надто великих обчислювальних та часових витрат для оцінювання функції пристосованості. В деяких реальних задачах, таких як, наприклад, задачі структурної оптимізації, одиничний запуск функціональної оцінки потребує на виконання необхідних розрахунків від декількох годин до декількох днів. Стандартні методи оптимізації, звичайно, не можуть подолати проблеми такого роду. В такому випадку, може бути доцільним (а подекуди і необхідним) нехтування точністю оцінювання функції пристосованості та використання відповідної процедури апроксимації, котра може бути вельми бажаною та ефективною.

Генетичні алгоритми погано масштабуються під складність задачі, що розв'язується. Це означає, що кількість елементів, які піддаються мутації може бути надто велика, якщо область пошуку розв'язку задачі надто багатовимірною. Це утруднює розв'язання таких задач, як, наприклад, проектування двигуна, будинку, літака та ін. Для того щоб використовувати еволюційні алгоритми для розв'язання подібних задач, останні мають бути розділені на простіші складові. Таким чином, ГА використовують, наприклад, при розробленні форми окремих складових двигуна, замість двигуна в цілому, форми фюзеляжу, замість розроблення усього літака одразу і т.д.

Слід зазначити, що за використанням ГА для розв'язку відповідної практичної задачі може бути більш привабливим лише у порівнянні з використанням інших методів. Як результат умова зупинки алгоритму при досягненні достатньої точності результату, або пошуку глобального розв'язку, для кожної задачі залишається рецептурною, часто будується на евристичних висновках.

В багатьох задачах ГА мають тенденцію сходиться не до глобального, а до локального оптимуму або навіть до «суперечних» точок. Це означає, що в ГА відсутні процедури, за допомогою яких можна було б сформувати критерій, згідно якого можна було б отримати інформацію про бажаність пожертвування короткотерміною високою пристосованістю чергового потомства для досягнення довготермінової (можна сказати глобальної) пристосованості потомства протягом визначеного періоду еволюції. Імовірність цього залежить від форми ландшафту функції пристосованості: окремі задачі можуть мати явне спрямування до глобального мінімуму, в той час як інші можуть вказувати напрямок для фітнес-функції, котрий приведе

лише до локального мінімуму. зазвичай така задача розв'язується використання декількох фітнес-функцій (на різних етапах еволюції), збільшенням імовірності мутацій, або використанням методів відбору, котрі підтримують різноманіття рішень в популяції, хоча, звичайно універсального розв'язку даної задачі не існує. Загальновизнаним методом підтримання популяційного різноманіття є встановлення обмеження за рівнем чисельності елементів з високою спорідненістю, котре знижує кількість представників схожих рішень в наступних поколіннях, дозволяючи іншим, менш схожим елементам залишатися в популяції. Однак дана процедура, може й не увінчатися успіхом, що визначається перш за все ландшафтом відповідної функції. Інший можливий метод, що часто використовується на практиці, полягає у простому заміщенні частини популяції випадково генерованими особинами (елементами), в момент, коли особини популяції стають надто схожі між собою.

Слід зазначити, що наведені недоліки ГА, в цілому не перешкоджають їх надалі більшому поширенню в алгоритмах навчання ШНМ найрізноманітнішої архітектури.

4.8. АЛГОРИТМ ІМІТАЦІЇ ПОВЕДІНКИ БДЖОЛИНОГО РОЮ

*Qui quaerit, reperit*⁴

Алгоритм імітації поведінки бджолиного рою (в англomовній технічній літературі цей алгоритм визначається як *Artificial Bee Colony Algorithm* або *Bees Algorithm*), як і генетичні алгоритми, відноситься до групи природних алгоритмів, тобто до алгоритмів, якими імітуються ті чи інші явища природного походження [31]. Для використання цього методу чисельної оптимізації не потрібне знання точного значення градієнта функції, мінімум якої розшукується. Алгоритм імітації поведінки бджолиного рою (алгоритм ПБР) був запропонований Кеннеді, Ебергартом и Ши та первісно призначався для імітації поведінки деякої спільноти (тобто для опису соціальної поведінки групи особин). Надалі алгоритм ПБР був спрощений, и була доведена його придатність для розв'язання задач пошуку екстремуму функції. Зазначимо, що в [31, 32] чимало уваги, серед іншого, приділяється опису різноманітних філософських аспектів алгоритму ПБР і, взагалі, так званого інтелекту рою, та проведено детальне дослідження можливості використання цього алгоритму у найрізноманітніших прикладних дослідженнях.

4.8.1. Базові засади алгоритму ПБР та основні визначення

В рамках алгоритму ПБР можливі варіанти розв'язку задачі пошуку екстремуму функції визначаються позицією (координатами) деяких

⁴ Хто шукає, той і знаходить (лат.)

артефактів, котрі зветься *агентами, частинками або бджолами*, а їх сукупність зазвичай називають *популяцією*.

Алгоритм ПБР розв'язує задачу пошуку, підтримуючи популяцію бджіл, та переміщуючи окремі бджоли цієї популяції в просторі розв'язків згідно деякої процедури. *Переміщення бджоли підкоряється принципу найкращого знайденого у відповідному просторі положення, котре постійно змінюється за умови, що бджола знайшла кращі положення.*

Принципи роботи алгоритму ПБР найпростіше і найефективніше розглядати на підставі його аналогії з реальним бджолиним роєм.

Уявимо рій бджіл на полі. Мета бджіл полягає у відшукуванні на полі області з найвищою щільністю (концентрацією) квітів. За умови відсутності будь-якого апіорного уявлення про поле (з точки зору розміщення квітів), бджоли починають пошук квітів з випадкових позицій та з випадковими (як за напрямком, так і за величиною) векторами швидкості.

Кожна бджола може пам'ятати ті позиції, де вона знайшла найбільшу кількість квітів, і якимось чином отримувати інформацію щодо областей, де найбільшу щільність квітів виявили інші бджоли. Вибираючи між поверненням до того місця, де бджола сама виявила найбільшу кількість квітів, або спрямуванням до того місця, яке визначене іншими бджолами, як найкращого з точки зору кількості квітів, бджола спрямовується в напрямку між двома точками в залежності від того, що більше вплине на її рішення - *персональний спомин* або *соціальний рефлекс*.

На шляху свого переміщення бджола може знайти місце з більш високою концентрацією квітів, ніж було виявлене (персонально нею або роєм) раніше. В подальшому це місце може бути позначене як нове місце з найбільшою концентрацією квітів, а також як місце найбільшої щільності квітів, що знайдене всім роєм.

Випадково бджола може пролетіти мимо місця, де концентрація (щільність) квітів більша, за ту, що була знайдена будь-якою іншою бджолою рою. Весь рій, за таких обставин, буде спрямовуватися в напрямку саме цього місця додатково до власних спостережень кожної бджоли.

Таким чином, досліджуючи поле, бджоли уповільнюють рух (зменшуючи величину вектора швидкості) в напрямку області з найбільшою концентрацією. Вони безперервно перевіряють (тестують) ті місця, які вже пролетіли, порівнюючи їх з тими місцями, що були знайдені раніше, сподіваючись знайти *абсолютну* найбільшу концентрацію квітів. В решті-решт бджола закінчує рух саме на ділянці (місці) поля з найвищою концентрацією квітів. Невдовзі й весь рій зосереджується навколо саме цієї *позиції*. Не маючи можливості виявити місця кращі за знайдене (тобто з більшою концентрацією квітів), бджоли безперервно рояться саме у вже знайденому районі, де концентрація квітів найвища, у порівнянні з усіма іншими «дослідженими» ділянками поля.

Описана поведінка бджіл і покладена в основу алгоритму ПБР.

Кожна бджола в рої розглядається як *частинка* або *агент*. Всі бджоли рою діють індивідуально у відповідності до одного *управляючого принципу*: прискорюватися в напрямку найкращої персональної та найкращої загальної позиції, постійно перевіряючи значення поточної позиції.

Положення (*позиція*) бджоли (частинки, агента) на полі подається координатами на площині x - y . Однак, в загальному випадку можна розширити це поняття на будь-який N -вимірний простір у відповідності з поставленою задачею. Кожен набір координат в цьому N -вимірному просторі є розв'язком або відповідним наближенням до розв'язку поставленої задачі.

За аналогією з поведінкою бджолиного рою в якості *функції придатності* приймається щільність (концентрація) квітів. При цьому, звичайно, чим більша концентрація квітів на виявленій ділянці, тим краща *позиція*. Але, враховуючи описану вище поведінку бджоли, слід розділити поняття позиції, принаймні на *дві* категорії.

Перша категорія – *персональна найкраща позиція*, яка відображає найкращу позицію (тобто відповідні координати), що відповідає найбільшій кількості квітів, виявлених бджолою самостійно. Кожна бджола рою має власну персональну найкращу позицію, яка визначається шляхом (розміщується на шляху), котрий ця бджола пролетіла. В кожній точці вздовж шляху руху бджола порівнює значення функції придатності поточної позиції зі значенням функції придатності, що відповідає персональній найкращій позиції. Координати персональної найкращої позиції замінюються на координати поточної позиції за умови, що остання має вище значення функції придатності.

Друга категорія - *глобальна найкраща позиція* – визначає область (координати) найбільшої концентрації квітів, яка виявлена всім роєм. Адже бджола має відповідний інструментарій, що дозволяє їй отримувати інформацію про обстеження поля іншими бджолами – роєм бджіл. Для усього рою бджіл глобальна найкраща позиція (тобто відповідні координати на полі) одна, і саме до цієї позиції спрямовується кожна бджола. В кожній точці протягом усього шляху свого переміщення кожна бджола порівнює значення функції придатності її власної поточної позиції з глобальною найкращою позицією. За умови, що якась із бджіл виявляє позицію з більшим значенням функції придатності, глобальна найкраща позиція замінюється на поточну позицію цієї бджоли.

4.8.2. Загальна схема реалізації алгоритму ПБР

Першим кроком в реалізації алгоритму ПБР є вибір параметрів, котрі необхідно оптимізувати (тобто які мають прийняти екстремального значення за результатами роботи алгоритму), та визначення допустимого інтервалу для пошуку цих оптимальних значень.

Надалі в допустимій області випадковим чином розміщуються бджоли,

кожній з яких надається вектор (напрямок та величина) швидкості руху. Під час реалізації алгоритму кожна бджола має переміщуватися крізь простір можливих розв'язків задачі так, нібито вона є однією з бджіл у рої. Алгоритм діє на кожну бджолу окремо, переміщуючи її кожного разу на невелику відстань від поточної позиції, циклічно рухаючи її крізь увесь рій.

Наступні кроки виконуються для кожної бджоли.

Оцінка придатності виявленої ділянки поля та порівняння з персональною найкращою позицією та з глобальною найкращою позицією.

Функція придатності, при обчисленні якої використовуються поточні координати бджоли в просторі розв'язків, дає можливість встановити чисельне значення, яке показує ступень придатності поточної позиції бджоли. Якщо це значення більше за значення персональної найкращої позиції даної бджоли, або глобальної найкращої позиції, то відповідні позиції замінюються поточною позицією.

Корегування швидкості бджоли.

Маніпуляції зі швидкістю бджоли є основним елементом всього процесу розв'язку задачі пошуку екстремуму функції. Точне усвідомлення рівняння, яке використовується для визначення швидкості, є ключем до усвідомлення всього процесу реалізації алгоритму ПБР.

Швидкість бджоли змінюється відповідно до взаємного розміщення персональної найкращої позиції та глобальної найкращої позиції. Бджола (а отже і вектор її швидкості) спрямовується в напрямку цих позицій найбільшої придатності згідно з рівнянням

$$v_n^{(i+1)} = w \cdot v_n^{(i)} + c_1 \cdot \text{rand}(\) \cdot (p_n - x_n) + c_2 \cdot \text{rand}(\) \cdot (g_n - \quad) \quad (4.20)$$

де

$v_n^{(i)}$ - швидкість бджоли вздовж n -тої координати на попередньому кроці;

x_n - n -та координата позиції бджоли;

p_n - n -та координата персональної найкращої позиції;

g_n - n -та координата глобальної найкращої позиції;

w - коефіцієнт масштабування швидкості бджоли на попередньому (i -му) кроці;

c_1, c_2 - масштабні коефіцієнти, котрі визначають відносне взаємне «тяжіння» до глобальної та персональної найкращої позиції.

Розрахунки за рівнянням (4.20), звичайно, проводяться для кожної n -тої координати позиції бджоли у багатовимірному (N -вимірному) просторі, в якому саме і відбувається процес розв'язку задачі.

З рівняння (4.20) видно, що «нова» (тобто на $(i+1)$ -му кроці) швидкість бджоли отримується з її «старої» (тобто на i -му кроці) швидкості шляхом

простого масштабування з коефіцієнтом w , та додаванням відповідної проекції векторів спрямованих у напрямку глобальної та персональної найкращої позиції.

Коефіцієнт w зазвичай називають «інерційною вагою», адже це число (яке вибирається в межах від 0 до 1) відображає в якій мірі бджола може залишитися на своєму поточному курсі, не піддаючись впливу персональної та глобальної найкращої позиції.

Коефіцієнти c_1, c_2 можна розглядати як пізнавальний та соціальний фактори, а саме:

c_1 - визначає як на поведінку бджоли впливає її власна пам'ять про персональну найкращу позицію;

c_2 - визначає як на поведінку бджоли впливають інші члени рою (через інформацію щодо глобальної найкращої позиції).

Збільшення c_1 передбачає дослідження простору розв'язків шляхом руху бджоли в напрямку своєї персональної найкращої позиції, тоді як збільшення c_2 передбачає дослідження простору розв'язків шляхом руху бджоли в напрямку глобальної найкращої позиції.

Функція $rand()$ відображає випадкове число, яке генерується (відповідно до вибраного закону) в інтервалі від -1 до 1. В загальному випадку кожна позначка $rand()$ відображає генерацію різних випадкових чисел. В рівнянні (4.20) таких позначок дві, а отже до складу цього рівняння входять два елемента генерації (в загальному випадку різних) випадкових чисел.

Більшість реалізацій алгоритму ПБР використовують дві незалежні випадкові величини (два незалежних випадково визначених коефіцієнти) для стохастичного змінення відносного тяжіння глобальної та персональної найкращої позиції.

Введення випадково визначених коефіцієнтів до процесу реалізації алгоритму ПБР призначене для моделювання незначного непередбачуваного компонента реальної поведінки бджолиного рою.

Граничні умови

Границі області, в якій відбувається реалізація алгоритму ПБР, були нами задані на початковому етапі реалізації алгоритму. Взагалі кажучи під час реалізації алгоритму ці границі мають тим чи іншим способом враховуватися, адже бджола може опинитися за межами вибраної для обстеження області. На сьогодні існують різні методи врахування границь, серед яких найпоширенішими є методи *поглинаючих*, *відбиваючих* та *невидимих* границь.

Згідно цих методів границі поля подаються як стіни, характер взаємодії яких із бджолами різний.

В методі поглинаючих границь вказані стіни вважаються повністю поглинаючими. Це означає, що швидкість бджоли у напрямку відповідної

координати має стати рівною нулю за умови, що ця бджола досягає (ударяється) границі поля (тобто границі простору розв'язків) в напрямку цієї координати. За таких умов бджола врешті-решт повернеться в заданий простір. Наведене означає, що границі-стіни поглинають енергію бджіл, які намагаються «пробити» дозволену область та вийти за її межі.

В методі відбиваючих границь стіни-границі відбивають бджолу, за умови, що вона в своєму переміщенні досягла «стінки». Це реалізується відповідною зміною вектора швидкості такої бджоли.

Метод «невидимих» границь на сьогодні вважається найефективнішим. Він полягає в тому, що будь-яка бджола може спокійно вилітати за межі границь, але, перебуваючи за межами дозволеної області, її інформація щодо концентрації квітів на полі (тобто відповідні значення функції придатності) не враховуються. Така ситуація зберігається до тих пір поки бджола не повернеться в дозволену область.

4.8.3. Варіант реалізації алгоритму ПБР

Алгоритм імітації поведінки бджолиного рою (як і будь-який інший алгоритм) має чимало варіантів та модифікацій узагальненої схеми. Ці варіанти та модифікації, що виникають під час та за результатами розв'язання конкретних задач, методологічно єдині, але різняться в окремих деталях.

Наведемо один з поширених на сьогодні варіантів реалізації алгоритму ПБР, схема якого дещо відрізняється від наведеної у попередньому параграфі.

Головне відрізнєння полягає в тому, що розрахунок швидкості бджоли (як довжина чергового кроку) не здійснюється, вводиться категорія *бджіл-розвідників*, а також виділяються категорії *найкращих* та *вибраних* ділянок у допустимій області аргументів функції придатності.

На першому кроці алгоритму в точки, з випадковими координатами, відправляється деяка наперед визначена кількість S *бджіл-розвідників*. В залежності від значення функції придатності (цільової функції), котре визначається згідно позиції (координат) бджоли, виділяються два види перспективних ділянок у дозволеній зоні, а конкретніше на поверхні рельєфу функції придатності. Поблизу саме цих ділянок за визначеними ознаками або на підставі евристичних даних вважається за можливе розміщення глобального максимуму. А саме:

- вибирається k *найкращих* ділянок, де значення функції придатності найбільше;
- вибирається t так званих *вибраних* ділянок, де значення функції придатності дещо менше, ніж на найкращих ділянках, але ці ділянки також є непоганими з точки зору значення функції придатності.

Звернемо увагу на одну особливість, котра може бути важлива при реалізації не тільки цього, але й будь-якого іншого варіанту алгоритму ПБР. Декілька бджіл можуть потрапити на одну й ту саму ділянку (розмір ділянки,

або допустима близькість бджіл, задається окремим параметром). За таких обставин можна виділити два варіанти подальшої реалізації алгоритму:

- вважати, що ці бджоли знайшли різні ділянки, які перехрещуються, а всі ці ділянки позначити як найкращі або вибрані;

- вважати, що знайдена одна ділянка, центр якої знаходиться в точці, яка відповідає позиції бджоли з найбільшим значенням функції придатності.

В реалізації, котра буде описана нижче, використовується (за випадковим вибором) другий варіант реалізації алгоритму.

В околицю k найкращих ділянок направляється K бджіл, а в околицю m вибраних ділянок направляється M бджіл, причому на кожную з найкращих ділянок має направлятися більше бджіл, ніж на кожную з вибраних ділянок. Можна зробити, наприклад, так, щоб при більшому значенні функції придатності більша кількість бджіл відправлялася на відповідну ділянку, а можна N та M зробити фіксованими величинами протягом усього процесу реалізації алгоритму.

Слід звернути увагу, що бджоли направляються не в точності на те місце, де бджоли-розвідники знайшли перспективні та найкращі місця, а в їх околицю. Крім того, область-околицю, в котру може бути відправлена бджола, можна зменшувати по мірі збільшення номера ітерації, щоб забезпечити поступове сходження до точки екстремуму. Однак, якщо вказану область зменшувати надто швидко, розв'язок може застрягти в точці локального екстремуму.

Після того як бджоли були відправлені на найкращі та вибрані ділянки, можна відправити тих самих бджіл-розвідників на інші випадкові точки (позиції).

Після всіх цих операцій знову знаходять k найкращих та m вибраних ділянок, на цей раз серед *всіх* бджіл рою, а не тільки серед бджіл-розвідників, і запам'ятовується найкраща точка (позиція), яка відповідає значенню функції придатності, більше якого до цього моменту не було знайдено. Це і буде проміжним розв'язком задачі.

Надалі описаний процес повторюється до тих пір поки не спрацює який-небудь з критеріїв зупинки.

Критеріїв зупинки може бути декілька.

Наприклад, якщо нам відоме значення функції придатності в глобальному екстремумі, то можемо повторювати процес реалізації алгоритму до тих пір, поки функція придатності не досягне деякого значення, близького до бажаного. Якщо значення функції в екстремумі невідоме, то можемо повторювати кроки алгоритму до тих пір, поки протягом якоїсь достатньо великої кількості ітерацій знайдено рішення практично перестане покращуватися.

4.8.4. Приклад використання алгоритму ПБР (реалізація ітерації процесу)

В наведемо нижче прикладі розв'язання задачі пошуку максимуму функції двох змінних використовуватимемо варіант алгоритму ПБР, наведений в п. 4.8.3.

Розглядатимемо процес пошуку максимуму функції

$$f(x, y) = -(x^2 + y^2) \quad (4.21)$$

Максимум цієї функції знаходиться в точці $(0, 0)$, причому $f(0, 0) = 0$

Етап ініціалізації полягає у визначенні всіх необхідних для реалізації алгоритму параметрів, а саме:

- кількість бджіл-розвідників: 10;
- кількість бджіл, що відправляються на *найкращі* ділянки: 5;
- кількість бджіл, що відправляються на *вибрані* ділянки: 2;
- кількість *кращих* ділянок: 2;
- кількість *вибраних* ділянок: 3;
- розмір області, що займається кожною ділянкою: 10.

Вважатимемо, що бджоли-розвідники потрапили на наступні ділянки-позиції (для зручності значення функції придатності для кожної позиції надані у порядку зменшення):

$$f(15, 18) = -549; \quad f(-30, 15) = -1125; \quad f(22, -31) = -1445;$$

$$f(18, 40) = -1924; \quad f(-25, 47) = -2834; \quad f(60, 86) = -10996;$$

$$f(-91, -99) = -18082; \quad f(17, -136) = -18785;$$

$$f(-152, -1) = -22501; \quad f(-222, 157) = -73933.$$

Спочатку, на підставі обчислених значень функції придатності (4.21), вибираємо 2 *найкращі* позиції (точки), а саме:

$$(15, 18) \Rightarrow f(15, 18) = -549;$$

$$(-30, -15) \Rightarrow f(-30, -15) = -1125.$$

Надалі вибираємо інші 3 перспективні позиції – *вибрані* позиції:

$$(22, -31) \Rightarrow f(22, -31) = -1445;$$

$$(18, 40) \Rightarrow f(18, 40) = -1924;$$

$$(-25, 47) \Rightarrow f(-25, 47) = -2834.$$

В околицю *найкращих* позицій-точок відправляємо по 5 бджіл.

Для першої найкращої точки значення координат, якими обмежується відповідна ділянка, визначаються як

інтервал вздовж першої координати

$$[15 - 10 = 5; 15 + 10 = 25];$$

інтервал вздовж другої координати

$$[18 - 10 = 8; 18 + 10 = 28];$$

Те саме робимо й для другої найкращої точки:

інтервал вздовж першої координати

$$[-30 - 10 = -40; -30 + 10 = -20];$$

інтервал вздовж другої координати

$$[-15 - 10 = -25; -15 + 10 = -5].$$

Аналогічно розраховуються інтервали (границі) для трьох *вибраних* ділянок, а саме:

для першої вибраної ділянки

інтервал вздовж першої координати $[12; 32]$;

інтервал вздовж другої координати $[-41; -21]$;

для другої вибраної ділянки

інтервал вздовж першої координати $[8; 28]$;

інтервал вздовж другої координати $[30; 50]$;

для третьої вибраної ділянки

інтервал вздовж першої координати $[-35; 15]$;

інтервал вздовж другої координати $[37; 57]$.

В даному випадку вздовж кожної координати розмір області однаковий і дорівнює 20 (тобто подвоєне значення, прийняте на етапі ініціалізації). В загальному випадку це, звичайно, не є обов'язковим.

На кожну з найкращих ділянок направляємо по 5 бджіл, а на кожну з вибраних - по 2 бджоли. При цьому *не будемо* змінювати положення бджіл, що знайшли найкращі та вибрані ділянки, адже інакше існує імовірність того, що на наступній ітерації максимальне значення функції придатності виявиться гіршим, ніж на попередньому кроці.

Припустимо тепер, що на першій ділянці в нас розміщуються бджоли, позиції яких відповідають наступним значенням функції придатності:

$$f(15, 18) = -549; \quad f(7, 12) = 193; \quad f(10, 10) = 100;$$

$$f(16, 24) = 832; f(18, 24) = 900.$$

Як видно, вже серед цих нових точок є кращі за попереднє рішення.

Таку саму процедуру реалізуємо і з другою найкращою ділянкою, а надалі аналогічно і з вибраними ділянками.

Серед всіх нових точок відмічаємо найкращі та вибрані й повторюємо розглянуті процедури.

Процес повторюється знову і знову до отримання очікуваних результатів.

Як і будь-який метод розв'язку задачі пошуку екстремуму функції, алгоритму імітації поведінки бджолиного рою притаманні визначені переваги та недоліки. Серед переваг, перш за все, слід виділити можливість ефективного розділення реалізації алгоритму ПБР на декілька паралельних процесів, за рахунок чого суттєво підвищується швидкість розв'язання відповідної екстремальної задачі. Також слід підкреслити, що алгоритм ПБР дозволяє достатньо легко визначити момент досягнення точки глобального мінімуму, в той час як в багатьох інших алгоритмах це суттєво ускладнене, а то й взагалі неможливе.

Зазначимо, що концепції алгоритму ПБР та ГА базуються на двох суттєво різних природних процесах: Алгоритм ПБР базується на соціальній поведінці рою (сукупності особин), а ГА імітує процес еволюції та природного відбору. Ця відмінність становить філософську та методологічну основу паралельного використання цих двох алгоритмів під час розв'язання задач пошуку глобального екстремуму функції в багатовимірному просторі.

4.9. АЛГОРИТМ ІМІТАЦІЇ ВІДПАЛУ МЕТАЛУ

Алгоритм імітації відпалу металу (англ. *Simulated annealing*) - загальний алгоритмічний метод розв'язання задачі глобальної оптимізації, який особливо прийнятний при розв'язанні задач в дискретному просторі. Цей алгоритм відомий також як алгоритм Метрополіса Н.К. (*N.C. Metropolis* - американський математик і фізик).

Алгоритм, являє собою математичний артефакт, що відповідним чином відображає (імітує) фізичний процес, який відбувається під час кристалізації речовини, в тому числі, наприклад, під час відпалу металів. При цьому вважається, що атоми вже вишукалися в кристалічну решітку, але ще допустимі переходи окремих атомів з одної комірки до другої. Вважається також, що процес відбувається за умови поступового зниження температури.

Припустимо, що метал нагрітий до високої температури. За таких умов атоми рухатимуться хаотично. Припустимо надалі, що метал швидко охолоджується. При цьому атоми застигатимуть у випадкових положеннях (тобто у комірках, до яких вони випадково потрапили).

Якщо ж навпаки, реалізувати процес повільного охолодження нагрітого металу, то атоми «намагатимуться» вишукатися деяким регулярним чином, який відповідатиме мінімуму енергії. Отже основою всього процесу відпалу металу є управління графіком зниження температури.

Перехід атома з одної комірки до другої відбувається з деякою імовірністю, причому ця імовірність зменшується з пониженням температури. Стала кристалічна решітка відповідає мінімуму енергії атомів. Отже атом або переходить до стану з меншим рівнем енергії, або залишається на місці.

Припустимо, що розшукується мінімум деякої числової функції

$$F(\mathbf{x}), \text{ де } \mathbf{x} = (x_1, \dots, x_m) \in \mathbf{X}.$$

Якщо вважати, що цією функцією описується запас енергії атомів кристалічної решітки, то моделювання наведеного вище процесу дозволяє знайти точку мінімуму цієї функції або множину точок, на якій цей мінімум досягається.

Цей процес можна подати у наступному вигляді.

Введемо послідовність точок $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$ простору \mathbf{X} .

Побудуємо алгоритм послідовного відшукування наступної точки на підставі даних поточної точки, починаючи з точки \mathbf{x}_0 , котра є початковим наближенням. Реалізація алгоритму має припинитися, за умови досягнення точки $\mathbf{x}_n \in \mathbf{X}$, в якій функція $F(\mathbf{x}_n)$ досягає мінімального значення.

Точку \mathbf{x}_{i+1} , згідно алгоритму, можна отримати на підставі реквізитів поточної точки \mathbf{x}_i наступним чином.

К реквізитам точки \mathbf{x}_i застосовується оператор, котрий випадковим чином модифікує її, за результатами чого отримується нова точка \mathbf{x}^* .

Точка \mathbf{x}^* стає точкою \mathbf{x}_{i+1} з імовірністю $P(\mathbf{x}^*, \mathbf{x}_{i+1})$, котра розраховується згідно розподілення Гібса, яке відображає стан ізотермічної макроскопічної термодинамічної системи частинок, що знаходяться в тепловій рівновазі з термостатом (оточуючим середовищем), а саме

$$P(\mathbf{x}^* \rightarrow \mathbf{x}_{i+1} | \mathbf{x}_i) = \begin{cases} 1, & F(\mathbf{x}^*) - F(\mathbf{x}_i) < 0 \\ \exp\left(-\frac{F(\mathbf{x}^*) - F(\mathbf{x}_i)}{Q_i}\right), & F(\mathbf{x}^*) - F(\mathbf{x}_i) \geq 0 \end{cases},$$

де

$Q_i > 0$ - елементи довільної спадаючої до нуля додатної послідовності, котра задається як аналог спадаючої температури кристалічної решітки

речовини (металу).

Підкреслимо, що швидкість та закон спадання температури можуть бути задані довільно, згідно досвіду та бажань дослідника.

Алгоритм імітації відпалу металу схожий на алгоритм градієнтного спуску, але за рахунок випадковості вибору проміжної точки характеризується суттєво меншою «схильністю» потрапляти до локальних мінімумів. Це обумовлене тим, що алгоритми, що базуються на градієнтному спуску, завжди намагаються просуватися до точки екстремуму, тоді як алгоритм імітації відпалу металу на початкових етапах реалізації часто «відходить» від точки екстремуму заради отримання надалі розширених можливостей щодо просування саме в сторону глобального екстремуму.

Слід підкреслити, що в загальному випадку алгоритм імітації відпалу металу не гарантує знаходження мінімуму функції, однак за умови правильного вибору методики генерації випадкової точки в просторі аргументів цільової функції, як правило, відбувається покращення початкового (попереднього) наближення.

КОНТРОЛЬНІ ПИТАННЯ

1. В чому полягає сутність розв'язання задачі оптимізації?
2. Як в загальному випадку формулюється задача оптимізації?
3. В чому полягає метод половинного ділення?
4. Яким чином реалізується алгоритм золотого перерізу?
5. Для яких функції використовується пошук екстремуму за методом половинного ділення?
6. Чи потрібне знання похідної функції при реалізації методу золотого перерізу?
7. В чому сутність градієнтного методу пошуку екстремуму?
8. Що характерне для детермінованих методів пошуку екстремуму?
9. Що характерне для імовірнісних методів пошуку екстремуму?
10. В чому суть методу Ньютона?
11. Які умови існування екстремуму вважаються необхідними?
12. Які умови існування екстремуму вважаються достатніми?
13. В чому особливість методу найшвидшого спуску?
14. Як визначається крок переміщення в методах оптимізації?
15. Як реалізується метод імовірнісного градієнту?
16. Як реалізуються методи проб?
17. Що таке градієнт функції?
18. Як визначити матрицю Гессе?
19. Як графічно подається багатовимірна функція в теорії оптимізації?
20. Чи можна комбінувати детерміновані та імовірнісні методи пошуку екстремуму функції?

-
21. Як реалізується алгоритм найкращої проби?
 22. Назвіть декілька можливих комбінацій використання детермінованих та імовірнісних методів для пошуку екстремуму?
 23. В чому полягає суть методу найкращої проби з направляючим гіперквадратом?
 24. В чому суть генетичних алгоритмів?
 25. Що таке хромосома?
 26. Як подається хромосома під час реалізації генетичних алгоритмів?
 27. Охарактеризуйте основні генетичні оператори?
 28. Охарактеризуйте оператор кросинговеру?
 29. Охарактеризуйте оператор мутації?
 30. Охарактеризуйте оператор інверсії?
 31. Як кодуються значення функції під час реалізації генетичних алгоритмів?
 32. Якими особливостями характеризується код Грея?
 33. Чим код Грея відрізняється від двійкового коду?
 34. Що таке метод рулетки?
 35. Що таке функція пристосованості?
 36. Як імітується процес схрещення під час реалізації генетичних алгоритмів?
 37. Як відображаються в генетичних алгоритмах властивості наслідування батьківських ознак?
 38. В чому полягає сутність алгоритму імітації поведінки бджолиного рою?
 39. В чому найсуттєвіші філософські розрізнення генетичних алгоритмів та алгоритмів імітації поведінки бджолиного рою?
 40. Як реалізується змінення аргументу функції за умови використання генетичних алгоритмів?
 41. Як реалізується змінення аргументу функції за умови використання алгоритму імітації поведінки бджолиного рою?
 42. Як «обслідуються» виділені ділянки в алгоритмі бджолиного рою?
 43. До класу детермінованих чи імовірнісних алгоритмів відносяться генетичні алгоритми?
 44. До класу детермінованих чи імовірнісних алгоритмів відносяться алгоритми імітації поведінки бджолиного рою?
 45. В чому полягає сутність алгоритму, що імітує відпал металу?