

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ**

**КАФЕДРА КОСМІЧНИХ СИСТЕМ КОМПЛЕКСІВ ТА  
СУПУТНИКОВИХ ТЕЛЕКОМУНІКАЦІЙ**

Проект

**Навчальний посібник  
ОСНОВИ ТЕОРІЇ СИСТЕМ**

Для студентів напрямку підготовки

8.05090102 Апаратура радіозв'язку, радіомовлення і телебачення,  
8.05090103 Радіоелектронні пристрої, системи та комплекси,  
8.05090302 Телекомунікаційні системи та мережі

Розробник: Ніколов М.О.

Київ – 2014

Основи теорії систем. Навчальний посібник з курсу «Загальна теорія систем»

Укладач: к.т.н., с.н.с. Ніколов М.О.

Рік видання - 2015

Категорія - Загальна теорія систем

Навчальний посібник присвячено оснивним темам занять по курсу «Загальна теорія систем. Розглядаються основні методи системного аналізу систем, шкали вимірювань та елементи експертного аналізу, базисні методи математичного аналізу нелінійних динамічних систем, елементи теорії бифуркації та кризи, фрактальної організації систем, теорії нечітких множин.

Для студентів та інженерів, що спеціалізуються у галузі зв'язку та телекомунікації.

## **ЗМІСТ**

### **ВСТУП**

#### **Розділ 1. ЗАГАЛЬНІ ВІДОМОСТІ ТЕОРІЇ СИСТЕМ І СИСТЕМНОГО АНАЛІЗУ**

- 1.1. Поняття «система»
  - 1.2. Загальні властивості і класифікація систем
  - 1.3. Інтеграція та диференціація наук. Системний підхід.
  - 1.4. Поняття «проблеми»
    - 1.4.1. Варіанти рішення проблем
    - 1.4.2. Типи ідеологій
    - 1.4.3. Чотири типи втручань
  - 1.5. Алгоритми системного аналізу
  - 1.6. Методи системного аналізу
  - 1.7. Шкали вимірів в експертному аналізі. Оцінка середніх величин
  - 1.8. Математичні методи аналізу експертних оцінок
- Питання для самоконтролю

#### **Розділ 2. МОДЕЛЮВАННЯ СИСТЕМ**

- 2.1. Властивості моделей
  - 2.2. Цілі моделювання
  - 2.3. Класифікація видів моделей систем
  - 2.4. Інші види моделювання
  - 2.5. Математичне моделювання
  - 2.6. Особливості системного моделювання
  - 2.7. Властивості системи з точки зору моделювання
  - 2.8. Етапи моделювання та загальні вимоги до моделей
  - 2.9. Якість моделювання
  - 2.10. Принцип подібності при математичному моделюванні
  - 2.11. Модель чорного ящика
  - 2.12. Труднощі побудови моделі чорного ящика
  - 2.13. Внутрішня неоднорідність систем
- Питання для самоконтролю

#### **Розділ 3. ТОПОЛОГІЧНІ МЕТОДИ І ГРАФІЧНІ РОЗВ'ЯЗКИ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ**

- 3.1. Інтегральні криві рівнянь другого порядку
  - 3.2. Класифікація особливих точок
  - 3.3. Метод ізоклін
  - 3.4. Приклади застосування методу ізоклін
- Питання для самоконтролю

## Розділ 4. БІФУРКАЦІЯ. КРИЗА. ХАОС

- 4.1. Перетворення системи
- 4.2. Кризи і загибель систем
- 4.3. Хаос і його роль у розвитку систем
- 4.4. Структурна стійкість і нестійкість функцій
- 4.5. Біфуркації стаціонарних станів
- 4.6. Біфуркація у простій дисипативній системі
- 4.7. Методи спрощення систем кінетичних рівнянь
- 4.8. Редукція систем и теория катастроф.

## Розділ 5. ФРАКТАЛИ

- 5.1. Поняття «фрактал»
  - 5.2. Класифікація фрак талів
    - 5.2.1. Геометричні фрак тали
    - 5.2.2. Алгебраїчні фрактали
    - 5.2.3. Стохастичні фрактали
  - 5.3. Самоподібність і фрактальна розмірність
  - 5.4. Довжина берегової лінії. Периметри фракталів
  - 5.5. «Гра в хаос»
  - 5.7. Фрактали та самоорганізація в природі
- Завдання для самоконтроля

## Розділ 6. ЕЛЕМЕНТИ ТЕОРІЇ НЕЧІТКИХ МНОЖИН

- 6.1. Два основні підходи до формалізації нечіткості
- 6.2. Мережі нечітких елементів
- 6.3. Теорія нечітких підмножин і теорія структурних функцій
- 6.4. Ступінь нечіткості
- 6.5. Нечіткі і можливі випадкові події

## Розділ 7. ПРИКЛАДНІ АСПЕКТИ

- 7.1. Чотириполосник, як приклад моделі чорного ящика
- 7.2. Фрактальні антени.

## ВСТУП

Система – грецьке слово, буквально означає ціле, складене з частин. В іншому значенні – порядок, визначений правильним розташуванням частин та їх взаємозв'язками. В даний час термін "система" відноситься до найбільш вживаних. Це пояснюється тим, що за ним стоїть розвинена методологічна традиція, яка характеризує склався протягом всієї інтелектуальної історії людства, і особливо в останні десятиліття, дуже ефективний стиль мислення. Системне мислення – це мислення сучасної людини. Що воно означає? Поглиблену відповідь на це запитання можна отримати, прочитавши цю книгу.

Якщо ж відповідати узагальнено, то системний стиль мислення, або системний підхід являє собою специфічний зміст, аспект, принцип мислення, при якому категорія "система" застосовується в якості методу, інструменту пізнання.

Термін " системний підхід " змістовно відображає групу методів, за допомогою яких реальний об'єкт описується як сукупність взаємодіючих компонентів. Ці методи розвиваються в рамках окремих наукових дисциплін і загальнонаукових концепцій, є результатом їх міждисциплінарного синтезу.

Використання системного підходу в науці стимулюється також успіхом приватних системних теорій в інших галузях знань, розвитком кібернетики і суспільних наук.

Системний підхід – ефективний спосіб розумової діяльності, що забезпечив значні відкриття в науці, винаходи в техніці і досягнення у виробництві в другій половині ХХ ст. Це зумовлює постійну увагу до нього з боку інтелектуалів. Без володіння цим методом неможливі творча самореалізація, професійна діяльність. Разом з тим зростаюча потреба в системному мисленні вимагає спеціального вивчення у вищих навчальних закладах дисципліни "Теорія систем і системний аналіз". Однак сама теорія систем недостатньо систематизована і підготовлена для викладання.

Даний навчальний посібник являє собою один з можливих варіантів викладу курсу "Теорія систем і системний аналіз". На нашу думку, в країнах СНД спостерігається серйозна криза освіти, яка полягає в тому, що вона орієнтоване на знання, вміння та навички, але не на розуміння. Схема хороша для підготовки промислового робітника, водія, льотчика, але малоефективна в навчанні інженера, менеджера, економіста, соціолога тощо, оскільки в цих професіях важливі не просто знання, а їх системність і розуміння. Швидше за все тут буде ефективна формула – знання, розуміння, вміння. Вона акцентує увагу не на тому, щоб знати, а щоб розуміти і розуміється застосовувати.

Сучасне ж навчання нагадує ходіння навколо предмета і нескінченний його опис. Від цього розуміння предмета змінюється дуже повільно, а знання залишається поверховим. І життєдіяльність студента перетворюється або на тривалий відпочинок, а з ним і в уповільнений інтелектуальний розвиток, або в каторгу запам'ятовування, на яку штовхає зростання професійної

інформації. У підручниках дуже мало уваги приділяється методам науки, особливостям її парадигми, домінує орієнтація на описові знання. Звідси в освіті постійно загострюється потреба в освоєнні студентами методів. Пропонована книга являє собою спробу зменшення методологічного дефіциту за допомогою докладного викладу тільки одного методу – системного підходу в якості головного досягнення методології.

Без системного підходу не обходиться нині жодна сфера високопрофесійної діяльності. Можна з упевненістю констатувати, що багато помилок в управлінні державою викликані тим, що державні службовці та службовці місцевого самоврядування не володіють ні теорією систем, ні системним аналізом.

Важливі рішення приймаються нерідко за принципом підкинутої монети, без бачення їх впливу на різні підсистеми складного і взаємозалежного суспільного організму. Економіка і її найважливіші складові бізнес і фінанси відрізняються незначним інноваційним тонутом, який стримується самим персоналом. Менеджери, керівники фірм, директори підприємств, фінансисти практично не знайомі з принципами управління складними системами, що саморозвиваються. Завдання, які ставить перед ними життя, не вирішуються тільки тому, що вони не можуть зрозуміти їх і сформулювати в системних категоріях. Трагічні наслідки природних, екологічних і техногенних катастроф значною мірою обумовлені не просто нерозумінням системності, а нездатністю втілити ідеї в такі дії, які не порушували б системні закони природи і суспільства.

Об'єктом осмислення в навчальному посібнику виступають системи, а предметом – основні ідеї теорії систем і системного аналізу. Світ систем вивчається людьми з глибокої давнини. Елементи системного світогляду виникли вже в античному світі. Протягом всієї історії розвитку науки, освіти та й культури загалом, людство накопичило величезне багатство системних ідей, які розкидані в різних сферах наукового знання. Це знання потребує переосмислення та інтеграції.

Володіння системним аналізом, системним моделюванням і конструюванням, системної практичної діяльності – вища характеристика розумової культури людини. Важливо, що будь-якому фахівцеві доводиться "мати справу" із систематизацією інформації, системними дослідженнями, які можна здійснювати тільки володіючи спеціальними знаннями та навичками.

Тому мета посібники зводиться не тільки до того, щоб методично представити вже готове знання про системи. Вона полягає в тому, щоб виділити всі аспекти системності, осмислити тенденції її розвитку, інтегрувати різні аспекти системного знання, які розкидані в наукових джерелах, а також описати ті положення теорії систем, які ще не отримали розвиток.

## Розділ 1.

### Загальні відомості теорії систем і системного аналізу

Теорія систем являє собою наукову дисципліну, котра вивчає різні явища, відволікаючись від їх конкретної природи, і ґрунтується лише на формальних взаємозв'язках між різними складовими їх чинниками і на характері їх змін під впливом зовнішніх умов. При цьому результати всіх спостережень пояснюються лише взаємодією їх компонентів, наприклад характером їх організації та функціонування, а не за допомогою безпосереднього звернення до природи залучених до явищ механізмів (будь вони фізичними, біологічними, соціологічними або чисто концептуальними).

#### 1.1. Поняття «система»

Система (від ст.-грец. – ціле, складене із частин; з'єднання) – множна елементів, що знаходяться у відносинах і зв'язках один з одним, яке утворює певну цілісність, єдність.

Згідно І. Канту, наукове знання є система, в якій ціле панує над частинами.

З одного боку термін «система» використовується для визначення **об'єктивно** існуючих феноменів, а з іншого боку – як метод вивчення та представлення феноменів, тобто як **суб'єктивна** модель реальності.

Серед прикладів визначень системи можна виділити:

- Система – комплекс взаємодіючих компонентів.
- Система – сукупність елементів, що перебувають у певних відносинах один з одним і з середовищем.
- Система – безліч взаємопов'язаних елементів, відокремлений від середовища і взаємодіє з нею, як ціле.
- Приклади конструктивних визначень:
- Система – комбінація взаємодіючих елементів, організованих для досягнення однієї або декількох поставлених цілей.
- Система – кінцеве безліч функціональних елементів і відносин між ними, виділене з середовища відповідно з певною метою в рамках певного часового інтервалу.
- Система – відображення в свідомості суб'єкта (дослідника, спостерігача) властивостей об'єктів і їх відносин у вирішенні завдання дослідження, пізнання.
- Система  $S$  на об'єкті  $A$  щодо інтегративної властивості (якості) є сукупність таких елементів, що знаходяться в таких відносинах, які породжують дану інтегративну властивість.

Повноцінного визначення поняття «система», ймовірно, немає тому, що до цих пір недооцінювалася роль поняття «мета». Будь-які властивості систем, в кінцевому підсумку, пов'язані з поняттям мети, тому що будь-яка

система відрізняється від інших систем постійністю своїх дій. А її прагнення зберегти цю постійність є відмітною якістю будь-якої системи – її метою.

Отже, мета визначає систему.

## 1.2 Загальні властивості і класифікація систем

До основних найбільш загальних властивостей систем можна віднести цілісність, синергетичність і ієрархічність.

**Цілісність** системи є абстрактна сутність та характеризує властивість об'єкту як сукупність його складових, організованих у відповідності з визначеними принципами. Цілісність системи – під цим розуміється, що в деякому істотному аспекті «сила» або «цінність» зв'язків елементів усередині системи вище, ніж сила або цінність зв'язків елементів системи з елементами зовнішніх систем або середовища.

**Синергічність, емерджентність** – поява у системі властивостей, не властивих елементам системи; принципова незвідність властивостей системи до суми властивостей складових її компонентів (неаддитивність). Можливості системи перевершують суму можливостей складових її частин; загальна продуктивність або функціональність системи краще, ніж у простій сумі елементів.

**Ієрархічність** – кожен компонент системи може розглядатися як система; сама система також може розглядатися як елемент деякої над системи.

При категоріальній класифікації системи поділяються за загальними характеристиками, властивими будь-яким системам, незалежно від їх матеріального втілення. Найбільш часто розглядаються наступні категоріальні характеристики:

- Кількісно всі компоненти систем можуть характеризуватися як монокомпоненти (один елемент, одне відношення) і полікомпоненти (багато властивостей, багато елементів, багато відносин).

- Для статичної системи характерно те, що вона знаходиться у стані відносного спокою, її стан з плином часу залишається постійним. Динамічна система змінює свій стан у часі.

- Відкриті системи постійно обмінюються речовиною, енергією або інформацією з середовищем. Система замкнута (замкнута), якщо до неї не надходять і з неї не виділяються речовина, енергія або інформація. У даному випадку ми об'єднали закриті та ізольовані системи під одним поняттям.

- Поведінка детермінованих систем цілком зрозуміла і передбачувана на основі інформації про їх стан. Поведінка ймовірностних систем визначається цією інформацією не повністю, дозволяючи лише говорити про ймовірність переходу системи в той чи інший стан.

- За походженням виділяють штучні, природні і змішані системи.



- За ступенем організованості виділяють клас добре організованих, клас погано організованих (дифузних) систем і клас систем які розвиваються (самоорганізованих).

- При розподілі систем на прості і складні спостерігається найбільші розбіжності точок зору, однак найчастіше складності системи надають такі характеристики як велика кількість елементів, різноманіття можливих форм їх зв'язку, множинність цілей, різноманіття природи елементів, мінливість складу і структури і т. д.

### 1.3 Інтеграція та диференціація наук. Системний підхід.

В основі загальної теорії систем, яка є найбільш високим рівнем організації науки і пізнання, лежить інтеграція і диференціація наук. Від лат. **differentia** – різниця, відмінність, **integratio** – відновлення, заповнення. Це два протилежних, але взаємопов'язаних процеси розвитку наукового знання. Диференціація – більш глибоке і ретельне дослідження окремих явищ і процесів певній галузі дійсності на певній стадії еволюції науки. Саме в результаті такого дослідження з'являються окремі наукові дисципліни зі своїм предметом і специфічними методами пізнання. Інтеграція наукового знання здійснюється в різних формах і зачіпає як емпіричні, так і фундаментальні теоретичні закони. Нерідко Інтеграція починається з застосування понять і методів однієї науки в іншій і завершується з побудовою широких міждисциплінарних теорій і напрямів дослідження. Зрозуміло, теорія систем прагне інтегрувати диференціальні наукові знання і методи в один інструмент, методологію, світогляд. Великий внесок у розуміння та значення диференціації та інтеграції наукових знань внесли К. Маркс і Ф. Енгельс.

Однак, теорія систем, намагаючись об'єднати розрізнені знання і методи різних наук, виробляє свої методи та індивідуальні особливості, диференціюється. Найбільш загальним інструментом ОТС є системний аналіз.

«Системний аналіз (системний підхід) – сукупність методологічних засобів, використовуваних для підготовки і обґрунтування рішень по складним проблемам політичного, військового, соціального, економічного, наукового, технічного характеру».

«... Найважливіші принципи системного аналізу зводяться до наступного: процес прийняття рішень повинен починатися з виявлення і чіткого формулювання кінцевих цілей; необхідно розглядати всю проблему як ціле, як єдину систему і виявляти всі наслідки і взаємозв'язки кожного приватного рішення; необхідні виявлення та аналіз можливих альтернативних шляхів досягнення мети; цілі окремих підрозділів не повинні вступати в конфлікт з цілями всієї програми...»

### 1.4. Поняття «проблеми»

Мета в пізнанні визначається поняттям «проблема». Саме для вирішення тієї чи іншої проблеми задаються цілі і відповідно до них вводиться конкретне розуміння і визначення системи.

Проблема – це суб'єктивне негативне ставлення суб'єкта до реальності.

Хоча конкретні системи та їх оточення об'єктивні за характером, вони в той же час і суб'єктивні, оскільки конфігурація утворюючих її елементів вибирається відповідно з інтересами дослідника. Різні спостерігачі одного і того ж явища можуть відобразити його в різних системах і оточеннях. Для одного дослідника системою може бути сам об'єкт, а для іншого – одна з його молекул, по відношенню до якої об'єкт грає роль оточення.

Система складається з безлічі індивідів, але вона до того ж сама є індивідом (не всякий індивід має сенс розглядати як систему). Тому структурний індивід може розглядатися як система, якщо нас цікавить взаємодія його частин (скажімо, механічних точок), а деякі системи можуть розглядатися як структурні індивіди.

У понятті проблеми нерозривно пов'язані два аспекти – об'єктивний (наявність реальної ситуації) і суб'єктивний (негативна оцінка реальності суб'єктом).

Що значить «вирішити проблему»? З визначення стає зрозумілим, що для цього слід зробити що завгодно, аби зменшити або зовсім зняти невдоволення суб'єкта.

#### **1.4.2. Варіанти рішення проблем**

Можливі варіанти вирішення проблеми природно розбиваються на дві групи: 1) впливати на суб'єкт з метою зменшити його невдоволення, не змінюючи реальності, 2) змінити реальність так, щоб невдоволення суб'єкта послабшав.

##### **1. Способи впливу на суб'єкт**

Є три можливості змінити на краще ставлення суб'єкта до реальності, не змінюючи самої реальності.

*По-перше.* Чим незадоволений суб'єкт? Тим, що йому відомо про ситуацію. Але ж він знає не все! І серед того, що він не знає, цілком може виявитися інформація позитивного характеру. Якщо її повідомити суб'єкту, його невдоволення зменшиться. Варто відзначити ще одну особливість даної можливості. Додаткова інформація, що повідомляється клієнту, обов'язково повинна бути позитивною, але не обов'язково правдивою. Зустрічаються випадки, коли проблема вирішується за допомогою неправдивої інформації.

*По-друге.* Наступна можливість вирішення проблеми без зміни реальності полягає в тому, щоб змінити сприйняття цієї дійсності суб'єктом.

Оскільки оцінка суб'єктом своїх взаємин з навколишнім середовищем є психічним явищем, то існує можливість впливу на психіку суб'єкта в потрібному напрямку. Форми впливу можуть бути різними: психічні (гіпноз, навіювання, пропаганда, реклама і т.д.); фізичні (вплив різних полів – акустичних, електричних, магнітних); хімічні (психотропні медикаменти, наркотики, алкоголь). Підкреслимо, що ми не оцінюємо, що добре, а що погано; ми лише констатуємо наявність фактичних можливостей.

*По-третє.* Третя можливість вирішення проблеми без зміни самої проблемної ситуації заснована на тому, що проблема виникла в результаті взаємодії суб'єкта з ситуацією. Тому іноді проблему можна вирішити, перервавши цю взаємодію. Тут теж є цілий спектр варіантів: від приємних для проблемоносія (підвищення по службі, направлення на навчання або у відпустку), за допомогою більш менш нейтральних (переведення в інший підрозділ, ротація), до хворобливих (звільнення тощо) та навіть до вкрай жорстокого, засуджуючого, але, на жаль, існуючого ("Є людина – є проблема, немає людини – немає проблеми" – вислів, приписуване Сталіну).

## 2. Втручання в реальність

Звернемося тепер до другої групи можливостей вирішення проблеми через втручання в саму проблемну ситуацію. Природно, втручання повинно так змінювати ситуацію, щоб невдоволення зменшилося або взагалі зникло.

Однак при цьому доводиться стикатися з досить суттєвою обставиною, яка, по суті, і дала поштовх для детальної розробки технології прикладного системного аналізу. Справа в тому, що в реальній (проблемній) ситуації беруть участь не тільки проблемоносій, а й багато інших суб'єктів, які оцінюють цю ж ситуацію зі своїх позицій

Будь-яка зміна ситуації в результаті втручання в неї буде помічено й оцінено всіма її учасниками, і зовсім не обов'язково вона буде схвалена усіма. Виникає принципово важливе питання: як слід діяти в зв'язку з цією обставиною?

Щоб відповісти на це питання, звернемося до принципової, кардинальної відмінності між об'єктом і суб'єктом.

Суб'єкт, будучи одночасно фізичним об'єктом, існує в реальному фізичному середовищі і, як всякий об'єкт, схильний до впливів цього середовища.

На відміну від об'єкта, суб'єкт не тільки підпорядкований природним закономірностям, а й наділений здатністю оцінювати свої взаємодії з середовищем: щось йому може подобатися, а може і не подобатися. Тут і закладена індивідуальність суб'єкта.

У цьому зв'язку корисною може виявитися наступна порада: Всякий раз, коли у вашій присутності прозвучить будь яке оцінне слово (добре – погано, корисно – шкідливо, правильно – неправильно і т.п.), насторожитесь і поставте запитання: " В якому сенсі? " суть поради в тому, що оцінки не бувають об'єктивними. Оцінки завжди суб'єктивні, і якщо ви хочете

зрозуміти справжній зміст сказаного, треба з'ясувати, які критерії застосовує оцінювач; одне і те ж різні суб'єкти можуть оцінювати по-різному.

#### 1.4.2. Типи ідеологій

Правильною вважається поведінка, яка максимально узгоджується з прийнятою суб'єктом **ідеологією**. Саме ідеологія і визначає, що погано, а що добре, що правильно, а що неправильно.

З'ясовується, що ідеології можуть бути різними. Вибір кожним «своєї» ідеології є комплексний результат особистого вибору, впливу виховання, культури, обставин. Ідеологи приносять велика кількість аргументів на користь саме своєї ідеології, обговорюють її багато відмінності від інших навчань. Однак можна вказати на одну, визначальну особливість, принципово відрізняє ідеології між собою. Це визначення того, яке відношення до інших суб'єктів вважати правильним.

Перший тип ідеології назвемо умовно **«принципом пріоритету меншості»**. У нашому випадку цей принцип призводить до того, щоб здійснити втручання, угодне клієнту, а інтереси інших учасників не приймаються до уваги. Комусь із них це може сподобатися, комусь — ні, це нас не цікавить. Можна навести життєві приклади реалізації такої ідеології (диктатура, ієрархічна організація, егоїзм, самолюбство і т.д.). Можна навіть явно визнати те, що в деяких обставинах така ідеологія дає найбільший шанс на виживання (армія, війна, надзвичайна ситуація та ін.)

Необхідно, однак, мати на увазі, що з реалізацією цієї ідеології пов'язаний ряд особливостей, які неминуче доведеться враховувати. По-перше, здійснення такого підходу до вирішення проблем «меншості» обов'язково викличе невдоволення певної частини решти приватників ситуації, що спонукає їх до дій у відповідь. Звідси у які взяли цю ідеологію повинна бути сила для придушення незадоволених і готовність застосовувати цю силу.

Друга ідеологія може бути названа **«принципом пріоритету групи»**. Згідно їй серед учасників ситуації, крім клієнта, є інші суб'єкти, не менш важливі і цінні, ніж клієнт.

Тому тепер втручання повинно проводитися з урахуванням інтересів всіх «наших». Це, з одного боку, ускладнює проектування втручання, але з іншого — відкриває можливість використання ресурсів не тільки клієнта, а й інших членів нашої групи. Відомо безліч прикладів реальної практики даної ідеології (расизм, націоналізм, фашизм, комунізм, будь-яка групова діяльність, в тому числі партійна, профспілкова, спортивна і т.п.). Однак не зайве відзначити деякі невід'ємні особливості цієї ідеології, які латентно вбудовані в неї і у відповідних умовах можуть проявитися негативно. Насамперед, це **подвійна мораль**: ділячи всіх на «своїх» і «чужих», вона дозволяє до них ставитися по-різному. Ще одна слабкість цієї ідеології полягає в тому, що виникає суперечність між проголошеною рівністю,

демократією в середині групи і необхідністю організації групової діяльності, тобто створення ієрархічних структур з їх принциповою нерівністю.

Третя ідеологія — «принципом пріоритету кожного». В основі її лежать два постулати :

- Немає жодного однакового суб'єкта, всі вони різні.
- Незважаючи на відмінності, всі суб'єкти рівноцінні і рівноправні.

Звідси випливає, що неправильно, аморально вирішувати проблеми одних за рахунок інших. Правильним, моральним визнається тільки поліпшує втручання .

**Поліпшує втручання** - це така зміна проблемної ситуації, яка позитивно оцінюється хоча б одним з її учасників і немає негативного впливу на всіх інших.

Безумовно поліпшує втручання є найбільш прийнятний варіант ідеології. Але слід звернути увагу, що дані ідеології запозичені з прикладного системного аналізу соціального життя суспільства. При вирішенні проблем у фізиці, хімії, біології, математики, електроніки ... поліпшує втручання тотожне тому, що необхідно розглядати нескінченно велика кількість ступенів свободи . Це призведе до того, що проблема стане нерозв'язною . На практиці доводиться виділяти головні і другорядні (мало значущі фактори). Іншими словами задача зводиться до спрощення, редукції системи, тобто включаються в роботу інші типи ідеологій .

### 1.4.3. Чотири типи втручань

Повертаючись до способів вирішення проблем, цікаво розглянути класифікацію, запропоновану Р. Акоффа . Він зауважив, що, незважаючи на колосальне розмаїття, несхожість проблем, способів їх вирішення всього чотири .

**1 . Невтручання** чи Absolution . Цим терміном в позначаються дії священика, який відпускає гріхи парафіянам: він вислуховує сповідь і нічого не робить. У професійній мові системного аналізу цим терміном позначено невторчання - у розрахунку на те, що природний хід подій приведе до вирішення проблеми. Підкреслимо, що невторчання володіє одним з ознак поліпшувального втручання: при цьому нікому не стає гірше. Щоб стати «поліпшеним втручанням», потрібно, щоб події дійсно вели до рішення проблеми, а щоб віддати перевагу саме такій поведінці, потрібно, щоб будь-які запропоновані втручання приводили до гірших результатів, ніж невторчання. Приклади: поведінка лікаря при неможливості вилікування пацієнта, дії сапера при зустрічі з незнайомим вибуховим пристроєм, ваша розумна поведінка в життєвому конфлікті ваших друзів - подружжя тощо

**2 . Часткове втручання** чи Resolution . У системному аналізі цей термін позначає часткове втручання, дія, яка знижує незадоволеність, що послаблює гостроту проблеми, але не знімає її повністю. Зазвичай цей спосіб застосовується при дефіциті ресурсів, що не дозволяє повністю вирішити

проблему: «всім сестрам по сережках», розподіл за жеребом або по черзі, підвищення пенсій і стипендій на тлі інфляції - приклади цього способу .

**3. Оптимальне рішення** чи Solution. На відміну від розмовного вживання цього слова («рішення» ), в професійній мові системного аналізу воно є терміном, що позначає найкраще в заданих умовах втручання. У російській мові цьому відповідає термін оптимальне рішення. Однак зрозуміло, що варіант, найкращий за одним критерієм, не обов'язково буде найкращим за іншим критерієм, тим більше якщо ресурси обмежені в даних конкретних умовах. Інша небезпечна сторона оптимізації - нечіткість завдань. Майже всі прикладні завдання мають кілька « розмиті межі »; ніколи невідомо повної інформації про систему.

**4. Розчинення** чи Dissolution. Цим терміном позначено втручання, що закінчується повним зникненням проблеми і не появою нових проблем.

Р. Акофф привів живий приклад застосування всіх чотирьох методів до вирішення однієї реальної проблеми .

В автобусній компанії великого міста виникла проблема: після введення надбавок за якість роботи почалися конфлікти між водіями та кондукторами. Справа в тому, що якість роботи водіїв оцінювалась по точності дотримання графіка руху, а кондукторів - по тому, наскільки якісно вони обслуговують пасажирів. У години пік кондуктори затримували сигнал відправки (ім треба було перевіряти на виході не тільки наявність квитка, а й правильність оплати, залежно від відстані, а на вході - продати квитки, кожному до його станції призначення ), а це негативно позначалося на надбавку водіїв.

Спочатку керівництво компанії ігнорувало проблему (ABSOLUTION), очікуючи, що все владнається само собою . Але проблема продовжувала загострюватися, в конфлікт були залучені профспілки. То ді керівництво спробувало повернутися до старої системи оплати (RESOLUTION); проте обидві профспілки запротестували, так як це означало б скасування надбавок . Потім керівництво запропонувало профспілкам самим домовитися про розділ фонду доплат ( SOLUTION ), але ті не бажали співпрацювати між собою.

Проблема була «розчинена» (DISSOLUTION) запрошеним системним аналітиком, який виявив, що в години пік число автобусів на лінії (регульоване протягом дня) перевищує число зупинок. На ці години кондукторів стали знімати з автобусів і закріплювати за зупинками. Вони продавали квитки ще до приходу автобуса, встигали перевірити квитки у пасажирів що виходять і стали вчасно подавати сигнал відправлення. По закінченні години пік кондуктори поверталися в автобуси, а зайві автобуси знімалися з лінії. До того ж компанії знадобилась менша кількість кондукторів.

Цей цікавий приклад не повинен створювати враження, ніби чотири типи рішень розташовані в порядку абсолютного переваги: в даному випадку

«розчинення» (зняття обмеження) виявилось кращим з них, але в інших проблемах кращим може виявитися будь-яке інше .

## 1.6. Алгоритми системного аналізу

Під алгоритмом розуміють певну логічну схему, послідовність діяльності для досягнення певної мети, тобто це модель, що відображає послідовний взаємозв'язок дій, кроків, застосування засобів, що в підсумку приводять до завчасно, наперед визначеної мети.

Одну з найбільш детальних схем здійснення системного аналізу запропонував Ю.Черняк, який виділив 12 етапів та відповідні методи їх здійснення:

1.Аналіз проблеми: виявлення, точне формулювання, дослідження логічної структури, ретроспективний аналіз, виявлення зовнішніх зв'язків, визначення реальності(можливості) вирішення проблеми в принципі.

2.Визначення системи: специфікація задачі, визначення позиції спостерігача, визначення об'єкта, виділення елементів (встановлення межі дезагрегації системи), визначення підсистем, характеристика середовища.

3.Аналіз структури системи: визначення рівнів ієрархії, визначення аспектів структури та способів їх вираження, виявлення процесів і функцій елементів, виявлення процесів управління (потоків інформації), специфікація підсистем та процесів діяльності й розвитку.

4.Формулювання мети і критерію системи: визначення цілей (вимог над системи та обмежень середовища), формулювання генеральної мети, визначення критерію, декомпозиція цілі та критерію за підсистемами, композиція загального критерію на основі часткових критеріїв підсистем.

5.Декомпозиція мети, виявлення потреб в ресурсах і процесах: формулювання цілей кожного ієрархічного рівня; формулювання зовнішніх цілей і обмежень; визначення потреби в ресурсах і процесах.

6.Виявлення ресурсів і процесів, композиція цілей: оцінка існуючих технологій і потужностей, оцінка сучасного стану ресурсів, оцінка існуючих планових проектів, оцінка можливостей взаємодії з іншими системами, оцінка суб'єктивних факторів, композиція цілей.

7.Прогноз і аналіз майбутніх умов: аналіз стійких тенденцій розвитку системи, прогноз стану середовища, передбачення появи принципово нових факторів, аналіз ресурсів майбутнього, комплексний аналіз факторів майбутнього, аналіз ймовірних змін цілей і критеріїв.

8.Оцінка цілей і засобів: кількісна оцінка критерію, оцінка взаємозалежності цілей, оцінка відносно ї важливості цілей, оцінка дефіцитності і вартості ресурсів, оцінка впливу зовнішніх факторів, обчислення комплексних розрахункових оцінок.

9.Відбір варіантів: оцінка цілей на сумісність, перевірка цілей на повноту, відкидання надлишкових цілей, планування варіантів досягнення

кожної окремої цілі, оцінка і порівняння варіантів, узгодження комплексу взаємопов'язаних варіантів.

10. Діагноз існуючої системи: моделювання технологічного та економічного процесів, розрахунок потенціальної і фактичної потужності, аналіз втрат потужності, виявлення недоліків організації виробництва і управління, обґрунтування і аналіз заходів вдосконалення.

11. Побудова комплексної програми розвитку: формулювання проектів і програм; визначення черговості цілей і засобів їх досягнення; розподіл сфер діяльності; розподіл сфер компетенції; розробка комплексного плану виконавцями.

12. Проектування організації досягнення цілей: визначення цілей організації, формулювання функцій організації, проектування організаційної структури, проектування інформаційних механізмів, проектування режимів роботи, проектування механізмів забезпечення (матеріального, психологічного, морального).

Інша схема алгоритму системного аналізу [2]

- визначення конфігуратора ;
- визначення проблеми та проблематики;
- виявлення цілей;
- формування критеріїв;
- генерування альтернатив;
- побудова і використання моделей;
- оптимізація;
- вибір;
- декомпозиція;
- агрегування;
- дослідження інформаційних потоків;
- дослідження ресурсних можливостей;
- спостереження та експерименти над досліджуваною

системою;

- реалізація, впровадження результатів аналізу.

На кожній із зазначених стадій системного дослідження застосовуються різноманітні методи, взаємозв'язок і послідовність яких визначаються конкретними задачами та специфікою об'єктів (проблем). Серед найбільш поширених методів такі: діагностичні, «дерева цілей», моделювання, економічного аналізу, експертних оцінок, математичні, статистичні.

## 1.6. Методи системного аналізу

З усієї великої сукупності методів системного аналізу кожен з них застосовується для вирішення певної конкретної задачі (операції) або сукупності задач. Єдиної, загально визнаної класифікації методів системного аналізу поки що немає. З наявних спроб класифікації видно, що в переважній



їх більшості до критеріїв класифікації належать спосіб реалізації та форма представлення. Це, зокрема, спостерігається в публікаціях Ю. Черняка, Ю. Сурміна, Г. Юна. Потрібно також зауважити, що власне класифікація методів системного аналізу має більш ілюстративне значення, ніж практичну спрямованість. З огляду на цю обставину узагальнений поділ методів системного аналізу можна здійснити наступним чином:

- експертно-інтуїтивні (неформальні) методи, до яких належать, зокрема, методи «мозкової атаки», сценаріїв, експертних оцінок);
- кількісні (формальні) методи (математичні, статистичні);
- графічні методи (дерева цілей, дерева взаємозв'язків);
- методи моделювання поєднують елементи зазначених методів (імітаційні, ігрові, макетні моделі).

Методи експертних оцінок базуються на отриманні, обробці та узагальненні інформації фахівців (експертів), які мають високу кваліфікацію та досвід у відповідній галузі знань (діяльності). Загалом ці методи поділяються на дві групи: індивідуальні експертні оцінки та колективні експертні оцінки. До першої групи належать, зокрема, методи психоінтелектуальної генерації ідей, метод інтерв'ю. До другої групи належать методи експертних комісій, метод Дельфі, метод колективної генерації ідей «мозкової атаки», метод керованої генерації ідей, метод деструктивної відносної оцінки, аналітичний метод (експертні оцінки моделі об'єкта).

Метод «мозкової атаки» базується на стимулюванні творчої продуктивної діяльності експертів шляхом спільного обговорення конкретної проблеми, яке регламентується певними правилами. При цьому «забороняється» оцінка висунутих ідей, обмежується час одного виступу, дозволяються багато разові виступи одного й того ж учасника; пріоритет виступу має експерт, що розвиває попередню ідею; обов'язково фіксуються всі висловлені ідеї, оцінка ідей здійснюється на наступних етапах.

Метод Дельфі полягає у виявленні узгодженої оцінки експертної групи шляхом їх автономного опитування в декілька турів, що передбачає повідомлення експертам результатів попереднього туру з метою додаткового обґрунтування оцінки експертів у наступному турі. Це найбільш формальний метод серед експертних методів, він найчастіше використовується для аналізу соціальних, економічних, політичних проблем, які не мають достатнього кількісного виразу.

Метод складання сценаріїв найчастіше застосовується для вирішення задач прогнозування. Сценарій – це опис ймовірного розвитку процесу чи стану об'єкта в майбутньому на основі правдоподібних припущень. Зазвичай складають три типи сценаріїв: оптимістичний, середній та песимістичний. Складання сценарію – це не якась довільна процедура, це насамперед науковий метод, що підкоряється певному алгоритму. Складання сценарію передбачає логічну послідовність операцій, які включають, зокрема:

формулювання проблеми; відбір і оцінку чинників розвитку процесу; обґрунтування показників і критеріїв, що характеризують стан об'єкта (процесу); формулювання припущень щодо ймовірного розвитку процесу; порівняння (зіставлення) бажаних показників майбутнього розвитку з показниками та характеристиками, що очікуються в результаті припущень (суджень, передбачень); виявлення ймовірних наслідків розвитку подій «за сценарієм», розробка заходів з регулювання процесу в бажаному напрямку.

### **1.7. Шкали вимірів в експертному аналізі. Оцінка середніх величин**

Думки експертів часто висловлені в *порядковій шкалі*, тобто експерт може сказати (і обґрунтувати), що один показник якості продукції важливіший, ніж інший, перший технологічний об'єкт небезпечніший, ніж другий, і так далі. Але він не в змозі сказати, *в скільки разів* або *на скільки* важливіший, відповідно, небезпечніший. Експертів часто просять дати ранжування (впорядкування) об'єктів експертизи, тобто розташувати їх в порядку зростання (чи убутання) інтенсивності експертизи характеристики, що цікавить організаторів. Ранг – це номер (об'єкту експертизи) у впорядкованому ряду.

Формально ранги виражаються числами 1, 2, 3,..., але з цими числами не можна робити звичні арифметичні операції. Наприклад, хоча  $1 + 2 = 3$ , але не можна стверджувати, що для об'єкту, що стоїть на третьому місці, інтенсивність характеристики, що вивчається, дорівнює сумі інтенсивностей об'єктів з рангами 1 і 2. Так, один з видів експертного оцінювання – оцінки учнів. Навряд чи хто-небудь стверджуватиме, що знання відмінника дорівнюють сумі знань двієчника і трієчника (хоча  $5 = 2 + 3$ ), хорошист відповідає двом двієчникам ( $2 + 2 = 4$ ), а між відмінником і трієчником така ж різниця, як між хорошистом і двієчником ( $5 - 3 = 4 - 2$ ). Тому очевидно, що для аналізу подібного роду якісних даних потрібна арифметика, що дає базу для розробки, вивчення і застосування конкретних методів розрахунку. Це є *репрезентативна теорія вимірів* (РТВ). Потрібно мати на увазі, що нині термін «теорія вимірів» застосовується для позначення цілого ряду наукових дисциплін: класична метрологія, РТВ, деяких інших напрямів, наприклад, алгоритмічна теорія вимірів.

**Основні шкали виміру.** Відповідно до РТВ при математичному моделюванні реального явища або процесу слід передусім встановити, *в яких типах шкал виміряні* ті або інші змінні. Тип шкали задає групу допустимих перетворень. Допустимі перетворення не міняють співвідношень між об'єктами виміру. Наприклад, при вимірі довжини перехід від аршин до метрів не міняє співвідношень між довжинами даних об'єктів, якщо перший об'єкт довший за другий, то це буде встановлено і при вимірі в аршинах, і при вимірі в метрах.

Вкажемо основні види шкал виміру і відповідні групи допустимих перетворень. У *шкалі найменувань* (інша назва *номінальної шкали*) **допустимими** є усі взаємно-однозначні перетворення. У цій шкалі числа використовуються лише як мітки. Приблизно так само, як при задачі білизни в пральню, тобто лише для розрізнення об'єктів. У шкалі найменувань виміряні, наприклад, номери телефонів, автомашин, паспортів, студентських квитків. Стать людей теж виміряна в шкалі найменувань, результат виміру набуває два значення – чоловічий, жіночий. Раса, національність, колір очей, волосся – номінальні ознаки. Номери букв в алфавіті – теж виміру в шкалі найменувань. Нікому при повному розумі не прийде в голову складати або множити номери телефонів, такі операції не мають сенсу. Порівнювати букви і говорити, наприклад, що буква П краще за букву З, також ніхто не буде. Єдине, для чого годяться виміри в шкалі найменувань – це розрізняти об'єкти. У багатьох випадках тільки це від них і вимагається. Наприклад, шафки в роздягальнях для дорослих розрізняють по номерах, тобто числам, а в дитячих садах використовують малюнки, оскільки діти ще не знають чисел.

У *порядковій шкалі* числа використовуються для встановлення порядку між об'єктами. Простим прикладом є оцінки знань учнів. Символічно, що в середній школі застосовуються оцінки 2, 3, 4, 5, а у вищій рівно той же сенс виражається словесно – незадовільно, задовільно, добре, відмінно. Цим підкреслюється «нечисловий» характер оцінок знань учнів. У *порядковій шкалі допустимими* є усі строго зростаючі перетворення.

Чому думки експертів природно висловлювати саме в *порядковій шкалі*? **Як показали численні дослідження, людина правильніше (і з меншими складнощами) відповідає на питання якісного, наприклад, порівняльного, характеру, ніж кількісного.** Так, йому легше сказати, яка з двох гирь важче, ніж вказати їх зразкову вагу в грамах.

Використовується багато інших відомих прикладів *порядкових шкал*. Так, наприклад, в мінералогії використовується шкала Мооса, згідно з якою мінерали класифікуються згідно з критерієм твердості. А саме: тальк має бал 1, гіпс – 2, кальцій – 3, флюорит – 4, апатит – 5, ортоклаз – 6, кварц – 7, топаз – 8, корунд – 9, алмаз – 10. *Порядковими шкалами* в географії є шкала вітрів Бофорта («штиль», «слабкий вітер», «помірний вітер» і так далі), шкала сили землетрусів. У медицині *порядковими шкалами* – є шкала стадій гіпертонічної хвороби (по Мясникову), шкала мір серцевої недостатності (по Стражеско-Василенко-Лангу), шкала міри вираженості коронарної недостатності (по Фогельсону). Номери будинків також виміряні в *порядковій шкалі*. При оцінці якості продукції і послуг, в т.з. кваліметрії (буквальний переклад: вимір якості) популярні *порядкові шкали*. А саме, одиниця продукції оцінюється як придатна або не придатна. При ретельнішому аналізі використовується шкала з трьома градаціями: є значні дефекти, є присутніми тільки незначні дефекти, немає дефектів.

При оцінці екологічних дій перша оцінка зазвичай порядкова: природне середовище стабільне, природне середовище пригноблене (деградує). Аналогічно в еколого-медичній шкалі: немає вираженої дії на здоров'я людей – відзначається негативна дія на здоров'я. Порядкова шкала використовується і в інших областях.

**Порядкова шкала і шкала найменувань – основні шкали якісних ознак.** Тому в багатьох конкретних областях результати якісного аналізу можна розглядати як виміри по цих шкалах.

**Шкали кількісних ознак – це шкали інтервалів, відношень, різниць, абсолютна.** За шкалою *інтервалів* вимірюють величину потенційної енергії або координату точки на прямій. У цих випадках на шкалі не можна відмітити ні природний початок відліку, ні природну одиницю виміру. Допустимими перетвореннями в шкалі інтервалів є лінійні зростаючі перетворення, тобто лінійні функції. Температурні шкали Цельсія і Фаренгейта пов'язані саме такою залежністю:  $\tilde{N}_0 = 5/9(\hat{O}_0 - 32)$ , де  $\tilde{N}_0$  – температура за шкалою Цельсія, а  $\hat{O}_0$  – температура за шкалою **Фаренгейта**.

З кількісних шкал найбільш поширеними в науці і практиці є шкали *відношень*. У них є природний початок відліку – нуль, тобто відсутність величини, але немає природної одиниці виміру. За шкалою відношень виміряні більшість фізичних одиниць: маса тіла, довжина, заряд, а також ціни в економіці. Допустимими перетвореннями в шкалі відношень є подібні (що змінюють тільки масштаб).

Час вимірюється за шкалою *різниць*, якщо рік приймаємо природною одиницею виміру, і за шкалою інтервалів в загальному випадку. Природного початку відліку вказати на сучасному рівні знань не можна. Дату створення світу різні автори розраховують по-різному, так само як і момент Різдва Христового. Так, згідно нової статистичної хронології Господь Ісус Христос народився в 1054 р. в Стамбулі (він же – Царьград, Візантія, Троя, Єрусалим, Рим) по прийнятому нині літочисленню.

Тільки для *абсолютної* шкали результати вимірів – числа в звичайному значенні слова. Прикладом є число людей в кімнаті. Для абсолютної шкали допустимим є тільки тотожне перетворення.

В процесі розвитку відповідної області знання тип шкали може мінятися. Так, спочатку температура вимірювалася за *порядковою* шкалою (холодніше, тепліше). Потім – по *інтервальній* (шкали Цельсія, Фаренгейта, Реомюра). Нарешті, після відкриття абсолютного нуля температуру слід вважати виміряною за шкалою *відношень* (шкала Кельвіна). Потрібно відмітити, що серед фахівців іноді є розбіжності з приводу того, по яких шкалах слід вважати виміряними ті або інші реальні величини. Іншими

словами, процес виміру включає і визначення типу шкали (разом з обґрунтуванням).

**Інваріантні алгоритми і середні величини.** Основна вимога до алгоритмів аналізу даних формулюється в РТВ так: **висновки, зроблені на основі даних, виміряних в шкалі певного типу, не повинні мінятися при допустимому перетворенні шкали виміру цих даних.** Іншими словами, висновки мають бути *інваріантні* по відношенню до допустимих перетворень шкали.

Таким чином, одна з основних цілей теорії вимірів – боротьба з суб'єктивізмом дослідника при приписуванні чисельних значень реальним об'єктам. Так, відстані можна вимірювати в аршинах, метрах, мікронах, милях, парсеках і інших одиницях виміру. Масу (вагу) – в пудах, кілограмах, фунтах та ін. Ціни на товари і послуги можна вказувати в юанях, рублях, тенге, гривнях, латах, кронах, марках, доларах США і інших валютах (за умови заданих курсів перерахунку). Підкреслимо дуже важливу, хоча і цілком очевидну обставину: вибір одиниць виміру залежить від дослідника, тобто суб'єктивний. *Статистичні висновки можуть бути адекватні реальності тільки тоді, коли вони не залежать від того, якій одиниці виміру віддасть перевазі дослідник, тобто коли вони інваріантні відносно допустимого перетворення шкали.*

В якості прикладу розглянемо обробку думок експертів, виміряних в порядковій шкалі. Нехай  $Y_1, Y_2, \dots, Y_n$  – сукупність оцінок експертів, «виставлених» одному об'єкту експертизи (наприклад, одному з варіантів стратегічного розвитку фірми),  $Z_1, Z_2, \dots, Z_n$  – другому (іншому варіанту такого розвитку).

Як порівнювати ці сукупності? Очевидно, найпростіший спосіб – по середніх значеннях. А як обчислювати середні? Відомі різні види середніх величин: середнє арифметичне, медіана, мода, середнє геометричне, середнє гармонійне, середнє квадратичне. Узагальненням декількох з перерахованих являється середнє по Колмогорову. Для чисел  $X_1, X_2, \dots, X_n$  середнє по Колмогорову обчислюється за формулою

$$G\{(F(X_1)+F(X_2)+\dots+F(X_n))/n\},$$

де  $F$  – строго монотонна функція,  $G$  – функція, зворотна до  $F$ . Серед середніх по Колмогорову, багато добре відомих персонажів. Так, якщо  $F(x) = x$ , то середнє по Колмогорову – це середнє арифметичне, якщо  $F(x) = \ln x$ , то середнє геометричне, якщо  $F(x) = 1/x$ , то середнє гармонійне, якщо  $F(x) = x^2$ , то середнє квадратичне, і так далі. З іншого боку, такі популярні середні, як медіана і мода, не можна представити у вигляді середніх по Колмогорову.

Нагадаємо, що загальне поняття середньої величини введене французьким математиком першої половини XIX ст. академіком О. Коши. Воно таке: середньою величиною є будь-яка функція  $f(X_1, X_2, \dots, X_n)$  така, що при усіх можливих значеннях аргументів значення цієї функції не менше, ніж мінімальне з чисел  $X_1, X_2, \dots, X_n$ , і не більше, ніж максимальне з цих чисел. Середнє по Колмогорову – окремий випадок середнього по Коши. Медіана і мода, хоча і не є середніми по Колмогорову, але теж середні по Коши.

При допустимому перетворенні шкали значення середньої величини, очевидно, міняється. Але висновки про те, для якої сукупності середнє більше, а для якої менше, не повинні мінятися (відповідно до вимоги інваріантності висновків, прийнятому як основна вимога в РТВ). Сформулюємо відповідне математичне завдання пошуку виду середніх величин, результат порівняння яких стійкий відносно допустимих перетворень шкали.

Нехай  $f(X_1, X_2, \dots, X_n)$  – середнє по Коши. Нехай середнє по першій сукупності менше середнього по другій сукупності:

$$f(Y_1, Y_2, \dots, Y_n) < f(Z_1, Z_2, \dots, Z_n) \quad (1)$$

Згідно РТВ для стійкості результату порівняння середніх необхідно, щоб для будь-якого допустимого перетворення  $g$  з групи допустимих перетворень у відповідній шкалі була справедливою також нерівність

$$f(g(Y_1), g(Y_2), \dots, g(Y_n)) < f(g(Z_1), g(Z_2), \dots, g(Z_n)) \quad (2)$$

тобто середнє перетворених значень з першої сукупності також було менше середнього перетворених значень для другої сукупності. Причому сформульована умова має бути вірна для будь-якої з двох сукупностей  $Y_1, Y_2, \dots, Y_n$  і  $Z_1, Z_2, \dots, Z_n$  і, нагадаємо, будь-якого допустимого перетворення  $g$ . Згідно з РТВ тільки такими середніми можна користуватися при аналізі думок експертів і інших даних, виміряних в даній шкалі.

За допомогою математичної теорії, розвиненою А. І. Орловим в 1970-х роках, вдається описати вид допустимих середніх в основних шкалах: в шкалі найменувань в якості середнього годиться тільки мода; з усіх середніх по Коши в порядковій шкалі в якості середніх можна використовувати тільки члени варіаційного ряду (порядкові статистики), зокрема, медіану (при непарному об'ємі вибірки; при парному ж об'ємі слід застосовувати один з двох центральних членів варіаційного ряду – як їх іноді називають, ліву медіану або праву медіану), але не середнє арифметичне, середнє геометричне і так далі; у шкала інтервалів з усіх середніх по Колмогорову можна застосовувати тільки середнє арифметичне; у шкалі відношень з усіх середніх по Колмогорову стійкими відносно порівняння є тільки степеневі середні і середнє геометричне.

Наведемо приклад, що показує некоректність використання середнього арифметичного  $f(X_1, X_2) = (X_1 + X_2) / 2$  в порядковій шкалі. Нехай  $Y_1 = 1, Y_2 = 11, Z_1 = 6, Z_2 = 8$ . Тоді  $f(Y_1, Y_2) = 6$ , що менше, ніж  $f(Z_1, Z_2) = 7$ . Нехай строго зростаюче перетворення  $g$  таке, що  $g(1) = 1, g(6) = 6, g(8) = 8, g(11) = 99$ . Таких перетворень багато. Наприклад, можна покласти  $g(x) = x$  при  $x$ , що не перевершують 8, і  $g(x) = 99(x - 8) / 3 + 8$  для  $x$ , більших 8. Тоді  $f(g(Y_1), g(Y_2)) = 50$ , що більше, ніж  $f(g(Z_1), g(Z_2)) = 7$ . Як бачимо, в результаті допустимого, тобто строго зростаючого перетворення шкали впорядкованість середніх змінилася.

Приведені результати про середні величини широко застосовуються, причому не лише в теорії експертних оцінок або соціології, але і, наприклад, для аналізу методів агрегації датчиків в АСУ ТП доменних печей. Велике прикладне значення РТВ в завданнях стандартизації і управління якістю, зокрема, в кваліметрії. Тут є і цікаві теоретичні результати. Так, наприклад, будь-яка зміна коефіцієнтів ваговитості одиничних показників якості продукції призводить до зміни впорядкування виробів по середньозваженому показнику (ця теорема доведена В. В. Подиновским).

**Методи середніх балів.** Нині поширені експертні, маркетингові, кваліметричні, соціологічні і інші опитування, в яких опитуваних просять виставити бали об'єктам, виробам, технологічним процесам, підприємствам, проектам, заявкам на виконання науково-дослідних робіт, ідеям, проблемам, програмам, політикам і тому подібне, а потім розраховують середні бали і розглядають їх як інтегральні оцінки, виставлені колективом опитаних. Якими формулами користуватися для обчислення середніх величин? Адже середніх величин, як ми знаємо, дуже багато різних видів. Зазвичай застосовують середнє арифметичне. Але *такий спосіб некоректний*, оскільки бали зазвичай виміряні в порядковій шкалі (див. вище). Обґрунтованим є використання медіан в якості середніх балів. Проте повністю *ігнорувати середні арифметичні недоцільно через їх звичність і поширеність*. Тому **доцільно використовувати одночасно обидва методи – і метод середніх арифметичних рангів (балів), і метод медіанних рангів**. Така рекомендація знаходиться у згоді з концепцією стійкості, що рекомендує використовувати різні методи для обробки одних і тих самих даних з метою виділити висновки, що отримуються одночасно при усіх методах. Такі висновки, мабуть, відповідають реальній дійсності, тоді як висновки, що міняються від методу до методу, залежать від суб'єктивізму дослідника, що вибирає метод обробки початкових експертних оцінок.

**Приклад порівняння восьми проектів.** Розглянемо конкретний приклад застосування тільки що сформульованого підходу. За завданням керівництва фірми аналізувалися вісім проектів, запропонованих для включення в план стратегічного розвитку фірми. Вони були позначені таким чином: Д, Л, М-К, Би, Г-Б, Сол, Стеф, До (по прізвищах менеджерів,

що запропонували їх для розгляду). Усі проекти були спрямовані 12 експертам, призначеним Правлінням фірми. У приведеній нижче таблиці 1 приведені ранги восьми проектів, присвоєні їм кожним з 12 експертів відповідно до представлення експертів про доцільність включення проекту в стратегічний план фірми. При цьому експерт привласнює ранг 1 найкращому проекту, який обов'язково потрібно реалізувати. Ранг 2 отримує від експерта другий по привабливості проект,..., нарешті, ранг 8 – найбільш сумнівний проект, який реалізовувати варто лише в останню чергу. Аналізуючи результати роботи експертів (тобто таблиця 1), члени Правління фірми були вимушені констатувати, що повної згоди між експертами немає, а тому дані, приведені в таблиці, слід піддати ретельнішому математичному аналізу.

Таблиця 1.

Ранги 8 проектів по мірі привабливості для включення в план стратегічного розвитку фірми.

№ експерта	Д	Л	М-К	Б	Г-Б	Сол	Стеф	К
1	5	3	1	2	8	4	6	7
2	5	4	3	1	8	2	6	7
3	1	7	5	4	8	2	3	6
4	6	4	2,5	2,5	8	1	7	5
5	8	2	4	6	3	5	1	7
6	5	6	4	3	2	1	7	8
7	6	1	2	3	5	4	8	7
8	5	1	3	2	7	4	6	8
9	6	1	3	2	5	4	7	8
10	5	3	2	1	8	4	6	7
11	7	1	3	2	6	4	5	8
12	1	6	5	3	8	4	2	7

*Примітка.* Експерт № 4 вважає, що проекти М-К і Б рівноцінні, але поступаються лише одному проекту – проекту Сол. Тому проекти М-К і Б повинні були б стояти на другому і третьому місцях і отримати бали 2 і 3. Оскільки вони рівноцінні, то отримують середній бал  $(2+3)/2 = 5/2 = 2,5$ .



**Метод середніх арифметичних рангів.** Спочатку був застосований метод середніх арифметичних рангів. Для цього передусім була підрахована сума рангів, присвоєних проектам (див. таблицю 1). Потім ця сума була розділена на число експертів, в результаті розрахований середній арифметичний ранг (саме ця операція дала назву методу). По середніх рангах будується підсумкове ранжування (у іншій термінології – впорядкування), виходячи з принципу – чим менше середній ранг, тим краще проект.

Найменший середній ранг, рівний 2,625, у проекту Б, – отже, в підсумковому ранжуванні він отримує ранг 1. Наступна за величиною сума, рівна 3,125, у проекту М-К, – і він отримує підсумковий ранг 2. Проекти Л і Сол мають однакові суми (рівні 3,25), значить, з точки зору експертів вони рівноцінні (при даному способі зведення разом думок експертів), а тому вони повинні б стояти на 3 і 4 місцях і отримують середній бал  $(3+4) / 2 = 3,5$ . Подальші результати приведені в таблицю 2 нижче. Отже, ранжування по сумах рангів (чи, що те ж саме, по середніх арифметичних рангах) має вигляд:

$$\text{Б} < \text{М-К} < \{\text{Л, Сол}\} < \text{Д} < \text{Стеф} < \text{Г-Б} < \text{До}. \quad (3)$$

Тут запис типу «А<Б» означає, що проект А перемагає проекту Б (тобто проект А краще за проект Б). Оскільки проекти Л і Сол отримали однакову суму балів, то по даному методу вони еквівалентні, а тому об'єднані в групу (у фігурних дужках). У термінології математичної статистики ранжування (3) має один зв'язок.

Таблиця 2. Результати розрахунків по методу середніх арифметичних і методу медіан для даних, приведених в таблиці 1.

	Д	Л	М-К	Б	Г-Б	Сол	Стеф	К
Сума рангів	60	39	37,5	31,5	76	39	64	85
Середнє арифметичне рангів	5	3,25	3,125	2,625	6,333	3,25	5,333	7,083
Підсумковий ранг по середньому арифметичному	5	3,5	2	1	7	3,5	6	8
Медіани рангів	5	3	3	2,25	7,5	4	6	7
Підсумковий ранг по медіанах	5	2,5	2,5	1	8	4	6	7

**Метод медіан рангів.** Значить, підсумок розрахунків – ранжування (3), і на її основі належить приймати рішення? Але тут найбільш знайомий з сучасною економетрикою член Правління згадав, що відповіді експертів виміряні в порядковій шкалі, а тому для них неправомірно проводити усереднювання методом середніх арифметичних. Потрібно використовувати метод медіан. Що це означає? Потрібно узяти відповіді експертів, що відповідають одному з проектів, наприклад, проекту Д. Це ранги 5, 5, 1, 6, 8, 5, 6, 5, 6, 5, 7, 1. Потім їх потрібно розташувати в порядку неубування (простіше було б сказати – "в порядку зростання", але оскільки деякі відповіді співпадають, то доводиться використовувати незвичний термін "неубування"). Отримаємо послідовність: 1, 1, 5, 5, 5, 5, 5, 6, 6, 6, 7, 8. На центральних місцях – шостому і сьомому – стоять 5 і 5. Отже, медіана дорівнює 5.

Медіани сукупностей з 12 рангів, які відповідають певним проектам, наведені в передостанньому рядку табл.2. (При цьому медіани обчислені за звичайними правилами статистики - як середнє арифметичне центральних членів варіаційного ряду.) Підсумкове впорядкування по методу медіан приведено в останньому рядку таблиці. Ранжировка (тобто упорядкування - підсумкове думка комісії експертів) по медіанам має вигляд:

$$Б < \{М-К, Л\} < Сол < Д < Стеф < К < Г-Б . \quad (4)$$

Оскільки проекти Л і М-К мають однакові медіани балів, то з даного методу ранжування вони еквівалентні, а тому об'єднані в групу (кластер), тобто з точки зору математичної статистики ранжировка (4) має один зв'язок.

**Порівняння ранжировок за методом середніх арифметичних і методу медіан.** Порівняння ранжировок (3) і (4) показує їх близькість (схожість). Можна прийняти, що проекти М-К, Л, Сол впорядковані як М-К < Л < Сол, але через похибки експертних оцінок в одному методі визнані рівноцінними проекти Л і Сол (ранжировка(3)), а в іншому - проекти М-К і Л (ранжирування (4)). Істотним є тільки розбіжність, що стосується впорядкування проектів К і Г-Б: в ранжируванні(3) Г-Б < К, а в ранжировці (4), навпаки, К < Г-Б. Однак ці проекти - найменш привабливі з восьми розглянутих, і при виборі найбільш привабливих проектів для подальшого обговорення і використання це розбіжність су-громадської.

Розглянутий приклад демонструє схожість і відмінність ранжировок, отриманих за методом середніх арифметичних рангів і по методу медіан, а також користь від їх спільного застосування.

## 1.8. Математичні методи аналізу експертних оцінок

При аналізі думок експертів можна застосовувати найрізноманітніші статистичні методи, описувати їх - значить описувати всю прикладну статистику. Проте можна виділити основні широко використовувані в даний час методи математичної обробки експертних оцінок - це перевірка узгодженості думок експертів (або класифікація експертів, якщо немає узгодженості) і усереднення думок експертів всередині узгодженої групи.

Оскільки відповіді експертів у багатьох процедурах експертного опитування - не числом, а такі об'єкти нечислової природи, як градації якісних ознак, ранжування, розбиття, результати парних порівнянь, нечіткі переваги і т.д., то для їх аналізу виявляються корисними методи статистики об'єктів нечислової природи.

**Перевірка узгодженості думок експертів і класифікація експертних думок.** Ясно, що думки різних експертів різняться. Важливо зрозуміти, наскільки велике це розходження. Якщо мало - усереднення думок експертів дозволить виділити те загальне, що є у всіх експертів, відкинувши випадкові відхилення в ту чи іншу сторону. Якщо велике - усереднення є чисто формальною процедурою. Так, якщо уявити собі, що відповіді експертів рівномірно покривають поверхню бублика, то формальне усереднення вкаже на центр дірки від бублика, а такої думки, не дотримується жоден експерт. Зі сказаного ясна важливість проблеми перевірки узгодженості думок експертів.

Розроблено ряд методів такої перевірки. Статистичні методи перевірки узгодженості залежать від математичної природи відповідей експертів. Відповідні статистичні теорії вельми важкі, якщо ці відповіді - ранжировки або розбиття, і досить прості, якщо відповіді - результати незалежних парних порівнянь. Звідси випливає рекомендація по організації експертного опитування: чи не намагайтеся відразу отримати від експерта ранжирування або розбиття, йому важко це зробити, та й наявні математичні методи не дозволяють далеко просунути в аналізі подібних даних. Наприклад, рекомендують перевіряти узгодженість ранжировок за допомогою коефіцієнта рангової конкордації Кендалла-Сміта. Але давайте згадаємо, яка статистична модель при цьому використовується. Перевіряється нульова гіпотеза, згідно з якою ранжировки незалежні і рівномірно розподілені на множині всіх ранжировок. Якщо ця гіпотеза приймається, то звичайно, ні про яку узгодженості думок експертів говорити не можна. А якщо відхиляється? Теж не можна. Наприклад, може бути два (або більше) центру, біля яких групуються відповіді експертів. Нульова гіпотеза відхиляється. Але хіба можна говорити про узгодженість?

Експерту набагато легше на кожному кроці порівнювати тільки два об'єкти. Нехай він займається парними порівняннями. *Непараметрична*

*теорія парних порівнянь (теорія люсіанів) дозволяє вирішувати більш складні завдання, ніж статистика ранжировок або розбиття. Зокрема, замість гіпотези рівномірного розподілу можна розглядати гіпотезу однорідності, тобто замість збіги всіх розподілів з одним фіксованим (рівномірним) можна перевіряти лише збіг розподілів думок експертів між собою, що природно трактувати як узгодженість їх думок. Таким чином, вдається позбутися від неприродного припущення рівномірності.*

За відсутності узгодженості експертів природно розбити їх на групи подібних за думку. Це можна зробити різними методами статистики об'єктів нечислової природи, що відносяться до кластер-аналізу, попередньо ввівши метрику в простір думок експертів. Ідея американського математика Джона Кемені про аксіоматичному введенні метрик (см. Нижче) знайшла численних продовжувачів. Проте методи кластер-аналізу звичайно є евристичними. Зокрема, неможливо з позицій статистичної теорії обґрунтувати "законність" об'єднання двох кластерів в один. Мається важливий виняток - для незалежних парних порівнянь (люсіанів) розроблені методи, що дозволяють перевіряти можливість об'єднання кластерів як статистичну гіпотезу. Це - ще один аргумент за те, щоб розглядати теорію люсіанів як ядро математичних методів експертних оцінок.

**Знаходження підсумкової думки комісії експертів.** Нехай думки комісії експертів або якоїсь її частини визнані узгодженими. Яке ж підсумкове (середнє, загальне) думка комісії? Згідно ідеї Джона Кемені слід знайти середнє думка як рішення *оптимізаційної задачі*. А саме, треба мінімізувати сумарну відстань від кандидата в середні до думок експертів. Знайдене таким способом середню думку називають "медіаною Кемені".

Математична складність полягає в тому, що думки експертів лежать в деякому просторі об'єктів нечислової природи. Загальна теорія подібного усереднення побудована в ряді робіт, зокрема, показано, що в силу узагальнення закону великих чисел середню думку при збільшенні числа експертів (чиї думки незалежні і однаково розподілені) наближається до деякого межі, який природно назвати *математичним очікуванням* (випадкового елементу, що має той же розподіл, що і відповіді експертів).

У конкретних просторах нечислових думок експертів обчислення медіани Кемені може бути досить складною справою. Крім властивостей простору, велика роль конкретних метрик. Так, в просторі ранжировок при використанні метрики, пов'язаної з коефіцієнтом рангової кореляції Кендалла, необхідно проводити досить складні розрахунки, в той час як застосування показника відмінності на основі коефіцієнта рангової кореляції Спірмена призводить до впорядкування за середніми рангах.

**Бінарні відношення і відстань Кемені.** Як відомо, бінарне відношення  $A$  на кінцевій множині  $Q = \{q_1, q_2, \dots, q_k\}$  - це підмножина *декартового*

квадрата  $Q^2 = \{ (q_m, q_n), m, n = 1, 2, \dots, k \}$ . При цьому пара  $(q_m, q_n)$  входить до  $A$  тоді і тільки тоді, коли між  $q_m$  і  $q_n$  має місце аналізоване ставлення. Кожне кластеризоване ранжирування, як і будь-яке бінарне відношення, можна задати матрицею  $\| x(a, b) \|$  з 0 і 1 порядку  $k \times k$ . При цьому  $x(a, b) = 1$  тоді і тільки тоді, коли  $a < b$  або  $a = b$ . У першому випадку  $x(b, a) = 0$ , а в другому  $x(b, a) = 1$ . При цьому хоча б одне з чисел  $x(a, b)$  і  $x(b, a)$  дорівнює 1.

Як використовувати зв'язок між ранжування і матрицями? Наприклад, з визначення суперечливості пари  $(a, b)$  (див. Вище, пункт про теорію вимірювань) випливає, що для знаходження всіх таких пар можна скористатися матрицями, відповідними ранжируванням. Досить по-елементно перемножити дві матриці  $\| x(a, b) \|$  і  $\| y(a, b) \|$ , що відповідають двом кластеризованим ранжируванням, і відібрати ті і тільки ті пари, для яких  $x(a, b) y(a, b) = x(b, a) y(b, a) = 0$ .

В експертних методах використовують, зокрема, такі бінарні відносності, як ранжирування (упорядкування, або розбиття на групи, між якими є строгий порядок), відношення еквівалентності, толерантності (відносини подібності). Як випливає зі сказаного вище, кожне бінарне відношення  $A$  можна описати матрицею  $\| a(i, j) \|$  з 0 і 1, причому  $a(i, j) = 1$  тоді і тільки тоді, коли  $q_i$  і  $q_j$  перебувають у відношенні  $A$ , і  $a(i, j) = 0$  в іншому випадку.

**Визначення.** Відстанню Кемені між бінарними відносинами  $A$  і  $B$ , описуваними матрицями  $\| a(i, j) \|$  і  $\| b(i, j) \|$  відповідно, називається число  $D(A, B) = \sum | a(i, j) - b(i, j) |$ , де підсумовування проводиться по всіх  $i, j$  від 1 до  $k$ , тобто відстань Кемені між бінарними відношеннями дорівнює сумі модулів різниць елементів, що стоять на одних і тих же місцях у відповідних їм матрицях

Легко бачити, що відстань Кемені - це число незбіжних елементів в матрицях  $\| a(i, j) \|$  і  $\| b(i, j) \|$ .

**Медіана Кемені і закони великих чисел.** За допомогою відстані Кемені знаходять підсумкове думка комісії експертів. Нехай  $A_1, A_2, A_3, \dots, A_p$  - відповіді  $p$  експертів, представлені у вигляді бінарних відносин. Для їх усереднення використовують т.зв. медіану Кемені  $\text{Arg min } \sum D(A_i, A)$ , де  $\text{Arg min}$  - то або ті значення  $A$ , при яких досягає мінімуму зазначена сума відстаней Кемені від відповідей експертів до поточної змінної  $A$ , по якій і проводиться мінімізація. Таким чином,  $\sum D(A_i, A) = D(A_1, A) + D(A_2, A) + D(A_3, A) + \dots + D(A_p, A)$ . Крім медіани Кемені, використовують середнє по Кемені, в якому замість  $D(A_i, A)$  коштує  $D^2(A_i, A)$ . Медіана Кемені - окремий випадок визначення емпіричного середнього в просторах нечислової природи. Для неї справедливий закон великих чисел, тобто емпіричне середнє наближається при зростанні числа складових (тобто  $p$  - числа доданків у сумі), до теоретичного середнього:  $\text{Arg min } \sum D(A_i, A) \rightarrow \text{Arg min } M D(A_i, A)$ . Тут  $M$  - символ математичного сподівання. Передбачається, що

відповіді р експертів  $A_1, A_2, A_3, \dots, A_p$  є підстави розглядати як незалежні однаково розподілені випадкові елементи (тобто як випадкову вибірку) у відповідному просторі довільній природи, наприклад, в просторі упорядкувань або відносин еквівалентності.

Закони великих чисел показують, по-перше, що медіана Кемені володіє *стійкістю* по відношенню до незначної зміни складу експертної комісії; по-друге, при збільшенні числа експертів вона *наближається до деякої межі*. Його природно розглядати як *істинну думку* експертів, від якого кожен з них кілька відхилявся з випадкових причин. Розглянутий тут закон великих чисел є узагальненням відомого в статистиці «класичного» закону великих чисел. Він заснований на іншій математичній базі - теорії оптимізації, в той час як «класичний» закон великих чисел використовує підсумовування. Впорядкування та інші бінарні відношення не можна складати, тому доводиться застосовувати іншу математику. Обчислення медіани Кемені - завдання цілочисельного програмування. Зокрема, для її знаходження використовується різні алгоритми дискретної математики, зокрема, засновані на методі гілок і меж. Застосовують також алгоритми, засновані на ідеї випадкового пошуку, оскільки для кожного бінарного відношення неважко знайти безліч його сусідів.

Таблиця 3.

Матриця попарних відстаней

0	2	13	1	7	4	10	3	11
2	0	5	6	1	3	2	5	1
13	5	0	2	2	7	6	5	7
1	6	2	0	5	4	3	8	8
7	1	2	5	0	10	1	3	7
4	3	7	4	10	0	2	1	5
10	2	6	3	1	2	0	6	3
3	5	5	8	3	1	6	0	9
11	1	7	8	7	5	3	9	0

Розглянемо приклад обчислення медіани Кемені. Нехай дана квадратна матриця (порядку 9) попарних відстаней для безлічі бінарних відносин з 9 елементів  $A_1, A_2, A_3, \dots, A_9$  (см. Табл. 3). Потрібно знайти в цій множині медіану для безлічі з 5 елементів  $\{A_2, A_4, A_5, A_8, A_9\}$ .

Відповідно до визначення медіани Кемені слід ввести в розгляд функцію

$$C(A) = \sum D(A_i, A) = D(A_2, A) + D(A_4, A) + D(A_5, A) + D(A_8, A) + D(A_9, A),$$

розрахувати її значення для всіх  $A_1, A_2, A_3, \dots, A_9$  і вибрати найменше. Проведемо розрахунки:  $C(A_1) = 24$ ,  $C(A_2) = 13$ ,  $C(A_3) = 21$ ,  $C(A_4) = 27$ ,  $C(A_5) = 16$ ,  $C(A_6) = 23$ ,  $C(A_7) = 15$ ,  $C(A_8) = 25$ ,  $C(A_9) = 25$ . З усіх обчислених сум найменша дорівнює 13, і досягається вона при  $A = A_2$ , отже, медіана Кемені - це  $A_2$ .

### Питання для самоконтролю

- 1 Які три способи впливу на суб'єкт без зміни реальності можуть привести до вирішення його проблеми?
- 2 Охарактеризуйте динамічні властивості систем.
- 3 Дайте визначення поняття втручання, що покращує
- 4 Яка основна відмінність суб'єкта від об'єкта?
- 5 Поясніть відмінність між ростом і розвитком системи.
6. Які є шкали вимірювань?
7. Що таке середня величина

### Основна література

1. Сурмін Ю.П. Теорія систем і системний аналіз: Навч.посібник.-К.: МАУП, 2003.-368с.
2. Ф.І.Перегудов, Ф.П.Тарасенко Основи системного аналізу Томськ - 1997ббк22.161П27.
3. Орлов А. І. Експертні оцінки. Навчальний посібник. М .: ІВСТЕ, 2002

## Розділ 2.

### МОДЕЛЮВАННЯ СИСТЕМ

При вивченні систем різної природи дослідник стикається з проблемою їх відображення, а також використання в пізнавальній і практичній діяльності. Об'єкт фіксується термінами мови, відображається на папері кресленнями, графіками, фотографіями, рівняннями і формулами, а також макетами, механізмами, пристроями. Потім ці відображення застосовуються для наукового дослідження (наприклад, спостереження, експерименту ) або для практичної діяльності. Відображення об'єктів називаються - моделями, процес їх створення - моделюванням, а використання, відповідно, в науці називаються модельним дослідженням (модельним експериментом, модельним спостереженням) і модельною практикою в практичній діяльності. Способи побудови моделей отримали назву методи моделювання. Вони дуже різноманітні. Практично кожна наука має свій арсенал методів моделювання. Розрізняють геометричне, фізичне, хімічне, біологічне, економічне, соціальне, політичне, культурологічне та математичне моделювання .

Моделювання - дослідження об'єктів пізнання на їх моделях; побудова і вивчення моделей реально існуючих об'єктів, процесів або явищ з метою отримання пояснень цих явищ, а також для передбачення явищ, що цікавлять дослідника.

Необхідними і достатніми ознаками моделі є сформульовані В. А. Штоффом такі умови:

- між моделлю та оригіналом маєтись відношення подібності, форма якого явно виражена і точно зафіксована (умови відбиття або уточненої аналогії);
- модель в процесі наукового пізнання є заступником досліджуваного об'єкта (умова репрезентації );
- вивчення моделі дозволяє отримати інформацію (відомості) про оригінал (умови екстраполяції ).

#### 2.1. Властивості моделей

При побудові моделі дослідник завжди виходить з поставлених цілей, враховує тільки найбільш істотні для їх досягнення фактори. Тому будь – яка модель нетотожна об'єкту - оригіналу і, отже, неповна, оскільки при її побудові дослідник враховував лише найважливіші з його точки зору чинники. Інші фактори, незважаючи на свій відносно малий вплив на поведінку об'єкта в порівнянні з вибраними факторами, у сукупності все може призводити до значних, розбіжностей між об'єктом і його моделлю.

Якщо результати моделювання задовольняють дослідника і можуть служити основою для прогнозування поведінки або властивостей досліджуваного об'єкта, то говорять, що модель адекватна (від лат.



Adaequatus - прирівняний) об'єкту. При цьому адекватність моделі залежить від цілей моделювання і прийнятих критеріїв.

В якості однієї з характеристик моделі може виступати простота (або складність) моделі. Очевидно, що з двох моделей, що дозволяють досягти бажаної мети та одержати необхідні результати з заданою точністю, перевага повинна бути віддана більш простій. При цьому адекватність і простота моделі далеко не завжди є суперечливими вимогами. Враховуючи безкінечну складність будь-якого об'єкта дослідження, можна припустити існування нескінченної послідовності його моделей, що розрізняються за ступенем повноти, адекватності та простоти.

В якості ще однієї властивості моделі можна розглядати потенційність моделі (від лат. *Potential* - міць, сила), або передбачувальний з позицій можливості отримання нових знань про досліджуваний об'єкт.

В наукових дослідженнях моделі, що не володіють визначеною «передбачуваністю», навряд чи можуть вважатися задовільними.

Відомо чимало випадків, коли вивчення або використання моделей дозволило зробити відкриття. Як приклад можна навести відкриття планети Нептун, положення якої було передбачене французьким астрономом Лаверов на підставі розрахунків, виконаних з використанням закону всесвітнього тяжіння (тобто моделі) і даних про рух планети Уран.

## 2.2. Цілі моделювання

Добре побудована модель, як правило, доступніша, інформативніша і зручніша для дослідника, ніж реальний об'єкт. Найважливішим і найбільш поширеним призначенням моделей є їх застосування при вивченні та прогнозуванні поведінки складних процесів і явищ. Потрібно враховувати, що деякі об'єкти і явища взагалі не можуть бути вивчені безпосереднім чином.

Інше, не менш важливе, призначення моделей полягає в тому, що з їх допомогою виявляються найбільш суттєві фактори, формують ті чи інші властивості об'єкта, оскільки сама модель відображає лише деякі основні характеристики вихідного об'єкта, облік яких необхідний при дослідженні того чи іншого процесу чи явища.

Модель дозволяє навчитися правильно управляти об'єктом шляхом апробування різних варіантів управління. Використовувати для цього реальний об'єкт часто буває ризиковано або просто неможливо. Наприклад, отримати перші навички в управлінні сучасним літаком безпечніше, швидше і дешевше на тренажері (тобто моделі), ніж піддавати себе і дорогої машину ризику.

Отже, модель потрібна для того, щоб:

- 1) зрозуміти, як влаштований конкретний об'єкт: якою є його структура, внутрішні зв'язки, основні властивості, закони розвитку, саморозвитку і взаємодії з навколишнім середовищем;

- 2) навчитися управляти об'єктом або процесом, визначати найкращі способи управління при заданих цілях і критеріях;
- 3) прогнозувати прямі і непрямі наслідки реалізації заданих способів і форм впливу на об'єкт.

### **2.3 . Класифікація видів моделей систем**

В основі моделювання лежить теорія подібності, яка стверджує, що абсолютна подібність може мати місце лише при заміні одного об'єкта іншим точно таким же. При моделюванні абсолютна подібність не має місця і прагне до того, щоб модель досить добре відображала досліджуваний бік функціонування об'єкта.

В якості одного з перших ознак класифікації видів моделювання можна вибрати ступінь повноти моделі і розділити моделі відповідно з цією ознакою на повні, неповні та наближені. В основі повного моделювання лежить повна подібність, яка проявляється як у часі, так і в просторі. Для неповного моделювання характерна неповна подібність моделі досліджуваного об'єкта. В основі наближеного моделювання лежить наближена подібність, при якій деякі сторони функціонування реального об'єкта не моделюються зовсім. Класифікація видів моделювання систем  $S$  наведена на рис. 2.1. Залежно від характеру досліджуваних процесів у системі  $S$  всі види моделювання можуть бути розділені на детерміновані та стохастичні, статичні і динамічні, дискретні, безперервні і дискретно-безперервні. Детерміноване моделювання відображає детерміновані процеси, тобто процеси, в яких передбачається відсутність усяких випадкових впливів; Стохастичне моделювання відображає імовірнісні процеси та події. У цьому випадку аналізується ряд реалізацій випадкового процесу й оцінюються середні характеристики, тобто набір однорідних реалізацій.

Статичне моделювання служить для опису поведінки об'єкта в який-небудь момент часу, а динамічне моделювання відображає поведінку об'єкта в часі . Дискретне моделювання служить для опису процесів, які передбачаються дискретними, відповідно безперервне моделювання дозволяє відобразити безперервні процеси в системах, а дискретно-безперервне моделювання використовується для випадків, коли хочуть виділити наявність як дискретних, так і безперервних процесів.

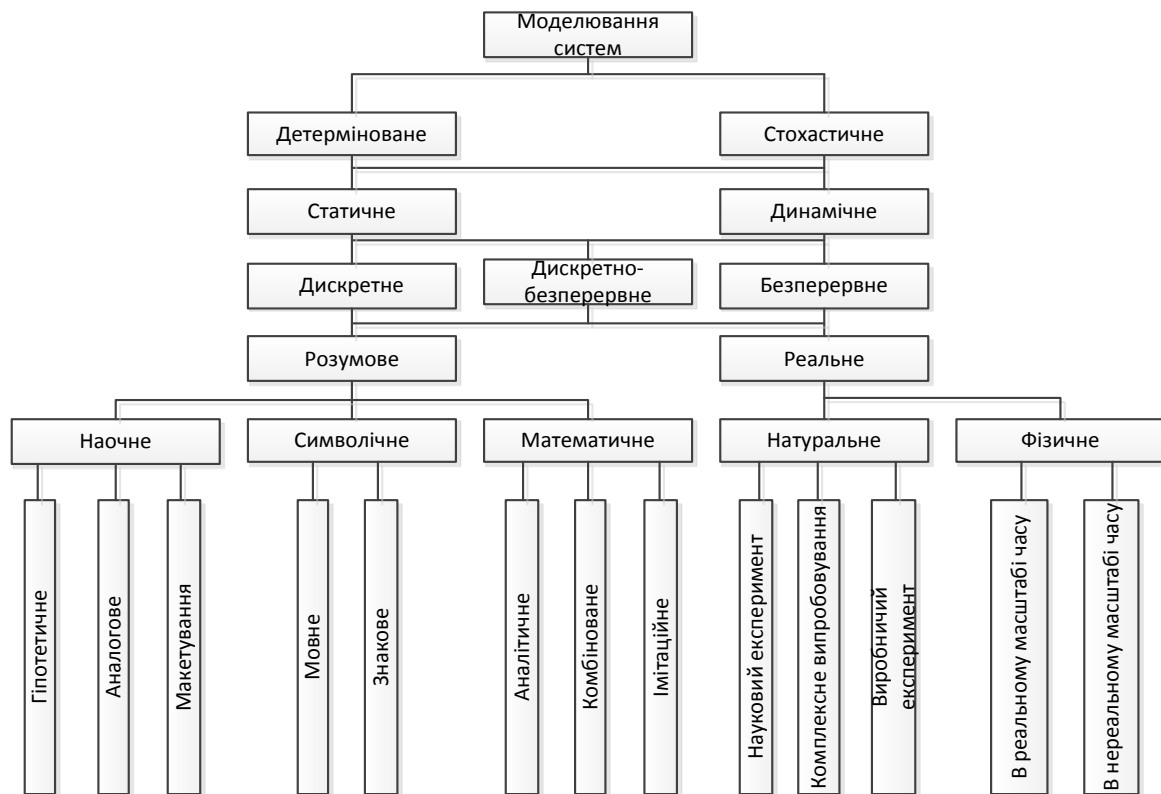


Рис. 2.1. Класифікація видів моделей систем

Залежно від форми подання об'єкт можна виділити уявне і реальне моделювання. Уявне моделювання часто є єдиним способом моделювання об'єктів, які або практично не реалізуються в заданому інтервалі часу, або існують поза умовами, можливих для їх фізичного створення. Наприклад, на базі уявного моделювання можуть бути проаналізовані багато ситуацій мікросвіту, які не піддаються фізичному експерименту. Уявне моделювання може бути реалізовано у вигляді наочного, символічного і математичного. При наочному моделюванні на базі уявлень людини про реальні об'єкти створюються різні наочні моделі, що відображають явища і процеси, що протікають в об'єкті .

В основу гіпотетичного моделювання дослідником закладається деяка гіпотеза про закономірності протікання процесу в реальному об'єкті, яка відображає рівень знань дослідника про об'єкт і базується на причинно-наслідкових зв'язках між входом і виходом досліджуваного об'єкта. Гіпотетичне моделювання використовується, коли знань про об'єкт недостатньо для побудови формальних моделей.

Аналогове моделювання ґрунтується на застосуванні аналогій різних рівнів. Найвищим рівнем є повна аналогія, що має місце тільки для досить простих об'єктів. З ускладненням об'єкта використовують аналогії наступних рівнів, коли аналогова модель відображає кілька або тільки одну сторону функціонування об'єкта. Істотне місце при уявному наочному моделюванні займає макетування. Уявний макет може застосовуватися у випадках, коли

протікають в реальному об'єкті процеси не піддаються фізичному моделюванню, або може передувати проведення інших видів моделювання. В основі побудови уявних макетів також лежать аналогії, проте зазвичай базуються на причинно-наслідкових зв'язках між явищами і процесами в об'єкті.

Якщо ввести умовне позначення окремих понять, тобто знаки, а також певні операції між цими знаками, то можна реалізувати знакове моделювання і за допомогою знаків відображати набір понять - скласти окремі ланцюжка з слів і пропозицій. Використовуючи операції об'єднання, перетину і доповнення теорії множин, можна в окремих символах дати опис якогось реального об'єкта. В основі мовного моделювання лежить деякий тезаурус. Останній утворюється з набору вхідних понять, причому цей набір повинен бути фіксованим.

Слід зазначити, що між тезаурусом і звичайним словником є принципові відмінності. Тезаурус – словник, який очищений від неоднозначності, тобто в ньому кожному слову може відповідати лише єдине поняття, хоча в звичайному словнику одному слову можуть відповідати кілька понять.

Символічне моделювання являє собою штучний процес створення логічного об'єкта, який заміщає реальний і виражає основні властивості його відносин з допомогою певної системи знаків або символів. Математичне моделювання. Для дослідження характеристик процесу функціонування будь-якої системи  $S$  математичними методами, включаючи і машинні, повинна бути проведена формалізація цього процесу, тобто побудовано математичну модель.

Під математичним моделюванням будемо розуміти процес встановлення відповідності даному реальному об'єкту деякого математичного об'єкта, званого математичною моделлю, і дослідження цієї моделі, що дозволяє отримувати характеристики розглянутого реального об'єкта. Вид математичної моделі залежить як від природи реального об'єкта, так і завдань дослідження об'єкта і необхідної достовірності і точності вирішення цього завдання.

Будь-яка математична модель, як і всяка інша, описує реальний об'єкт лише з деякою мірою наближення до дійсності. Математичне моделювання для дослідження характеристик процесу функціонування систем можна розділити на аналітичне, імітаційне і комбіноване. Для аналітичного моделювання характерно те, що процеси функціонування елементів системи записуються у вигляді деяких функціональних співвідношень (алгебраїчних, інтегро-дифференціальних, звичайно - різницевого тощо) або логічних умов.

Аналітична модель може бути досліджена такими методами:

- аналітичним, коли прагнуть отримати в загальному вигляді явні залежності для потрібних характеристик;

- чисельним, коли, не вміючи вирішувати рівнянь у загальному вигляді, прагнуть отримати числові результати при конкретних початкових даних;

- якісним, коли, не маючи рішення в явному вигляді, можна знайти деякі властивості рішення.

Найбільш повне дослідження процесу функціонування системи можна провести, якщо відомі явні залежності, що зв'язують потрібні характеристики з початковими умовами, параметрами і змінними системи  $S$ . Однак такі залежності вдається отримати тільки для порівняно простих систем. При ускладненні систем дослідження їх аналітичним методом нашоухується на значні труднощі, які часто бувають непереборними. Тому, бажаючи використати аналітичний метод, в цьому випадку йдуть на істотне спрощення первинної моделі, щоб мати можливість вивчити хоча б загальні властивості системи. Таке дослідження на спрощеній моделі аналітичним методом допомагає отримати орієнтовні результати для визначення більш точних оцінок іншими методами. Чисельний метод дозволяє досліджувати порівняно з аналітичним методом більш широкий клас систем, але при цьому отримані рішення носять приватний характер. Чисельний метод особливо ефективний при використанні ЕОМ.

В окремих випадках дослідження системи можуть задовольнити і ті висновки, які можна зробити при використанні якісного методу аналізу математичної моделі. Такі якісні методи широко використовуються, наприклад, в теорії автоматичного управління для оцінки ефективності різних варіантів систем управління. В даний час поширені методи машинної реалізації дослідження характеристик процесу функціонування великих систем. Для реалізації математичної моделі на ЕОМ необхідно побудувати відповідний моделюючий алгоритм. При імітаційному моделюванні реалізує модель алгоритм відтворює процес функціонування системи  $S$  в часі, причому імітуються елементарні явища, що становлять процес, зі збереженням їх логічної структури і послідовності протікання в часі, що дозволяє за вихідними даними отримати відомості про стани процесу в певні моменти часу, що дають можливість оцінити характеристики системи  $S$ .

Основною перевагою імітаційного моделювання порівняно з аналітичним є можливість вирішення більш складних завдань. Імітаційні моделі дозволяють досить просто враховувати такі фактори, як наявність дискретних і безперервних елементів, нелінійні характеристики елементів системи, численні випадкові впливи та інші, які часто створюють труднощі при аналітичних дослідженнях. В даний час імітаційне моделювання - найбільш ефективний метод дослідження великих систем, а часто і єдиний практично доступний метод отримання інформації про поведінку системи, особливо на етапі її проектування. Коли результати, отримані при відтворенні на імітаційній моделі процесу функціонування системи  $S$ , є реалізаціями випадкових величин і функцій, тоді для знаходження характеристик процесу потрібно його багаторазове відтворення з подальшою

статистичною обробкою інформації і доцільно в якості методу машинної реалізації імітаційної моделі використовувати метод статистичного моделювання.

Спочатку був розроблений метод статистичних випробувань, що представляє собою чисельний метод, який застосовувався для моделювання випадкових величин і функцій, імовірнісні характеристики яких збігалися з рішеннями аналітичних завдань (така процедура отримала назву методу Монте -Карло ). Потім цей прийом стали застосовувати і для машинної імітації з метою дослідження характеристик процесів функціонування систем, схильних випадковим впливам , тобто з'явився метод статистичного моделювання. Таким чином, методом статистичного моделювання будемо надалі називати метод машинної реалізації імітаційної моделі, а методом статистичних випробувань (Монте -Карло) – чисельний метод рішення аналітичної задачі.

Метод імітаційного моделювання дозволяє вирішувати завдання аналізу великих систем  $S$ , включаючи завдання оцінки: варіантів структури системи, ефективності різних алгоритмів керування системою, впливу зміни різних параметрів системи.

Імітаційне моделювання може бути покладено також в основу структурного, алгоритмічного і параметричного синтезу великих систем, коли потрібно створити систему, із заданими характеристиками при певних обмеження, яка є оптимальною за деякими критеріями оцінки ефективності.

При вирішенні завдань машинного синтезу систем на основі їх імітаційних моделей крім розробки моделюючих алгоритмів для аналізу фіксованої системи необхідно також розробити алгоритми пошуку оптимального варіанту системи. Далі в методології машинного моделювання будемо розрізняти два основні розділи: статику і динаміку, основним змістом яких є відповідно питання аналізу та синтезу систем, заданих моделюючими алгоритмами.

Комбіноване (аналітико - імітаційне) моделювання при аналізі та синтезі систем дозволяє об'єднати гідності аналітичного та імітаційного моделювання. При побудові комбінованих моделей проводиться попередня декомпозиція процесу функціонування об'єкта на складові під процеси і для тих з них , де це можливо, використовуються аналітичні моделі, а для решти під процесів будуються імітаційні моделі. Такий комбінований підхід дозволяє охопити якісно нові класи систем, які не можуть бути досліджені з використанням тільки аналітичного та імітаційного моделювання окремо.

#### **2.4. Інші види моделювання**

При реальному моделюванні використовується можливість дослідження різних характеристик або на реальному об'єкті цілком, або на її частині. Такі дослідження можуть проводитися як на об'єктах, що працюють в нормальних режимах , так і при організації спеціальних режимів для оцінки характеристик, які цікавлять дослідника( при інших значеннях змінних і

параметрів , в іншому масштабі часу і т. д.). Реальне моделювання є найбільш адекватним , але при цьому його можливості з урахуванням особливостей реальних об'єктів обмежені.

*Приклад.* Проведення реального моделювання АСУ ( Автоматична Система Управління ) підприємством зажадає, по-перше, створення такої АСУ, а по-друге, проведення експериментів з керованим об'єктом, тобто підприємством, що в більшості випадків неможливо.

Натурним моделюванням називають проведення дослідження на реальному об'єкті з подальшою обробкою результатів експерименту на основі теорії подібності . При функціонуванні об'єкта відповідно до поставленої мети вдається виявити закономірності перебігу реального процесу. Потрібно відзначити, що такі різновиди натурального експерименту, як виробничий експеримент і комплексні випробування, володіють високим ступенем вірогідності.

З розвитком техніки і проникненням в глиб процесів, що протікають в реальних системах, зростає технічна оснащеність сучасного наукового експерименту. Він характеризується широким використанням засобів автоматизації проведення, застосуванням досить різноманітних засобів обробки інформації, можливістю втручання людини в процес проведення експерименту, та відповідно з цим з'явився новий науковий напрям - автоматизація наукових експериментів.

Відмінність експерименту від реального протікання процесу полягає в тому, що в ньому можуть з'явитися окремі критичні ситуації і визначатися межі стійкості процесу. У ході експерименту вводяться нові чинники і впливи в процесі функціонування об'єкта. Один з різновидів експерименту - комплексні випробування, які також можна віднести до натурального моделювання, коли внаслідок повторення випробувань виробів виявляються загальні закономірності про надійність цих виробів, про характеристики якості і т.д. У цьому випадку моделювання здійснюється шляхом обробки та узагальнення відомостей, що проходять в групі однорідних явищ .

Поряд із спеціально організованими випробуваннями можлива реалізація натурального моделювання шляхом узагальнення досвіду, накопиченого в ході виробничого процесу, тобто можна говорити про виробничий експеримент. Тут на базі теорії подібності обробляють статистичний матеріал по виробничому процесу і отримують його узагальнені характеристики .

Іншим видом реального моделювання є фізичне, відмінне від натурального тим, що дослідження проводиться на установках, які зберігають природу явищ і володіють фізичною подібністю. У процесі фізичного моделювання задаються деякі характеристики зовнішнього середовища і досліджується поведінка або реального об'єкта, або його моделі при заданих або створюваних штучно впливах зовнішнього середовища.

Фізичне моделювання може протікати в реальному і нереальному (псевдо реальному) масштабах часу, а також може розглядатися без

урахування часу. В останньому випадку вивченню підлягають так звані "заморожені" процеси, які фіксуються в деякий момент часу. Найбільша складність і інтерес з точки зору вірності одержуваних результатів являє фізичне моделювання в реальному масштабі часу. З точки зору математичного опису об'єкта і залежно від його характеру моделі можна розділити на моделі аналогові (безперервні), цифрові (дискретні) і аналого - цифрові (комбіновані).

Під аналоговою моделлю розуміється модель, яка описується рівняннями, що зв'язують безперервні величини. Під цифровою розуміють модель, яка описується рівняннями, що зв'язують дискретні величини, представлені в цифровому вигляді. Під аналого - цифровою розуміється модель, яка може бути описана рівняннями, що зв'язують безперервні і дискретні величини.

Особливе місце в моделюванні займає кібернетичне моделювання, в якому відсутня безпосередня подібність фізичних процесів, що відбуваються в моделях, реальним процесам. У цьому випадку прагнуть відобразити лише деяку функцію і розглядають реальний об'єкт як "чорний ящик", що має ряд входів і виходів, і моделюють деякі зв'язки між виходами і входами. Найчастіше при використанні кібернетичних моделей проводять аналіз поведінкової сторони об'єкта при різних впливах зовнішнього середовища.

Таким чином, в основі кібернетичних моделей лежить відображення деяких інформаційних процесів управління, що дозволяє оцінити поведінку реального об'єкта. Для побудови імітаційної моделі в цьому випадку необхідно виділити досліджувану функцію реального об'єкта, спробувати формалізувати цю функцію у вигляді деяких операторів зв'язку між входом і виходом і відтворити на імітаційній моделі дану функцію, причому на базі зовсім інших математичних співвідношень і, природно, іншій фізичній реалізації процесу.

## 2.5 . Математичне моделювання

Історія розвитку та застосування системного підходу пов'язана з методами опису систем. Вони класифікуються в порядку зростання формалізованості - від якісного опису на початковому етапі до кількісного системного моделювання із застосуванням ЕОМ. Поділ методів на якісні та кількісні носить досить умовний характер. У якісних методах основна увага приділяється гіпотезам і аналогіям. Вони застосовуються тоді, коли відсутні описи закономірностей систем у вигляді аналітичних залежностей. Якісний аналіз ведеться, як правило, на природній мові.

**Гіпотеза** - певні передбачення, що ґрунтуються на певній кількості досвідчених даних, спостережень, здогадів і т.ін.

**Аналогією** називають судження про будь-яку приватну схожість двох об'єктів, яке не завжди може бути виражене в кількісних оцінках. Сучасна наукова гіпотеза створюється, як правило, за аналогією з перевіреними на



практиці науковими положеннями. Таким чином, аналогія пов'язує гіпотезу з експериментом.

До якісних методів належать: методи типу мозкової атаки або колективної генерації ідей, методи типу сценарії, методи експертних оцінок, морфологічні методи, коли систематично знаходяться всі «мислимі» варіанти вирішення проблеми або реалізації системи шляхом комбінування виділених елементів і їх ознак.

Кількісні методи передбачають певні експериментальні дослідження, аналіз результатів з метою оцінки коректності прийнятих на першому етапі гіпотез. Таким чином здійснюється перехід від словесного опису системи до формального подання процесу та його функціонування у вигляді деякої математичної моделі. Якщо експерименти проводяться на ЕОМ, то їх називають машинними експериментами, а сам процес проведення експериментів імітаційним моделюванням. Якщо експерименти проводяться безпосередньо на реальному об'єкті, то такі експерименти називаються натурними експериментами. Численні дослідження і розрахунки пов'язані з вибором оптимальної структури системи і моделі, оптимальних значень параметрів. Виконання таких досліджень можливе лише за наявності математичного опису процесу функціонування системи, тобто її математичної моделі. У цьому випадку розглядають математику не тільки як метод розрахунку, але і як метод мислення.

При побудові математичних моделей процесів функціонування систем існують такі основні підходи: безперервно-детермінований (наприклад, диференціальні рівняння, рівняння стану); дискретно-детермінований (кінцеві автомати); дискретно-стохастичний (імовірнісні автомати); безперервно-стохастичний (системи масового обслуговування); узагальнений або універсальний (агрегативна системи).

Класифікація моделей та видів моделювання об'єктів і систем у відповідності з теорією подібності повинна виділити в них найбільш загальні ознаки і властивості реальних систем. Нижче наведена одна з можливих класифікацій.

<b>Ознаки класифікації</b>	<b>Види математичних моделей</b>
1. Приналежність до ієрархічного рівня	<ul style="list-style-type: none"> <li>▪ Моделі мікрорівня</li> <li>▪ Моделі макрорівня</li> <li>▪ Моделі мета рівня</li> </ul>
2. Характер взаємовідносин з середовищем	<ul style="list-style-type: none"> <li>▪ Відкриті (непереривний обмін)</li> <li>▪ Закриті (слабкий зв'язок)</li> </ul>
3. Характер відображуваних властивостей об'єкта	<ul style="list-style-type: none"> <li>▪ Структурні</li> <li>▪ Функціональні</li> </ul>
4. Спосіб подання властивостей об'єкта	<ul style="list-style-type: none"> <li>▪ Аналітичні</li> <li>▪ Алгоритмічні</li> <li>▪ Імітаційні</li> </ul>
5. Спосіб отримання моделі	<ul style="list-style-type: none"> <li>▪ Теоретичні</li> <li>▪ Емпіричні</li> </ul>
6. Причинна обумовленість	<ul style="list-style-type: none"> <li>▪ Детерміновані</li> <li>▪ Імовірнісні</li> </ul>
7. По відношенню до часу	<ul style="list-style-type: none"> <li>▪ Динамічні</li> <li>▪ Статичні</li> </ul>
8. За типом рівнянь	<ul style="list-style-type: none"> <li>▪ Лінійні</li> <li>▪ Нелінійні</li> </ul>
9. За безліччю значень змінних	<ul style="list-style-type: none"> <li>▪ Безперервні</li> <li>▪ Дискретні</li> <li>▪ Дискретно-безперервні</li> </ul>
10. За призначенням	<ul style="list-style-type: none"> <li>▪ Технічні</li> <li>▪ Економічні</li> <li>▪ Соціальні і т.п.</li> </ul>

## **2.6 . Особливості системного моделювання**

Системне моделювання являє собою сукупність конкретних різновидів моделювання, найбільш важливі серед яких:

- атрибутивний, спрямоване на систематизацію інформації про властивості об'єктів. При цьому використовуються різного роду класифікації, матриці, таблиці, які дозволяють систематизувати властивості об'єктів, виділити головні і другорядні;
- структурний, що забезпечує подання структури об'єкта або процесу моделювання;
- організаційний, що припускає вивчення організації системи;

- функціональний, орієнтоване на побудову і дослідження функцій досліджуваного явища;
- структурно - функціональний, що ставить своєю ціллю дослідженню взаємозв'язку структури і функції досліджуваного об'єкта чи процесу;
- вітальний, спрямоване на представлення та вивчення тих чи інших етапів життєвого шляху системи.

Системне моделювання не обмежується задоволенням простої цікавості по відношенню до моделі. Воно дуже прагматичне. Його найважливішим призначенням виступає не просто отримання знань про систему, а її оптимізація. Це пошук оптимуму характеристик системи відповідно з деякими критеріями оптимальності. Математика оперує поняттям "оптимуму функції". Оптимум функції  $f(x)$  на множині  $M$  приватне значення  $f(x_0)$  цієї функції, що задовольняє одному з співвідношень:  $f(x_0)$  більше і дорівнює  $f(x)$  для всіх  $x$  змін (глобальний максимум) або для всіх  $f(x_0)$  менше і дорівнює  $f(x)$  для всіх  $x$  з  $M$  (глобальний мінімум). Точка оптимума функції  $f(x)$  на безліч  $M$  є однією з точок екстремуму цієї функції на множині  $M$ .

Системне моделювання орієнтоване на пошук в системній моделі оптимальних характеристик з метою перетворення за принципом оптимальності реальних об'єктів практичної діяльності людей.

## 2.7 . Властивості системи з точки зору моделювання

**Стан системи** - це сукупність значень її показників. Всі можливі стани системи утворюють її безліч станів. Якщо в цій безлічі визначено поняття близькості елементів, то воно називається простором станів.

**Рух (поведінка) системи** - це процес переходу системи з одного стану в інший, з нього в третє і т.д. Якщо перехід системи з одного стану в інший відбувається без проходження будь-яких проміжних станів, то система називається дискретною. Якщо при переході між будь-якими двома станами система обов'язково проходить через проміжний стан, то вона називається динамічною (безперервною).

Розрізняють такі режими руху системи :

- **рівноважний**, коли система знаходиться весь час в одному і тому ж стані;
- **періодичний**, коли система через рівні проміжки часу проходить одні й ті ж стани;
- **перехідний режим** - рух системи між двома періодами часу, в кожному з яких система перебувала в стаціонарному режимі;
- **аперіодичний режим** - система проходить безліч станів, однак закономірність проходження цих станів є більш складною, ніж періодичні, наприклад, змінний період ;

- **ергодичний режим** - система проходить весь простір станів таким чином, що з плином часу проходить скільки завгодно близько до будь-якого заданого стану .

Якщо система знаходиться в рівноважному або періодичному режимі, то говорять , що вона знаходиться в сталому або стаціонарному режимі.

Підсумовуючи сказане вище , перерахуємо властивості систем :

- **Цілісність** - поява нової якості в об'єднанні саме цього набору елементів. Важливо довести цілісність втратою системних якостей при виключенні будь-якого з виділених елементів системи.

- **Різноманітність** - наявність якісно різних елементів системи, несучих різні функції .

- **Зв'язність** - здійснення обміну інформацією між елементами системи, неможливість включення в систему елементів без інформаційного обміну.

- **Цілеспрямованість** - можливість управління системою шляхом зміни параметрів в одному елементі для перетворення стану інших.

- **Стійкість** - (гомеостаз) здатність збереження властивостей системи при досить широкій зміні параметрів середовища .

## 2.8. Етапи моделювання та загальні вимоги до моделей

Моделювання в цілому включає в себе ряд етапів, що базуються на системному підході :

### 1.Змістовна постановка задачі:

- вироблення загального підходу до досліджуваної проблеми; визначення під задач;

- визначення основної мети і шляхів її досягнення.

### 2.Вивчення та збір інформації про об'єкт - оригінал:

- аналіз або підбір відповідних гіпотез, аналогій, теорій;
- облік досвідчених даних, спостережень і т.д.; визначення вхідних і вихідних змінних, зв'язків, прийняття спрощуюючих припущення.

### 3.Формалізація:

- приймаються умовні позначення і з їх допомогою описуються зв'язки між елементами об'єкта у вигляді математичних виразів. Намічується перехід до кількісного аналізу .

4.Вибір методу рішення. Для поставленої математичної задачі обґрунтовується метод її вирішення з урахуванням знань і уподобань користувача і розробника. При проектуванні доводиться вирішувати як лінійні, так і нелінійні задачі , використовувати ручні та машинні методи проектування, розрахунку і досліджень.

5.Реалізація моделі. Приймається критерій оцінки ефективності моделі, розробляється алгоритм, пишеться та налагоджується програма , щоб здійснити системний аналіз і синтез.

*б.Аналіз отриманих результатів.* Зіставляється передбачуване і отримане рішення, проводиться оцінка адекватності та похибки моделювання. Процес моделювання є ітеративним. У разі незадовільних результатів, отриманих на етапах 5 або 6, здійснюється повернення до одного з ранніх етапів, який міг призвести до розробки невдалої моделі. Уточнення моделі відбувається до тих пір, поки не будуть отримані прийнятні результати .

Таким чином, після проходження цих етапів найбільш повно можуть бути виконані вимоги, що пред'являються до моделей:

- **Універсальність** - характеризує повноту відображення моделлю досліджуваних властивостей реального об'єкта;
- **Адекватність** - здатність відобразити потрібні властивості об'єкта з похибкою не вище допустимої;
- **Точність** - оцінюється ступенем збігу значень характеристик реального об'єкта зі значеннями цих характеристик, отриманих за допомогою моделей;
- **Економічність** - визначається витратами ресурсів ЕОМ (пам'яті і часу на її реалізацію та експлуатацію).

## **2.9. Якість моделювання**

Якість моделювання може бути оцінена характеристикою його споживчих властивостей:

- *ефективність використання його за призначенням (мети) ;*
- *ресурсоемність;*
- *вартість .*

Ці характеристики (показники) у розгорнутому вигляді представлені на рис.2.9.1

Математичний підхід до моделювання має ряд недоліків:

- *низька адекватність математичної моделі реальному об'єкту ;*
- *проблеми, пов'язані з вирішуванням математичних моделей через наявність у них розривних функцій;*
- *непридатність математичних моделей для більшості об'єктів із змінною структурою;*
- *наближені методи реалізації моделей зі змінними коефіцієнтами вимагають значних витрат і не володіють достатньою точністю рішення.*

В даний час імітаційне моделювання в основному реалізується на ЕОМ. Оригінальний математичний опис будь-якої динамічної системи являє собою сукупність диференціальних, алгебраїчних, логічних, різницевих рівнянь, що описують фізичні процеси в окремих функціональних елементах системи



Рис. 2.9.1. Характеристики якості моделювання

## 2.10. Принцип подібності при математичному моделюванні

У процесі наукових досліджень роботи механічних систем представляється доцільним замінити об'єкти дослідження їх моделями (розрахунковими схемами). Моделювання досліджуваних процесів і елементів систем можливо при дотриманні законів схожості. Подібними явищами і елементами систем називають такі, при яких існує відповідність схожість величин. У найпростішому вигляді співвідношення подібності можна записати в такому вигляді:

$$\frac{H_i}{R_i} = v_i \quad (2.10.1)$$

де  $H_i$  і  $R_i$  — схожі параметри процесів та елементів досліджуваних систем;  $v_i$  — коефіцієнт схожості, або масштаб схожих параметрів. Слід зазначити, що величина  $v_i$ , може бути змінною, що залежить від будь-якого параметра процесу. Однак у більшості випадків досліджень величину  $v_i$  приймають постійною.

Для механічних систем розрізняють: геометричну, матеріальну, кінематичну та динамічну схожість. В загальному виді, якщо досліджуваний об'єкт виражений однорідними функціями виду

$$\varphi_1 + \varphi_2 + \varphi_3 + \dots + \varphi_n = \sum_1^n \varphi_i = 0, \quad (2.10.2)$$

а подібний об'єкт, також описується однорідними функціями

$$\Psi_1 + \Psi_2 + \Psi_3 + \dots + \Psi_n = \sum_1^n \Psi_i = 0, \quad (2.10.3)$$

та при цьому  $\varphi_i = f(H_1, H_2, \dots, H_m)$ , а  $\Psi_i = f(R_1, R_2, \dots, R_m)$  можливо міркувати таким чином. Уявімо вирази (2.10.2) і (2.10.3) в безрозмірній формі

$$1 + \frac{\varphi_1}{\varphi_n} + \frac{\varphi_2}{\varphi_n} + \dots + \frac{\varphi_{n-1}}{\varphi_n} = 1 + \sum_{i=1}^{n-1} \frac{\varphi_i}{\varphi_n} = 0; \quad (2.10.4)$$

$$1 + \frac{\Psi_1}{\Psi_n} + \frac{\Psi_2}{\Psi_n} + \dots + \frac{\Psi_{n-1}}{\Psi_n} = 1 + \sum_{i=1}^{n-1} \frac{\Psi_i}{\Psi_n} = 0; \quad (2.10.5)$$

Величини  $H_1, H_2, \dots, H_m$  и  $R_1, R_2, \dots, R_m$  подібні параметри, і оскільки системи подібні, можна записати

$$H_1 = v_1 R_1; H_2 = v_2 R_2; \dots; H_m = v_m R_m.$$

Підставляючи виписані співвідношення у вираз (2.10.2) і виносячи масштаби за знак функції у вигляді множників отримаємо

$$\varphi_i = f(H_1, H_2, \dots, H_m) = f(v_i R_i, v_2 R_2, \dots, v_m R_m),$$

$$\text{або } \varphi_i = N_i f(R_i, R_2, \dots, R_m) = N_i \psi_i, \quad (2.10.6)$$

де  $i=1, 2, 3, \dots$

Тепер, підставляючи вираз (2.10.6) у співвідношення (2.10.4).

$$1 + \frac{N_1}{N_n} \frac{\Psi_1}{\Psi_n} + \frac{N_2}{N_n} \frac{\Psi_2}{\Psi_n} + \dots + \frac{N_{n-1}}{N_n} \frac{\Psi_{n-1}}{\Psi_n} = 0.$$

Внаслідок однорідності виразу (2.10.2) загальні множники  $N_i$ , - для кожного члена  $\Psi_i$  рівні, тобто  $N_1 = N_2 = N_3 = \dots = N_n$  або  $\frac{N_1}{N_n} = \frac{N_2}{N_n} = \dots = \frac{N_{n-1}}{N_n} = 1$ . Тоді,

очевидно, вирази (2.10.4) і (2.10.5) тотожні і існують співвідношення

$$\frac{\varphi_1}{\varphi_n} = \frac{\Psi_1}{\Psi_n}; \frac{\varphi_2}{\varphi_n} = \frac{\Psi_2}{\Psi_n}; \dots; \frac{\varphi_{n-1}}{\varphi_n} = \frac{\Psi_{n-1}}{\Psi_n}. \quad (2.10.7)$$

Або узагальнюючи на число подібних об'єктів дослідження, маємо:

$$\frac{\varphi_i^{(1)}}{\varphi_n^{(1)}} = \frac{\varphi_i^{(2)}}{\varphi_n^{(2)}} = \dots = \frac{\varphi_i^{(j)}}{\varphi_n^{(j)}} = Idem^*, \quad (2.10.8)$$

де  $j=1, 2, 3, \dots$

Наведені вище безрозмірні співвідношення (2.10.7), чисельно однакові для подібних об'єктів дослідження, називаються критеріями подібності. Властивість критеріїв подібності  $K$  як умови подібності досліджуваних об'єктів можна коротко записати у вигляді  $K = Idet$ . Критерії подібності можуть перетворюватися в критерії інших форм шляхом їх перемноження, ділення, зведення в ступінь, множення на постійну величину і т. п. Ці дії можна записати у вигляді:

$$K_n K_{n+i} = Idem; K_n^n = Idem;$$

$$\frac{K_n}{K_{n+i}} = Idem; \frac{1}{K_n} = Idem; K_n^m = Idem.$$

При вираженні досліджуваних параметрів у вигляді неоднорідних функцій коефіцієнти подібності не можна висловлювати, як це робилося вище, у вигляді загальних множників, так як аргументи неоднорідних функцій повинні бути рівні н бути також критеріями подібності. Нехай  $\varphi_1 = D \sin abc$  и

$\Psi_1 = D \sin ABC$ . Подоба буде існувати тільки при рівності  $abc=ABC$ . Якщо аргументи неоднорідних функції складаються з суми величина  $a+b+c$  и  $A+B+C$ , то умови подоби буде  $a=A, b=B, c=C$ . Символи диференціювання та інтегрування, що входять до вихідні рівняння, оскільки вони не мають розмірності, можуть бути при встановленні подібності відкинуті і замінені наступними виразами:

$$\frac{d^n y}{dx^n} = \frac{y}{x^n}; \quad \int y dx = yx.$$

Більш загальні положення теорії подібності коротко сформулюємо наступним чином:

1. Необхідною умовою подібності двох систем є рівність відповідних критеріїв подібності цих систем;
2. Будь-яке повне рівняння того чи іншого явища, записане в певній системі одиниць, може бути представлено залежністю між критеріями подібності;
3. Умовою подібності явищ має бути існування рівностей найбільш важливих й незалежних критеріїв та подібність умов однозначності (початкові та граничні умови, вплив зовнішнього середовища та інш.).

Крім перерахованих загальних положень теорії подібності, існує ряд додаткових положень, застосовуваних при розгляді складних систем з підсистемами, нелінійних і анізотропних систем, при імовірнісний характер досліджуваних явищ і т. п. У механіці для моделювання процесів використовують ряд класичних критеріїв подібності. Висновок деяких наводимо нижче. Відомий закон Ньютона, що описує рух матеріальної точки при дії сили  $F$ , коротко формулюється в такому вигляді:

$$F = m \frac{d^2 x}{dt^2}, \quad (2.10.9)$$

де  $m$  — маса матеріальної точки;  $x$  — її координата;  $t$  — час.

Ділячи ліву частину виразу (2.10.9) на праву, отримаємо критерій подібності

$$K_{Ne} = \frac{Ft^2}{mx} = Idem, \quad (2.10.10)$$

де  $K_{Ne}$  — критерій Ньютона.

Для обертового руху, аналогічно  $M = I \frac{d^2 \varphi}{dt^2}$ , де  $M$  — обертовий момент;  $\varphi$  та  $I$  — кут повороту і момент інерції маси обертової системи.

Критерій подоби  $K_{Ne} = \frac{Mt^2}{I\varphi}$ . Для систем матеріальних точок при наявності між ними зв'язків існує правило: якщо швидкості тіл з різною масою, переміщаються на однакові відстані, рівні, то діючі сили прямо пропорційні масам тіл. Привільному падінні тіл сила тяжкості

$$F = m \frac{d^2 x}{dt^2} = mg,$$



де  $m$  — маса тіла;  $g$  — прискорення вільного падіння. Критерій подоби буде

$$K = \frac{x}{gt^2} = Idem.$$

Множачи  $K$  на квадрат критерію гомохронності \*, отримаємо критерій подібності, відомі в літературі під назвою критерію Фруда,

$$K_{Fr} = \frac{V^2}{g_x} = Idem. \quad (2.10.11)$$

## 2.11. Модель чорного ящика

Виділена, відмінна від усього іншого, система не ізолювана від навколишнього середовища. Навпаки, вони пов'язані і обмінюються між собою будь-якими видами ресурсів (речовиною, енергією, інформацією та ін.). Позначимо цю особливість терміном "відкритість" системи і обговоримо цю властивість докладніше.

Відзначимо, що зв'язки системи з середовищем мають спрямований характер: за однією середовище впливає на систему (їх називають **входами** системи), за іншими система здійснює вплив на середовище, щось робить в середовищі, щось видає в середовище (такі зв'язки називають **виходами** системи) (рис. 2). Перелік входів і виходів системи називають **моделлю чорного ящика**. У цій моделі відсутня інформація про внутрішні особливості системи. Незважаючи на (уявну) простоту і бідність змісту моделі чорного ящика, ця модель часто цілком достатня для роботи з системою.

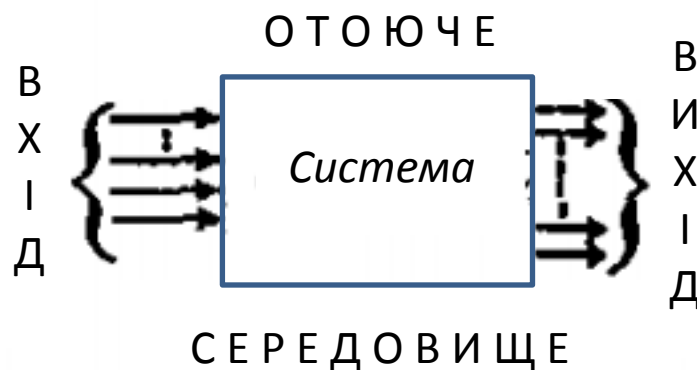


Рис. 2.3. Узагальнене графічне представлення моделі «чорного ящика»

У багатьох випадках керування технікою (автомобілем, радіоапаратурою, комп'ютером, приладом) або людьми (наприклад, в менеджменті) інформація тільки про вхід та вихід керованої системи дозволяє успішно досягати мети. Однак для цього модель повинна відповідати певним вимогам. Ви можете відчувати труднощі, якщо не знаєте,

що у деяких моделях телевізорів кнопку включення треба не натискати, а витягати, або що в деяких готелях вимикач в темному приміщенні суміщений з засувкою, і взагалі зустрівшись з приладом, не всі входи якого вам відомі. Ясно, що для успішного управління системою модель чорного ящика повинна містити всю інформацію, необхідну для досягнення мети. При спробі задовольнити цю вимогу виконавець зустрінеться з труднощами, які слід мати на увазі. Перерахуємо ці труднощі.

## 2.12. Труднощі побудови моделі чорного ящика

Всі вони виникають з того, що модель завжди містить кінцевий список зв'язків, тоді як їх число у реальної системи не обмежена. Виникає питання: які з них включати в модель, а які - ні? Відповідь ми вже знаємо: в моделі повинні бути відображені всі зв'язки, суттєві для досягнення мети. Але слово "суттєві" - оціночне! Оцінку може дати тільки суб'єкт. Але крім здатності оцінювати, суб'єкт володіє ще однією властивістю - здатністю іноді помилятися у своїх оцінках. Помилка в оцінці призведе до того, що модель не цілком відповідатиме вимогу адекватності, а значить, її використання призведе до ускладнень в роботі з системою.

Можливі чотири типи помилок при побудові моделі чорного ящика:

*Помилка першого роду* відбувається, коли суб'єкт розцінює зв'язок як істотну і приймає рішення про включення її в модель, тоді як насправді по відношенню до поставленої мети вона несуттєва і могла б бути неврахованою. Це призводить до появи в моделі "зайвих" елементів, по суті непотрібних.

*Помилка другого роду*, навпаки, відбувається суб'єктом, коли він приймає рішення, що даний зв'язок несуттєвий і не заслуговує бути включений в модель, тоді як насправді без нього наша мета не може бути досягнута в повній мірі або навіть зовсім.

Втрати, що наносяться помилкою першого роду, пов'язані з тим, що інформація, внесена нею, зайва. При роботі з такою моделлю доведеться витратити зайві ресурси на фіксацію та обробку зайвої інформації, наприклад, витратити на неї пам'ять машини і час обробки. На якості рішення це не позначиться, а на вартості та своєчасності - обов'язково.

Втрати від помилки другого роду - це втрата від того, що інформації для повного досягнення мети не вистачить, мета не може бути досягнута в повній мірі.

Гірше та помилка, втрати від якої більше. А це залежить від конкретних обставин. Наприклад, якщо час є критичним фактором, то помилка першого роду стає набагато більш небезпечною, ніж друга: вчасно прийняте, хай не найкраще, рішення переважніше оптимального, але запізненого.

**Помилка першого роду та помилка другого роду** в математичній статистиці — це ключеві поняття задач перевірки статичної гіпотези. Тим паче, данні поняття часто використовуються в інших областях, коли мова

йде про прийняття «бінарного» рішення (так/ні) на основі деякого критерію (тесту, перевірки, вимірювання), який з деякою ймовірністю може дати неправдивий результат.

Нехай дана вибірка  $\mathbf{X} = (X_1, \dots, X_n)^\top$  з невідомого спільного розподілу  $\mathbb{P}^{\mathbf{X}}$ , і поставлене бінарне завдання перевірки статистичних гіпотез:  $H_0$ ,  $H_1$ , де  $H_0$  — нульова гіпотеза, а  $H_1$  — альтернативна гіпотеза. Припустимо, що задано статистичний критерій  $f: \mathbb{R}^n \rightarrow \{H_0, H_1\}$ , зпівставляючи кожній реалізації вибірки  $\mathbf{X} = \mathbf{x}$  одну із наявних гіпотез. Тоді можливі наступні чотири ситуації:

Розподілення  $\mathbb{P}^{\mathbf{X}}$  вибірки  $\mathbf{X}$  відповідає гіпотезі  $H_0$ , та вона дійсно визначена статистичним критерієм, тобто  $f(\mathbf{x}) = H_0$ .

1. Розподілення  $\mathbb{P}^{\mathbf{X}}$  вибірки  $\mathbf{X}$  відповідає гіпотезі  $H_0$ , але вона мабуть знехтувана астатичними критеріями, тобто  $f(\mathbf{x}) = H_1$ .

2. Розподілення  $\mathbb{P}^{\mathbf{X}}$  вибірки  $\mathbf{X}$  відповідає гіпотезі  $H_1$ , і вона точно визначена статичними критеріями, тобто  $f(\mathbf{x}) = H_1$ .

3. Розподілення  $\mathbb{P}^{\mathbf{X}}$  вибірки  $\mathbf{X}$  відповідає гіпотезі  $H_1$ , але вона невірно знехтувана статичними критеріями, тобто  $f(\mathbf{x}) = H_0$ .

У другому і четвертому випадку говорять, що сталася статистична помилка, і її називають **помилкою першого і другого роду** відповідно.

		Вірна гіпотеза	
		$H_0$	$H_1$
Результат застосування критерію	$H_0$	$H_0$ вірно прийнята	$H_0$ невірно прийнята (Помилка <i>другого</i> роду)
	$H_1$	$H_1$ невірно відхилена (Помилка <i>першого</i> роду)	$H_1$ вірно відхилена

Як видно з вищенаведеного визначення, **помилки першого і другого роду** є взаємно - симетричними, тобто якщо поміняти місцями гіпотези  $H_0$  та  $H_1$ , то *помилки першого роду* перетворяться на *помилки другого роду* та навпаки. Тим не менш, в більшості практичних ситуацій плутанини не відбувається, оскільки прийнято вважати, що *нульова гіпотеза*  $H_0$  відповідає стану «за замовчуванням» (природному, найбільш очікуваному стану речей) - наприклад, що обстежуваний людина здорова, або що проходить через рамку метало детектора пасажир не має заборонених металевих предметів. Відповідно, *альтернативна гіпотеза*  $H_1$  позначає протилежну ситуацію, яка зазвичай трактується як менш вірогідна, неординарна, що вимагає будь-якої реакції.

З урахуванням цього *помилку першого роду* часто називають **помилковою тривоگوю, помилковим спрацьовуванням або хибно позитивним спрацьовуванням** - наприклад, аналіз крові показав наявність захворювання, хоча насправді людина здорова, або метало детектор видав сигнал тривоги, спрацювавши на металеву пряжку ременя. Слово «позитивний» в даному випадку не має відношення до бажаності або небажаності самої події .

В інших областях зазвичай використовують словосполучення зі схожим змістом, наприклад, «**помилкове спрацьовування**», «**помилкова тривога**» і т.п. В інформаційних технологіях часто використовують англійський термін *falsepositive* без перекладу.

Відповідно, *помилку другого роду* іноді називають **пропуском події або хибно негативним спрацьовуванням** - людина хвора, але аналіз крові цього не показав, або у пасажира є холодна зброя, але рамка метало детектора його не виявлено (наприклад, через те, що чутливість рамки відрегульована на виявлення тільки дуже масивних металевих предметів). Наведемо кілька прикладів практичного використання оцінок помилок моделі першого і другого роду.

**Приклад 1.** У задачі радіолокаційного виявлення повітряних цілей, насамперед, у системі ППО помилки першого і другого роду, з формулюванням «помилкова тривога» і « пропуск мети» є одним з основних елементів як теорії , так і практики побудови радіолокаційних станцій. Ймовірно, це перший приклад послідовного застосування статистичних методів в цілій технічній галузі.

**Приклад 2.** У завданнях фільтрація спаму *помилка першого роду* відбувається, коли механізм блокування/фільтрації спаму помилково класифікує легітимне e-mail-повідомлення як спам і перешкоджає його нормальній доставці. У той час як більшість «антиспам» - алгоритмів здатні блокувати/фільтрувати великий відсоток небажаних e-mail-повідомлень, набагато більш важливим завданням є мінімізувати число «помилкових тривог» ( помилкових блокувань потрібних повідомлень).

*Помилка другого роду* відбувається , коли антиспам- система помилково пропускає небажане повідомлення, класифікуючи його як « не спам ». Низький рівень таких помилок є індикатором ефективності антиспам-алгоритму . Поки не вдалося створити антиспамову систему без кореляції між імовірністю помилок першого і другого роду. Імовірність пропустити спам у сучасних систем коливається в межах від 1 % до 30 %. Імовірність помилково відкинути валідне повідомлення - від 0,001 % до 3 %. Вибір системи та її установок залежить від умов конкретного одержувача: для одних одержувачів ризик втратити 1 % гарною пошти оцінюється як незначний , для інших же втрата навіть 0,1 % є неприпустимою.

**Приклад 3.** У завданнях антивірусного тестування програмного забезпечення поняття помилки першого роду використовується , коли

антивірусне програмне забезпечення помилково класифікує нешкідливий файл як вірус . Невірне виявлення може бути викликано особливостями евристики , або неправильної сигнатурою вірусу в базі даних. Подібні проблеми можуть відбуватися також і з антитроянською і антишпигунською програмами.

*Помилкою третього роду* прийнято вважати наслідки незнання. Для того щоб оцінювати істотність деякому зв'язку, треба знати , що вона взагалі є. Якщо це невідомо , питання про включення або не включення її в модель взагалі не варто: в моделях є тільки те, що ми знаємо. Але від того, що ми не підозрюємо про існування якоїсь зв'язку, вона не перестає існувати і проявлятися в реальній дійсності. А далі все залежить від того, наскільки вона істотна для досягнення нашої мети. Якщо вона несуттєва, то ми в практиці і не помітимо її наявності в реальності і відсутності в моделі. Якщо ж вона істотна , ми будемо відчувати ті ж труднощі , що і при помилку другого роду . Різниця полягає в тому, що помилку третього роду важче виправити: треба здобувати нові знання.

*Помилка четвертого роду* може виникнути при неправильному віднесення відомою і визнаною істотного зв'язку до числа входів або виходів. Наприклад, жорстку кореляцію між урожайністю зернових і яйценіскістю курей можна тлумачити як вхід - то з них, що відомо, а вихід - те, що треба оцінити . Але ж можна злічити носіння певного головного убору входом , оскільки було точно встановлено, що в Англії минулого століття здоров'я чоловіків, що носять циліндри, було набагато краще, ніж здоров'я що носять кепки. А проблема симптомів і синдромів в медицині?

### 2.13. Внутрішня неоднорідність систем

Якщо заглянути всередину " чорного ящика", то з'ясується, що система не однорідна, не монолітна: можна виявити , що різні якості в різних місцях відрізняються. Опис внутрішньої неоднорідності системи зводиться до відокремлення щодо однорідних ділянок, проведення кордонів між ними. Так з'являється поняття про *частини системи* . При більш детальному розгляді виявляється, що виділені великі частини теж не однорідні, що вимагає виділяти ще більш дрібні частини. В результаті (рис.2.13.1) виходить **ієрархічний список** частин системи , який ми будемо називати *моделлю складу системи* .

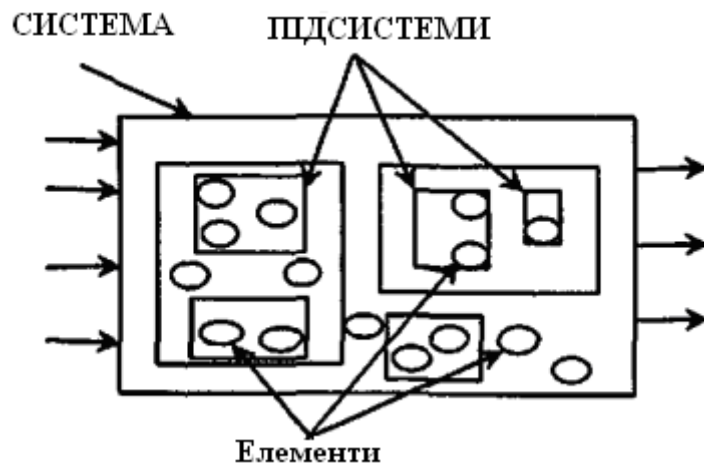


Рис. 2.13.1. Внутрішня неоднорідність систем

Інформація про склад системи може використовуватися для роботи з системою. Цілі взаємодії з системами можуть бути різними, у зв'язку з чим можуть різнитися і моделі складу однієї і тієї ж системи. Корисну, придатну для роботи модель створити непросто.

Поняття структури системи досить близьке до поняття її форми, але вони не тотожні. Структура - це лише один аспект поняття форми, внутрішньої організації змісту. Крім цього форма виступає і як вираження змісту, і як характеристика зовнішнього виду об'єкта (зовнішня форма).

Структуру системи можна характеризувати за такими загальними ознаками:

- кількість зв'язків та характер взаємодії елементів;
- інтенсивність зв'язків, частота (кількість зв'язків, що припадають на один елемент);
- кількість зовнішніх зв'язків, співвідношення між кількістю внутрішніх та зовнішніх зв'язків як характеристика відкритості системи.

Складність може бути: структурною, функціональною, динамічною, описовою (пізнавальною, гносеологічною) (рис. 2.13.2).

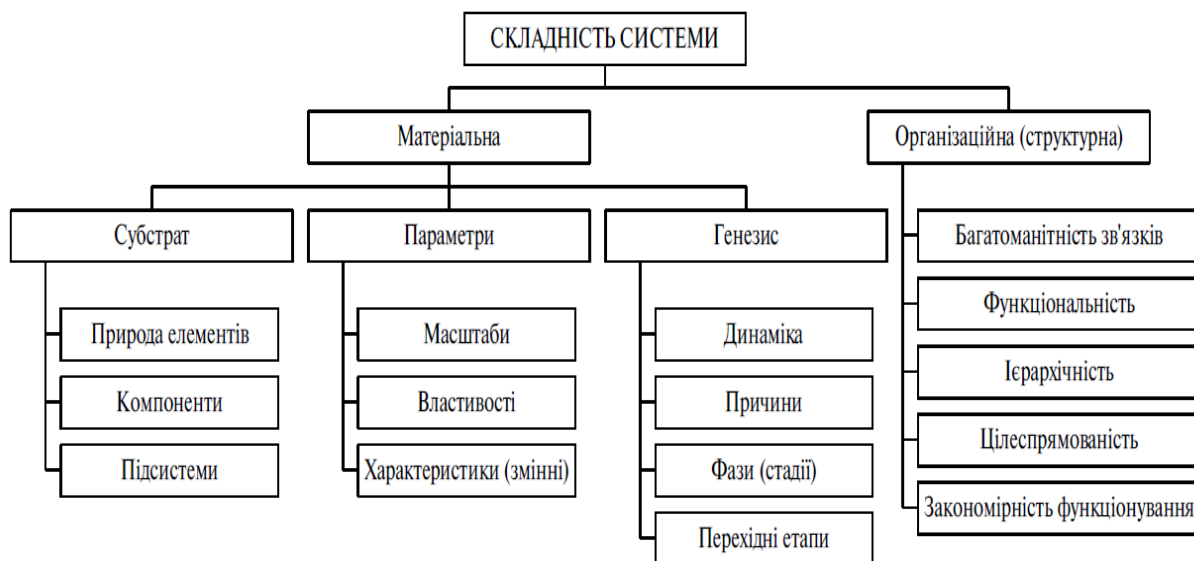


Рис. 2.13.2. Системне уявлення складності

**Труднощі побудови моделі складу.** Деякі системи диференціюються на частини мимовільно в процесі природного росту і розвитку ( організми , соціуми , планетні системи , молекули , родовища корисних копалин і т.д.). Штучні системи завідомо збираються з раніше окремих частин ( механізми , будівлі , тексти , мелодії та ін.). Є й змішані типи систем (заповідники , сільськогосподарські системи , природно охоронні організації , тягловий транспорт).

З іншого боку, запитаєте , з яких частин складається університет у ректора, студента , бухгалтера , господаря, - і кожен видасть свою, відмінну від інших модель складу. Так само по-різному визначають склад літака льотчик, стюардеса, пасажир. Можна сказати, що тіло складається з правої і лівої половинок, а можна - з верхньої і нижньої. Так з чого ж воно складається "насправді"?

Труднощі побудови моделі складу, які кожному доводиться долати, можна представити трьома положеннями.

*Перше.* Ціле можна ділити на частини по-різному (як розрізати булку хліба на скибки різного розміру і форми). А як саме треба? Відповідь: так, як вам треба для досягнення вашої мети.

Тоді природно повернутися до питання: а чи існують частини "насправді"? Зверніть увагу на акуратну формулювання розглянутого властивості: *розрізнення* частин, а не роздільність на частини. Ми з ще однією сторони вийшли на проблему *цілісності* систем: можна розрізняти потрібні вам для вашої мети частини системи.

*Друге.* Кількість частин в моделі складу залежить і від того, на якому рівні зупинити дроблення системи . Частини на кінцевих гілках отриманого ієрархічного дерева називаються *елементами* . У різних обставинах припинення декомпозиції виробляється на різних рівнях. Наприклад, при описі майбутніх робіт доводиться давати досвідченому працівникові і

новачкові інструкції різного ступеня подробиці. Таким чином, модель складу залежить від того, що вважати *елементарним*, а оскільки це слово оцінне, то це не абсолютне, а відносне поняття. Проте зустрічаються випадки, коли елемент носить природний, абсолютний характер (клітина - найпростіший елемент живого організму; індивід – останній елемент суспільства, фонемі - найдрібніші частини усного мовлення) або визначається нашими можливостями (наприклад, можна припускати, що електрон теж з чогось складається, але поки фізики не змогли виявити його частини з дробовим зарядом).

*Третє*. Будь-яка система є частиною якоїсь більшої системи (а нерідко частиною відразу декількох систем). А цю мета систему теж можна ділити на підсистеми по-різному. Це означає, що *зовнішня межа системи має відносний, умовний характер*. Навіть "очевидна" межа системи (шкіра людини, огорожа підприємства тощо) за певних умов виявляється недостатньою для визначення кордону в цих умовах. Наприклад, під час трапези я беру виделкою з тарілки котлету, відкушую її, пережовує, ковтаю, переварюю. Де та межа, перетинаючи яку котлета стає моєю частиною? Інший приклад з кордоном підприємства. Працівник впав на сходах і зламав ногу. Після лікування при оплаті бюлетеня виникає питання: яка це була травма - побутова або виробнича (вони оплачуються по-різному)? Немає сумніву, якщо це були сходи підприємства. Але якщо це були сходи будинку, де живе працівник, то все залежить від того, як він йшов додому. Якщо прямо з роботи і ще не дійшов до дверей квартири, травма вважається виробничою. Але якщо він по дорозі зайшов в магазин чи кінотеатр - травма побутова. Як бачимо, закон визначає межі підприємства *умовно*.

Умовність границь системи знову повертає нас до проблеми цілісності, тепер уже цілісності всього світу. Визначення межі системи виробляється з урахуванням цілей суб'єкта, який буде використовувати моделі системи.

### **Питання для самоконтролю**

1. Охарактеризуйте динамічні властивості систем
2. Дайте визначення поняття Помилка першого-третього роду при застосуванні моделі чорного ящика
3. Які критерії якості моделювання?
4. Імпульсна характеристика електронної системи

### **Література**

1. Комаров М.С. Основы научных исследований. – Львов: Вища школа. Изд-во при Львов. ун-те. 1982. – 128 с.



2. Сурмин Ю. П. Теория систем и системный анализ: Учеб.пособие. — К.: МАУП, 2003. — 368 с.
3. Т.Т. Газизов Моделирование систем. Учебное пособие. Томский государственный педагогический университет. - [http://koi.tspu.ru/koi\\_books/gazizov/](http://koi.tspu.ru/koi_books/gazizov/)

### Розділ 3.

## Топологічні методи і графічні розв'язки диференціальних рівнянь

### 5.1. Інтегральні криві рівнянь другого порядку

Топологічний метод аналізу диференціальних рівнянь є одним з найважливіших засобів дослідження різних явищ, що пов'язані з нелінійними коливаннями та процесами.

Цей метод використовується для вивчення автономних систем. З його допомогою розв'язки диференціальних рівнянь шукаються не як явні функції часу, а у вигляді інтегральних кривих у фазовому просторі або, у більш загальному випадку, у просторі станів. Дослідження інтегральних кривих дозволяє отримати достатнє уявлення про якісний характер розв'язку, а також деякі кількісні дані. Графічні методи представлення розв'язків використовують в основному для систем малого порядку.

Умови рівноваги, коли система знаходиться у стані спокою, відповідають *особливій точці* у просторі станів; періодичним розв'язкам відповідають *граничні цикли*. Таким чином, інтегральні криві, які прямують до таких особливих точок або граничним циклом, можуть зображувати перехідні режими системи.

Для автономної системи, яка описується двома диференціальними рівняннями

$$\begin{aligned}\frac{dx}{dt} &= X(x, y); \\ \frac{dy}{dt} &= Y(x, y)\end{aligned}$$

інтегральна крива має вигляд:

$$\frac{dy}{dx} = \frac{Y(x, y)}{X(x, y)}$$

Точка  $(x_0, y_0)$ , у якій функції  $X(x_0, y_0) = Y(x_0, y_0) = 0$ , називається особливою точкою чи точкою рівноваги.

### 5.2. Класифікація особливих точок

Для дослідження особливих точок, розглядають лінеаризовану систему диференціальних рівнянь, що описують рух поблизу положення рівноваги (особливої точки).

$$\frac{dx}{dt} = X(x, y) = a_1x + a_2y + X_2(x, y);$$

$$\frac{dy}{dt} = Y(x, y) = b_1x + b_2y + Y_2(x, y)$$

де,  $X_2(x, y)$   $Y_2(x, y)$  функції порядку вище за перший. Відмітимо, що кожна особлива точка досліджується окремо.

Розкладемо праві частини системи за степенями  $\xi$  та  $\eta$  – малим відхиленням від стаціонарних значень  $x_0, y_0$ , нехтуючи похідними високого порядку:

$$\xi = x - x_0; \quad \eta = y - y_0$$

$$\frac{d\xi}{dt} = \frac{\partial X}{\partial x} \Big|_{x_0, y_0} \xi + \frac{\partial X}{\partial y} \Big|_{x_0, y_0} \eta = a_1\xi + a_2\eta;$$

$$\frac{d\eta}{dt} = \frac{\partial Y}{\partial x} \Big|_{x_0, y_0} \xi + \frac{\partial Y}{\partial y} \Big|_{x_0, y_0} \eta = b_1\xi + b_2\eta;$$

Ця «нова» система рівнянь має **нетривіальне рішення**

$$\xi = \xi_0 e^{\lambda t}, \quad \eta = \eta_0 e^{\lambda t},$$

якщо  $\lambda$  є коренем характеристичного рівняння

$$\begin{vmatrix} a_1 - \lambda & a_2 \\ b_1 & b_2 - \lambda \end{vmatrix} = 0$$

Для систем диферінційних рівнянь з двома змінними корені характеристичного рівняння:

$$\lambda_{1,2} = \frac{1}{2} \left[ a_1 + b_2 \pm \sqrt{(a_1 + b_2)^2 - 4(a_1 b_2 - a_2 b_1)} \right];$$

$$\lambda_{1,2} = \frac{1}{2} \left[ a_1 + b_2 \pm \sqrt{(a_1 - b_2)^2 + 4a_2 b_1} \right]$$

У залежності від характеру інтегральних кривих біля особливих точок, тобто в залежності від коренів характеристичних рівнянь особливі точки є такими, що класифікуються на:

- вузлова точка (вузол), якщо корені характеристичного рівняння обидва дійсні та одного знаку:

$$(a_1 - b_2)^2 + 4a_2 b_1 \geq 0;$$

$$a_1 b_2 - a_2 b_1 > 0$$

- сідло, якщо обидва корені дійсні, але мають різні знаки:

$$(a_1 - b_2)^2 + 4a_2 b_1 > 0;$$

$$a_1 b_2 - a_2 b_1 < 0$$

- фокус, якщо корені комплексно-спряжені:

$$(a_1 - b_2)^2 + 4a_2 b_1 < 0;$$

- центр або фокус, якщо обидва корені уявні:

$$(a_1 - b_2)^2 + 4a_2b_1 < 0;$$

$$a_1 + b_2 = 0$$

Для стійких особливих точок дійсна частина коренів характеристичного рівняння від'ємна, а для нестійких – додатна.



Рис. 11. Типичные фазовые портреты в окрестности точки равно-

Межа стійких фокусів та вузлів визначається виразом:

$$a_1 + b_2 = 0$$

Межа сідел:

$$a_1 b_2 - a_2 b_1 = 0$$

Межа фокусів:

$$(a_1 + b_2)^2 - 4(a_2b_1 - a_1b_2) = 0$$

### 5.3 Метод ізоклін

Метод ізоклін дає можливість досліджувати поле інтегральних кривих графічним методом, не розв'язуючи диференціальних рівнянь.

$$\frac{dx}{dt} = \frac{Y(x, y)}{X(x, y)} \equiv F(x, y) = const,$$

де  $F(x, y) = C$  – ізокліна, тобто крива на площині  $x, y$ , вздовж якої нахил  $\frac{dy}{dx}$  інтегральних кривих постійний.

Після того, як крива  $F(x, y) = C_i$  проведена, уздовж неї наносять короткі відрізки прямих, що мають заданий нахил  $\frac{dy}{dx} = C$ . Повторюючи ці побудовання для інших значень  $C_i$  отримують ряд кривих  $F(x, y) = C_i$  із відповідними нахилами. Ці нахили визначають поле напрямів дотичних та інтегральних кривих. Починаючи з точки  $A(x, y)$  та прямуючи завжди згідно напрямлення ліній поля можна накреслити деяку неперервну криву. Крива, що отримана, є інтегральною кривою, яка проходить крізь початкову точку  $A(x, y)$ .

### 5.4. Приклади застосування методу ізоклін

Розглянемо звичайне диференціальне рівняння першого порядку,

$$\dot{x} = \frac{dx}{dt} = t + \frac{t}{x} = t \cdot \left(1 + \frac{1}{x}\right),$$

яке легко розв'язується методом розділення змінних.

$$\frac{dx}{\left(1 + \frac{1}{x}\right)} = t \cdot dt,$$

$$\int \frac{x}{1+x} dx = \int t \cdot dt,$$

$$\int \frac{x+1-1}{1+x} dx = \int t \cdot dt$$

$x - \ln|x+1| = \frac{1}{2}t^2 + C$  – рішення даного диференціального рівняння, яке достатньо важко аналізувати.

Отримаємо рішення цього ж рівняння методом ізоклін.

$$\frac{dx}{dt} = t \cdot \left(1 + \frac{1}{x}\right) = C_i$$

1. Знаходимо рівняння ізоклін

$$t \cdot \left(1 + \frac{1}{x}\right) = C_i$$

$x = \frac{t}{C_i - t}$  - рівняння ізокліни, тобто криві на кожній з яких дотична до

шуканої інтегральної кривої рівна  $\frac{dx}{dt} = C_i$

2. З рівняння ізоклін визначаємо асимптоти

$$\begin{array}{ll} \text{при } t \rightarrow \infty & x \rightarrow -1 \\ \text{при } x \rightarrow \infty & C_i = t \end{array}$$

3. Визначаємо області, в яких інтегральна крива ввігнута ( $\frac{dx}{dt} \geq 0$ ) і

випукла, якщо ( $\frac{dx}{dt} \leq 0$ ) легко знайти другу похідну шуканого рівняння.

$$\begin{aligned} \ddot{x} &= \frac{d^2x}{dt^2} = 1 + \frac{d}{dt} \left( \frac{t}{x} \right) = 1 + \frac{x - t \cdot dx/dt}{x^2} = \frac{x^2 + x - t - t^2/x}{x^2} = \frac{x^3 + x^2 - x \cdot t^2 - t^2}{x^3} = \\ &= x^{-3} [x^2 \cdot (x+1) - t^2 \cdot (x+1)] = x^{-3} [(x+1) \cdot (x^2 - t^2)] = x^{-3} [(x+1) \cdot (x-t) \cdot (x+t)] \end{aligned}$$

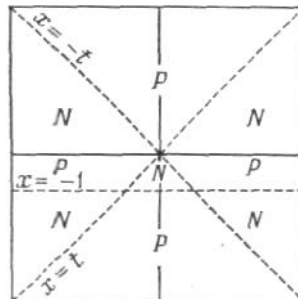


Рис. 5.1. Області випуклості (N) і ввігнутості (P) інтегральних кривих

$$\text{рівняння } \frac{dx}{dt} = t + \frac{t}{x}$$

4. Будуємо деякі ізокліни

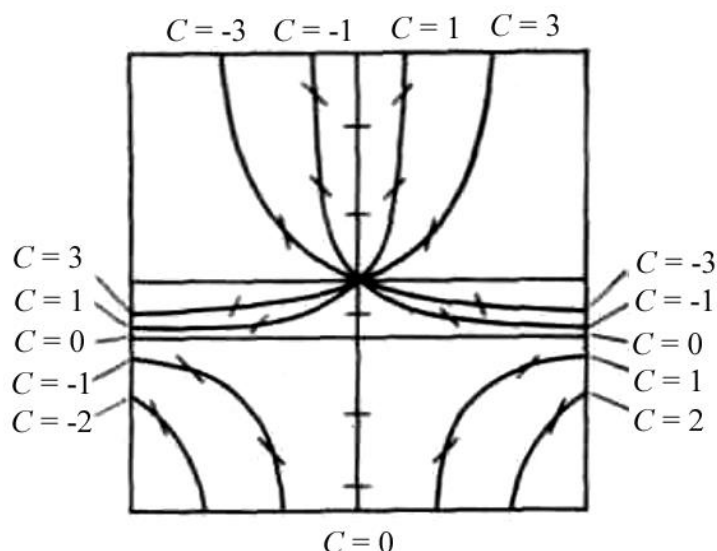


Рис. 5.2. Ізокліни рівняння

- а)  $C_1 = 0 \quad \frac{dx}{dt} = 0 \quad x = -1 \quad (t \rightarrow 0, \text{ вісь } x \text{ ізокліна } \frac{dx}{dt} = 0);$
- б)  $C_2 = 1 \quad \frac{dx}{dt} = 1 \quad x = \frac{t}{1-t} \quad (t = 1 - \text{асимптота});$
- в)  $C_3 = -1 \quad \frac{dx}{dt} = -1 \quad x = \frac{t}{-1-t} \quad (t = -1 - \text{асимптота});$
- г)  $C_4 = 3 \quad \frac{dx}{dt} = 3 \quad x = \frac{t}{3-t} \quad (t = 3 - \text{асимптота});$
- д)  $C_5 = -3 \quad \frac{dx}{dt} = -3 \quad x = \frac{t}{-3-t} \quad (t = -3 - \text{асимптота});$

5. Будуємо сімейство шуканих інтегральних кривих, які дотикаються побудованих ізоклін з заданим нахилом.

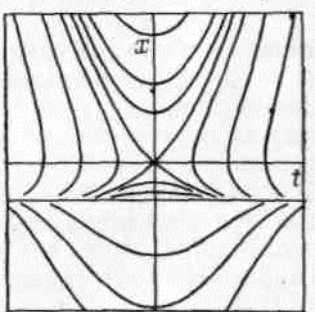


Рис. 5.3. Інтегральні криві диференціального рівняння

В якості прикладу застосування методу ізоклін для диференціального рівняння другого порядку розглянемо класичне нелінійне диференціальне рівняння Ван дер Поля, яке описує коливальну систему зі змінним коефіцієнтом демпфірування

$$\frac{d^2x}{dt^2} - \varepsilon(1-x^2)\frac{dx}{dt} + x = 0, \quad \varepsilon > 0$$

Це рівняння другого порядку може бути зведено до системи двох рівнянь першого порядку.

$$\frac{dx}{dt} = y$$

$$\frac{dy}{dt} = \varepsilon(1-x^2) \cdot y - x$$

Інтегральна крива визначається рівнянням

$$\frac{dy}{dx} = \frac{\varepsilon(1-x^2) \cdot y - x}{y}$$

Єдиною особливою точкою цього рівняння є початок координат  $x_0 = y_0 = 0$ , звідси випливає, що метод ізоклін застосовується усюди, крім цієї точки.

Коефіцієнти лінійного розкладу системи в околі особливої точки:

$$a_1 = 0, \quad a_2 = 1$$

$$b_1 = -1, \quad b_2 = \varepsilon$$

Корені характеристичного рівняння:

$$\lambda_{1,2} = \frac{1}{2} \cdot [\varepsilon \pm \sqrt{\varepsilon^2 - 4}]$$

При  $\varepsilon < 2$  особлива точка нестійкий фокус

При  $\varepsilon > 2$  особлива точка нестійкий вузол

Рівняння ізокліни:

$$y = \frac{x}{\varepsilon(1-x^2) - C_i}$$

На рис.5.4 показано графічна побудова для  $\varepsilon = 1$ .

$$1. \quad C_1 = 0 \quad \frac{dy}{dx} = 0 \quad y = \frac{x}{1-x^2}$$

$$2. \quad C_2 = 1 \quad \frac{dy}{dx} = 1 \quad y = -\frac{1}{x}$$

$$3. \quad C_3 = -1 \quad \frac{dy}{dx} = -1 \quad y = \frac{x}{2-x^2}$$

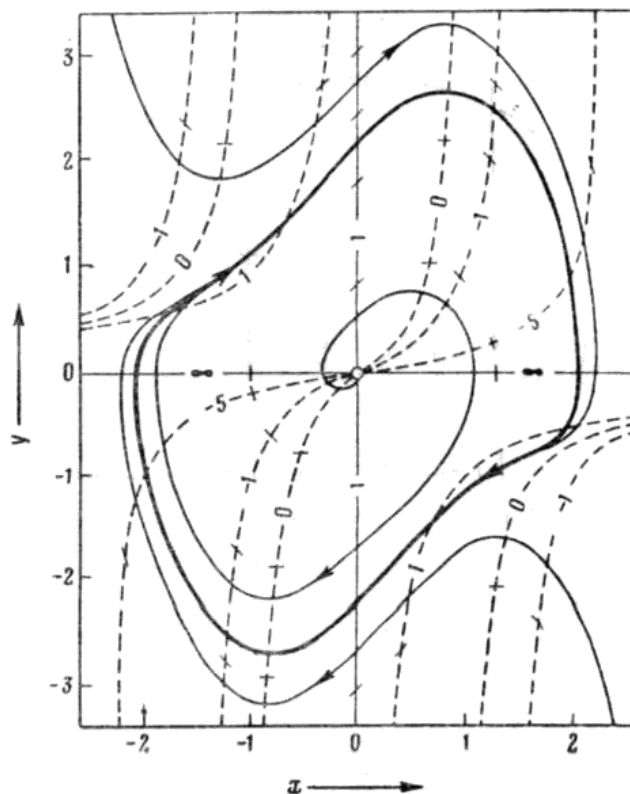


Рис. 5.4 Фазовий портрет рівняння Ван дер Поля

Враховуючи умови Бендиксона, що якщо вираз  $\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y}$  міняє знак, то можлива наявність стійкого граничного циклу. Визначаємо область зміни знаку виразу

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = 0 + (1 - x^2) \begin{matrix} > \\ < \end{matrix} 0$$

Знак змінюється в області  $x = \pm 1$ .

Всередині області  $x < \pm 1$  траєкторії уходять із особливої точки  $x_0 = y_0 = 0$  нестійкого фокуса і зсередини навиваються на граничний цикл. Траєкторії, які знаходяться поза граничним циклом, прагнуть до нього із зовнішньої сторони.

## 5.5. Завдання

### 5.5.1 Знайти розв'язки рівнянь методом ізоклін

1.  $\frac{dx}{dt} = x + 5t - 2;$   $x_0 = 5, \quad t_0 = 0$
2.  $\frac{dx}{dt} = -2t - x;$   $x_0 = -1, \quad t_0 = 0$
3.  $\frac{dx}{dt} = -x + t;$   $x_0 = 0,2, \quad t_0 = 0$
4.  $\frac{dx}{dt} = 3x + 2t - 5;$   $x_0 = 4, \quad t_0 = 0$



5.  $\frac{dx}{dt} = -4x - 5t;$   $x_0 = 0, t_0 = 0$
6.  $\frac{dx}{dt} = x + 3t;$   $x_0 = 0, t_0 = 0$
7.  $\frac{dx}{dt} = 2x + 3t;$   $x_0 = 0, t_0 = 0$
8.  $\frac{dx}{dt} = -4x - 5t;$   $x_0 = 0, t_0 = 0$

### 5.5.2 Знайти розв'язки рівнянь

1.  $\frac{dx}{dt} = x^2 + t^2 - 16;$  з мінімальним значенням  $X_{min} = 2$
2.  $\frac{dx}{dt} = x^2 + t^2 - 25;$  з максимальним значенням  $X_{max} = 4$
3.  $\frac{dx}{dt} = x^2 + t^2 - 100;$  з мінімальним значенням  $X_{min} = 6$
4.  $\frac{dx}{dt} = x^2 + t^2 - 106;$  з максимальним значенням  $X_{max} = 5$
5.  $\frac{dx}{dt} = x^2 + t^2 - 64;$  з мінімальним значенням  $X_{min} = 5$

### 5.5.3 Дослідити особливі точки системи рівнянь

1. Знайти межу сідел системи:  
 $\dot{x} = 1 - y \ln |x|;$   
 $\dot{y} = x(x^2 + 2y - a^2);$
2. Дослідити особливі точки системи:  
 $\dot{x} = x + 8 - y;$   
 $\dot{y} = xy + 7$
3. Знайти межу стійких вузлів та фокусів системи:  
 $\dot{x} = -x + xy^2;$   
 $\dot{y} = x + y$
4. Знайти межу сідел системи:  
 $\dot{x} = x + 5y^2;$   
 $\dot{y} = x^2 + xy$
5. Знайти межу фокусів системи:  
 $\dot{x} = y(x + y + 4) + x^2;$   
 $\dot{y} = 5x + 4y - 6$
6. Знайти умови стійкості системи:  
 $\dot{x} = y(5 + x + x^2) - 2x;$   
 $\dot{y} = (y + 2y^2) + x + x^2$
7. Знайти межу фокусів системи:  
 $\dot{x} = 1 - y \ln |x|;$   
 $\dot{y} = x(x^2 + 2y - a^2);$
8. Дослідити стаціонарні точки системи:

$$\dot{x} = x + y + 6;$$

$$\dot{y} = x(x + y^2);$$

9. Дослідити стаціонарні точки системи:

$$\dot{x} = xy;$$

$$\dot{y} = y - 6$$

10. Знайти межу стійкості вузлів та фокусів системи:

$$\dot{x} = 1 - y \ln |x|;$$

$$\dot{y} = x(x^2 + 2y - a^2);$$

## Розділ 4.

### БІФУРКАЦІЯ. КРИЗА. ХАОС

#### Перетворення системи

У процесі перетворення системи, її зміни діють внутрішні і зовнішні причини, що виявляються з більшою або меншою силою в різноманітних системах. Серед зовнішніх причин найбільш значущі:

- зміна зовнішнього середовища як середовища проживання системи, що викликає функціональну зміну її елементів. Тому неможливо тривале існування незмінної системи: будь-яка зміна, як би повільно і непомітно вона не протікала, неминуче призводить до якісної зміни системи. Причому зміна зовнішнього середовища може відбуватися як незалежно від системи, так і під впливом самої системи. Прикладом може служити діяльність людського суспільства, що сприяє зміні навколишнього середовища не тільки на користь, але і на шкоду (забруднення водойм, атмосфери і пр.);

- активний вплив середовища на систему, що призводить до проникнення в систему далеких об'єктів, що призводять до функціональними змінами окремих елементів (перетворення атомів під впливом космічних променів);

- вплив системи на зовнішнє середовище, яке викликає сильну реакцію на систему. У цьому випадку зміни системи відбуваються завдяки впливу системи на середовище і здійсненню впливу середовища на систему.

Внутрішні причини змін:

- 1) безперервний кількісний ріст диференційованих елементів системи в обмеженому просторі, у результаті чого загострюються протиріччя між ними;

- 2) поява автономних елементів, підструктур, підсистем, які можуть створити потужні вузли внутрішніх протиріч як з системою, так і одних з іншим;

- 3) накопичення "помилки" і "збоїв" у відтворенні собі подібних (мутації в живих організмах). Якщо елемент-"мутант" більш відповідає середовищу, то він починає розмножуватися. Це і є виникнення нового, що вступає в протиріччя зі старим;

- 4) припинення росту і відтворення складових елементів системи, у результаті система гине.

Виходячи з розуміння зрілої системи як єдності і сталості структури, визначимо різні форми перетворення, безпосередньо пов'язані зі зміною кожного з перерахованих атрибутів системи:

- перетворення, що приводить до знищення всіх взаємозв'язків елементів системи (руйнування кристала, розпад атома і т.п.);
- перетворення системи в якісно інше, але рівний по ступені організації стан. Це відбувається через: а) зміни складу елементів системи (заміщення одного атома в кристалі на інший), б) функціональної зміни окремих елементів і (або) підсистем у системі (перехід ссавців від сухопутного способу життя до водяного);
- перетворення системи в якісно інше, але нижче за ступенем організованості стану і пояснюється: а) функціональними змінами елементів і (або) підсистем у системі (приспособлення тварин до нових умов середовища), б) структурними змінами (модифікаційні перетворення в неорганічних системах, наприклад перехід алмазу в графіт);
- перетворення системи в якісно інше, але вище за ступенем організації стану. Воно реалізується як в рамках однієї форми руху, так і при переході від однієї форми до іншої. Цей тип перетворення пов'язаний із прогресивним, поступальним розвитком системи. Зауважимо, що на перетворення системи накладаються природні обмеження, пов'язані з природою системи, механізмом її спадкування, ресурсними можливостями системи і середовища.

Отже, перетворення - неминучий етап у розвитку системи . Вона вступає в нього в силу наростаючих протиріч між новим і старим , між функціями елементів і характером зв'язку між ними , між протилежними елементами . Перетворення може відбивати як завершальний кінцевий етап у розвитку системи , так і перехід систем ( стадій) одна в іншу. Перетворення є період дезорганізації системи , коли старі зв'язки між елементами рвуться , а нові ще тільки створюються . Перетворення може означати і реорганізацію системи , а також перетворення системи як цілого в елемент іншої , вищої системи .

### **Кризи і загибель систем**

Сучасна наука виробила по відношенню до криз кілька концептуальних позицій. Криза розглядається як закономірне стан розвитку системи, яке формується на попередніх її етапах. Соціальна криза, наприклад не падає з неба, якщо виключити з розгляду наслідки для суспільства космічних та екологічних катастроф, а спонтанно складається в суспільному організмі.

Криза представляє собою ослаблення життєвих сил системи , її нездатність зберігати стан рівноваги , досягати поставлені раніше цілі. Система втрачає свою місію , виявляється нездатною до цілепокладання і стає неідеальною. Окремі її підсистеми і елементи відхиляються від мети , формують свої цілі або живуть безцільно , так що місія системи стає декларативною , втрачає природну опору . Система починає " плавитися " ,

втрачати колишню цілісність . Наприклад , усі епохи соціальних криз зазвичай характеризувалися загальною розслабленням населення , зростанням егоїзму , прагненням до насолоди , розкоші і т. п.

Криза виступає як діалектичну єдність двох процесів: руйнування і творення. Завдяки їм відбувається руйнування одних елементів, структур, функцій системи і виникнення інших.

У реальному житті спостерігаються різноманітні співвідношення між собою руйнування і творення. Умовно виділимо три варіанти: перший передбачає домінування творчих процесів, другий відрізняється деяким рівновагою руйнування і творення, а третій - домінуванням руйнування.

Кризу не можна розглядати як суто негативне явище. Цю думку переконливо підкреслює Ю. В. Яковець: "Кризи прогресивні при всій їх хворобливості. Криза виконує в динаміці хвилеподібного, суперечливого руху систем три найважливіші функції:

- різкого ослаблення та усунення (або якісного перетворення) застарілих елементів панівною, переважною, але вже вичерпала свій потенціал системи;
- розчистку дороги для затвердження спочатку слабких елементів нової системи, майбутнього циклу;
- випробування на міцність і передача у спадок тих елементів системи (зазвичай це одночасно зберігаються елементи надсистеми і суперсистеми), які акумулюються, накопичуються, переходять у майбутнє (іноді частково модифікуючись) " .

Криза створює найбільш сприятливі умови для мутацій, з одного боку, і забезпечення наступності та успадкування в розвитку - з іншого. Чим глибше і руйнівніше криза, тим сильніше він впливає на генетичне ядро системи і тим серйозніше може перебудуватися система. Якщо ж криза, як вітер, проноситься над верхівками дерев, то система обробляється тільки легким переляком, майже повністю відновлюючи себе.

Криза не тільки неминуча, але й необхідна фаза розвитку системи. Він поєднує в собі не тільки згасання колишньої якості системи, але і поява нового. Якщо визначати точніше, то суть кризи становить лихоліття в системі, коли стара система вже зникла, а нова ще не з'явилася. Досить часто кращою категорією для характеристики природи кризи є категорія "хаос". Не випадково на Русі, як правило, соціальна криза називали "смутним часом".

Криза - це етап перебудови системи, здобуття нею сталого вектора просування до нової якості. Криза представляється епохою втрати і набуття системою самої себе. Він є неминучою фазою оновлення. Підкреслимо, що інноваційний підхід до розуміння природи кризи набагато ефективніше катастрофічного до нього підходу, бо катастрофічний підхід проповідує

порятунку, а інноваційний - діяльність по свідомому оновленню системи і тим самим прискоренню її виходу зі стану якісної невизначеності.

Криза явище конкретно-історичне, що розвивається від епохи до епохи. Якщо кризи аграрного суспільства охоплювали переважно сферу сільського господарства і приводили до голоду, а кризи промислові зупиняли виробництво товарів, виступаючи як товарні, то кризи постіндустріального суспільства будуть охоплювати сферу інформації та інтелекту.

Разом з цивілізацією відбувається розвиток екологічних, політичних, релігійних, ментальних та інших криз, які знаходять специфіку епохи.

Криза має розглядатися як складне, комплексне явище, що охоплює систему. Він являє собою полиструктурне і полифункціональне освіту. У якості його найважливіших характеристик виступають: сфера прояву, масштаби, ступінь охоплення об'єкта або процесу, фактори породження, характер ураження системи, наслідки для суспільства та ін.

Сучасна цивілізація породила величезну різноманітність соціальних криз. Спираючись на основні параметри криз, можна побудувати їх класифікацію, яка представлена в табл. 20.

Основні класифікації	Соціальні кризи	
	Тип	Характеристика
1	2	3
Сфера прояву кризи	Економічний	Криза перевиробництва, недовиробництво тощо
	Екологічний	відрізняється руйнуванням екологічного середовища, різким погіршенням природних умов для життя людини
	Соціальний	Проявляється у деградації та руйнації соціальної структури та інститутів, погіршення життя людей
	Політичний	Об'єднує сімейство криз, що протікають у сфері політики
	Етнічний	Виступає як криза політичного, економічного і культурного розвитку етносів та етнічної структури.
	Криза культури	Являє собою культурну деградацію суспільства, руйнування цінностей і інститутів культури
	Ідеологічний	Виражається у відсутності ідей і цілей суспільного розвитку, земної їх

		сурогатами.
	Духовний, моральний	Виражається у втрати духовності, деградації моралі
Масштаби кризи	Глобальний	Охоплює планетарну цивілізацію
	Національний	Проходить в масштабах держави, під загрозою національна безпека.
	Регіональний	Розгортається в межах регіону.
	Локальний	Охоплює мікросоціальний простір
Охоплення сторін об'єкта або процесу	Моноаспектний	Вражає тільки одну сторону, характеристику, аспект об'єкту або процесу
	Поліаспектний	Вражає декілька аспектів
	Комплексний	Охоплює всі сторони об'єкта або процеса.
Тривалість	Короткостроковий	Проходить в короткі періоди часу
	Середньостроковий	Продовжується декілька років.
	Довгостроковий	Триває десятки років
	Наддовгостроковий	Триває на протязі століть і навіть тисячоліть
Фактори що породжують	Внутрішній	Народжен внутрішніми факторами.
	Зовнішній	Народжен взаємовідносинами з навколишньою середою
	Змішаний	Народжен як зовнішніми, так і внутрішніми характеристиками
Характер враження системи	Генетичний	Виявляється завдяки мутації соціального спадкування
	Змісту	Враження життєво важливих елементів системи
	Організаційний	Порушення організації системи
	Функціональний	Суттєві порушення функціонування системи
	Структурний	Порушення структури системи
Характер дозволу	Вибухові	Завершуються соціальним вибухом
	Замираючі	Приводять до поступового згасанню коливального процесу
Характер діючих причин	Циклічний	Діючі причини носять закономірний циклічний характер
	Випадковий	Викликаний стихійними лихами, помилками, аваріями
Наслідки	Інноваційний	Забезпечуючий оновлення системи
	Руйнуючий	Руйнуючий систему, переводячи її в інобуття
	Консервуючий	Консервує відсталу, віджившу соціальну форму.

Одна з ключових проблем народжується науки про кризи - крізісології - пояснює причини криз та шляхи їх подолання. Виниклі в сучасній науці кризові та антикризові теорії дають картину причин і шляхів подолання криз, представлену в табл. 21, з якої можна зробити такі висновки:

- причини криз різноманітні і взаємопов'язані між собою. Вони немов "викликають" до дії один одного. Наприклад, відсутність антикризового та інноваційного мислення персоналу фірми неминуче викликає до життя помилки в управлінні, що призводять до структурних порушень системи, тобто психологічні причини породжують організаційні, структурні і т.п.;

- існує відома кореляція між причинами кризи і способами його подолання, як існує зв'язок між хворобою і ліками;

- якими б причинами криза не викликався до життя, він являє собою об'єктивне явище, яке розгортається в просторі і в часі і не може бути подолане відразу. Криза характеризується своїм неповторним типом соціальної інерції. Суспільству, подолав його, він ще довго бачиться в кошмарних снах і позначається у всіх сферах життєдіяльності;

- досі мало досліджена проблема структури кризи. При цьому структуру кризи не слід плутати зі структурою системи, яка переживає кризу. Оскільки криза виступає як процес, то його структура - це структура процесу.

#### Концепції причин кризи та шляхів їх подолання

Теорія		Пояснення причин кризи	Шляхи подолання кризи
Назва	Автори		
1	2	3	4
Теорія катастроф	Т.Постон І.Стюарт В.І.Арнольд	Викликається природними та соціальними	Прогнозування катастроф, вироблення засобів по зниженню та подоланню їх наслідків
Теорія циклів	К.Маркс Т.Мальтус	Виступає етапом в циклічному	Розрахунок тривалості та



	Н.Д.Кондарт'єв Л.Н.Гумілев А.Тоффлер П.А. Сорокін А.Л.Чіжевський	розвитку	етапів циклу, вироблення засобів по пом'якшенню удару кризи
Теорія систем	А.А.Богданов Л.Берталанфи Г.Хаген	Пояснюється виникнення необхідності перебудови системи	Структурно- функціональна перебудова системи, яка пом'якшує кризу
Іноваційна теорія	Й.Шумптер А.І.Пригожин	Пояснюється або не достатком нововведень, або ростом їх числа та неосвоїності	Підтримання постійного тону нововведень, скорочення термінів їх освоєння
Теорія еліт	М.Вебер	Викликаний помилками та злочинами старої еліти і необхідністю її зміни	Підготовка нової еліти, забезпечення зміни еліт
Теорія антикризового керування	Р.Акофф Е.А.Уткин	Викликаний не ефективністю керування	Реформування системи керування за допомогою його орієнтування на антикризове керування
Психологічна теорія	Л.Н.Гумілев П.А. Сорокін	Виникають в свідомості, психіці людей	Вироблення нової системи цінностей, парадигм, норм, психології і підготовка в світі персоналу, шокова терапія

### Хаос і його роль у розвитку систем

Один з варіантів розвитку - руйнування системи , зниження впорядкованості і організованості аж до виникнення хаосу . У давньогрецькій

міфології хаос - це зяюча безодня , наповнена туманом і мороком , з якої відбулося все суще ; стан невпорядкованості , визначальне не лише руйнування , але народження систем . В іншій трактуванні хаос - сутнісний безлад , безлад - логос , велика безодня , повна творчих сил і божественного насіння , як єдина безладна маса , важка й темна , суміш землі , води , вогню і повітря. Ця наповнена безодня , невизначена з точки зору спокою , постійно кипить , видозмінюючи не тільки свою форму , але й елементи генотипу. У більшості народів хаос виступав у вигляді водної стихії. Так , асиро - вавілоняни представляли світ як первісний хаос , що складається з проматерія Тіамат і " первородного Сотворителя " Апсу . А древні єгиптяни уявляли хаос Нун безкрайніх водним простором , зануреним у темряву . У ряду народів поняття про хаос відсутнє в міфології , але репрезентована первинний стан всього сущого таки нагадує його . Наприклад , згідно з японською міфології , земля існує з самого початку , але перебуває у воді в напіврідкому стані , нагадуючи медузу.

Якщо відкинути наївність древніх щодо божественної природи хаосу , то перед нами залишиться дивно діалектичне його розуміння . Хаос в розумінні стародавніх греків набагато евристичності для науки , ніж такі терміни ХХ ст. , Як " невизначеність " , " ентропія " , " негоентропія " та ін Формулюючи їх , людська думка лише ковзнула по поверхні хаосу , забувши наївну геніальність античності про те , що хаос не зводиться до розпаду і дегградації. Він разом з ними неминуче містить елементи творення, гени нового. Хаос - це не анігіляція об'єкта , а балансування на деякій грані між буттям і небуттям . Він забезпечує расковиваніе жорстких структур , набуття елементами більшого числа ступенів свободи , що в кінцевому підсумку є найважливішою умовою природного оновлення .

Для хаосу властиві такі риси:

- "розм'якшення", "розпушення", руйнування і дегградація колишніх систем, структур, організацій;
- невизначена перехідність, бо, як вважали античні греки, "семябогов" ещенеплодотвориловсюреальностьхаоса, Атам, де це сталося, беспорядочно мечутся зародки будущею міра;
- творчі початку і джерела розвитку в умовах хаосу мінімальні, вони "не включені". Рух елементів безладно, не підкоряється ніякому закону, немає протиріч і розвитку. Однак ці джерела завжди готові прокинутися і забезпечити ривок у розвитку системи;

- розрив зв'язків між елементами. Хаос нагадує фізичну плазму - потік електронів та іонів, де йдуть свої поляризації, плазмові коливання і вибухи.

Відсутність в понятті хаосу-якої впорядкованості заважало його науковому дослідженню. Досить було сказати слово "хаос", і всім ставало зрозуміло, що мова йде про відсутність порядку, втрати якої-небудь предметності.

Хаос - порожнеча характеризується мінімальною кількістю елементів . Це , що називається , " розряджений хаос". В. І. Вернадський вважав , що вакуум не їсти порожнеча з температурою абсолютного нуля , а є активна область максимальної енергії нам доступного Космосу , тобто порожнечі немає. Під ці роздуми підходить гіпотеза , що припускає мимовільне народження атомів у космічному вакуумі. Вона добре пояснює деякі природні явища , але вимагає відмови від закону збереження енергії (точніше , незначних за величиною відхилень від закону). Однак ніхто не заважає припустити , що ця енергія , зосереджена в вакуумі , має принципово іншу природу. З цих позицій дуже вчасно звучать слова Вернадського : " Про ці просторах з розсіяними атомами і молекулами правильніше мислити не як про матеріальну порожнечі " вакууму " , але як про концентрацію своєрідною енергії , в розсіяному вигляді містить колосальні запаси матерії та енергії ..." .

Наповнений хаос являє собою невпорядковану сукупність елементів , він нагадує броунівський рух з його загальної безладністю . Хаос як невпорядкованість характеризується тим , що постійно нарастають ентропійні процеси , рвуться зв'язки , руйнуються структури . В результаті виникає всеохоплюючий безлад . Але хаос може виступити і у вкрай протилежній формі - у вигляді хаосу - заорганізованості , коли структур стає так багато , що система перестає діяти в напрямку до мети , починається її дезорганізація . Тут проявляється ефект появи безладу в умовах жорсткого детермінізму . Типовим прикладом є деградація зрівняльного соціалізму , в якому жорстко контролювалися процеси розподілу , які в кінцевому результаті привели до хаотизації споживання.

Значний внесок у розуміння хаосу вніс І. І. Пригожин, який розробив його концепцію на прикладі організації фізико-хімічних систем. Суть поглядів І. І. Пригожина полягає в наступних положеннях:

1. Хаос розглядається як носій можливої впорядкованості, як творче начало, конструктивний механізм еволюції. Звідси процес розвитку

виступає формуванням порядку з хаосу, яке являє собою процес самоорганізації під впливом різноманітних факторів. Хаотичні коливання, що у системах, - провісники і супутники змін укладу системи.

2. Хаос - динамічне мінливе явище. У ньому постійно утворюються флуктуації, які представляють собою випадкові відхилення величин, що характеризують систему, що складається з великого числа частинок, від їх середнього значення. Флуктуації прагнуть вивести систему з рівноваги, намагаються заволодіти нею, що призводить до руйнування колишніх структур і переходу системивное стан.
3. Перехід у новий стан здійснюється через точки біфуркації , які виступають як ситуації роздвоєння , коли перед системою відкриваються різні варіанти розвитку . У точці біфуркації система як би робить вибір , який визначає її подальшу еволюцію. При цьому перехід через біфуркацію випадковий. І. І. Пригожин зазначав: " Коли система , еволюціонуючи , досягає точки біфуркації , детерміністичних опис стає непридатним . Флуктуація змушує систему вибирати ту гілку , по якій буде проходити подальша еволюція системи .

Перехід через біфуркацію - такий же випадковий процес, як кидання монети. Існування нестійкості можна розглядати як результат флуктуації, яка спочатку була локалізована в малої частини системи, а потім поширилася і призвела до нового макроскопічному стану ". Перехід на більш високий рівень впорядкованості отримав назву диссипативної структури.

4. Концепція І. І. Пригожина дає переконливі пояснення з точки зору зміни ходу часу в системах і прискорення розвитку. Виникнення флуктуації, по суті, являє собою зародження нової системи, її часу і тимчасових характеристик. Розвиток нелінійних процесів, зростання ролі нових структур, заповнення ними простору системи викликає процеси прискорення розвитку.
5. Процеси саморозвитку в складних системах досліджуються загальнонаукової теорією самоорганізації - синергетикою , яка спрямована на пошук законів еволюції відкритих нерівноважних систем будь-якої природи . Термін " синергетика " введений в обіг німецьким дослідником Г. Хаген , який розглядає її як

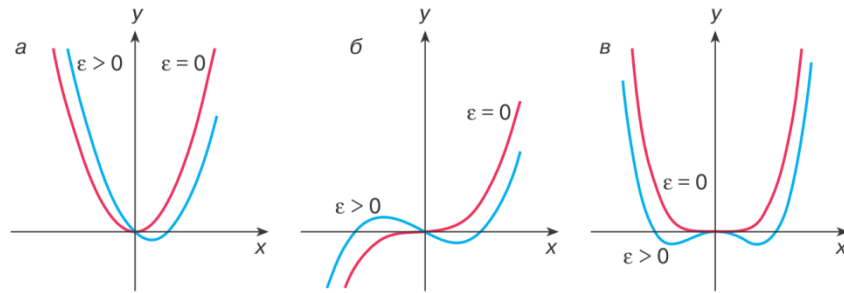
міждисциплінарну науку, пов'язану з різними областями фізики, хімії, біології, кібернетики. Синергетика досліджує такі взаємодії елементів системи, які призводять до виникнення просторових, тимчасових або просторово-часових структур у макроскопічних масштабах.

Ідеї І. І. Пригожина і Г. Хагена володіють значним творчим потенціалом. Призначені для пояснення процесів самоорганізації складних систем і складових їх ефектів, вони досить широко застосовуються для пояснення соціальних систем. Вони ефективні при осмисленні соціальних катастроф, перехідних періодів, реформування суспільства, управління в умовах соціальних заворушень.

### **Структурна стійкість і нестійкість функцій**

Один із аспектів задач на екстремум, який довгий час залишався поза зором математиків і фізиків, тісно пов'язаний із сучасним означенням структурної стійкості функцій. Якщо ми розглянемо, наприклад, функції  $y = x^2$ ,  $y = x^3$  та  $y = x^4$ , то всі вони мають нульову першу похідну на початку координат (у таких випадках кажуть, що  $x = 0$  – критична точка). Перша та третя функції мають в критичній точці мінімальне значення, а друга – точку згину, і в традиційних рамках задач на екстремум ця різниця найбільш важлива. Але виберемо дещо іншу точку зору. Спробуємо трохи поворушити розглянуті функції, увівши слабке обурення: 1)  $y = x^2 - \varepsilon x$ ; 2)  $y = x^3 - \varepsilon x$ ; 3)  $y = x^4 - \varepsilon x^2$ , де параметр  $\varepsilon$  може бути скільки завгодно малим по величині (рис.1).

В результаті такого збурення у випадку (1) ніяких принципових змін не відбувається: зберігається єдина критична точка, що зміщена лише на малу величину  $x_0 = \varepsilon/2$ , причому значення функції в цій точці (єдиний мінімум) змінюється на величину  $y_0 = -\varepsilon^2/4$  (рис.1, а). В другому та третьому випадку зовсім інша ситуація. Друга функція, для якої початок координат був точкою згину, набуває дві екстремальні точки  $x_{1,2} = \pm\sqrt{\varepsilon/3}$ , одна з яких відповідає мінімуму, а друга – максимуму (рис. 1, б). Функція  $y = x^4$ , яка мала єдиний мінімум на початку координат, у результаті малого зворушення, має вже три критичні точки (рис.1, в). При цьому початок координат стає точкою максимуму, а у двох нових критичних точках, як завгодно близьких до точки  $x = 0$ , функція приймає мінімальні значення.



**Рис.1.** Деформація функцій в околі критичної точки при зміні параметра

Побудова математичної моделі будь-якого процесу зв'язана із знехтуванням малими членами. У нашому першому випадку це цілком виправдано: врахування малого відхилення функції від квадратичної параболи призводить не до якісних, а до невеликих кількісних змін. У другому та третьому випадках поведінка при врахуванні малих поправочних членів якісно інша. Таким чином, функції  $y = x^3$  та  $y = x^4$ , попри те, що друга функція має екстремум, а перша – ні, об'єднує спільна властивість, яку, не вдаючись до строгих означень, назовемо структурною нестійкістю. Цей термін відображає те, що при невеликій зміні структури функції її поведінка в околі критичної точки різко змінюється. А функція  $y = x^2$  навпаки – структурно стійка.

Властивість структурної стійкості (нестійкості) функції не було включено до арсеналу математичних понять аж до 30-х років ХХ століття, коли воно було вперше сформульовано А.А.Андроновим (1901 - 1952). Через декілька десятиліть поняття структурної стійкості стало одним із ключових для теорії катастроф.

### **Біфуркації стаціонарних станів**

Здавалось би, згідно до сказаного, структурно нестійкі у критичних точках функції непридатні для опису дійсності. Але, як правило, функції, що виникають у фізичних задачах, містять деякі параметри, значення яких можуть змінюватися у певному діапазоні (подібно до параметру  $\varepsilon$  у наших прикладах). В таких випадках ми маємо справу з сімейством функцій, які залежать від параметра. Може трапитись, що при зміні останнього з неминучістю досягається значення (в нашому випадку  $\varepsilon = 0$ ), відповідне структурно нестійкій критичній точці, яка тим самим набуває цілком реального змісту. Більш того, саме ця точка, будучи одною з реалізацій

сімейства критичних точок, являється найбільш важливою, оскільки з нею пов'язані якісні зміни в поведінці системи.

Термін “біфуркація”, який став одним із ключових в теорії катастроф, буквально означає “роздвоєння”, але зазвичай застосовується в більш широкому сенсі для позначення різних якісних перебудов різних об'єктів при зміні параметрів, від яких вони залежать. У прикладі з сімейством  $y = x^4 - \varepsilon x^2$  значення параметра  $\varepsilon = 0$  відповідає також точці біфуркації, оскільки при переході  $\varepsilon$  від негативних значень до позитивних, єдиний стійкий стаціонарний стан  $x = 0$ , стаючи нестійким, доповнюється парою стійких станів  $x = \pm\sqrt{\varepsilon/2}$ . У прикладі із сімейством функцій  $y = x^3 - \varepsilon x$ , при негативних  $\varepsilon$  стаціонарні стани взагалі відсутні, а в точці  $\varepsilon = 0$  відбувається народження пари таких станів, один з яких стійкий, а другий - нестійкий. В обох випадках значення  $\varepsilon = 0$  відповідає точкам біфуркації, хоча і різних типів.

Спільна мета дослідження точок біфуркації, як математична проблема, складається з їх класифікації і аналізу поведінки сімейств функцій поблизу структурно нестійких критичних точок.

Ще один, аспект задач на мінімум і максимум, також тісно пов'язаний із структурною нестійкістю і вирішальним чином вплинув на формування теорії катастроф, відноситься до поняття "особливість відображення".

### **Біфуркація у простій дисипативній системі**

Виберемо спочатку просту абстрактну математичну модель, на якій простіше зобразити деякі механізми, відповідальні за якісні зміни поведінки.

Розглянемо єдину змінну  $x$ , яка змінюється відповідно з рівнянням:

$$\frac{dx}{dt} = f(x, \lambda) = -x^3 + x\lambda \quad (7.1)$$

Ця проста система має єдиний управляючий параметр  $\lambda$ . Причому швидкість зміни  $x$ , позначена як функція  $f$ , залежить від цього параметра суворо лінійно. Знаходимо стаціонарні стани:

$$0 = -x_s^3 + \lambda x_s$$

Це алгебраїчне рівняння завжди має один тривіальний розв'язок:

$$x_s = 0,$$

$$x_{s_1} \equiv x_0 = 0.$$

Але можуть бути і нетривіальні розв'язки. Вони задовольняють рівнянню, яке отримують з кубічного після скорочення на  $x_s$ :

$$-x_s^2 + \lambda = 0.$$

При від'ємному  $\lambda$  – це рівняння має уявний розв'язок, що не може відповідати будь-якій фізично можливій ситуації. При позитивному  $\lambda$  це рівняння допускає два розв'язки:

$$x_{\pm} = \pm\sqrt{\lambda}$$

$$x_{S_{2,3}} \equiv x_{\pm} = \pm\sqrt{\lambda}.$$

Ці розв'язки зливаються з  $x_0$  при  $\lambda = 0$ , та відхиляються від нього при  $\lambda > 0$ . Це і є явище біфуркації. На рис. 7.1 зображено графік стосовно нашого завдання. Суцільною та штриховою лініями позначені асимптотично стійкі ( $S$ ) та нестійкий ( $U$ ) розв'язки.

Симетрична біфуркація (типу камертона), при якій стандартний стан  $x_0$  утрачає стійкість при  $\lambda > 0$  та призводить до двох гілок стійких розв'язків  $x_+$  та  $x_-$ , що зливаються з  $x_0$ .

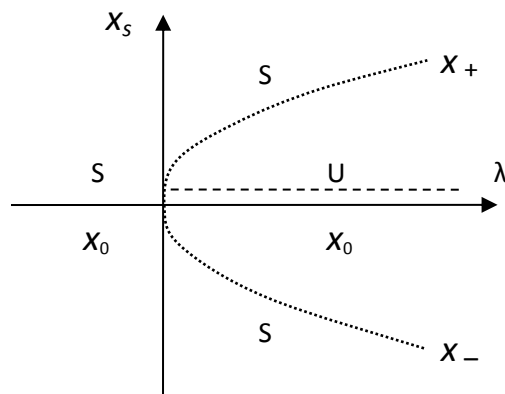


Рис. 7.1. Залежність  $x_s$  від  $\lambda$

У розглянутій тут простій моделі проблема стійкості досить тривіальна, оскільки рівняння (7.1) допускає точний розв'язок за допомогою стандартних методів.

Іншими словами, гілки  $x_{\pm}$  з'являються в результаті біфуркації в той момент, коли стандартний стан  $x_0 = 0$  утрачає стійкість. Таку біфуркацію інколи називають *надкритичною*.

Пригадаємо, що ми на початку мали просте кубічне рівняння, яке плавно залежало від параметру  $\lambda$ . Тепер видно, що повільна динаміка призводить до особливості. Дійсно в околі точки біфуркації  $\lambda = 0$  розв'язки  $x_{\pm}$  не можуть бути розкладені в ряд по степеням параметра  $\lambda$ . Отже, вони залежать від  $\lambda$  *не аналітично*. Це є математичним відображенням якісної зміни поведінки, яка обумовлена біфуркацією.

Розглянемо ще більш простий приклад дисипативної системи:



$$\frac{dx}{dt} = -x^2 + \mu$$

Керуючий параметр позначимо через  $\mu$ . Вирази для нерухомих точок  $x_s$  мають вигляд:

$$x_{\pm} = \pm\sqrt{\mu}$$

Цю залежність зображено на рис. 7.2.

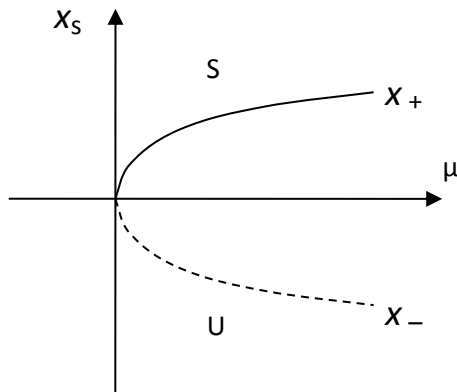


Рис. 7.2. Залежність  $x_s$  від  $\pm\sqrt{\mu}$

В цьому випадку також можна знайти точний розв'язок динамічного рівняння, та показати, що гілка  $x_-$  не стійка, а  $x_+$  – асимптотично стійка. Зазначимо, що при зменшенні позитивних  $\mu$  у точці  $\mu=0$  виникає зіткнення та анігіляція стійкої та нестійкої гілки. За цією причиною точка  $\mu=0$  називається *граничною точкою*, чи згорткою.

Тепер об'єднаємо два попередніх приклади, розглянувши дисипативну систему, яка описується рівнянням:

$$\frac{dx}{dt} = -x^3 + \lambda x + \mu \tag{7.2}$$

Тут немає квадратичного члену, який фігурує в рівнянні (7.2), оскільки його завжди можна усунути відповідною заміною змінних. Таким чином ми приходимо до завдання з двома параметрами ( $\lambda$  та  $\mu$ ). Нерухомі точки тепер визначаються канонічною формою кубічного рівняння загального виду:

$$-x_s^3 + \lambda x_s + \mu = 0$$

Це рівняння може мати три реальні розв'язки. Більш того, при зміні параметрів відбувається злиття трьох розв'язків, в результаті чого залишається єдиний реальний розв'язок.

Область існування трьох реальних коренів закінчується в точці (початок координат), в якій залежність  $\lambda$  від  $\mu$  має особливість. Це відома *особливість типу вістря*.

На рис. 7.3 та 7.4 приведені залежності розв'язків від параметрів.

На рис. 7.3 зображена залежність  $x_s$  від  $\mu$  при фіксованому  $\lambda$ .  $S$  – подібна форма кривої обумовлена безліччю розв'язків в відповідній області значень параметра  $\mu$ . Більш того, дві з гілок виявляються стійкими одночасно.

Область бістабільності закінчується у двох граничних точках  $\mu_1$  та  $\mu_2$ .

При цьому спостерігається *цикл гістерезису*.

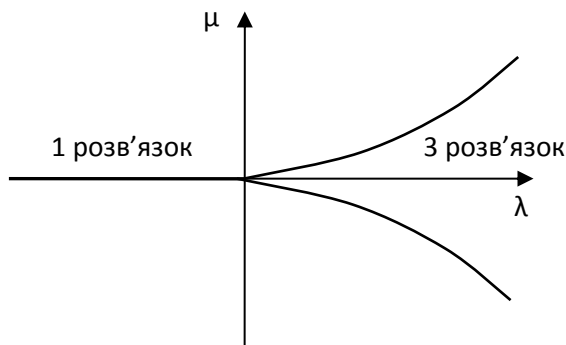


Рис. 7.3. Залежність  $\mu$  від  $\lambda$

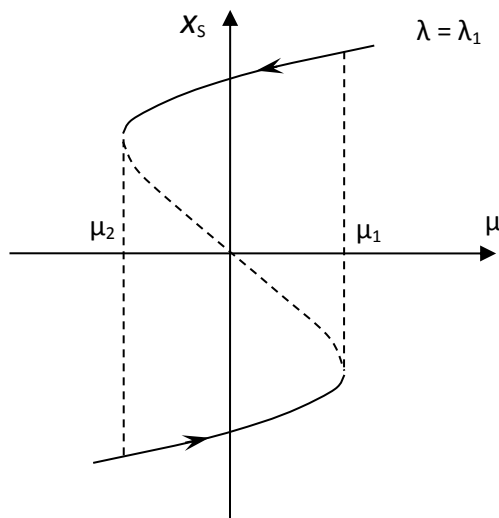


Рис. 7.4. Залежність  $x_s$  від  $\mu$

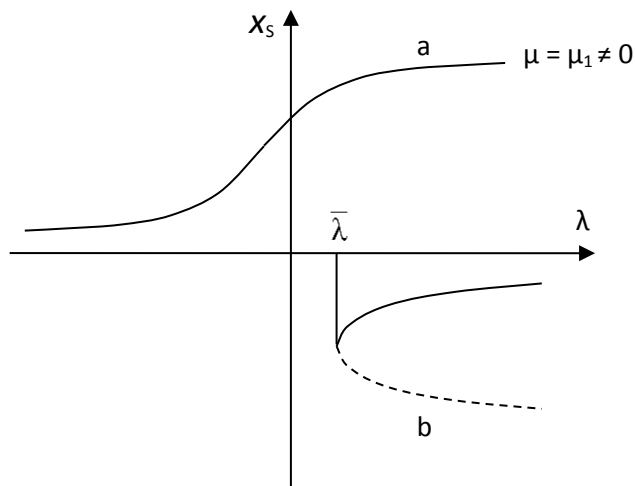


Рис. 7.5. Залежність  $x_s$  від  $\lambda$

На рис. 7.5 побудована залежність розв'язку  $x_s$  від  $\lambda$  при фіксованому  $\mu$ . Тепер є дві окремі криві, одна з яких (крива  $a$ ) визначена для всіх значень  $\lambda$ , а інша (крива  $b$ ) визначена лише при  $\lambda \geq \bar{\lambda}$  та має особливість типу граничної точки при  $\bar{\lambda}$ .  $\lambda < \bar{\lambda}$  має лише стійкий розв'язок, а при  $\lambda > \bar{\lambda}$ , як і раніше має місце бістабільність.

Що ми досягли з точки зору фізичних застосувань? Встановили деякі “максимальні” математичні моделі, відтворюючі одне з експериментально спостережуваних перехідних явищ, що супроводжують складну поведінку, а саме виникнення множинних одночасно стійких розв'язків. Це явище напряму зв'язане зі здатністю до переключень, і, тим самим, до виконання регуляторних задач.

### Методи спрощення систем кінетичних рівнянь

Спрощення математичної моделі полягає в зменшенні числа рівнянь і, разом з цим, числа параметрів, що визначають поведінку системи. Навіть порівняно прості біохімічні процеси складаються з багатьох стадій і містять багато проміжних речовин. Математична модель, буквально відповідна ланцюга біохімічних реакцій, містить багато (кілька десятків) змінних і, відповідно, рівнянь. З іншого боку, більшість вдалих і змістовних математичних моделей складається з двох-трьох нелінійних рівнянь. В цьому параграфі ми обговоримо методи редукції системи п рівнянь ( $n \sim 10^1 - 10^2$ ) до системи значно нижчого порядку.

Нехай нам вдалося після низки перетворень і вибору відповідних масштабів представити систему (1.1) у вигляді

$$\varepsilon^2 \frac{dx_i}{dt} = F_i(x_1, x_2, \dots, x_n) \quad i = 1, \dots, l; \quad (1.7a)$$

$$\varepsilon \frac{dx_j}{dt} = F_j(x_1, x_2, \dots, x_n) \quad j = l+1, \dots, l+m; \quad (1.7\acute{a})$$

$$\frac{dx_k}{dt} = F_k(x_1, x_2, \dots, x_n) \quad k = l+m+1, \dots, n; \quad (1.7\hat{a})$$

тобто розташувати її за ступенями малого параметра  $\varepsilon$  при похідній .

Легко помітити , що коефіцієнти  $\varepsilon$  і  $\varepsilon^2$  фактично визначають швидкості зміни концентрацій . Справді , систему ( 1.7) можна представити також у вигляді

$$\frac{dx_i}{dt} = \frac{1}{T_1} F_i, \quad \frac{dx_j}{dt} = \frac{1}{T_2} F_j, \quad \frac{dx_k}{dt} = \frac{1}{T_3} F_k,$$

де  $T_1 = \varepsilon^2, T_2 = \varepsilon, T_3 = 1$  . Якщо ми цікавимося поведінкою всіх змінних , як на малих відрізках часу порядку  $\varepsilon^2$  , так і на часах порядку одиниці ; то нам необхідно досліджувати повну систему ( 1.7). Якщо ж нас цікавлять явища, що відбуваються в системі на середніх часах  $T_2 \sim \varepsilon$  , то рівняння ( 1.7a ) із постійною часу  $\varepsilon^2$  будуть описувати дуже швидкі процеси , а рівняння (1.7В), навпаки , - дуже повільні (порівняно з часом  $T_2$ ) процеси. Щодо останніх можна сказати, що за час  $T_2$  початкові значення концентрацій  $x_k$  не встигають помітно змінитися , тобто в останніх рівняннях ці повільні змінні можна замінити постійними ( початковими ) значеннями. Тим самим порядок системи ( 1.7) знижується.

Залишилася систему  $l+m$  рівнянь можна редукувати далі . Оскільки  $T_1$  час встановлення змінних  $x_i$  , - багато менше характерного часу  $T_2$  системи (1.7б ) , то ці змінні встигнуть досягти своїх стаціонарних значень раніше, ніж змінні  $x_j$  встигнуть помітно змінитися (для цього , звичайно, обов'язково, щоб система ( 1.7a ) , що описує швидкі процеси , мала стійке стаціонарний стан ) . Замінивши тепер в рівняннях ( 1.7б )  $x_i$  на їх стаціонарні значення , ми знову знизимо порядок системи і залишимо лише  $gp$  диференціальних рівнянь , характерні часи яких одного порядку ( $\sim T_2$ ) .

Залишається відкритим питання про те , як в реальній системі визначити характерні часи. Це питання аж ніяк не простий і вирішується різним чином залежно від виду рівнянь , що ми побачимо на конкретних прикладах , розглянутих у нашій книзі .

Слід відзначити , що зазначеним методом редукції давно користуються хіміки , називаючи його методом стаціонарних концентрацій .

Математичні аспекти цього методу були досліджені в роботах Понтрягнна [3], Тихонова і його учнів [9, 10]. Ми наведемо тут формулювання відомої теореми Тихонова і обговоримо ряд наслідків з неї.

Запишемо систему  $N$  рівнянь, частина з яких містить малий параметр  $\varepsilon$  перед похідної:

$$\varepsilon \frac{dx_p}{dt} = F_p(x_1, \dots, x_r, x_{r+1}, \dots, x_N), \quad p = 1, \dots, r; \quad (1.8a)$$

$$\frac{dx_q}{dt} = F_q(x_1, \dots, x_r, x_{r+1}, \dots, x_N), \quad q = r+1, \dots, N; \quad (1.8a')$$

Назвемо систему (1.8a) приєднаної, а (1.8a') при  $\varepsilon = 0$  - виродженої, або скороченої.

Теорема. Рішення повної системи (1.8) прагне до вирішення виродженої при  $\varepsilon \rightarrow 0$ , якщо виконуються наступні умови:

а) рішення  $\bar{x}_1 = \varphi_1(x_1, \dots, x_N), \dots, \bar{x}_r = \varphi_r(x_1, \dots, x_N)$  - ізольований корінь алгебраїчної системи

$$F_p(x_1, \dots, x_r, x_{r+1}, \dots, x_N) = 0 \quad p = 1, \dots, r; \quad (1.9)$$

(у його  $\varepsilon$  - округа немає інших коріння);

б) рішення  $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_r$  - стійка ізольована особлива точка приєднаної системи (1.8A) при всіх значеннях  $x_{r+1}, \dots, x_N$ ;

в) початкові умови  $x_1^0, x_2^0, \dots, x_r^0$  потрапляють в області впливу стійкої особливої точки приєднаної системи;

г) рішення повної і приєднаної систем єдині, а праві частини безперервні. У моделях біологічних процесів умови (а), (в) і (г), як правило, виконуються (випадки їх порушення зустрічаються рідко). Проте умова (б) порушується в широкому класі моделей релаксу цінних автоколивальних процесів. Зауважимо, що число початкових умов виродженої системи менше, ніж повній: всі початкові значення швидких змінних виявляються «зайвими» і ніяк не фігурують в виродженій системі. Теорема стверджує, що якщо виконана умова (в), то результат не залежить від початкових умов для змінних приєднаної системи. Розглянемо наступний приклад. Нехай ми маємо систему двох рівнянь

$$\frac{dx}{dt} = P(x, y), \quad \frac{dy}{dt} = Q(x, y). \quad (1.10)$$

Досліджуємо спочатку поведінку системи (1.10) на фазовій площині. Рівняння для фазових траєкторій буде

$$\frac{dy}{dx} = \frac{Q(x, y)}{P(x, y)}. \quad (1.11)$$

Проведемо головні ізокліни системи: ізокліни вертикалей, рівняння якої  $P(x, y) = 0$ , і ізокліни горизонталей  $Q(x, y) = 0$ . Припустимо, що ці криві мають в позитивному квадранті одну точку перетину  $(\bar{x}, \bar{y})$  і ця точка стійка (рис. 1.1).

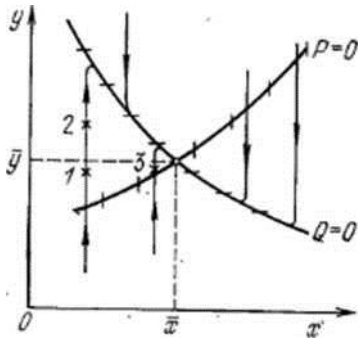


Рис. 1.1. Фазовий портрет системи (1.10).

Траєкторії системи в будь-якій точці площини, за винятком околиці лінії  $Q(x, y) = 0$ , мають нахил  $dy/dx \sim \varepsilon^{-1}$  тобто розташовані майже вертикально. На малюнку лінії зі стрілками схематично зображують хід інтегральних кривих. Розглянемо рух до точки  $(\bar{x}, \bar{y})$  з різних початкових точок (1, 2 і 3 на рис. 1.1). Швидкість руху по вертикальному участку порядку  $\varepsilon^{-1}$ . При наближенні до лінії  $Q=0$  рух уповільнюється і вздовж цієї ізокліни точка рухається зі швидкістю порядку одиниці. Тому час руху з точок 1 і 2, що відрізняються тільки початковими значеннями  $y^0$ , буде практично одне і теж і траєкторії будуть відрізнятися лише протягом часу  $\sim \varepsilon$ . Якщо ж розглянути рух з точки 3, в якій буде інше початкове значення  $x^0$ , то траєкторія буде сильно відрізнятися від перших двох і час руху по ній буде істотно менше. Звідси слід, що для опису поведінки системи протягом часу істотні лише початкові значення  $x^0$  і не істотні початкові значення  $y^0$ . Застосуємо тепер до системи (1.10) теорему Тихонова і порівняємо результати з проведенням якісних досліджень. Будемо вважати, що  $P(x, y)$  і  $Q(x, y)$  такі, що умови (а) - (г) виконані. Приєднаною системою буде в нашому випадку друге рівняння (1.10). Його стаціонарний стан визначається рівнянням  $Q(x, y) = 0$ , коріння якого

$$y = \bar{y}(x). \quad (1.12)$$

Ця функція, як ми вже знаємо, являє собою ізокліни горизонталей. Підставляючи (1.12) у перше з рівнянь системи (1.10), отримуємо вироджену систему, яка зводиться до рівняння

$$dx/dt = P(x, y(x)). \quad (1.13)$$

Це рівняння описує одновимірний рух зображає точки, а саме, повільний рух уздовж інтегральної кривої  $Q = 0$ . Швидкі рухи по вертикалях вироджена система взагалі не описує, у зв'язку з чим початкова умова для швидкої змінної

$y_0$  в виродженій системі не фігурує. Звідси видно, якою мірою  $\epsilon$  в яких інтервалах аргументу / вирішення виродженої системи близько до вирішення повній. У цьому прикладі стаціонарний стан як повною, так і виродженої системи - стійкий вузол. Розглянемо один конкретний приклад, в якому вироджена система допускає автоколивальний режим. Нехай маємо систему трьох рівнянь:

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = -x + \delta(1-z)y, \quad \epsilon \frac{dz}{dt} = -z + x^2 \quad (1.14)$$

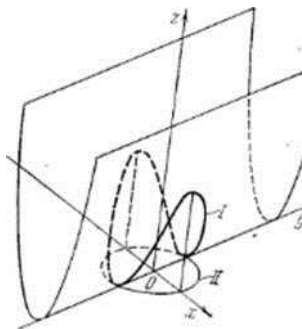


Рис. 1.2. Пространственный предельный цикл — кривая / на поверхности  $z = x^2$ . Кривая /7 — проекция предельного цикла / на плоскость

При  $\delta = 0$  з (1.14) папау сподівається класична система рівнянь Ван-дер-Поля, яка описує автоколивання змінної  $x(t)$  (або  $y(t)$ ) (1.1,6) і неодноразово буде нами використовуватися в далекішому викладі. На площині перемінних  $x, y$  система Ван-дер-Поля має фазовий портрет у вигляді стійкого граничного циклу, на який ззовні і зсередини навиваючи траєкторії дві точки (особлива точка - нестійкий фокус).

Нехай тепер  $\epsilon \neq 0$ , але це означає, що час встановлення змінної  $z$  в системі (1.14) багато менше всіх інших характерних часів системи. У тривимірному просторі приєднане рівняння буде

$$\epsilon dz/dt = -z + x^2,$$

стаціонарне рішення якого - парабола  $z = x^2$ . У тривимірному просторі граничний цикл системи (1.14) буде розташований на параболічній поверхні  $z = x^2$ . Проекція просторового граничного циклу на площину  $z = 0$ ,  $x, y$  буде наближено збігатися з граничним циклом системи при  $\epsilon = 0$  (рис. 1.2). Нехай

тепер зображає точка виявляється збитої з трьохмірного граничного циклу . Так як приєднане рівняння має стійке рішення при всіх дійсних  $x$  , то точка знову потрапляє на поверхню  $z = x^2$  . Це відбудеться практично миттєво через малість  $\epsilon$  , після чого точка знову буде рухатися по просторових спіралях , розташованих на цій поверхні , наближаючись до граничного циклу. Отже , практично можна замість дослідження руху зображає точки в просторі провести його в площині  $x, y$ .

Корисно згадати випадок, в якому порушується умова (б) теореми Тихонова. Це має місце, наприклад, коли притягаюча ізокліни має S-подібну форму, так що стаціонарні стану на її проміжної гілки (де  $dy/dx > 0$ ) нестійкі. У цьому випадку редукція по Тихонову можлива лише в обмежених областях фазового простору. Релаксаційні системи з S-характеристикою вельми поширені; ми обговоримо їх властивості в гол. 7, 9, 10.

### **Редукция систем и теория катастроф.**

У попередньому параграфі передбачалося , що малий параметр (відношення характерних часів ) існує апіорі. Дійсно , в багатьох фізичних завданнях є чітка тимчасова ієрархія , що , за нашим переконанням , не випадково , а визначається специфікою систем . Однак навіть в скороченій ( виродженої ) системі , де всі часи мають однаковий порядок , можна виділити таку область параметрів і змінних , в кото  $\tau$  рій є можливість додатково знизити порядок системи .

Це має місце поблизу точки біфуркації , коли одне з характеристичних чисел  $\rho$  ( або його речова частина) звертається в нуль і змінює знак . Одночасно змінюється топологія фазового портрету , зокрема , число особливих рішень . Можна сформулювати дві « теореми відомості » : першу - для біфуркацій седлового типу і другу - для біфуркацій фокусної типу. Підкреслимо , що теореми відомості носять локальний характер , вони справедливі в обмеженій області фазового простору , де відхилення змінних від стаціонарних значень достатньо малі , на відміну від теореми Тихонова , має глобальний характер. Почнемо з сідлової біфуркації .

Нехай система  $n$ -го порядку (1.1) має стійку особливу точку ( всі характеристичні числа мають негативні речові частини). Нехай при зміні параметрів одне з характеристичних чисел  $\rho_i = \epsilon$  (  $\epsilon \ll 1$  , будучи чисто речовим, проходить через нуль і стає позитивним, тоді як речові частини решти характеристичних чисел залишаються від'ємного ( $\text{Re } \rho_i < 0, i > 1$ ) і великими ( $|\text{Re } \rho_i| \gg |\epsilon|$ ).



Ми не припускаємо тут, що особлива точка ізольована, більше того, в загальному випадку при зміні параметрів в окрестності особливої точки обов'язково є (або з'являються) інші особливі точки. Розглянута особлива точка при цьому може вийти чи зникнути. У цьому зв'язку розкласти праві частини (1.1) за відхиленнями змінних від особливої точки не коректно. Можна, однак, розкласти поблизу якої іншої точки, що знаходиться в околиці особливої. Враховуючи це, систему (1.1) можна представити у вигляді

$$\frac{dx'_1}{dt} = \alpha - \varepsilon x'_1 + \{x'^2\}_1 + \{x'^2\}_1 + \dots \quad (1.15)$$

$$\frac{dx'_i}{dt} = \alpha x'_i + \{x'^2\}_i + \{x'^2\}_i + \dots \quad (i = 2, 3, \dots, n)$$

Тут  $x'_i (i > 1)$  — відхилення змінних від їх стаціонарних значення:  $x'_1 = x_1 - \bar{x}_1$ ; відхилення  $X_1$  від деякого значення, від якого від стаціонарного на величину порядку  $\varepsilon$ ;  $\alpha$ -параметр, пов'язаний з  $\varepsilon$  (він має порядок  $\varepsilon^2$ ). Фігурними дужками позначені форми з алгебраїчними членами типу  $\{x'^2\} = \sum_{i=1}^n b_{ki} x'_i x'_i$ ;  $\{x'^3\}$  визначається аналогічно. Розділимо всі рівняння (1.15) на  $\varepsilon$  і введемо новий час  $t' = \varepsilon t$

$$\frac{dx'_1}{dt'} = \alpha' - x'_1 + \frac{1}{\varepsilon} \{x'^2\}_1 + \frac{1}{\varepsilon} \{x'^3\}_1 + \dots \quad (1.16a)$$

$$\varepsilon \frac{dx'_i}{dt'} = \alpha x'_i + \{x'^2\}_i + \{x'^3\}_i + \dots \quad (1.16b)$$

Система (1.16b) є приєднаною із стійкою особливою точкою  $x_i = 0$  (1-2, 3, .. п).

Тому, згідно теореми Тихонова, можна покласти  $\varepsilon = 0$  і

$$x'_i = [\{x'^2\}_i + \{x'^3\}_i + \dots] \alpha_i^{-1} \quad (1.17)$$

Підставимо в рівняння (1.16a) і перейдемо знову до часу

При цьому отримаємо

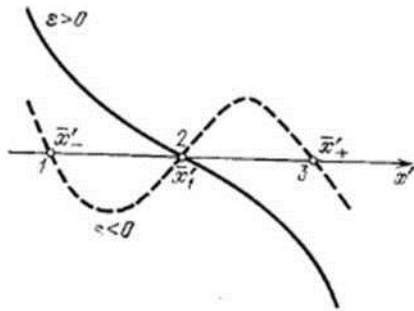
$$\frac{dx'_1}{dt} = \alpha - \varepsilon x'_1 + b x'^2_1 + c x'^3_1 + o(x'^4_1) \quad (1.18)$$

де під  $o(x'^4_1)$  маються на увазі члени малості  $x'^4_1$  і вище.

Тепер, щоб дослідити характер біфуркацій в повній системі (1.16), потрібно вивчити, як змінюється тип особливої точки в залежності від параметрів  $\varepsilon, \alpha, b$  і з цього лише одного рівняння (1.18). У цьому і полягає теорема відомості.

З'ясуємо, які класи біфуркацій можливі в даному випадку, а) Поблизу біфуркаційних значень параметрів рівняння

$$\alpha - \varepsilon x_1' + bx_1'^2 + cx_1'^3 = 0 \quad (1.19)$$



має один дійсний корінь  $\bar{x}'_1$  і два комплексних кореня з малими уявними частинами і однаковими речовими частинами, близькими до  $\bar{x}'_1$ . У межі така ситуація описується поліномом (1,19) при  $a = b = 0$ . У першому важливому випадку ( $c < 0$ ) при  $\varepsilon > 0$  мається один стійкий стан рівноваги  $\bar{x}'_1 = 0$  (рис. 1.3). При  $\varepsilon < 0$  це

Рис. 1.3. Біфуркація в рівнянні (1.18), відповідна катастрофі типу «збірка».

стан стає нестійким, але поруч з ним з'являються два стійких:  $\bar{x}'_{1,\pm} = \pm \sqrt{\varepsilon/c}$ . Це означає, що система (1.18) стає триггерной і описує

поведінку повної системи (1.16) в області зміни змінних  $|\bar{x}'_1| \leq \sqrt{|\varepsilon|/c}$  при необмежений часи Еволюція системи при переході  $\varepsilon$  через нуль вчиняється плавно. Такий процес можна назвати м'яким переходом в новий стан. В іншому випадку ( $c > 0$ ) при  $\varepsilon > 0$ ) є три близьких стаціонарних стану, з них середнє ( $\bar{x}'_1 = 0$ ) устійливо. При  $\varepsilon < 0$  бічні стану зникають, а середнє стає нестійким; при цьому область застосовності редукованою системи (1.18)

вузька - вона обмежена значеннями  $x_1' \leq \sqrt{\varepsilon/c}$  і часами порядку  $t \sim \varepsilon^{-1}$

б) Нехай тепер реальні частини комплексних коренів і вещественний корінь  $\bar{x}'_1$  істотно розрізняються. Тоді в результаті біфуркації в системі (1.16) виникають або зникають два близьких стаціонарних стану розташованих далеко від третього  $\bar{x}'_+$ . У цьому випадку в (1.19) можна знехтувати членом  $cx_1'^3$ .

Тоді, за л Ожен я точок 1 і 2 наближено визначаються корінням квадратного рівняння

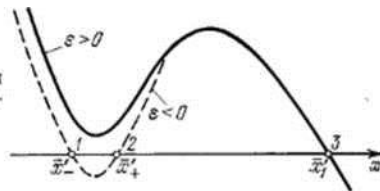
$$\bar{x}'_{+,-} = (\varepsilon \pm \sqrt{\varepsilon^2 - 4ab}) / (2b)$$

Простежимо біфуркацію при зміні знака  $\varepsilon$ . Якщо  $a\bar{b} > 0$ , то при  $\varepsilon > 2\sqrt{ab}$  існує один стійкий стан  $(\bar{x}^-)$  і одне нестійке  $(\bar{x}^+)$  поблизу нього на відстані  $\sim \frac{\varepsilon}{b}$  ( мал. 1.4).

В інтервалі  $-2\sqrt{ab} < \varepsilon < 2\sqrt{ab}$  обидва стаціонарних стану зникають і з'являються знову при  $\varepsilon \leftrightarrow 2\sqrt{ab}$ . В останньому випадку ліве стан стійко, а праве нестійка. При еволюції параметрів відбувається «анігіляція» двох стаціонарних станів і система рухається до третього стійкого стану.

Останній, однак, розташоване далеко від області  $x'_1 \sim \varepsilon (\bar{x}'_1 \sim -\frac{b}{c})$ , і цей процес вже не описується одним рівнянням (1.1). Таким чином, відбувається «жорстке» порушення нової режиму. Область застосовності теореми відомості при цьому дуже вузька - вона обмежена значеннями  $x'_1 \sim \varepsilon$  і часом  $t \sim \varepsilon^{-1}$

Рис. 1.4. Біфуркація в уравнении (1.18), соответствующая катастрофе типа «складка».



Найбільшою областю застосовності теореми відомості при біфуркації седлового типу володіють системи, які приблизенно описуються рівнянням виду

$$\frac{dx'_1}{dt} = -\varepsilon x'_1 - c x'^3_1 \quad (c > 0)$$

Вище передбачалося, що вдалині від біфуркації характерні часи зміни  $x'_1$  і  $x'_1$  одного порядку. У приватному, але важливому випадку змінна  $x'_1$  може виявитися «швидкої», тобто її характерний час  $T_1 \ll 1$ . Тоді викладене вище справедливо в області  $\varepsilon \ll T_1$  і втрачає силу при  $\varepsilon \gg T_1$ . Однак твердження про те, що завдання зводиться до одновимірної, залишається вірним і в останньому випадку, тобто в області фазового простору, де точка рухається по координаті  $x_1$  а решта розглядаються як параметри.

Розглянемо тепер теорему відомості для біфуркації фокусного типу. Нехай за допомогою лінійної заміни змінних система (1.1) наведена до виду

$$\begin{aligned} \frac{dx'_1}{dt} &= a_{11}x'_1 + a_{12}x'_2 + \{x'^2\}_1 + \{x'^2\}_1 + \dots, \\ \frac{dx'_2}{dt} &= a_{21}x'_1 + a_{22}x'_2 + \{x'^2\}_2 + \{x'^2\}_2 + \dots, \end{aligned} \quad (1.20)$$

$$\frac{dx'_i}{dt} = \sum_{j=3}^n a_{ij}x'_j + \{x'^2\}_i + \{x'^2\}_i + \dots \quad (i=3,4,\dots,n) \quad (1.260)$$

При цьому  $a_{12} = -a_{21} = a$ ,  $|a_{11}| \sim |a_{22}| \sim \varepsilon$ , Тоді характеристичний коріння лінеаризованної системи (1.20) представляються у вигляді  $p_{1,2} = -\varepsilon \pm ia$ . Користуючись децицею величин  $A1$  і  $a22$ , можна поки  $\sim$  зати, що і в цьому випадку підсистема (1.206) є приєднуваль  $\sim$  ненной, а вся система (1.20) редукується до двох рівнянь, що містять тільки змінні  $x'_1$  і  $x'_2$ . Доказ цього твердження належить Шошітайшвілі [5, 12].

На відміну від попереднього, в даному випадку стаціонарне стан є ізольованим незалежно від наявності або відсутності квадратичних членів  $\{x'^2\}_{1,2}$ . Область застосовності двомірного наближенні залежить від нелінійних членів. Якщо  $\varepsilon < 0$ , то в системі виникає м'яке збудження. За наявності предель  $\sim$  ного циклу (як це було в прикладі § 3) область значень  $x'_{1,2}$ , в якій справедлива теорема відомості, обмежена величинами порядку амплітуди граничного циклу і не обмежена за часом. В іншому випадку вона обмежена умовами  $|x'_{1,2}| < \sqrt{\varepsilon}$  и  $t \leq \varepsilon^{-1}$

Розглянуті біфуркації не зводяться один до одного. Поблизу інших, більш складних біфуркацій  $l$ -мірна модель може бути замінена еквівалентною трьох-, чотирьох- і т. д. - мірної в залежно  $\sim$  сті від складності біфуркацій.

Як вже згадувалося, теорія біфуркацій близька в ідейному відносини до теорії «катастроф». Сам термін, а також ряд основних понять цього напрямку були запропоновані Рене Томом. Теорія катастроф має як методологічний, так і чисто практичний аспекти і їй присвячена багата література (див. [12] і бібліографію там). Наша мета - обговорити тут найпростіші катастрофи у зв'язку з математичними моделями, якими ми займаємося.

Нехай модель має лише одну змінну  $x$ :

$$dx/dt = P(x) \text{ или } dx/dt = -dV(x)/dx \quad (1.21)$$

Де  $V(x) = -\int P(x)dx$ . Величину  $V(x)$  називають потенціалом,

оскільки (1.21) описує рух частинки в потенційному полі  $U(x)$  в сильно в'язкому середовищі. Чгоби вивчити рух системи поблизу локальних екстремумів  $V(x)$ , функцію  $P(x)$  розкладають в ряд близько стаціонарних станів і обмежуються кількома молодшими членами розкладання. Мінімальне число членів  $p$  і число параметрів  $L$ , які потрібно враховувати, визначаються ступенем виродження стаціонарного стану. Вони пов'язані

простим відносинами :  $K = n - 1$  ; число до називають коразмірністю катастрофи. Розглянемо найпростіші випадки .

1 . Особлива точка ізольована , що не вироджена і стійка. У цьому випадку мінімальна модель має вигляд

$$P(x) = -x, \quad V(x) = x^2/2 \quad (n=1, k=0) \quad (1.22)$$

Тут і нижче  $x$  - це відхилення від стаціонарного стану . Форма  $dx/dt = -x$  є в цьому випадку загальною і до неї локально може бути приведена будь-яка одномірна модель шляхом вибору масштабів  $x, t$  і зсуву  $x = x' + a$

2 . Злиття двох особливих точок . Тоді мінімальна модель

$$P(x) = u_1 + x^2 \quad V(x) = -u_1 x - x^3/3 \quad (n=2, k=1) \quad (1.23)$$

Форма  $P(x)$  в ( 1.23 ) також є загальною і до неї наводиться будь квадратичний поліном . При  $u_1 < 0$  є два стану , одне з яких стійко , а інше нестійка. При  $u_1 = 0$  вони зливаються , і при  $u_1 > 0$  обидва зникають. При зміні параметра в зворотньому напрямку , навпаки , з'являються дві особис точки . ця ситуація іменується катастрофою типу «складки», саме вона розглядається в п. 1, а попереднього розділу.

Пояснимо тепер, яким чином можна користуватися образами теорії катастроф при вивченні математичних моделей. Зробимо це на прикладі моделі другого порядку, яка містить змінні  $x$  і  $u_1$ . Нехай  $x$  є швидкої змінної, але виключити її неможна, так як швидкий процес не скрізь стійкий. «Складка» відпоствует моделі:

$$\tau \frac{du_1}{dt} = \tilde{Q}(u_1, x) \quad (1.24a)$$

$$\frac{dx}{dt} = \tilde{P}(u_1, x) = -u_1 + x^2 \quad (1.24b)$$

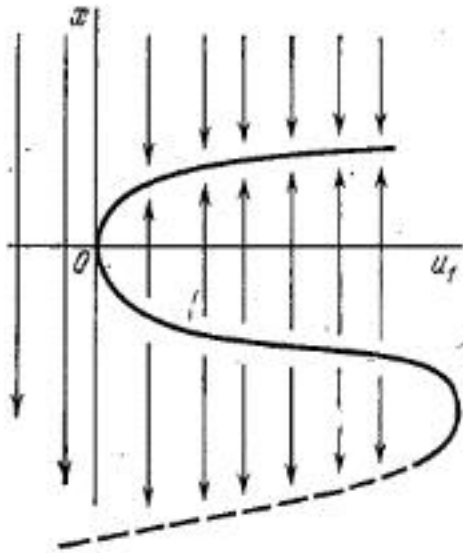


Рис. 1.5. «Складка» на площині  $x, u_1$ .

де  $\tau \gg 1$ ; характерний час зміни  $x$  прийнято за одиницю. Фазовий портрет системи представлений на рис. 1.5. Ізокліни  $P = 0$  має стійку гілку - аттрактор у формі складки (звідки ясно походжу  $\neg$  дення цього терміну). При повільному зміні  $u_1$  відповідно до (1.24а) при досягненні значення  $u_1 = 0$  відбуватися зрив зображає точки і пере  $\neg$  скок на інший стійкий аттрактор («катастрофа»), цей аттрактор зобра  $\neg$  дружин на рис. 1.5 умовно (пунктиром), оскільки він не описується квадрата  $\neg$  тичної формою (1.24б) і визначається глобальними властивостями моделі. Заметім, однак, що в «коректних» моделях, адекватних реальним об'єктам, він завжди присутній.

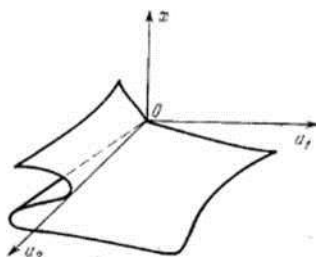
Підкреслимо, що форма функції  $P(x)$  (і потенціалу  $V$ ) є загальною або універсально деформовуваною. Вона не змінюється при малих деформаціях. Іншими словами, характер катастрофи залишаються структурно стійким, оскільки зливаються дві (і тільки дві) особливі точки. Поняття структурної стійкості в теорії катастрофи має важливе методологічне значення.

В ідейному відношенні воно відповідає поняттю «грубості» моделі, введеному Андроном ще задовго до появи «теорії катастроф». Значення його полягає в тому, що якісні ви  $\neg$  води, отримані на основі грубої (або структурно стійкою) моделі, є загальними і залишаються справедливими, навіть якщо параметри моделі визначені не точно або варіюють від випадку до випадку. У біології це властивість особливо важливо.

Катастрофи типу складки з'являються в моделях, що описують релаксаційні автоколивання, так звані «що чекають» режими і тригерні системи (див. рис. 9.3). У розподілених системах моделі типу складки

використовуються для опису автоволнових процесів і дисипативних структур. Слияние трех особых точек. Тогда минимальная модель имеет вид

$$\begin{aligned} P(x) &= x^3 + u_1 x + u_2 \\ V(x) &= -x^4/4 - u_1 x^2/2 - u_2 x \end{aligned} \quad (n=3, k=2) \quad (1.25)$$



1.6. «Сборка» в пространстве  $(x, u_1, u_2)$ .

Тут  $P$ - кубічний поліном, що залежить від двох параметрів  $u_1$  і  $u_2$ . Відповідна математична модель містить три змінні. Ізокліни поверхню  $dx/dt = 0$  (аттрактор) представлена на рис. 1.6. Видно, що на ній є збірка, вершина якої відповідає злиттю трьох особливих точок, що має місце при 0. На ребрах збірки мають іще катастрофи типу складки. Таким чином, в тривимірному фазовому просторі складці відповідає більш потужне безліч, ніж збірці. Моделі, що містять катастрофу типу збірки, використовуються для опису релаксаційних автоколивань малої амплітуди, коливальних режимів зі зміщенням «середньої точки» і дисипативних структур ступеневої типу.

4. Злиття чотирьох і п'яти особливих точок; відповідні катастрофи називаються «ластівчин хвіст» і «метелик». Фазові простору при цьому чотири-і п'ятимерних і геометричні предствалення цих катастроф не настільки наочні. Підкреслимо суттєву відмінність катастроф типу складки і збірки. У разі Форма (1.23) не описує поведінку системи при більших часах. Зображає точка йде з розглянутої локальної області Фазового простору (де справедлива форма т іншими словами, катастрофа типу складки не локалізуєма. Те ж відноситься і до катастрофи «ластівчин хвіст» з парною коразмірністю.

У разі «збірки» форма (1.25) описує поведінку системи і при великих часах, оскільки зображає точка залишається поблизу колишнього стаціонарного стану (на відстані. Можна сказати, що катастрофа типу збірки локалізуєма; це відносно стосується і до катастрофи «метелик» з непарної коразмірністю.

В закінчення цього параграфа згадаємо знову про «дивних аттракторах» у зв'язку з теорією біфуркацій.

Перш за все, дивний аттрактор може виникнути у разі біфуркації типу збірки, при цьому «перемішування» траєкторій здійснюється за рахунок зривів зображає точки з країв збірки. Інший шлях виникнення дивного аттрактора - це слідує один за одним біфуркації подвоєння періоду коливань в автономній автоколебательній системі. Послідовність біфурційних значень деякого параметра  $\mu_i$  сходиться; при цьому для самих різних систем ставлення послідовних різниць біфуркаційних параметрів прагне до універсальної константи Фейгенбаума  $F=4,6692$ .

Зазначимо, що в обох випадках розмірність дивного аттрактора менше розмірності всього фазового простору (у разі системи третього порядку - це деяка поверхню або навіть частина її). Можна навести приклад дивного аттрактора, що займає все фазовий простір. Це система типу «більярд Синая» (див. рис. 12.2), в якій розглядається поведінка кульки на ділянці площині, обмеженій відбивають опуклими стінками. Система консервативна і фазовий простір її чотирирівимірному (дві координати і два імпульси). Кулька здійснює випадковий рух по площині, відбиваючись від криволінійної стінки, і вимірюється свої координати і імпульси так, що фазові траєкторії заповнюють рівномірно всі фазовий простір (відбувається повне перемішування траєкторій). Біфуркація в цій системі, приводячи до появи дивного аттрактора, виникає в результаті зміни параметра - кривизни стінки. Використання моделі.



## Розділ 5.

# ФРАКТАЛИ

### 5.1. Поняття «фрактал»

Поняття *фрактал* і *фрактальна геометрія* з'явилися в кінці 70-тих років минулого віку. Слово «*фрактал*» утворене від латинського *fractus* і в перекладі означає *зібраний з фрагментів*. Воно було запропоноване Бенуа Мандельбротом у 1975 році для позначення нерегулярних, але самоподібних структур. Народження фрактальної геометрії прийнято пов'язувати з виходом в 1977 році книги Мандельброта «*The Fractal Geometry of Natur*». В його роботах використані наукові результати інших вчених, працюючих в період 1875-1925 років – Пуанкаре, Фату, Жюліа, Кантор, Хаусдорф. Але тільки в наш час вдалося об'єднати їх роботу в єдину систему.

Визначення фрактала, яке дав Мандельброт, звучить так: «*Фракталом називається структура, яка складається з частин, які в деякому сенсі подібні цілому*».

Однією з основних властивостей фракталів є самоподібність. В самому простому випадку невелика частина фрактала містить інформацію про весь фрактал. Найбільш відомими фрактальними об'єктами, які з'явилися задовго до самого слова "фрактал", можна назвати криву Коха (аналог в природі – сніжинка), трикутну і універсальну криву Серпинського, фігуру Ліхтенберга, яка з'являється в вигляді узора на пластинці діелектрика після його пробою.

**Фігури Ліхтенберга** – картина розподілення іскрових каналів, розмішених на поверхні твердого діелектрика при так званому ковзкому іскровому розряді (рус. *скользящий искровой разряд*). Вперше подібні розряди спостерігались Г.К. Ліхтенбергом в 1777. В сильному розряді високі тиски і температури в іскрових каналах деформують поверхню діелектрика, утворюючи фігури Ліхтенберга. При малих розрядах фігури Ліхтенберга відповідають вибірковій поляризації діелектрика, і їх можливо зробити помітними, посипаючи поверхню діелектрика спеціальним порошком або проявляючи фотопластинку, підкладену під час розряду під шар діелектрика. Фігури Ліхтенберга біля аноду і катоду різко відрізняються за зовнішнім виглядом (рис. 5.1); тому за їх допомогою можливо встановити, від якого з цих електродів утворювались іскрові канали (т. з. полярність іскрового розряду).

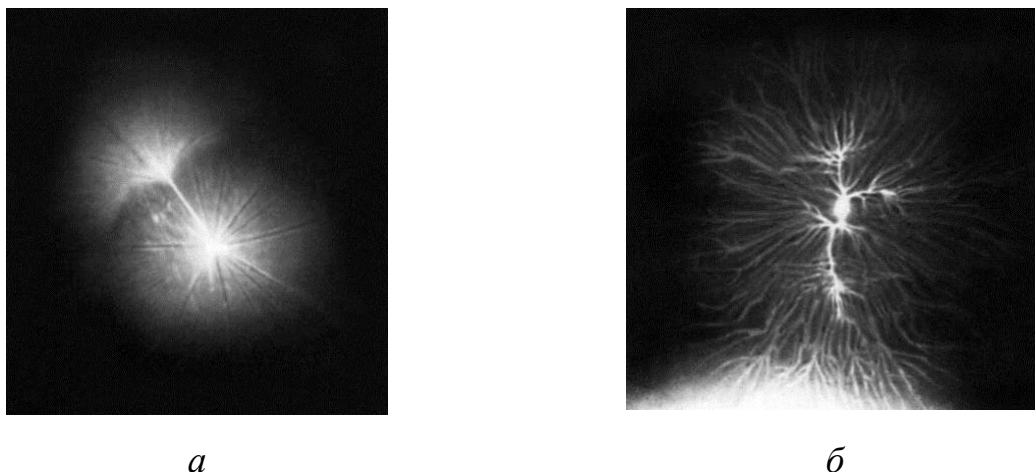


Рис. 5.1. Ліхтенберга фігури:

*a* – для від’ємного заряду; *б* – позитивного заряду

Велику роль фрактали відіграють в машинній графіці. Вони приходять на допомогу, наприклад, коли необхідно, за допомогою декількох коефіцієнтів, задати лінії і поверхні дуже складної форми. З точки зору машинної графіки, фрактальна геометрія незамінна при генерації штучних хмаринок, скель, поверхні моря.

## 5.2. Класифікація фрак талів

### 5.2.1. Геометричні фрак тали

Фрактали даного класу найбільш наочні. В двохвимірному випадку їх отримують за допомогою деякої ламаної (або поверхні в трьохвимірному випадку), яка називається *генератором*. За один крок алгоритма кожний із відрізків, які складають ламану, замінюються на ламану-генератор, в відповідному масштабі. В результаті нескінченного повторення цієї операції, утворюється геометричний фрактал.

Одним із стандартних прикладів фракталів цього класу є триадна крива Коха. Побудова кривої Коха починається з прямолінійного відрізка одиничної довжини  $L(1) = 1$ . Цей вихідний відрізок називається *затравка* і може бути замінений будь-яким багатокутником, наприклад рівностороннім трикутником, квадратом.

Затравка – це нульове покоління кривої Коха. Побудова кривої Коха продовжується: кожна ланка затравки замінюється утворюючим елементом, який позначений на рис. 5.2 через  $n = 1$ . В результаті такої заміни ми отримуємо 1-е покоління – криву з чотирьох прямолінійних ланок, кожна довжиною по  $1/3$ . Довжина всієї кривої 1-го покоління складає величину  $L(1/3) = 4/3$ . Наступне покоління отримуємо при заміні кожної прямолінійної ланки зменшеним утворюючим елементом. В результаті ми отримуємо криву 2-го покоління, яка складається з  $N = 4^2 = 16$  ланок, кожна довжиною  $\delta = 3^{-2} =$

1/9. Довжина кривої 2-го покоління складає  $L(1/9) = (4/3)^2 = 16/9$ . Замінюючи всі ланки попереднього покоління кривої зменшеним утворюючим елементом, отримуємо нове покоління кривої. При  $n$  спрямованому до нескінченності крива Коха стає фрактальним об'єктом.

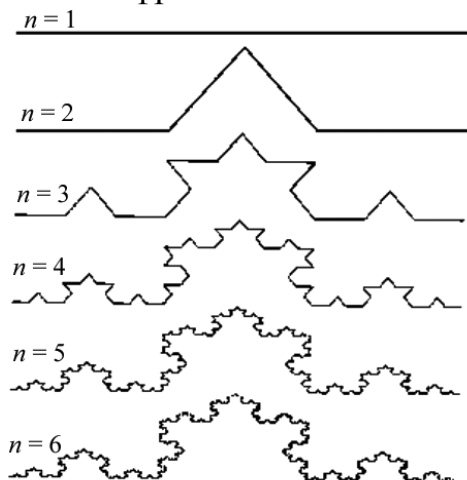


Рис. 5.2. Побудова тріадної кривої Коха

### 5.2.2. Алгебраїчні фрактали

Це сама велика група фракталів. Отримують їх за допомогою нелінійних процесів в  $n$ -вимірних просторах. Найбільш вивчені двовимірні процеси. Визначаючи нелінійний ітераційний процес, як дискретну динамічну систему, використовувати термінологію теорії нелінійної динаміки: *фазовий портрет, усталений, атрактор* та ін.

Відомо, що нелінійні динамічні системи мають декілька стійких станів. Той стан, в якому опинилась динамічна система після деякого числа ітерацій, залежить від її початкового стану. Тому кожний стійкий стан (або як говорять – атрактор) володіє деякою областю початкових станів, з яких система обов'язково опиниться в розглянутих кінцевих станах. Таким чином фазовий простір системи ділиться на *області притягання* атракторів. Якщо фазовим є двовимірний простір, то розфарбовуючи області притягання різними кольорами, можливо отримати *кольоровий фазовий портрет* цієї системи (ітераційного процесу). Змінюючи алгоритм вибору кольору, можна отримати складні фрактальні картини з примхливими багатокольоровими візерунками (рис. 15.3).

Класичними прикладами алгебраїчних фракталів є множини Жюліа і Мандельброта, які отримують за допомогою простої ітераційної функції:

$$z_{n+1} = z_n + c, \quad (5.1)$$

де  $z$  – комплексна функція двох змінних  $x$  та  $y$ ;  $c$  – комплексне число ( $a$  і  $b$  – дійсні числа);  $n = 0, 1, 2, \dots$  – натуральні числа.

$$z = x + iy$$

$$c = a + ib$$

Поведінка послідовності (5.1) залежить від параметра  $c$  і початкової точки  $Z(0)$ . Якщо зафіксувати  $c$  і змінювати  $z(0)$  в полі комплексних чисел, то одержимо множину Жюліа, а якщо зафіксувати  $z(0) = 0$  і змінювати параметр  $c$ , то одержимо множину Мандельброта.

Якщо узяти  $z(0)$  далеко від нуля, то послідовність буде швидко прямувати до нескінченності. Це, вірно також і тоді, коли точка  $z(n)$  для деякого  $n$  знаходиться далеко від нуля. Але існують і такі значення  $z(0)$ , для яких послідовність  $z_n$  ніколи не віддаляється, а завжди залишається обмеженою.

Чітке математичне визначення множини Жюліа формулюється наступним чином: множина Жюліа функції  $f$ , яка позначається  $J(f)$ , визначається як

$$J(f) = \partial \left\{ z : f^{(n)} \rightarrow \infty \text{ при } n \rightarrow \infty \right\}.$$

Таким чином, множина Жюліа функції  $f$  є границя множини точок  $z$ , які прямують до нескінченності при ітеруванні функції  $f(z)$ .

Для множини Мандельброта  $M$  для полінома  $f_c(z) = z^2 + c$  визначається як множина усіх  $c \in \mathbb{C}$ , для яких орбіта нульової точки обмежена:

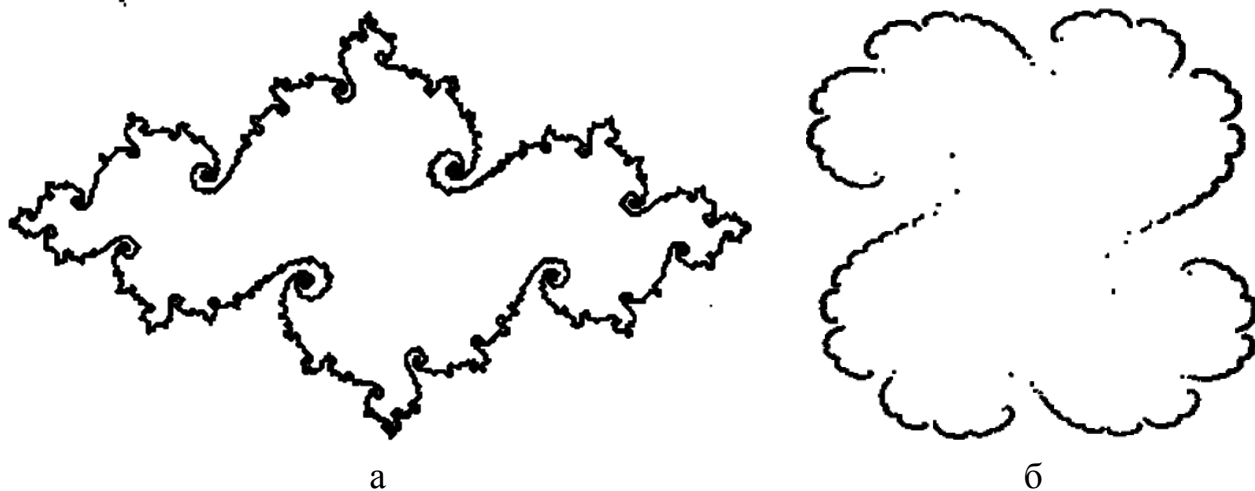
$$M = \left\{ c \in \mathbb{C} : \left\{ f_c^{(n)}(0) \right\}_{n=0}^{\infty} \text{ обмежена} \right\}$$

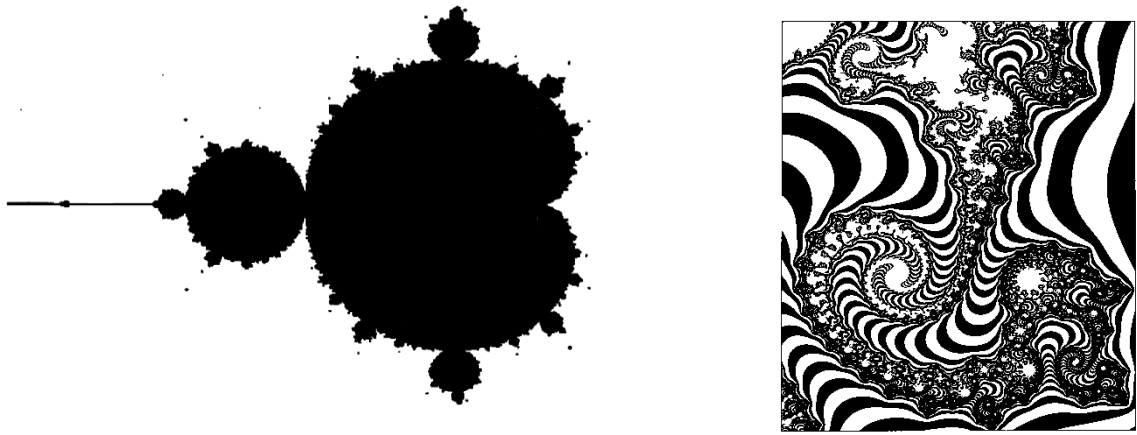
чи

$$M = \left\{ c \in \mathbb{C} : f_c^{(n)}(0) \neq \infty \text{ при } n \rightarrow \infty \right\}$$

Вид множини Жюліа і Мандельброта суттєво залежить від вибору параметра  $c$  чи початкової точки  $z(0)$ , відповідно. Змінюючи ці параметри, можна одержати неймовірну різноманітність цих множин.

Є два основні типи множини Жюліа: деякі з них є цілісними (зв'язані множини), а інші є хмарами з крапок (Канторови чи незв'язані множини). Приклади таких множин приведені на рис. 15.3.





В

Рис. 5.3 Множини Жюліа та Мандельброта:

$a$  – зв’язана множина Жюліа для  $z^2 - 0.7382 + 0.0827i$ ;

$b$  – незв’язана множина Жюліа для  $z^2 + 0.50$ ;

$c$  – множина Мандельброта для  $z^2 + c$ ;

$z$  – фрагмент множини Мандельброта для  $z^2 + c$  при кольоровому кодуванні значень точок  $z_n$

### 5.2.3. Стохастичні фрактали

Ще одним відомим класом фракталів є стохастичні фрактали, які можливо отримати в тому випадку, коли в ітераційному процесі випадковим способом поміняти будь-які параметри. При цьому отримуємо об’єкти дуже схожі на природні несиметричні дерева, порізані берегові лінії та ін. Двовимірні стохастичні фрактали використовують при моделюванні рельєфу місцевості і поверхні моря .

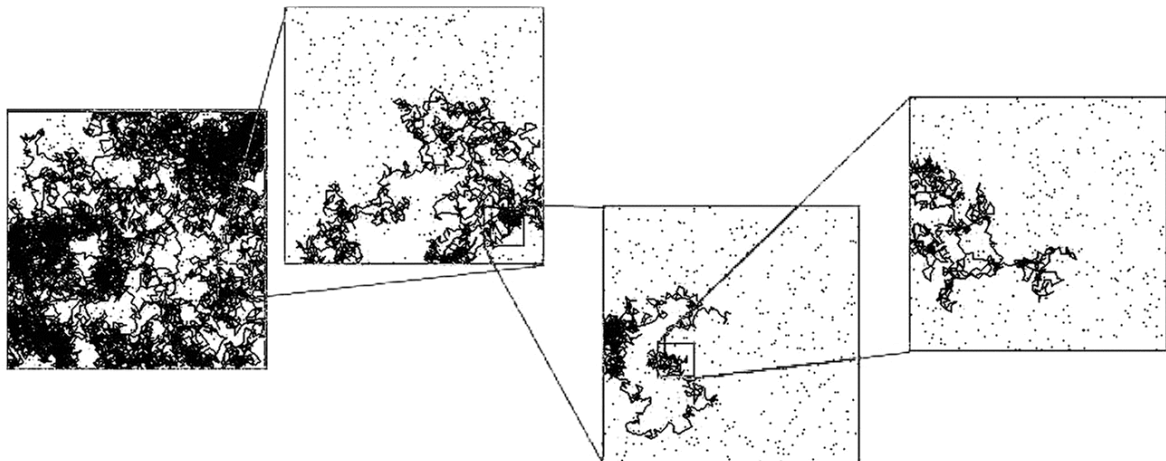


Рис. 5.4. Особливий вид стохастичних фракталів складають так звані *фрактальні кластери*. Класична модель для дослідження цих видів структур: модель агрегації, обмеженої дифузією (diffusion limited aggregation, DLA) описує агрегацію частинок в умовах їх випадкового руху (Witten, Sander,

1981; Сандер. 1987).

Стохастичні фрактали часто отримують за допомогою додавання деякого випадкового параметру в детермінований фрактал, що і найчастіше зустрічається в природі (рис.5.4 ).

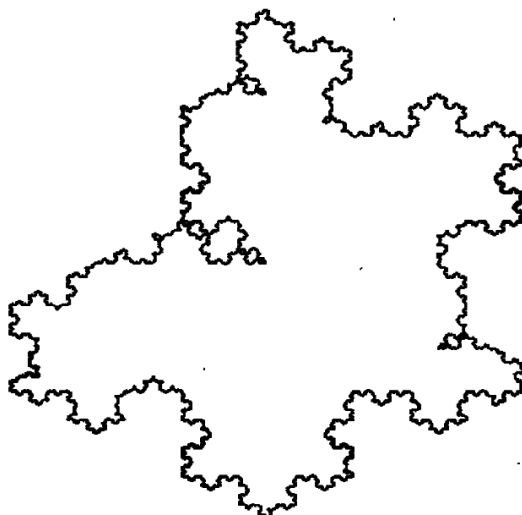


Рис. 5.5. Рандомізована крива Коха

Існують і інші класифікації фракталів, наприклад, ділення фракталів на детерміновані (алгебраїчні і геометричні) і не детерміновані (стохастичні).

### 5.3. Самоподібність і фрактальна розмірність

Поділимо відрізок прямої на  $N$  рівних частин. Потім кожен частину можна вважати копією всього відрізка, зменшеною в  $1/r$  раз. Отже,  $N$  і  $r$  пов'язані співвідношенням  $Nr = 1$  (рис. 15.2). Якщо квадрат розділити на  $N$  рівних квадратів ( з площею, в  $1/r^2$  раз менше площі вихідного ), то співвідношення матиме вигляд  $Nr^2 = 1$ . Якщо куб розділити на  $N$  рівних кубів (з об'ємом, в  $1/r^3$  раз менше об'єму вихідного ), то співвідношення матиме наступний вигляд:  $Nr^3 = 1$ . Помітно, що розмірність  $d$  об'єкту будь-то одновимірний відрізок, двовимірний квадрат або трьохвимірний куб, з'являється як ступінь  $r$  в співвідношенні між  $N$ , числом рівних підоб'єктів, і коефіцієнтом подібності  $r$ . А саме:

$$Nr^d = 1 \quad (5.2)$$

Множини, побудовані на рис. 15.6, мають цілу розмірність. Задася питанням, чи можлива така побудова, при якій показник  $d$  в формулі (5.2) не являється цілим, тобто таке, що при розбитті вихідної множини на  $N$  підмножин, що не перетинаються, отриманих масштабування оригіналу з коефіцієнтом  $r$ . Як ми переконаємося, відповідь – так! Таку множину називають *самоподібним фракталом*. Величину  $d$  називають *фрактальною (дробною) розмірністю* або *розмірністю подібності*. Явний вираз для  $d$  через  $N$  і  $r$  знаходиться логарифмуванням обох частин (5.2):

$$d = \frac{\log N}{\log 1/r}$$

Логарифм можна взяти по будь-якій додатній основі, відмінній від одиниці, наприклад, по основі 10 або по основі  $e=2,7183$ .

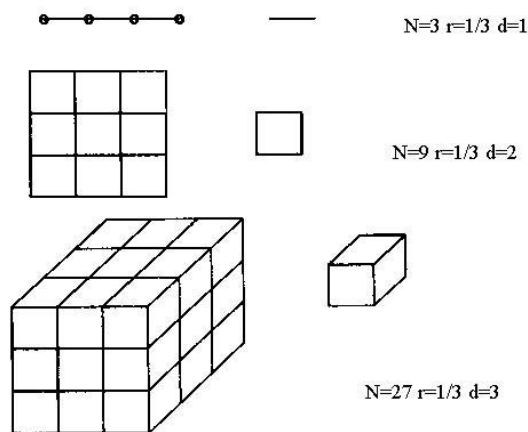


Рис. 5.6. Розмірність простих геометричних фігур

Але, в більш загальному випадку, коефіцієнти подібності не обов'язково одні й ті самі для всіх підмножин фрактала. В цьому випадку формула для розмірності (5.2) не підходить.

Також необхідно відмітити, що деякі множини цілої розмірності також являються фракталами.

Для ілюстрації існування дрібної розмірності наведемо вже частково розглянуту сніжинку Коха та килим Серпинського.

**Сніжинка Коха.** Границя сніжинки, вигаданої Гельгогом фон Кохом в 1904 році (рис. 15.2.), описується кривою, яка складається з трьох однакових фракталів розмірності  $d = 1,2618$ . Кожна третина сніжинки будується ітеративно, починаючи з одної із сторін рівностороннього трикутника. Нехай  $L_0$  — початковий відрізок. Заберемо середню частину і додамо два нових відрізка такої ж довжини, як показано на рис. 5.7. Назвемо отриману множину  $L_2$ . Повторимо дану процедуру багаторазово, на кожному кроці замінюючи середню третину двома новими відрізками. Позначимо через  $L_n$  фігуру, отриману після  $n$ -го крока.

Інтуїтивно зрозуміло, що послідовність кривих  $\{L_n\}_{n=1}^{\infty}$  зводиться до деякої граничної кривої  $L$ . Розглянемо деякі її властивості. Якщо взяти копію  $L$ , зменшену в три рази ( $r = 1/3$ ), то всю множину  $L$  можна зібрати з  $N = 4$  таких копій. Отже, співвідношення самоподібності виконується при вказаних  $N$  і  $r$ , а розмірність фрактала буде:

$$d = \log(4) / \log(3) = 1,2618.$$

Також важливою властивістю, якою володіє границя сніжинки Коха – її нескінченна довжина. Звичайно гладкі або хоча б частково гладкі криві завжди мають кінцеву довжину (можна переконатися інтегруванням). Мандельброт в цьому зв'язку опублікував ряд робіт, в яких досліджується питання про вимірювання довжини берегової лінії Великобританії. В якості моделі він використовував фрактальну криву, яка нагадувала границю сніжинки за виключенням, що в неї ввели елемент випадковості, який враховував випадковість в природі. В результаті виявилось, що крива, яка описує берегову лінію, має нескінченну довжину.

Доведемо це. Нехай вихідний відрізок  $L_0$  має одиничну довжину. Тоді довжина кривої  $L_1$  дорівнює  $4/3$ . Довжина кривої  $L_2$  дорівнює  $16/9$ . Продовжуючи таким чином отримуємо, що крива  $L_n$  після  $n$ -го кроку має довжину  $4^n/3^n$ . Отже, довжина граничної кривої  $L$  дорівнює нескінченності:

$$\lim_{n \rightarrow \infty} 4^n / 3^n = \infty$$



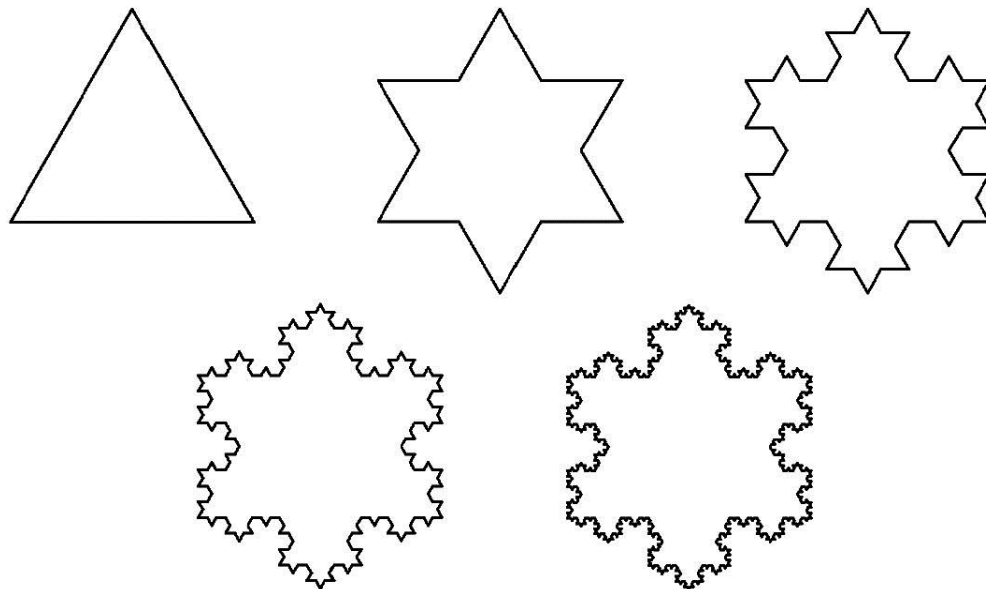


Рис. 5.7. Сніжинка Коха

**Килим Серпинського.** Ще один приклад простого самоподібного фрактала – килим Серпинського (рис. 15.8), вигаданий польським математиком Вацлавом Серпинським у 1915 році. Сам термін килим (gasket) належить Мандельброту. В способі побудови, який слідує нижче, ми розпочинаємо з деякої області і послідовно викидаємо внутрішні підобласті.

Нехай початкова множина  $S_0$  — рівносторонній трикутник разом з областю, яку він замикає. Розіб'ємо  $S_0$  на чотири менші трикутні області, з'єднавши відрізками середини сторін вихідного трикутника. Видалимо вміст маленької центральної трикутної області. Назвемо область, яка залишилась  $S_1$  (рис. 5.8). Потім повторимо процес для кожного із трьох маленьких трикутників, що залишилися, і отримаємо наступне наближення  $S_2$ . Продовжуючи таким чином, отримаємо послідовність вкладених множин  $S_n$ , їхній перетин і утворює килим  $S$ .

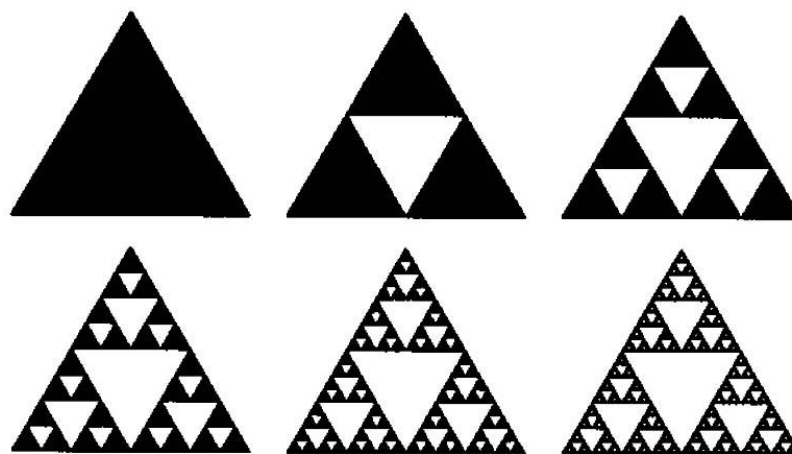


Рис. 15.8 Побудова килима Серпинського

Із побудови видно, що весь килим являє собою об'єднання  $N=3$  зменшених в два рази копій, які не перетинаються; коефіцієнт подібності  $r = 1/2$  (як по горизонталі, так і по вертикалі). Отже,  $S$  — самоподібний фрактал з розмірністю :

$$d = \log(3)/\log(2) \approx 1,5850.$$

Очевидно, що сумарна площа частин, викинутих при побудові, в точності дорівнює площі вихідного трикутника. На першому кроці ми викинули  $1/4$  частину площі. На наступному кроці ми викинули три трикутника, причому площа кожного дорівнює  $1/4^2$  площі вихідного. Розмірковуючи таким чином, ми переконуємось, що повна доля викинутої площі складає:

$$1/4 + 3(1/4^2) + 3^2(1/4^3) + \dots + 3^{n-1}(1/4^n) + \dots$$

При  $n \rightarrow \infty$  ця сума дорівнює 1. Отже, ми можемо стверджувати, що множина  $S$ , яка залишилась, тобто килим, має площу міри нуль.

**Фрактали з цілою розмірністю.** Як вже зазначалось, фрактальна розмірність не завжди дрібна. Таке явище може спостерігатися, коли для побудови фрактала використовується декілька коефіцієнтів подібності. В якості приклада наведемо фрактал – модифікований трикутний невід, де фрактальна розмірність  $d = 2$  (рис. 15.9). Для побудови цієї фігури використовується утворений елемент, наведений в лівому верхньому куті рис.15.9, який покриває одиничний відрізок і перетворюється з двома коефіцієнтами подібності: тлінний катет змінюється з коефіцієнтом подібності  $r_1 = 2/\sqrt{5}$ , короткий -  $r_2 = 1/\sqrt{5}$ . В цьому випадку вже не можна використовувати формулу (15.1). Мандельброт визначив розмірність подібності  $d$  як розмірність, для якої виконується співвідношення:

$$\sum_i r_i^d = 1 \tag{15.3}$$

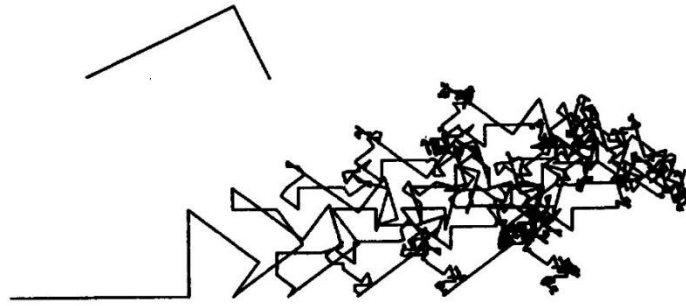


Рис. 15.9 Модифікований трикутний невід. Фрактальна розмірність  $d=2$ .

Слід звернути увагу, що для розглянутого фракталу властиве утворення «кластерних» структур, подібних стохастичному фракталу для броунівського руху частинки (рис. 15.9).

#### 15.4. Довжина берегової лінії. Периметри фракталів

На скільки велика довжина берегової лінії, наприклад, Норвегії (рис. 15.10)? Перед тим як відповісти на питання, з якого розпочинається цей підрозділ, необхідно вирішити, чи варто враховувати в берегову лінію острови. Як бути з річками? В якому місці фіорд перестає бути фіордом і де саме він переходить в річку? Відповісти на це запитання іноді легко, іноді ні. Але навіть якщо ми зможемо задовільно відповісти на всі запитання такого характеру, одна складність все ж залишається. Діло в тому, що якщо взяти курвіметр і надто малий масштаб карти, що не знадобилось би хвилюватись навіть за самі глибокі фіорди, і прийняти за довжину берегової лінії величину  $L$ . А можна вибрати більш великий масштаб карти. Цього разу в довжину берегової лінії ввійшли б найбільш глибокі фіорди. Для ще більш точного підрахунку довжини берегової лінії знадобились би такі карти, якими користуються сусіди при вирішенні питань про те, де повинна проходити огорожа між земельними ділянками. Зрозуміло, що уточнення можна вносити нескінченно. Кожного разу, коли ми будемо збільшувати роздільну здатність, довжина берегової лінії буде розростатися. Крім того, при визначенні довжини берегової лінії будуть виникати проблеми з островами і річками. Альтернативний спосіб вимірювання довжини берегової лінії складається з того, щоб нанести на карту сітку, як показано в верхній частині рис. 15.10. Нехай квадратні комірки сітки мають розміри  $\delta \times \delta$ . Число  $N(\delta)$  таких комірок, необхідних, щоб покрити берегову лінію на карті, приблизно дорівнює числу кроків, за які можна пройти по карті вздовж берега лінійними відрізками довжиною  $\delta$ . Зменшення  $\delta$  призводить до збільшення числа комірок, необхідних для покриття берегової лінії. Якби берегова лінія Норвегії мала б визначену довжину  $L_n$ , то можна було б чекати, що число квадратних комірок  $N(\delta)$ , необхідних для покриття берегової лінії на карті, буде обернено пропорційне  $\delta$ , а величина  $L(\delta) = N(\delta) \times \delta$  при зменшенні  $\delta$  буде прямувати до сталої  $L_n$ . Але, таким чином, очікування не здійсняться. Як видно з рис. 15.10, при зменшенні довжини  $\delta$  крока вимірювання довжини

збільшується. Графік на цьому рисунку виконаний в двічі логарифмічному масштабі і показує, що при зменшенні  $\delta$  виміряна довжина берегової лінії не прямує до сталої величини. Навпаки, виміряна довжина описується наближеною формулою

$$L(\delta) = a\delta^{1-d} \quad (15.4)$$

Для звичайної кривої можна було б очікувати, що  $a=L_n$  (по крайній мірі при досить малих  $\delta$ ) і показник  $d$  дорівнює одиниці. Але для берегової лінії Норвегії, як видно з графіку (рис. 15.10),  $d \approx 1,52$ . Берегова лінія - фрактал з фрактальною розмірністю  $d$ .

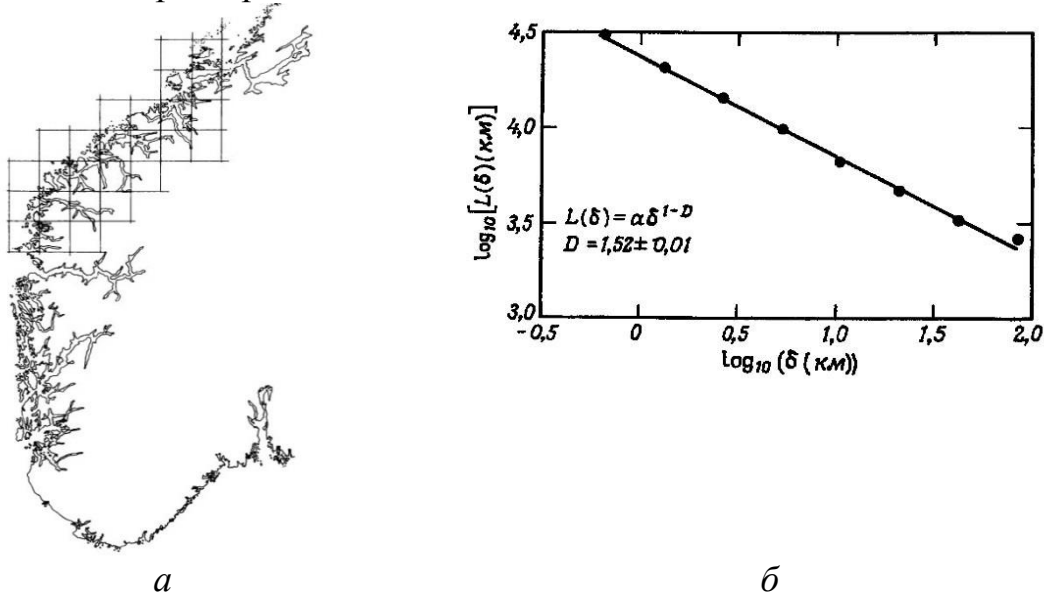


Рис. 15.10. Визначення довжини берегової лінії:

*a* – берегова лінія Норвегії; *б* – виміряна довжина берегової лінії, як функція кроку  $\delta$  (км) – довжини сторони  $\delta \times \delta$  квадратних комірок, що покривають берегову лінію по карті

Аналогічні залежності властиві і для інших країн (рис.15.11) Кутовий коефіцієнт цих прямих дорівнює  $1-d$ , де  $d$  – фрактальна розмірність.

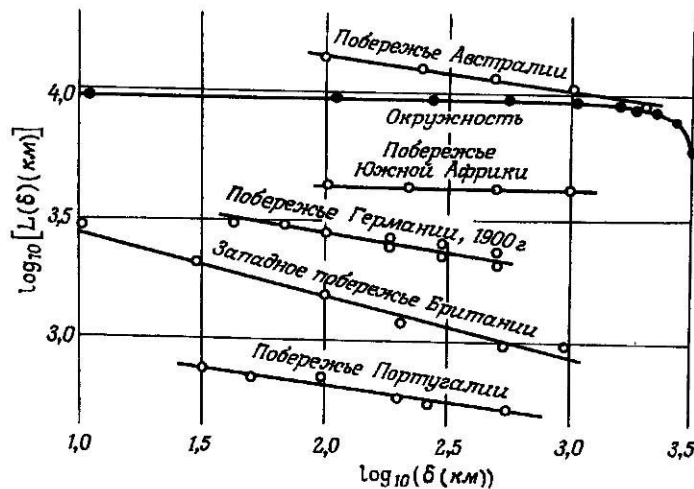


Рис. 15.11. Довжина берегової лінії як функція кроку  $\delta$  (км) – довжини сторони  $\delta \times \delta$  квадратних комірок, що покривають берегову лінію

по карті

### 15.5. «Гра в хаос»

Розглянемо наступну нехитру гру, яку М. Барнслі назвав грою в хаос (chaos game). Візьмемо вже знайомий нам рівносторонній трикутник з вершинами в точках А, В і С (рис. 15.12). Виберемо у середині цього трикутника довільним чином початкову крапку. Кинемо тепер гральну кістку, що є кубиком, на 6 гранях якого проставлені букви А, В і С. Нехай кожна буква присутня на двох з них, тоді вірогідність випадання будь-якої букви однакова і рівна  $1/3$ .

Припустимо, що в результаті першого кидка випала буква А. З'єднаємо в думках нашу початкову крапку з вершиною трикутника А відрізком прямої і на його середині поставимо крапку (рис. 15.12 а). Тепер нова крапка виконуватиме роль початкової. Після чого повторимо вищеописану процедуру з киданням кубика, проставляючи крапки у середині відповідного відрізка. Припустимо, на другому кроці випала буква С, потім В, потім знову С і т.д. В результаті на кожному кроці ми одержуватимемо все нові і нові крапки. Ставиться питання, як розподіляться усередині трикутника ці крапки після достатньо великого числа кроків?

На рис. 15.12 (зліва направо), показані результати цієї гри відповідно з 5 000, 10 000 і 50 000 крапок. Неймовірно, але факт – по мірі збільшення числа крапок все виразніше проступає структура трикутника Серпинського. Видно, що, хоча кожного разу вибір вершини трикутника відбувається чисто випадковим чином, розподілення крапок на площині зовсім не випадкова і володіє яскраво вираженою фрактальною структурою.

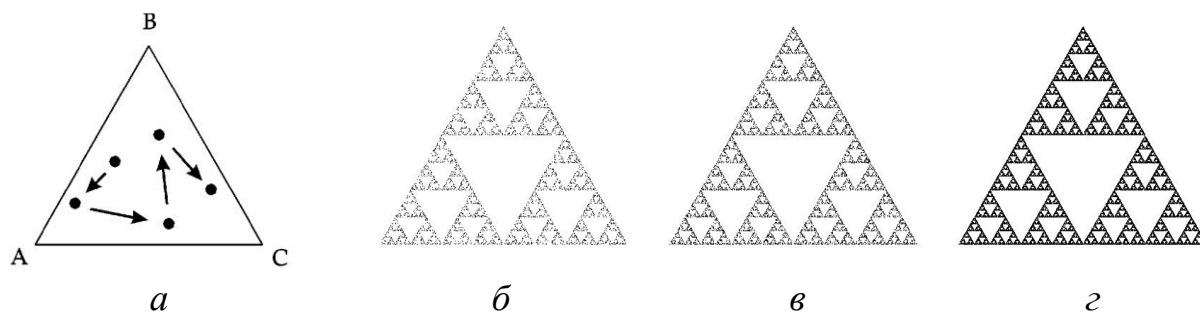


Рис. 15.12. Побудова килима Серпінського за алгоритмом «гра в хаос»:

- a* – фактично випадкове визначення координати крапок;
- б* – просторове розподілення крапок після 5 000 ітерацій;
- в* – просторове розподілення крапок після 10 000 ітерацій;
- г* – просторове розподілення крапок після 50 000 ітерацій

Аналогічним способом можна побудувати багато інших фракталів.

### 15.6. Мультифрактали

Фрактали неймовірно розширили наші можливості опису природи. Абстрактні конструкції Кантора, Коха, Больцано забезпечили нас моделями реальності набагато реалістичнішими, ніж Евклідова геометрія. Проте у фізиці, хімії, геології (і в деяких інших науках) ми стикаємося з багатьма явищами, що вимагають розширення поняття фрактала на складніші структури. Для повної характеристики цих неоднорідних об'єктів потрібна вже не одна, на відміну від регулярних фракталів, а цілий спектр фрактальних розмірностей, число яких в загальному випадку нескінченне. Причина цього полягає у тому, що разом з чисто геометричними характеристиками, які визначаються розмірністю Хаусдорфа, такі фрактали володіють деякими статистичними властивостями.

Багато дивних аттракторів нелінійних динамічних систем також володіють яскраво вираженою мультифрактальною структурою.

Простіше всього пояснити, що мається на увазі під мультифракталом (неоднорідним фракталом) на прикладі трикутника Серпінського, одержаного за допомогою методу випадкових ітерацій (рис. 15.13). Система проітерованих функцій для цього фрактала складається з трьох рівноймовірних лінійних перетворень (імовірність = 1/3). Перерозподілимо імовірність таким чином: на одне перетворення хай доводиться 90%, на інші два по 5%. Таким чином, вийде трикутник Серпінського, точки якого розподілені нерівномірно. Велика частина точок зібрана у одній вершині. Тоді як у інших вершин їх мало. Проте, фрактальна розмірність цього об'єкту дорівнює фрактальній розмірності класичного трикутника Серпінського (тобто з рівноймовірними перетвореннями) і дорівнює  $d = \ln 3 / \ln 2$ . Такий збіг і примушує займатися пошуком нових характеристик, які відрізняли б нерівномірний розподіл від рівномірного. Для цього на ряду з узагальненою фрактальною розмірністю часто доцільно використовувати інші - інформаційну фрактальну розмірність, кореляційну фрактальну розмірність, ентропію та багато інших характеристик.

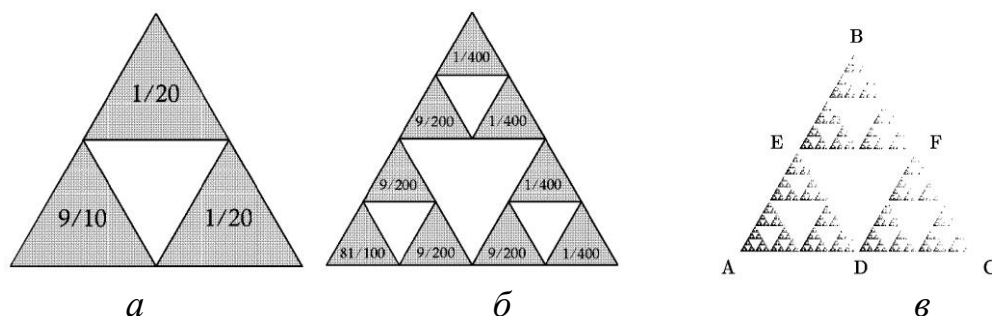


Рис.15.13. Мультифрактальний трикутник Серпінського: *а, б* – принцип побудови мультифрактального трикутника Серпінського й імовірність розподілення точок; *в* – результат побудови мультифрактального трикутника Серпінського за допомогою алгоритму «гра в хаос»

### 15.7. Фрактали та самоорганізація в природі

Теоретично можливість виникнення коливального, періодичного режиму в хімічних і біологічних системах було доведено у 1910 році А. Лотка. У 1937 році Колмогоров математично описав виникнення подібних неоднорідностей. Математичний доказ виникнення неоднорідності – структуризації, порушення просторової симетрії в початково однорідній системі – навів А. Т'юрінг (1952 році). Реакційно-дифузійний механізм виникнення неоднорідностей, теоретично описаний Т'юрінгом, послужив основою численного сімейства моделей біологічного морфогенезу.

**Морфогенез** – від грец. *morphe* - вигляд, форма, генез, морфогенія, формоутворення (біологічне), виникнення і розвиток органів, систем і частин тіла організмів як в індивідуальному, так і в історичному, або еволюційному, розвитку.

Класичним прикладом може слугувати поява *комірок Бенара* – поява складної просторової організації з узгодженим, когерентним переміщенням множини молекул і утворенням конвективних комірок в формі геометричних вельми правильних шестигранних структур в підігрітій знизу досить в'язкої рідини, наприклад, в шарі силіконового масла (рис. 15.14).

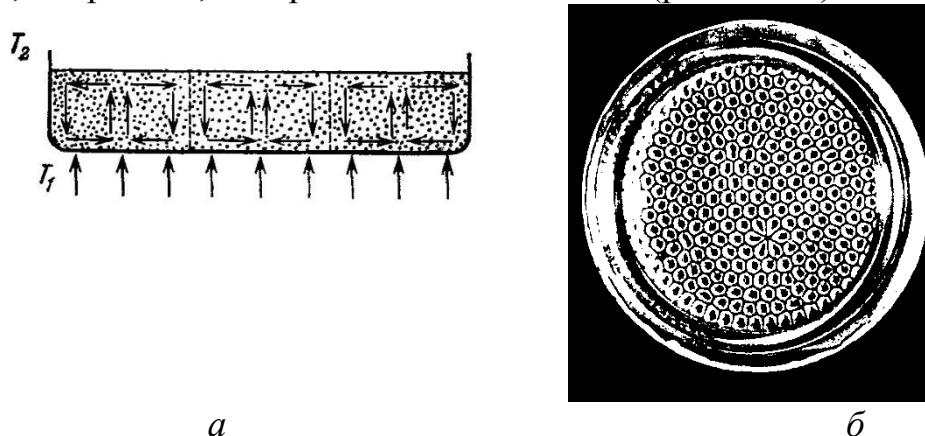


Рис. 15.14. Комірки Бенара:

*a* – схематичне зображення виникнення конвекційних комірок ( $T$  – температура;  $T_2 < T_1$ ); *б* – фотографія (вид зверху) комірок Бенара

Виникнення комірок Бенара відбувається при умові, що градієнт температур  $\Delta T = T_1 - T_2$  перевищує деяке критичне значення для даної рідини  $\Delta T_{крит}$ . Наявність подібних точок біфуркації загалі властиво для самоорганізації як біологічних структур та системам, так і фізичним (рис. 15.15).

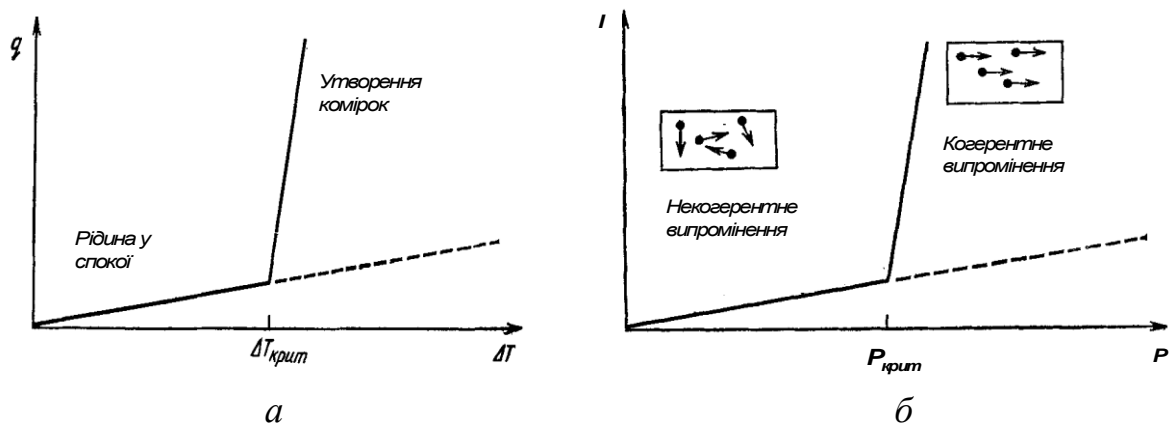


Рис. 15.15. Наявність точок біфуркації при самоорганізації: *a* – потік тепла ( $q$ ) при докритичних і надкритичних значеннях температури і умова виникнення комірок Бернара; *б* – генерація лазерного випромінювання ( $I$  – інтенсивність випромінювання) в залежності від потужності накачки ( $P$ )

Слід звернути увагу, що за сучасними уявленнями, організм людини чи будь-який інший біологічний об'єкт є фрактальною структурою, для якої образуючим елементом є материнська клітина. Всі клітини організму подібні одна до одної, змінюється чи «масштабуються», підлягають «афінним перетворюванням» лише деякі структурно-функціональні властивості, що приводить до спеціалізації клітин. Саме такий погляд на біологічні об'єкти пояснює, як одна молекула ДНК може закодувати розвиток та структуру всього організму. Завдяки цьому об'єм інформації, яка необхідна для відтворення хоча б однієї елементарної біологічної структури (наприклад, клітини) різко скорочується – ДНК необхідно лише запам'ятати деяку основоположну структуру та правило її перетворення. Такі перетворення потребують деякий час. Тому повне відтворення біологічного об'єкта з материнської клітини потребує години, дні, місяці та роки тощо, в залежності від складності повної структури. Саме тому, використання фрактальної геометрії в техніці, наприклад, при розробці комп'ютерних ігор, передачі інформації по каналам зв'язку та інші потребує відносно швидкісних обчислювальних можливостей на фоні суттєвого зменшення вихідної інформації.

Розглядаючи біологічні об'єкти чи інші реальні фізичні системи як фрактальні структури необхідно підкреслити, що в переважаючій більшості їх необхідно розглядати як динамічні системи, правила перетворення образуючого елемента фракталу кожен момент час змінюються за визначеними законами з урахуванням випадкових стохастичних зовнішніх процесів. В якості приклада наведемо еволюцію спіральної доменної структури в ультратонкому шару магніто-одноосьовому кристалі за одну хвилину під дією зовнішнього перемінного магнітного поля 300 Гц (рис. 15.16).



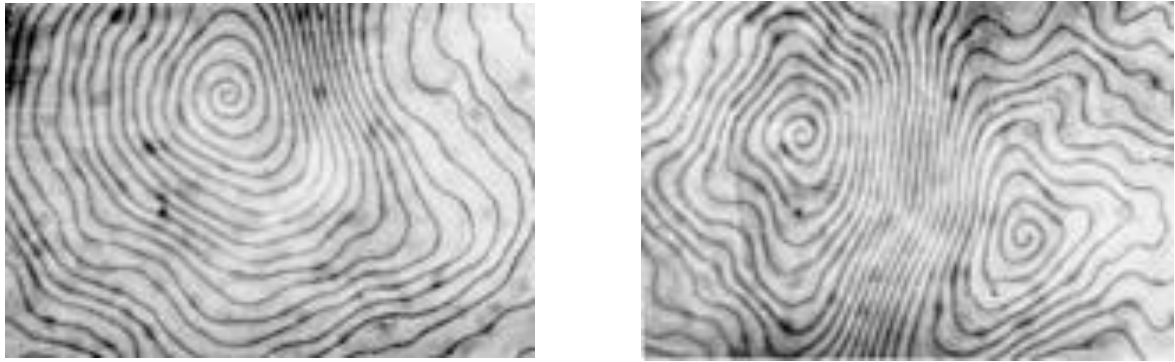


Рис. 15.16. Динамічна фрактальна структура

## 16. 6. Фрактальний аналіз

Головним для фракталів в універсальному науковому плані є те, що при зовнішньому дуже складному вигляді вони, по суті, надзвичайно прості, оскільки є безліччю взаємозв'язаних елементів, однакових по структурі, але відмінних один від одного, головним чином, своїми розмірами. Про такі фігури часто говорять, що вони володіють «масштабною подібністю» або «масштабною інваріантністю», «самоподібні». Це означає, що якщо збільшити масштаб фрактала, щоб виявилися найдрібніші його деталі, то останні матимуть таку ж конфігурацію, що і найкрупніші елементи, які містяться в структурі фрактала.

Прості фрактали, такі, як канторовський пил, сніжинки і ламані Коха, килим Серпинського, криві дракона, криві Пеано і Гільберта і багато інших, володіють регулярною геометрично правильною структурою. Кожен фрагмент такого геометрично правильного фрактала в точності повторює всю конструкцію в цілому. При менш точній транслокації (самоафінності або самоподібності) виникають інші, наприклад, випадкові фрактали. Прикладами випадкових фракталів можуть служити прикордонні і берегові ліній, пори в хлібі, дірки в деяких сортах сиру, частинки в порошках і т.і.

Для фракталів, характерною ознакою є неціла розмірність, яка визначається співвідношенням:

$$N(\varepsilon)\varepsilon^d = c \Rightarrow d = \frac{\log N(\varepsilon)}{\log(1/\varepsilon)} \Rightarrow d = -\frac{\log N(\varepsilon)}{\log(\varepsilon)} \quad (16.14)$$

де  $N(\varepsilon)$  - кількість рівних частин на які розбивається вихідне зображення, де кожен частину можна вважати копією всього зображення, що зменшено в  $1/\varepsilon$  разів;  $c$  - константа.

Структурні властивості біологічних тканин на медичних зображеннях також можуть бути оцінені за допомогою фрактальної геометрії. Найчастіше така оцінка проводиться саме за фрактальною розмірністю. Зазвичай, при такому аналізі медичне зображення переводять у бінарний формат, наприклад, за критерієм перевищення деякого порогу яскравості пікселів.

Комп'ютерні алгоритми обчислення розмірності Мінковського  $D$  найчастіше спираються на співвідношення аналогічне (16.15):

$$\log N(\varepsilon) = \log c - d \log \varepsilon. \quad (16.15)$$

Як видно, графік залежності  $\log N(\varepsilon)$  від  $\log \varepsilon$  – пряма з кутовим коефіцієнтом  $-d$ . В якості  $N(\varepsilon)$  можна використати мінімальне число кліток із стороною  $\varepsilon$ , необхідних для покриття фрактала. Для визначення невідомих параметрів  $c$  і  $d$  (хоча значення  $c$  звичайно не представляє інтересу) необхідно оцінити  $N(\varepsilon)$  для декількох значень  $\varepsilon$ .

Якщо використовувати клітки тільки двох розмірів,  $\varepsilon_1$  та  $\varepsilon_2$ , то невідомі  $c$  і  $d$  можна визначити з системи рівнянь:

$$\begin{aligned}\log N(\varepsilon_1) &= \log c - d \log \varepsilon_1 \\ \log N(\varepsilon_2) &= \log c - d \log \varepsilon_2\end{aligned}$$

Проте, враховуючи, що величини  $N(\varepsilon)$  можуть бути знайдені лише приблизно, має сенс оцінити  $N(\varepsilon)$  для більшого числа різних значень  $\varepsilon$ . В цьому випадку ми одержимо перевизначену систему (число рівнянь більше числа невідомого), яка, швидше за все, не матиме точного рішення. Стандартний прийом в подібній ситуації полягає в мінімізації суми квадратів відхилень (метод найменших квадратів – МНК-прямої). Так, якщо експериментальними даними є набір точок  $(x_i, y_i)$ , де  $i=1,2,\dots,n$  необхідно розв'язати систему рівнянь:

$$\begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix} \quad (16.16)$$

Невідомі  $b$  і  $m$  є коефіцієнтами прямої:  $y = mx + b$ .

## 16.6. 2. Yardstick метод

Рівняння Річардсона-Мандельброта надає математичну основу для розуміння геометричної і просторової структури фракталів.

$$L(\varepsilon) = N(\varepsilon) \cdot (\varepsilon), \quad (16.17)$$

де  $L(\varepsilon)$  – довжина ламаної лінії, якою можна покрити об'єкт;  $\varepsilon$  – довжина відрізка ламаної.

$$N(\varepsilon) = l_0^d \varepsilon^{-d} = \left( \frac{l_0}{\varepsilon} \right)^d, \quad (16.18)$$

де  $N(\varepsilon)$  – кількість відрізків довжини  $\varepsilon$ , які потрібно щоб покрити  $L(\varepsilon)$ ;  $\frac{l_0}{\varepsilon}$  – безрозмірна величина;  $l_0$  – шкала відліку, яка не впливає на визначення  $d$ .

Взявши  $l_0 = 1$ , виразимо з останньої рівності  $d$  і перейдемо до границі при  $\varepsilon \rightarrow 0$ , тоді будемо мати:

$$d = \lim_{\varepsilon \rightarrow 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)}, \quad (16.19)$$

де  $d$  – фрактальна розмірність, яка змінюється від 1 до 2.

Для кривої Коха  $d = 1,26185$ .

Спочатку вибирається відрізок довжини  $\varepsilon$ , а потім вибирається початкова точка на контурі і підраховується число  $N(\varepsilon)$  – кількість відрізків довжини  $\varepsilon$ , якими повністю можна покрити контур досліджуваного об'єкту. Покрити контур означає, що ми використовуємо відрізки сталої довжини  $\varepsilon$ , які наближають контур багатокутником. Ця процедура повторюється для кожної початкової точки на контурі та різних довжин  $\varepsilon$  відрізка, яким покривається контур. Далі будується «найбільш лінійна» ділянка залежності  $\log N(\varepsilon)$  від  $\log \varepsilon$ . На цій ділянці будується лінійна апроксимація методом найменших квадратів, це буде пряма (рис. 16.7). Кутовий коефіцієнт отриманої прямої буде прямувати до фрактальної розмірності даної клітини:

$$d = -\frac{\log N(\varepsilon)}{\log \varepsilon}. \quad (16.20)$$

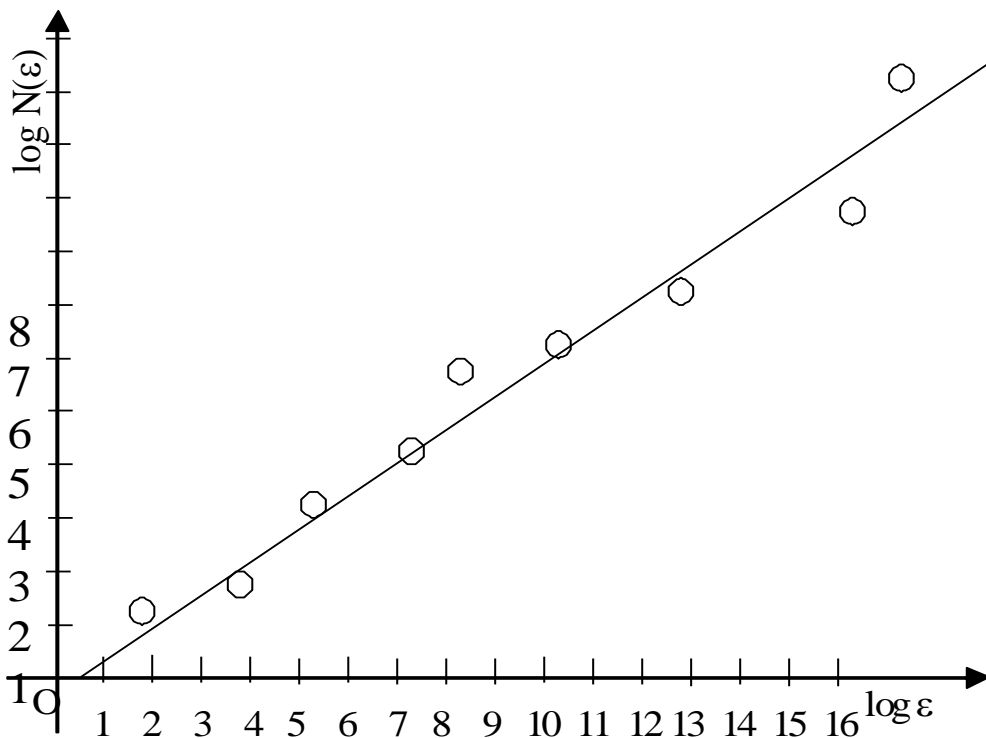


Рис. 16.7. Приклад оцінки фрактальної ємності

Порівняймо фрактальні розмірності тестових фігур та їх комбінації (рис. 16.8). Для цих фігур було обчислено фрактальну розмірність при  $\varepsilon = 5, 10, 20$ .

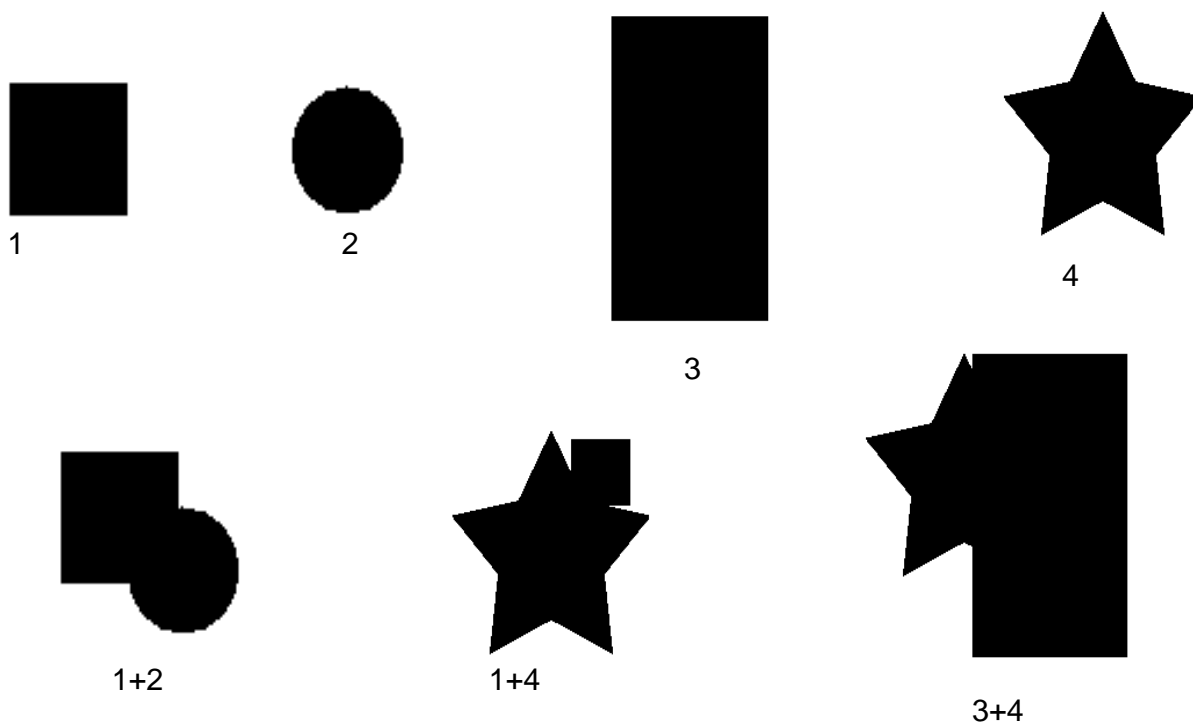


Рис. 16.8. Фігури та комбінації з поєднанням цих фігур, на яких було проведено тестові дослідження фрактальної розмірності

В табл. 16.3 наведені результати обчислення фрактальної розмірності, отримані за допомогою Yardstick методу.

Таблиця 16.3. Фрактальна розмірність тестових фантомних зображень

№ фігури	Максимальний розмір комірки при розбитті зображення, пікселі					
	5		10		20	
	Фрактальна розмірність	Дисперсія	Фрактальна розмірність	Дисперсія	Фрактальна розмірність	Дисперсія
1	1,0046	0,0000250	1,0062	0,0000568	1,0159	0,000227
2	1,1728	0,0009390	1,0730	0,0009620	1,0224	0,000790
3	1,0027	0,0000004	1,0027	0,0000035	1,0036	0,000014
4	1,1707	0,0011300	1,0892	0,0008760	1,0634	0,000511
1+2	1,0859	0,0001500	1,0551	0,0001490	1,0386	0,000176
1+4	1,1526	0,0006340	1,0943	0,0004730	1,0829	0,000272
3+4	1,0692	0,0001670	1,0417	0,0001150	1,0369	0,000065
Середнє значення	1,0941	0,0004351	1,0517	0,0003765	1,0377	0,000294

З одержаних даних видно, що у квадрата і прямокутника, тобто у 1 та 3 фігури, фрактальна розмірність близька до 1. Для 2 та 4 фігури – кола та п'ятикутної зірки фрактальна розмірність дещо більша за 1. Тобто фрактальна розмірність залежить від контуру фігури, при фіксованому  $\varepsilon$ .

Розглянемо зміну фрактальної розмірності при зміні довжини відрізка  $\varepsilon$ . В таблиці наведено дані для  $\varepsilon = 5; 10; 20$ . З отриманих даних видно, що для 1 і 3 фігури фрактальна розмірність зростає при збільшенні  $\varepsilon$ . А для 2 та 4 фігури фрактальна розмірність зменшується при збільшенні  $\varepsilon$ . Тобто для менш іррегулярних фігур залежність фрактальної розмірності  $d$  від  $\varepsilon$  є прямо пропорційною, а для інших фігур ця залежність є оберненопропорційною. Це пов'язано з тим, що при покриванні контуру фігури відрізками, ми отримуємо багатокутник, який наближається до контуру фігури. Коли покривати менш іррегулярний контур фігури, наприклад прямокутник, відрізками при досить малих  $\varepsilon$ , то в результаті одержимо майже такий прямокутник, а при покриванні цієї ж фігури відрізками при більших  $\varepsilon$ , в результаті одержимо не прямокутник, а багатокутник з більшим числом сторін і як наслідок фрактальна розмірність зростає. При покриванні більш іррегулярного контуру фігури все навпаки. При досить малих  $\varepsilon$  одержимо багатокутник, який наближається до контуру реальної фігури. При збільшенні  $\varepsilon$  в цьому багатокутнику стає менше сторін і він стає більш опуклим, тобто менш іррегулярним, і як наслідок його фрактальна розмірність зменшується.

Розглянемо тепер випадок, коли є фігури, що отримані перетином розглянутих вище фігур. Дослідимо зміну їх фрактальних розмірностей. Якщо розглянути фігуру, що одержана перетином квадрата і круга (1+2 фігури), то фрактальна розмірність збільшиться, оскільки контур квадрата з малою іррегулярністю доповнився деякою неоднорідністю (додали коло). Розглянемо як змінюється розмірність 4 фігури при перетині з 1 та 3 фігурами. Контур 4 фігури є більш іррегулярний в порівнянні з контуром квадрата (1 фігура) і контуром прямокутника (3 фігура). При доповненні 4 фігури 1 фігурою фрактальна розмірність дещо зменшиться в порівнянні з фрактальною розмірністю самої 4 фігури. Це пов'язано з тим, що квадрат (фігура 1) має менш іррегулярний контур ніж 4 фігура, але оскільки контур 4 фігури значно більший ніж 1 фігури, то розмірність 4 фігури зменшиться не суттєво. Якщо ж доповнювати 4 фігуру прямокутником (3 фігура) з досить великим контуром, то розмірність 4 фігури значно зменшиться (це впливає з вищенаведених міркувань). З наведеного аналізу дещо осторонь стоять розрахунки контуру круга, який має велику фрактальну розмірність.

Таким чином, Yardstick метод може бути використано для кількісної оцінки асиметрії(іррегулярності) контурів медичного зображення.

### 16.6. 3. Box-counting метод

Опишемо алгоритм пошуку фрактальної розмірності, який в літературі отримав назву «Box counting» чи клітинний метод.

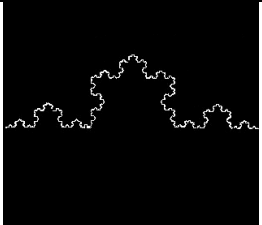
Розіб'ємо область, що містить фрактал  $A$ , на квадратні клітки (двовимірний випадок) декількох розмірів. Підрахуємо число кліток,

необхідних для покриття  $A$  у кожному випадку, і підставимо набуті значення в співвідношення (16.19).

Очевидно, якщо фрактал  $A$  є підмножиною прямої, то замість квадратів треба використовувати відрізки. Якщо ж  $A$  — підмножина тривимірного простору, то квадрати замінюються кубами.

Проведемо деякі дослідження на тестових фантомних зображеннях. Це є необхідним для встановлення певних закономірностей. Також це допоможе виділити класи задач, до яких можливо застосовувати методи, пов'язані з оцінкою фрактальною розмірністю “структурного хаосу” медичних зображень, що досліджуються. Приведемо декілька наборів тестових фантомних зображень, для виявлення закономірностей (беручи  $\varepsilon$  від 1 до 10, розмір картинки 400x400 пікселів). Візьмемо як приклади зображення, що містять 50 довільним чином розташованих точок; 2000 довільним чином розташованих точок; квадрат розміром 100x100 пікселів; повністю залитий малюнок; діагональ та крива Коха. Отримані результати містяться в табл. 16.4.

Таблиця 16.4. Фрактальна розмірність тестових фантомних зображень

Назва тесту	Фрактальна розмірність	Дисперсія	Середнє квадратичне відхилення
50 точок	0,000	0,000000	0,000
2 000 точок	0,277	0,008470	0,092
Квадрат	1,945	0,000611	0,025
Все біле	2,004	0,000058	0,007
Діагональ	1,002	0,000014	0,004
 Крива Коха	1,262	0,010803	0,104

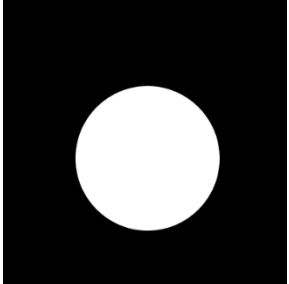
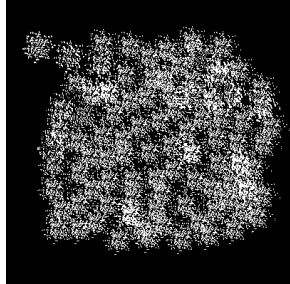
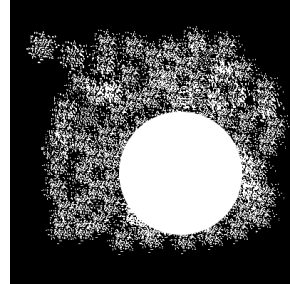
Аналізуючи отримані дані можна помітити, що за рахунок того, що розміри рисунків відносно невеликі, деякі приклади висвітлюють помилку при підрахунку. Так, наприклад розмірність повністю заповненого рисунку більша ніж 2, що в ідеалі неможливо – повинно бути рівно 2.

Приклад підрахунку фрактальної розмірності кривої Коха дає результат дуже близький до теоретичної розмірності: теоретично  $\frac{\ln 4}{\ln 3} = 1,261886\dots$ , практично  $1,262 \pm 0,104$ , хоча, як бачимо, присутня доволі велика дисперсія.

На перший погляд здається, що фрактальна розмірність росте з площею об'єкту. Тому приведемо приклади в табл. 16.5, які показують, що це не

завжди є вірним (зрозуміло що є багато випадків, коли фрактальна розмірність зростатиме зі збільшенням площі об'єкту, що вивчається).

Таблиця 16.5. Фрактальна розмірність тестових фантомних зображень

Параметри	№ малюнка		
	1	2	3
			
Фрактальна розмірність	1,951	1,573	1,711
Площа фігури	32 645	32 807	47 836
Дисперсія	0,000135	0,019667	0,006200

Фігури на малюнках 1 та 2 (табл. 16.5) мають приблизно однакову площу (32645 та 32807 відповідно), але при цьому їхня фрактальна розмірність суттєво розрізняється. Також можна бачити, що на 3 малюнку, що є об'єднанням двох перших, площа значно зростає (47836), але розмірність зменшується в порівнянні з малюнком 1 та збільшується в порівнянні з 2 відповідно на 14 % та 8 %. Також треба зазначити, що на малюнку 1 була отримана найменша дисперсія, а найбільша – на малюнку 2.

#### Завдання для самоконтроля

1. Використовуючи лінійну і кубічну сплайн-інтерполяцію, побудувати фазовий портрет функції  $y = f(t)$ , де  $f(t) = \sin(t)$ ,  $t = 0 \div 6\pi$   $\Delta t = 0.01\pi$ .
2. За дискретними значеннями  $y_i = f(t_i)$  побудувати фазові портрети  $y_i = f(y_{i+j})$ , де  $j = 1, 2, 10, 20$ .
3. Додати до кожного значення функції  $y_i = f(t_i)$  випадкове значення в межах від 0 до 1. За отриманими значеннями побудувати фазові портрети  $y_i = f(y_{i+j})$ , де  $j = 1, 2, 10, 20$ .
4. Виконати п.1-3 при нерівномірній випадковій дискретизації функції  $y = f(t)$ ,  $\Delta t = 0.001\pi \div 0.01\pi$ .
5. Проаналізувати отримані графічні зображення.
6. Розрахувати фрактальну розмірність крижинки Коха методом BoxCounting

## Розділ 6.

### Елементи теорії нечітких множин

#### Два основні підходи до формалізації нечіткості

Теорія нечітких множин, що розвивається після публікації в 1965 р. основоположною роботи Л. Заде , являє собою узагальнення і переосмислення найважливіших напрямків класичної математики. Біля її витоків лежать ідеї та дослідження багатозначної логіки ( тризначною логіки Лукасевича , до значної логіки Поста ) .Яка вказала на можливості переходу від двох до довільного числа значень істинності і поставила проблему оперування поняттями із змінним змістом; теорії ймовірностей , яка породивши велику кількість різних способів статистичної обробки експериментальних даних (наприклад , гістограми , функції розподілення). Відкрила шляхи визначення та інтерпретації функції приналежності ; дискретної математики (теорії матриць , теорії графів , теорії автоматів і т. д.). Запропонувала інструмент для побудови моделей багатовимірних і багаторівневих систем . Зручний при вирішенні практичних завдань.

Подальші кроки в цьому напрямку зв'язуються зі створенням строгих і гнучких математичних методів дослідження нечітко визначених об'єктів. При цьому нечіткість образів , уявлень і понять людини вводиться в формальні меделі розрізнити способами.

Можна виділити наступні основні класифікаційні ознаки способів формалізації нечіткості :

- 1 ) по виду представлення нечіткої суб'єктивної оцінки будь-якої величини ( непарного безлічі ) ;
- 2 ) з вигляду області значень функції приналежності ;
- 3 ) з вигляду області визначення функції приналежності ;
- 4 ) з вигляду відповідності між областю визначення і об-ласті значень ( однозначне , багатозначне ) ;
- 5 ) за ознакою однорідності або неоднорідності області значень функції приналежності .

Перший підхід до формалізації нечіткості полягає в наступному. *Нечітка множина* (НМ) утворюється шляхом введення узагальненого поняття приналежності, тобто розширення двоелементною безлічі значень характеристичної функції  $\{0, 1\}$  до континууму  $[0, 1]$ . Це означає, що перехід

від повної приналежності об'єкту класу до повної його не приналежності відбувається не скачкою, а плавно, поступово, причому приналежність елементу множини виражається числом з інтервалу  $[0, 1]$ . *Нечітка множина*

$$A = \{(x, \mu_A(x))\} \quad (1.1)$$



визначається математично як сукупність упорядкованих пар, складених з елементів  $x$  універсальної множини  $X$  та відповідних ступенів належності  $\mu_A(x)$  або (по-скільки функція приналежності є вичерпною характеристикою НМ) безпосередньо у вигляді функції  $\mu_A: X \rightarrow [0,1]**$ . Можливі види функцій  $\mu_A$  зображені на рис. 1.1. Приклад запису НМ:

$$A = \{(x_1, 0,2), (x_2, 0,6), (x_3, 1), (x_4, 0,8)\}, \quad (1.2)$$

або

$$A = 0,2 | x_1 + 0,6 | x_2 + 1 | x_3 + 0,8 | x_4$$

або

$$A = 0,2 | x_1 \quad 0,6 | x_2 \quad 1 | x_3 \quad 0,8 | x_4$$

Поняття НМ тісно пов'язане з центральним поняттям так альтернативної теорії множин - поняттям напівмножини. У той же час як безліч припускає наявність

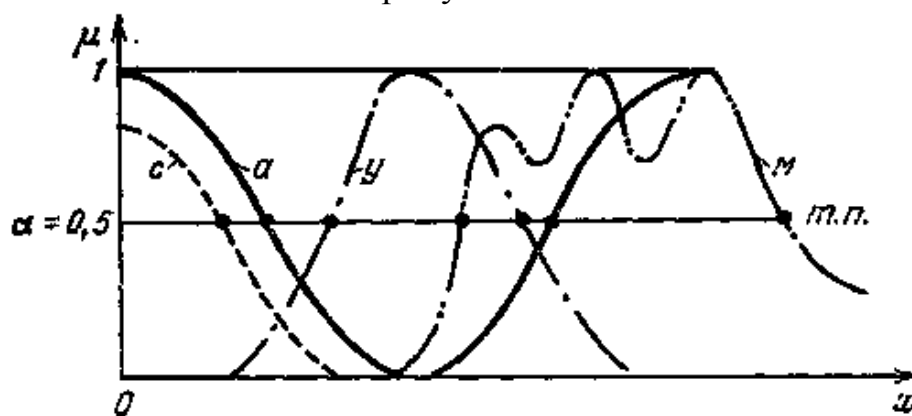


Рис. 1.1 Види функцій приналежності: с— субнормальна; а— амодальна; м— багатомодальна; у— унімодална; т.п.— точки переходу.

визначених меж приналежності і неналежності, пів-множини є більш широким поняттям, які не мають максимальних або мінімальних елементів, а отже,

фіксованих значень приналежності. В альтернативній теорії множин чітко розмежовуються поняття множини і класу. Множество - це сукупність чітко помітних елементів, які можна перерахувати, представити у вигляді списку. Поняття класу є більш загальним, ніж поняття множини. Поняття класу є більш загальним, ніж поняття множини. Властивість об'єктів  $\mu(x)$ , що розглядається як об'єкт, визначає клас  $\{x, \mu(x)\}$ . *Напівмножество* називається власний клас (не безліч), що є підкласом деякої безлічі  $X: A = \{x, \mu(x)\} \subset X$ . Оскільки при визначенні напівмножества не використовується відношення приналежності між елементом і множиною, цей математичний об'єкт є більш загальним, ніж НМ. Але для практичних застосувань напівмножин слід ввести функціональні обмеження на приналежність і апроксимувати полумножества нечіткими множинами. Способи наближення напівмножеств нечіткими множинами описані в [ 49-50 ].

Основні операції над НМ з класу всіх НМ  $F(X) = \{\mu \mid \mu: X \rightarrow [0,1]\}$  універсальної множини  $X$  представлені в табл. 1.1. Нижче наводяться найбільш важливі поняття теорії нечітких множин. *Нормальність НМ.* Нечітке множиство  $A$  нормально, якщо верхня межа його функції приналежності дорівнює одиниці, тобто  $\sup_{x \in X} \mu_A(x) = 1$ . При  $\sup_{x \in X} \mu_A(x) < 1$ . НМ називається субнормальним. Нечітка множина порожня, якщо  $\mu_A(x) = 0, x \in X$ . Непорожнє субнормальне НМ можна привести до нормального (нормалізувати) за формулою

$$\mu_A(x) = \frac{\mu_A(x)}{\sup_{x \in X} \mu_A(x)}. \quad (1.3)$$

Множина рівня  $\alpha$  НМ. Множина рівня  $\alpha$  ( $\alpha$ -зрізом) НМ  $A$  називається чітка підмножина універсальних множині  $X$ , визначається у вигляді

$$A_\alpha = \{x \in X \mid \mu_A(x) \geq \alpha\}, \quad (1.4)$$

де  $\alpha \in [0,1]$  (див. рис. 1.2). Наприклад, для (1.2) і  $\alpha = 0,6$  множинство рівня  $\alpha$  має вигляд  $A_\alpha = \{x_2, x_3, x_4\}$ . З іншого боку,  $A_\alpha$  є образ інтервалу  $[\alpha, 1]$  при зворотному відображенні  $\mu^{-1}$ , тобто  $A_\alpha = \mu^{-1}([\alpha, 1])$ . Множина суворого рівня визначається

$$A_\alpha^s = \{x \in X \mid \mu_A(x) > \alpha\}. \quad (1.5)$$

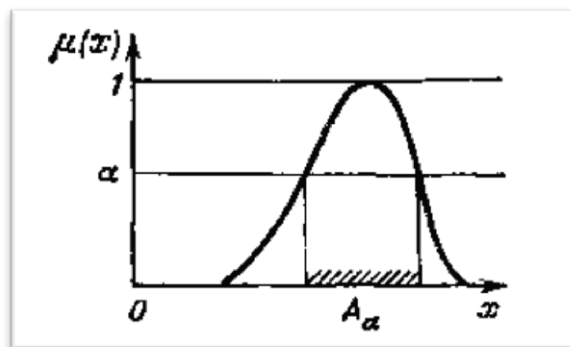


Рис (1.2) Нечіткі множини і його множинні рівні.

Наприклад, для (1.2) і  $\alpha = 0,6$  множиною суворого рівня буде  $A_\alpha^s = \{x_3, x_4\}$ . Зокрема, носієм НМ  $A$  (позначається)

$\text{supp } A$ ) є безліч елементів  $x \in X$ , для яких  $\mu_A(x) > 0$ , тобто  $\text{supp } A = \{x \in X \mid \mu_A(x) > 0\}$ . Поняття множини рівня є розширеним поняттям інтервалу  $\{x \mid a \leq x \leq b\}$ . Воно являє собою об'єднання не більше ніж рахункового числа інтервалів [ 2 ]. Відповідно, алгебра інтервалів є окремий випадок алгебри множин рівня.

Точка переходу НМ  $A$  - це такий елемент  $x \in X$  для якого  $\mu_A(x) = 0,5$ .

Чітка множина, наближена до НМ, визначається як:

*Опуклість* НМ . Нечітка множина  $A$  в просторі  $x = R^n$  називається опуклою нечіткою множиною тоді і тільки тоді , коли її функція приналежності опукла , тобто для кожної пари точок  $x, y \in X$  задовольняє нерівність

$$\mu_A(\lambda x + (1-\lambda)y) \geq \min(\mu_A(x), \mu_A(y)) \quad (1.7)$$

для всіх  $\lambda : 0 \leq \lambda \leq 1$

Приклад нечіткої функції - відображення  $\varphi : X \xrightarrow{H} Y$ , котре кожному  $x \in X$  ставить у відповідність  $y \in Y$  зі ступенем  $\mu_\varphi(x, y)$ . Інші варіанти - це функція з нечітким аргументом і функція з нечіткою областю визначення . Нечітка функція визначає нечітку поверхню приналежності в  $X \times Y$

( $X, Y$  - довільні множини) .

*Принцип узагальнення* [ 4 , 5 , 7 , 13 ] . Принцип узагальнення як одна з основних ідей теорії НМ носить евристичний характер і дозволяє розширювати область визначення вихідного відображення  $\varphi$  на клас НМ , а також узагальнити визначення операцій над НМ типу 1 , на НМ типу 2 і вище \*). Нехай  $\varphi : X \rightarrow Y$  - задане відображення , а  $A$  - НМ в  $X$  . Тоді образ НМ  $A$  при відображенні  $\varphi \in$  НМ в  $Y$  з функцією приналежності

$$\mu_B(y) = \sup_{x \in \varphi^{-1}(y)} \mu_A(x), \quad y \in Y, \quad (1.8)$$

де  $\varphi^{-1}(y) = \{x \in X \mid \varphi(x) = y\}$ . У разі нечіткого відображення  $X \xrightarrow{H} Y$  маємо

$$\mu_B(y) = \sup_{x \in X} \min\{\mu_A(x), \mu_\varphi(x, y)\}. \quad (1.9)$$

Згідно [5] нечіткою множиною типу  $n$  називається НМ , у якого значеннями функції приналежності є НМ типу  $n-1$ . Наприклад , НМ типу 1 є  $\mu : X \rightarrow [0, 1]$  , а НМ типу 2 можна визначити як  $\mu : X \times [0, 1] \rightarrow [0, 1]$  і т.д.

На відміну від булевої алгебри , де  $\forall A \subset X : A \cap \bar{A} = \emptyset, A \cup \bar{A} = X$  в  $F(X)$  закони вилученого третього не виконуються :  $A \cap \bar{A} \neq \emptyset, A \cup \bar{A} \neq X$ . В [ 28 ] показано , що при побудові операцій об'єднання або перетину в  $F(X)$  треба відкинути або закони виключеного третього , або властивості дистрибутивності і ідемпотентності . Вищевикладений підхід є найбільш поширеним при моделюванні нечітких понять.

Будь-яке НМ можна розкласти по множинам рівня згідно теоремі декомпозиції :

$$\mu_A = \bigcup_{\alpha \in [0,1]} (\alpha \cdot \mu_{A_\alpha}),$$

де

Дане розкладання лежить в основі другого способу формалізації нечіткості, коли нечіткість виражається за допомогою набору ієрархічно впорядкованих чітких множин.

Отже, для кінцевого числа  $n$  градацій розглянутих властивостей  $n$  - нечітких множин задається через  $n$ -ку звичайних множин  $F = (M_1, \dots, M_n)$ , де  $M_i \subseteq X$ ,  $i = 1, \dots, n$ , і  $M_1 \subseteq \dots \subseteq M_n$  [ 31 ]. Для нескінченного числа градацій маємо нескінченне сімейство множин  $F = (M_\alpha)$ ,  $\alpha \in [0, 1]$ , тобто відображення виду  $M : [0, 1] \rightarrow 2^X$ , де будь-якому числу (індексу)  $\alpha \in [0, 1]$  ставиться у відповідність чітка підмножина множини  $X$ . Тоді розмитість моделюється відображеннями  $M$  з класу

$$\Phi([0,1]) = \{M \mid M : [0,1] \rightarrow 2^X\}$$

з властивостями

а)  $M(0) = X$  ;

б)  $\forall \alpha_1, \alpha_2 \in [0,1] : \alpha_1 \leq \alpha_2 \Rightarrow M(\alpha_1) \supseteq M(\alpha_2)$ ,

та відповідними операціями над ними ( табл. 1.1).

Зв'язок між виділеними альтернативними способами формалізації нечіткості встановлюється на основі теореми представлення, згідно з якою класи  $F(X)$  і  $\Phi([0,1])$  ізоморфні щодо операцій перетину і об'єднання ( табл. 1.1). При цьому будь-якій бінарній операції в  $F(X)$  відповідає об'єднання пересічний різних зрізів в  $\Phi([0,1])$  [ 15 ] :

$$\forall x \in X : \mu_3(x) = \mu_1(x) * \mu_2(x) \Leftrightarrow M_3(\alpha) = \bigcup_{\beta_1 * \beta_2 \geq \alpha} (M_1(\beta_1) \cap M_2(\beta_2)) \forall \alpha \in [0,1]$$

(1.12)

Завдання імовірнісної міри на  $M$  обумовлює перехід до теорії випадкових множин [ 12 ]. В особливу групу ми виділяємо різні комбіновані підходи, що враховують як нечіткість, так і стохастичність в системах управління і штучного інтелекту. Зокрема, розподіл усіх нечітке безліч [ 36 ] визначається рандомизированной функцією приналежності  $\mu_A : X \times \Omega \rightarrow [0, 1]$ . Тут береться до уваги випадкова похибка при експертному оцінюванні функції приналежності (рис. 1.3).

Цей підхід, як і випадкові нечіткі множини [ 8 ], застосуємо зокрема при груповому експертному оцінюванні.

Поряд з виразом (1.11) є й інші варіанти формалізації нечіткості даних за допомогою сімейства

звичайних множин, наприклад *наближені множини* (rough sets) [51]. Нехай  $X$  -множина, а  $R \subset X \times X$  - співвідношення еквівалентності на  $X$ . Впорядковану пару  $\Psi = \langle X, R \rangle$  будемо називати простором наближень, а  $R$ -відношенням непомітності в просторі  $\Psi$ . Класи еквівалентності по відношенню  $R$  називаються елементарними множинами в  $\Psi$ , а будь-яке об'єднання елементарних множин утворює складну множину в  $\Psi$ .

Нехай  $Y \subset X$ . Під нижнім наближенням безлічі  $Y$  в просторі  $\Psi$  будемо розуміти найбільш складну множину в  $\Psi$ , що міститься в  $Y$ , а під верхнім

наближенням  $Y$  в  $\Psi$  - найменш складну множину в  $\Psi$ , що містить  $Y$ . Позначимо нижню і верхню наближення безлічі  $Y$  в просторі  $\Psi$  через  $\underline{\Psi}Y$  і  $\bar{\Psi}Y$  відповідно. Довільні дві множини  $Y, Z \subset X$  наближено (знизу) рівні в  $\Psi = \langle X, R \rangle$ , якщо  $\underline{\Psi}Y = \underline{\Psi}Z$ , і наближено (зверху) рівні в  $\Psi = \langle X, R \rangle$ , якщо  $\bar{\Psi}Y = \bar{\Psi}Z$ . Позначимо наближену рівність множин знизу і зверху в  $\Psi$  у формі  $Y \underset{\Psi}{\approx} Z$  і  $Y \overset{\Psi}{\approx} Z$  відповідно.

Дві множини  $Y, Z \subset X$  приблизноно рівні в  $\Psi = \langle X, R \rangle$ , якщо вони приблизно рівні знизу і зверху, тобто  $Y \underset{\Psi}{\approx} Z$ , якщо  $Y \underset{\Psi}{\approx} Z$  і  $Y \overset{\Psi}{\approx} Z$ . Тоді поняття приближеної множини вводиться у вигляді сімейства наближено рівних множин в просторі приближень. Відповідно, множина  $(\Psi)Y = \{Z \subset X \mid Z \underset{\Psi}{\approx} Y\}$  називається нижньою наближеною множиною, породженим множиною  $Y$  в просторі  $\Psi$ , множина  $(\bar{\Psi})Y = \{Z \subset X \mid Z \overset{\Psi}{\approx} Y\}$  - верхньою наближеною множиною, а  $(\Psi)Y = \{Z \subset X \mid Z \underset{\Psi}{\approx} Y\}$  наближеною множиною.

Данні три типи наближених множин являють собою сімейства сімейств звичайних множин.

В [ 52 ] запропоновано цікаве узагальнення простору  $\Psi$  і відносини непомітності  $R$ . У НМ  $\{(x, \mu, \sigma)\}$  визначається нечітка еквівалентність  $\sigma$ . Вводиться поняття абсолютно нечіткої множини як трійки  $A = (X, \mu, \sigma)$ , де  $\mu: X \rightarrow [0,1]$  є функція приналежності, а  $\sigma: X \times X \rightarrow [0,1]$  - функція нерозрізнуваності.

Близьким до ідей альтернативної теорії множин є недоопределенное безліч, що описується четвіркою  $N = \langle {}^+A, {}^-A, M_x, M_n \rangle$  [11]. Тут множини  ${}^+A$  і  ${}^-A$  суть кінцеві підмножини універсальної множини  $X$ , причому  ${}^+A$  є безліч елементів  $x \in X$ , які точно належать множені (денотату)  $A$ , а  ${}^-A$  є множиною елементів  $x \in X$ , котрі точно не належать множині  $A$ . Натуральні числа  $M_x$  і  $M_n$  висловлюють відповідно верхню і нижню оцінку потужності множини  $A$ . Це визначення, що моделює неповні відомості про конкретну сукупності  $A$  елементів деякого універсуму  $X$ , неявно задає тризначну функцію приналежності

Природним узагальненням  $N$  є перехід до пари  $N = \langle \mu_A, |\bar{\mu}_A| \rangle$ , де  $\mu_A$  є безперервна функція приналежності  $x \in X$  елементів множини (денотату)  $A$ , а  $|\bar{\mu}_A|$  характеризує

можливість для елементів натурального ряду бути значенням потужності даної множини  $A$ .

Отже, короткий огляд ряду способів формалізації нечіткості показує, що в цьому напрямі розвиваються два основні підходи. Перший базується на узагальненні поняття приналежності елемента множині, що приводить до розмивання меж множини, а в граничному випадку, до появи об'єкта з невизначеними межами - напівмножини. Другий підхід передбачає опис нечіткості за допомогою ієрархії - сімейства упорядкованих чітких множин. Простежується взаємозв'язок цих підходів, що вказує на існування глибокого внутрішнього зв'язку проблем математичної обробки нечіткої інформації та побудови моделей складних, ієрархічних систем. Нижче в рамках першого підходу ми обговоримо різні варіанти завдання області визначення і області значень функції приналежності НМ, а також відповідності між ними. Що виникає при цьому різноманітність видів НМ відкриває широкі перспективи їх застосування в моделях керування і штучного інтелекту.

### § 1.2. Види областей значень функцій приналежності

Всі розглянуті вище нечіткі об'єкти можна далі класифікувати по виду області значень функції приналежності. Крім інтервалу  $[0, 1]$  функція приналежності може приймати свої значення в інтервалі  $[-1, +1]$  (тоді граничне значення повної неналежності дорівнює  $-1$ , а  $0$  береться як точка переходу нечіткої множини [43]), на числовій прямій  $R$ , а також у різних множинах, наділених певною структурою. Наприклад, в [24] область значень функції приналежності  $\mu: X \rightarrow (-\infty; +\infty)$  складається з трьох ділянок:

- а)  $\mu(x) > 0$ ,
- б)  $-1 \leq \mu(x) < 0$ ,
- в)  $\mu(x) < -1$ .

При інтерпретації  $X$  як безлічі деталей підмножина  $\{x \in X \mid \mu(x) > 0\}$  охоплює всі придатні деталі, що не виходять за межі необхідних допусків, підмножина  $\{x \in X \mid -1 \leq \mu(x) < 0\}$  - непридатні деталі, які можна переробити, а  $\{x \in X \mid \mu(x) < -1\}$  - браковані деталі.

Історично першим узагальненням поняття НМ  $\mu: X \rightarrow [0, 1]$  стали *L-нечіткі множини*  $\mu: X \rightarrow L$  [32], тобто функції, приймаючі свої значення в кінцевій або нескінченній дистрибутивній решітці  $L$  (решітка - частково впорядкована множина з точною нижньою і точною верхньою межами). Области приналежності моделюються також повною, ґратково упорядкованою напівгрупою [32, 33], півкільцем [21, 59], категорією [54]. Важливим для практичних додатків в плані вираження якісних уявлень і оцінок людини в процесі вирішення завдання є випадок *S-нечітких множин*, які задаються парою  $(X, \mu)$ , де

$$\mu: X \rightarrow S \quad (1.13)$$

- відображення з  $X$  в лінійно упорядкований безліч  $S$  [32, 62]. На  $S$  природно накласти вимоги кінцівки і повноти. Приклад кінцевої лінійної впорядкованої множини - набір лінгвістичних значень лінгвістичної змінної «ЯКІСТЬ» = (погане, середнє, хороше, відмінне)\*. Властивість лінійної впорядкованості не несе інформації про відстанні між елементами множини

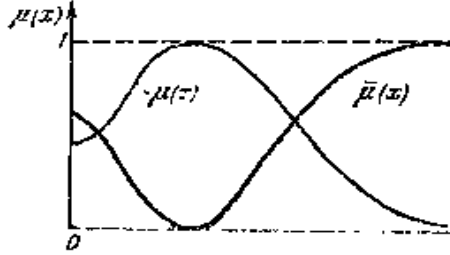
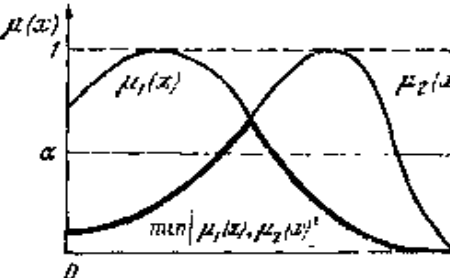
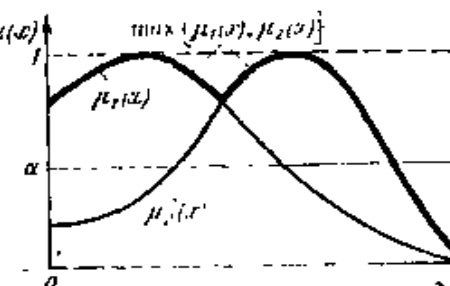
$S$ , тобто, якщо вказується, що один елемент переважніше іншого :  $s_i > s_j$ , то не можна виразити кількісно, наскільки  $s_i$  краще  $s_j$ .

*Лінгвістична змінна - це змінна, значеннями якої можуть бути не тільки числа, але і слова і словосполучення природної або штучної мови. Сенс кожного лінгвістичного значення висловлюється у вигляді нечіткої підмножини універсальної множини  $X$ .*

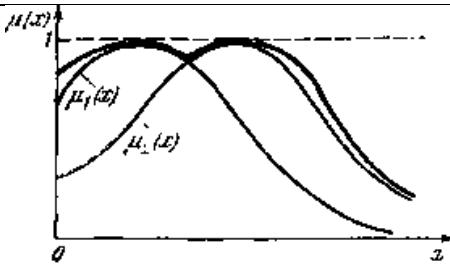
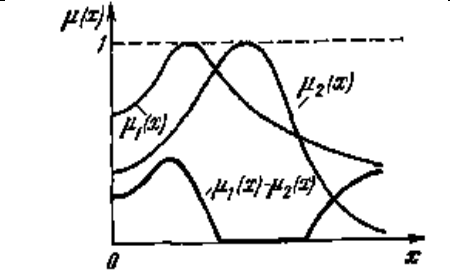
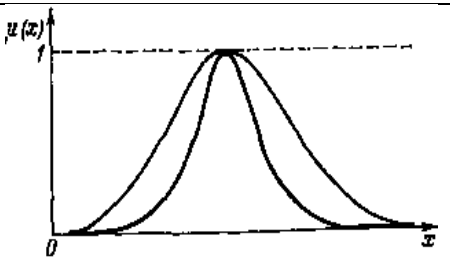




**Таблиця 1.1**  
**Основні операція в алгебрі нечітких множин**

№	Назва операції	Символічний запис (в класі F(X))	Графічне відображення	Символічний запис (в класі Ф([0,1]))	Модифікатор або зв'язка
1	Доповнення	$\mu_3(x) = \bar{\mu}(x) = 1 - \mu(x)$ $\forall x \in X$		$M_3(\alpha) = M(\alpha) = X \setminus UM(1 - \beta), \beta \in [0, \alpha]$ $\forall \alpha \in [0, 1]$	<b>НЕ</b>
2	Перетин I (мінімум: не взаємодіючих перемінних)	$\mu_3(x) = (\mu_1 \wedge \mu_2)(x) = \min\{\mu_1(x), \mu_2(x)\}$ $\forall x \in X$		$M_3(\alpha) = M_1(\alpha) \cap M_2(\alpha),$ $\forall \alpha \in [0, 1]$	<b>I</b> <b>(I, ..., I)</b>
3	Об'єднання I (максимум: не взаємодіючих перемінних)	$\mu_3(x) = (\mu_1 \vee \mu_2)(x) = \max\{\mu_1(x), \mu_2(x)\}$ $\forall x \in X$		$M_3(\alpha) = M_1(\alpha) \cup M_2(\alpha),$ $\forall \alpha \in [0, 1]$	<b>Чи</b> <b>(або, ..., або)</b>

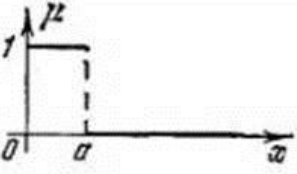
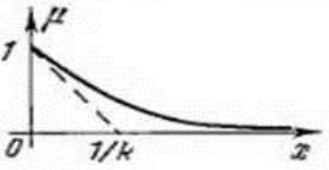
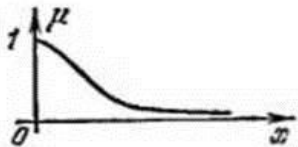
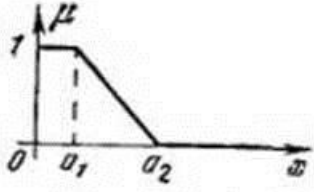
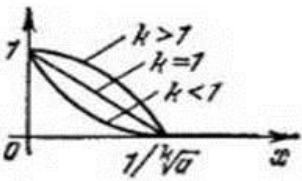
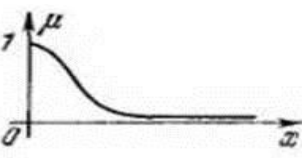
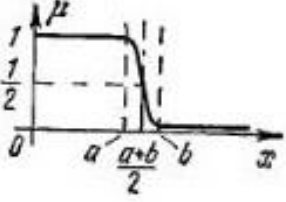
№	Назва операції	Символічний запис (в класі $F(X)$ )	Графічне відображення	Символічний запис (в класі $\Phi([0,1])$ )	Модифікатор або зв'язка
4	Перетин II (обмежене відтворення)	$\mu_3(x) = (\mu_1 \wedge \mu_2)(x) =$ $= \max\{0, \mu_1 + \mu_2(x) - 1\}$ $\forall x \in X$		$M_3(\alpha) =$ $= \bigcup_{\substack{\beta_1 + \beta_2 - 1 \geq \alpha \\ \beta_1, \beta_2 \in [0,1]}} (M_1(\beta_1) \cap M_2(\beta_2))$ $\forall \alpha \in [0,1]$	
5	Об'єднання II (обмежена сума)	$\mu_3(x) = (\mu_1 \vee \mu_2)(x) =$ $= \min\{1, \mu_1(x) + \mu_2(x)\}$ $\forall x \in X$		$M_3(\alpha) =$ $= \bigcup_{\substack{1 \geq \beta_1 + \beta_2 \geq \alpha \\ \beta_1, \beta_2 \in [0,1]}} (M_1(\beta_1) \cap M_2(\beta_2))$ $\forall \alpha \in [0,1]$	Або
6	Перетин III (алгебраїчне відтворення)	$\mu_3(x) = (\mu_1 \cdot \mu_2)(x) =$ $= \mu_1(x) \cdot \mu_2(x)$ $\forall x \in X$		$M_3(\alpha) =$ $= \bigcup_{\substack{\beta_1 \cdot \beta_2 \geq \alpha \\ \beta_1, \beta_2 \in [0,1]}} (M_1(\beta_1) \cap M_2(\beta_2))$ $\forall \alpha \in [0,1]$	

№	Назва операції	Символічний запис (в класі F(X))	Графічне відображення	Символічний запис (в класі Ф([0,1]))	Модифікатор або зв'язка
7	Об'єднання III (алгебраїчна сума)	$\mu_3(x) = (\mu_1 + \mu_2)(x) = \mu_1(x) + \mu_2(x) - \mu_1(x) \cdot \mu_2(x)$ $\forall x \in X$		$M_3(\alpha) = \bigcup_{\substack{\beta_1 + \beta_2 - \beta_1 \cdot \beta_2 \geq \alpha \\ \beta_1, \beta_2 \in [0,1]}} (M_1(\beta_1) \cap M_2(\beta_2))$ $\forall \alpha \in [0,1]$	Або
8	Різниця	$\mu_3(x) = \mu_1(x) - \mu_2(x) = \max\{0, \mu_1(x) - \mu_2(x)\}$ $\forall x \in X$		$M_3(\alpha) = \bigcup_{\substack{\beta_1 - \beta_2 \geq \alpha \\ \beta_1, \beta_2 \in [0,1]}} (M_1(\beta_1) \cap M_2(\beta_2))$ $\forall \alpha \in [0,1]$	
9	Концентрування	$\mu_3(x) = \mu^3(x)$ $\forall x \in X$			Дуже

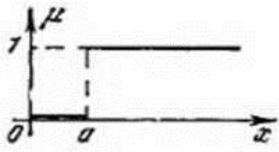
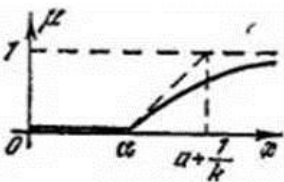
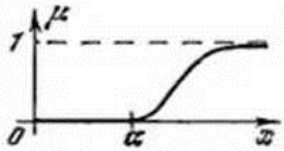
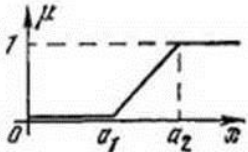
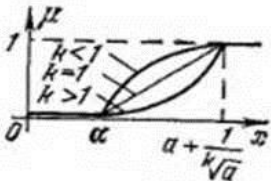
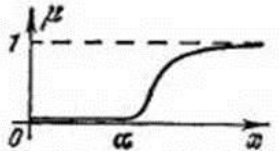
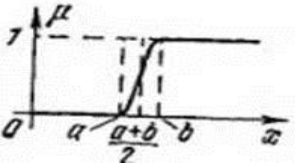
## 29. Огляд найпростіших функцій приналежності

Універсальні множини:  $\mathbf{R}^+, \mathbf{N}$

Функція приналежності утвердження «величина  $x$  мала»

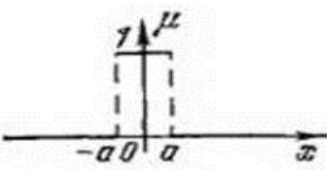
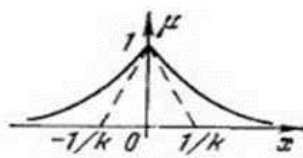
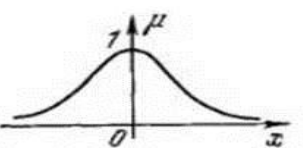
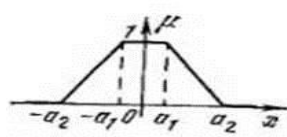
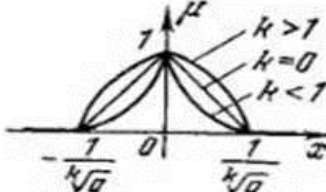
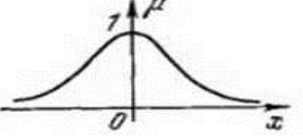
График	Функция	
	$\mu(x) = 1, 0 \leq x < a,$ $\mu(x) = 0, x > a.$	29.1
	$\mu(x) = e^{-kx}, k > 0.$	29.2
	$\mu(x) = e^{-kx^2}, k > 0.$	29.3
	$\mu(x) = 1, 0 \leq x \leq a_1,$ $\mu(x) = (a_2 - x) / (a_2 - a_1), a_1 \leq x \leq a_2,$ $\mu(x) = 0, a_2 \leq x.$	29.4
	$\mu(x) = 1 - ax^k, 0 \leq x \leq 1/\sqrt[k]{a},$ $\mu(x) = 0, 1/\sqrt[k]{a} \leq x.$	29.5
	$\mu(x) = 1 / (1 + kx^2), k > 1.$	29.6
	$\mu(x) = 1, 0 \leq x \leq a,$ $\mu(x) = \frac{1}{2} - \frac{1}{2} \sin \frac{\pi}{b-a} \left( x - \frac{a+b}{2} \right), a \leq x \leq b$ $\mu(x) = 0, b \leq x$	29.7

**Універсальні множини:  $\mathbf{R}^+, \mathbf{N}$**   
**Функція приналежності утвердження «величина  $x$  велика»**

	<p style="text-align: center;"><math>\mu(x) = 0, 0 \leq x &lt; a,</math>  <math>\mu(x) = 1, a \leq x.</math></p> <p><b>29.8</b></p>	
	<p><math>\mu(x) = 0, 0 \leq x \leq a,</math>  <math>\mu(x) = 1 - e^{-k(x-a)}, a &lt; x, k &gt; 0.</math></p> <p><b>29.9</b></p>	
	<p><math>\mu(x) = 0, 0 \leq x \leq a,</math>  <math>\mu(x) = 1 - e^{-k(x-a)^2}, a \leq x, k &gt; 0.</math></p> <p><b>29.10</b></p>	
	<p><math>\mu(x) = 0, 0 \leq x \leq a,</math>  <math>\mu(x) = \frac{x - a1}{a2 - a1}, a1 \leq x \leq a2</math>  <math>\mu(x) = 1, a2 \leq x</math></p> <p><b>29.11</b></p>	
	<p><math>\mu(x) = 0, 0 \leq x \leq a,</math>  <math>\mu(x) = a(x-a)^k, a \leq x \leq a + \frac{1}{\sqrt[k]{a}},</math>  <math>\mu(x) = 1, a + \frac{1}{\sqrt[k]{a}} \leq x</math></p> <p><b>29.12</b></p>	
	<p><math>\mu(x) = 0, 0 \leq x \leq a,</math>  <math>\mu(x) = k(x-a)^2 / (1 + k(x-a)^2), a \leq x &lt; \infty.</math></p> <p><b>29.13</b></p>	
	<p><math>\mu(x) = 0, 0 \leq x \leq a,</math>  <math>\mu(x) = \frac{1}{2} + \frac{1}{2} \sin \frac{\pi}{b-a} \left( x - \frac{a+b}{2} \right), a \leq x \leq b,</math>  <math>\mu(x) = 1, a \leq x</math></p> <p><b>29.14</b></p>	

**Універсальні:  $\mathbf{R}, \mathbf{Z}$**

Функція приналежності утвердження «величина  $|x|$  мала»

	$\begin{aligned} \mu(x) &= 0, & -\infty < x < -a, \\ \mu(x) &= 1, & -a \leq x \leq a, \\ \mu(x) &= 0, & a \leq x. \end{aligned}$	<p><b>29.15</b></p>
	$\begin{aligned} \mu(x) &= e^{kx}, & -\infty < x \leq 0, \\ \mu(x) &= e^{-kx}, & 0 \leq x < \infty, k > 1. \end{aligned}$	<p><b>29.16</b></p>
	$\mu(x) = e^{-kx^2}.$	<p><b>29.17</b></p>
	$\begin{aligned} \mu(x) &= 0, & -\infty < x \leq -a_2 \\ \mu(x) &= \frac{a_2 + x}{a_2 - a_1}, & -a_2 \leq x \leq -a_1, \\ \mu(x) &= 1, & -a_1 \leq x \leq a_1 \\ \mu(x) &= \frac{a_2 - x}{a_2 - a_1}, & a_1 \leq x \leq a_2, \\ \mu(x) &= 0, & a_2 \leq x < \infty \end{aligned}$	<p><b>29.18</b></p>
	$\begin{aligned} \mu(x) &= 0, & -\infty < x \leq \frac{1}{\sqrt[k]{a}} \\ \mu(x) &= 1 - a(-x)^k, & -\frac{1}{\sqrt[k]{a}} \leq x \leq 0, \\ \mu(x) &= 1 - a(x)^k, & 0 \leq x \leq \frac{1}{\sqrt[k]{a}} \\ \mu(x) &= 0, & \frac{1}{\sqrt[k]{a}} \leq x < \infty. \end{aligned}$	<p><b>29.19</b></p>
	$\mu(x) = \frac{1}{1+kx^2}, k > 1.$	<p><b>29.20</b></p>

	$\mu(x) = 0, \quad -\infty < x \leq -b,$ $\mu(x) = \frac{1}{2} + \frac{1}{2} \sin \frac{\pi}{b-a} \left( x + \frac{a+b}{2} \right), \quad -b \leq x \leq -a$ $\mu(x) = 1, \quad -a \leq x \leq a,$ $\mu(x) = \frac{1}{2} - \frac{1}{2} \sin \frac{\pi}{b-a} \left( x - \frac{a+b}{2} \right), \quad a \leq x \leq b,$ $\mu(x) = 0, \quad b \leq x < \infty$ <b>29.21</b>	<b>29.21</b>
--	--	--------------

**Універсальні: R, Z**

**Функція приналежності утвердження «величина |x| велика»**

	$\mu(x) = 1, \quad -\infty < x < -a,$ $\mu(x) = 0, \quad -a < x < a,$ $\mu(x) = 1, \quad a < x < \infty.$	<b>29.22</b>
	$\mu(x) = 1 - e^{kx}, \quad -\infty < x \leq 0,$ $\mu(x) = 1 - e^{-kx}, \quad 0 \leq x < \infty, \quad k > 1.$	<b>29.23</b>
	$\mu(x) = 1 - e^{-kx^2}, \quad k > 1.$	<b>29.24</b>
	$\mu(x) = 1, \quad -\infty < x < -a_2,$ $\mu(x) = -\frac{x+a_1}{a_2-a_1}, \quad -a_2 \leq x \leq -a_1,$ $\mu(x) = 0, \quad -a_1 \leq x < a_1,$ $\mu(x) = \frac{x-a_1}{a_2-a_1}, \quad a_1 \leq x \leq a_2,$ $\mu(x) = 1, \quad a_2 \leq x < \infty$	<b>29.25</b>

	$\mu(x) = 1, -\infty < x < -\frac{1}{\sqrt[k]{a}},$ $\mu(x) = a(-x)^k, -\frac{1}{\sqrt[k]{a}} < x < 0,$ $\mu(x) = ax^k, 0 \leq x \leq \frac{1}{\sqrt[k]{a}}$ $= 1, \frac{1}{\sqrt[k]{a}} \leq x < \infty.$	<p>29.26</p>
	$\mu(x) = \frac{kx^2}{1+kx^2} = \frac{1}{1+\frac{1}{kx^2}}, k > 1.$ <p>29.27</p>	<p>29.27</p>
	$\mu(x) = 1, -\infty < x \leq -b,$ $\mu(x) = \frac{1}{2} - \frac{1}{2} \sin \frac{\pi}{b-a} \left( x + \frac{a+b}{2} \right), -b \leq x \leq -a,$ $\mu(x) = 0, -a \leq x \leq a,$ $\mu(x) = \frac{1}{2} + \frac{1}{2} \sin \frac{\pi}{b-a} \left( x - \frac{a+b}{2} \right), a \leq x \leq b,$ $\mu(x) = 1, b \leq x < \infty.$	<p>29.28</p>

### 38. Мережі нечітких елементів

Подібно до того, як це робиться в теорії контактних мереж [2F], теорії надійності [7K] та інших, мережеве уявлення послідовно-паралельно з'єднаних елементів було б цікаво використовувати для аналізу функцій нечітких змінних.

**Нечіткий елемент мережі.** З кожною нечіткою змінною  $a \in [0,1]$  ми будемо пов'язувати елемент  $a$  позначається тим же символом. Нам потрібно побудувати мережу, що складається з таких елементів  $a$ .

З функцією  $a \wedge b$  зв'яжемо мережу, представлену на рис. 38.1, з функцією  $a \vee b$  зв'яжемо мережу, представлену на рис. 38.2. Першу будемо називати послідовною мережею, другу - паралельною.

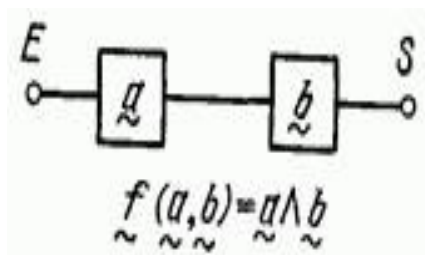


Рис. 38.1

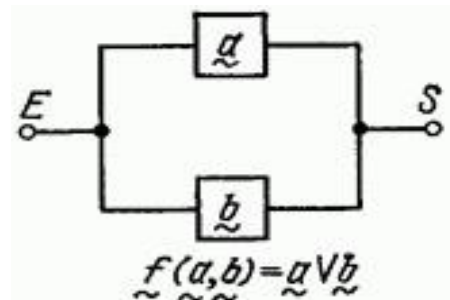


Рис. 38.2



У таких мережах необхідно ще вказувати вхід  $E$  і вихід  $S$ . Результат виконання операції на елементах мережі називається *потокот мережі*.

Так, якщо  $a = 0,7$  и  $b = 0,4$ , то для мережі на **рис. 38.1** потік із  $E$  у  $S$  дорівнює **0,4**, а для мережі на **рис.38.2** при тих же значеннях  $a$  і  $b$  потік дорівнює **0,7**.

**Теорема 1.** Кожній аналітичній функції нечітких змінних  $f(a, b, \dots)$  можна поставити у відповідності *мережу нечітких елементів*, з послідовним розташуванням яких пов'язана операція  $\wedge$ , а паралельним – операція  $\vee$ .

*Доведення.* Ми вже бачили, що кожною аналітичною функцією  $f(a, b, \dots)$  можна за визначенням поставити відповідно наведену поліноміальну форму відносно  $\wedge$  або  $\vee$ . Кожній з форм потім можна поставити у відповідність мережу.

**Приклад.** Функції (38.1), представленій у наведеній поліноміальній формі відносно  $\vee$

$f(a,b,c) = (a \wedge \bar{b} \wedge c) \vee (a \wedge b) \vee (\bar{a} \wedge b \wedge c),$	(38.1)
--	--------

можна поставити мережу, зображену на **рис. 38.3**.

Мережа, відповідна тій самій функції, але вираженої у наведеній поліноміальній формі відносно  $\wedge$

$f(a,b,c) = (a \vee \bar{a}) \wedge (a \vee b) \wedge (a \vee c) \wedge (b \vee \bar{b}) \wedge (b \vee c),$	(38.2)
--	--------

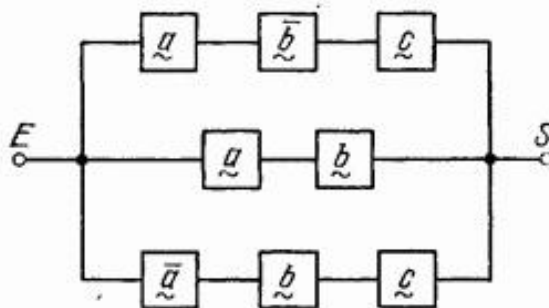
зображена на **рис. 38.4**.

**Маршрути.** Послідовність елементів, з'єднаних один за іншим зв'язкою  $\wedge$  від  $F$  до  $S$ , буде називатися *маршрутом* або *шляхом* \*).

Таким чином, на **рис. 38.3** послідовність елементів

$(a, \bar{b}, c) \in \text{маршрут}$	(38.3)
--------------------------------------	--------

Маршрут називається *простим*, якщо він не містить одного й того ж елемента  $x$  або елемента  $\bar{x}$  більше одного разу \*\*).



**Рис. 38.3**

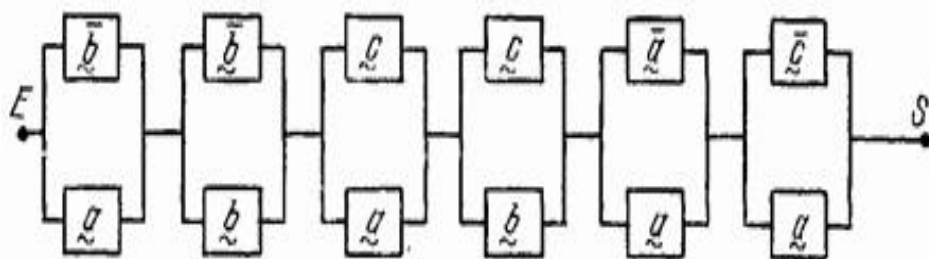


Рис. 38.4

Так, на рис. 38.4 послідовність елементів а

$(a, a, c, \bar{b}, b)$ становить маршрут,	(38.4)
$(a, c, \bar{b}, b)$ - простий маршрут.	(38.5)

Оскільки маршрут розглядається відносно операції  $\wedge$  - асоціативної і комутативної, то порядок, в якому елементи розташовані в послідовності, несуттєвий.

**Максимально простий маршрут.** Нехай  $I$  – звичайна множина простих маршрутів мережі, тоді будь-який простий маршрут, який не містить ніякого іншого маршруту з  $I$ , називається *максимально простим маршрутом*. Якщо  $n$  – число нечітких змінних у  $f$ , то очевидно, що максимально простий маршрут містить не більше  $2n$  елементів.

**Основна властивість.** Розташувавши усі максимально прості маршрути паралельно, отримаємо мережу, еквівалентну наведеному поліному відносно  $f$ .

Властивість стає очевидною, якщо зіставити способи побудови поліноміальних форм з побудовою послідовно-паралельних мереж на максимально простих маршрутах, які з'єднуються паралельно.

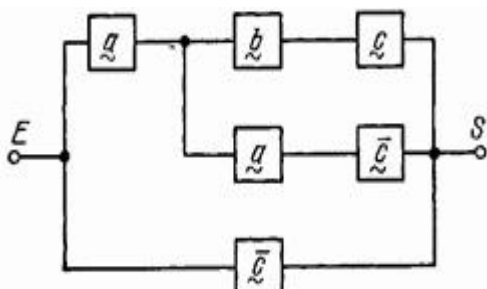


Рис. 38.5

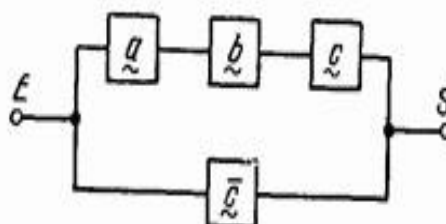


Рис. 38.6

**Приклад 1.** Розглянемо мережу на рис. 38.5, відповідну функції.

$f(a, b, c) = [a \wedge [(b \wedge c) \vee (a \wedge \bar{c})]] \vee \bar{c}$	(38.6)
Випишемо множину маршрутів $[(a, b, c), (a, a, \bar{c}), (\bar{c})],$	(38.7)
Множина простих маршрутів $[(a, b, c), (a, \bar{c}), (\bar{c})],$	(38.8)

Множина максимально простих маршрутів [ (a, b, c), (c̄)].	(38.9)
Останнє відповідає наведеній поліноміальній формі в: <u>f(a, b, c) = (a ∧ b ∧ c) ∨ (c̄)</u>	(38.10)

і найпростішій мережі, зображеній на **рис. 38.6**.

**Приклад 2.** Розглянемо більш складний випадок (**рис. 38.7**).  
Випишемо множину маршрутів<sup>\*)</sup>

[(a, b, c̄), (a, c, a), (a, c, c, ā, c̄), (a, b, ā, c, a), (a, b, ā, c, c, b, c̄), (b, c, a, ), (b, ā, c̄), (b, ā, b, c, a), (b, c, c, b, c̄), (b, c, c, b, ā, c, a)],	(38.11)
Множина простих маршрутів [(a, b, c̄), (a, c), (a, ā, c, c̄), (a, ā, b, c), (a, ā, b, c, c̄), (a, b, c), (ā, b, c̄), (a, ā, b, c), (b, c, c̄), (a, ā, b, c, )],	(38.12)
Множина максимально простих маршрутів [(a, b, c̄), (ā, b, c̄), (b, c, c̄), (a, c)]	(38.13)
і відповідну <b>(38.13)</b> наведену поліноміальну форму у ∨ <u>f(a, b, c) = (a ∧ b ∧ c̄) ∨ (ā ∧ b ∧ c̄) ∨ (b ∧ c ∧ c̄) ∨ (a ∧ c),</u>	(38.14)

**Плоскі мережі.** Якщо в мережі не існує зв'язку між двома елементами, що перетинає інший зв'язок, коли мережа між *E* та *S* намальована на площині, то говорять, що мережа реалізована в площині або планарна. В іншому випадку говорять, що мережа неплоска. Так, мережа на **рис. 38.5** плоска, а на **рис. 38.7** – неплоска.

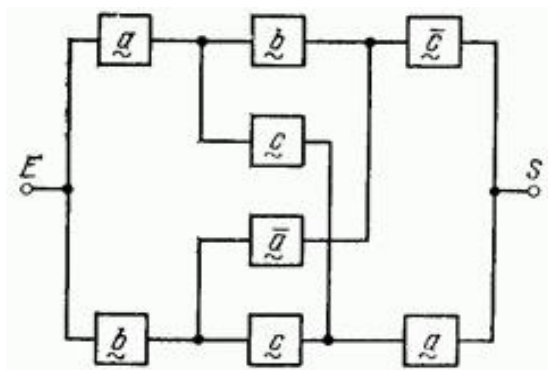


Рис. 38.7

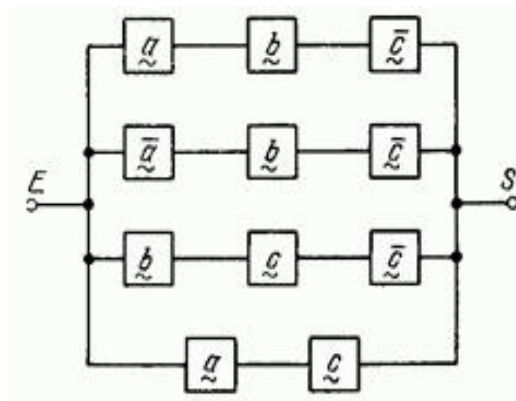


Рис. 38.8

Відзначимо наступну властивість: мережі, відповідні поліноміальним формам  $u \wedge$  або  $u \vee$ , - плоскі. Дійсно, будь-якій поліноміальній формі  $u \wedge$  відповідає паралельно-послідовна мережа, яка є плоскою, і аналогічно будь-якій поліноміальній формі  $u \vee$  відповідає послідовно-паралельна мережа, яка є плоскою (див., приклад, рис. 38.1 и 38.2).

**Двоякість плоскої мережі.** Нехай  $R$  – плоска мережа. Оскільки мережа плоска, то можна визначити межі  $\alpha, \beta, \dots$  як частини площини, обмежені зв'язками і елементами (див. рис. 38.9.), усередині яких не міститься жодного елемента. У кожній з цих меж виберемо точку, яка стане точкою перетину зв'язків нової мережі. Виберемо ще по точці і в двох зовнішніх межах: вище і нижче лінії  $ES$ .

Слідуючи правилу: кожному з вибраних точок з'єднати зв'язком з кожним з елементів, суміжним з межею, в якій знаходиться точка, - побудуємо нову мережу  $R^!$ . Мережа  $R^!$  називається двоякісною  $R$ . На рис. 38.9 штриховий лінією зображено мережу  $R^!$ , двоякісна  $R$ . На рис. 38.10 мережа  $R^!$  зображена безпосередньо.

Для мережі та двоїстої мережі легко перевірити наступну властивість:

$(R^!)^! = R,$	(38.15)
----------------	---------

тобто двоякісна мережа до мережі, яка сама є двоякісна мережа мережі  $R$ , співпадає з мережею  $R$ .

**Метод антимаршрутів.** Розглянемо плоску \*) мережу  $R$  і двоїсну мережу  $R^!$ . Маршрути, відповідні  $R^!$ , називаються *антимаршрутами* мережі  $R$ .

Максимально прості маршрути  $R^!$  дадуть максимально прості антимаршрути  $R$ , а пізніше приведуть до поліноміальній формі функції  $f(a, b, \dots, c)$  у  $R^!$ , представленої мережею  $R$ .

З цією поліноміальною формою в  $R$  буде пов'язана паралельно-послідовна мережа, еквівалентна даній мережі.

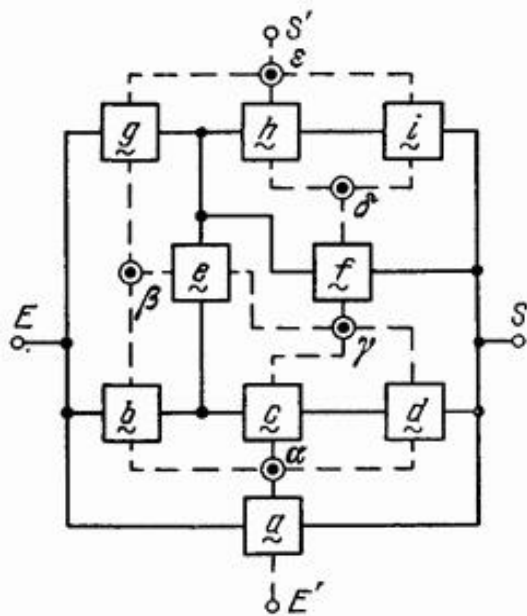


Рис. 38.9

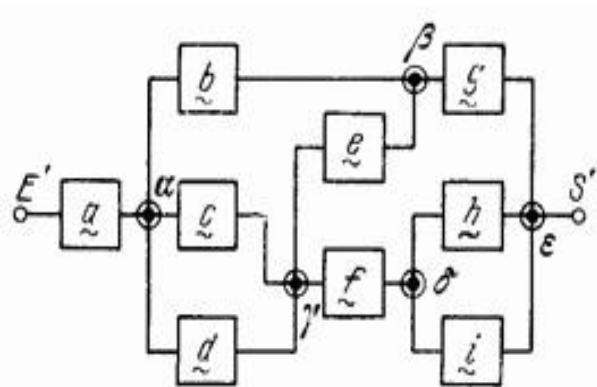


Рис. 38.10

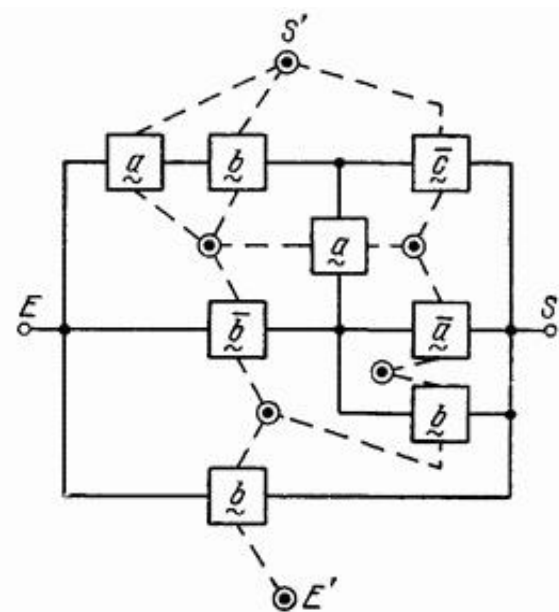


Рис. 38.11

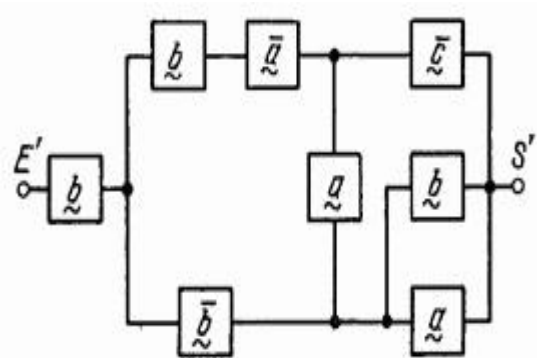


Рис. 38.12

**Приклад.** Розглянемо мережу  $R$  на рис. 38.11, двоякісна якої мережа  $R^I$  представлена на рис. 38.12

Маршрути мережі  $R^I$  - це антимаршрути мережі  $R$ . Впишемо їх

$[(b, b, \bar{a}, \bar{c}), (b, b, \bar{a}, a, b), (b, b, \bar{a}, a, a), (b, \bar{b}, \bar{a}, \bar{c}), (b, \bar{b}, b), (b, \bar{b}, a)].$	(38.16)
---	---------

Впишемо множину<sup>\*)</sup> простих антимаршрутів

$[(b, \bar{a}, \bar{c}), (b, \bar{a}, a)1, (b, \bar{a}, a)2, (b, \bar{b}, \bar{a}, \bar{c}), (b, \bar{b}, (b, \bar{b}, a)],$	(38.17)
--	---------

яка скорочується до множини простих антимаршрутів:

$[(\bar{a}, b, \bar{c}), (a, \bar{a}, b), (b, \bar{b})].$	(38.18)
---	---------

--	--

Таким чином, приведена форма відносно  $\wedge$ , відповідна паралельно-послідовній мережі на **рис. 38.13**, має вигляд

$f(a,b,c) = (\bar{a} \vee b \vee \bar{c}) \wedge (a \vee \bar{a} \vee b) \wedge (b \vee \bar{b})$	(38.19)
---	---------

Методом маршрутів можна знайти поліноміальну форму відносно  $\vee$ :

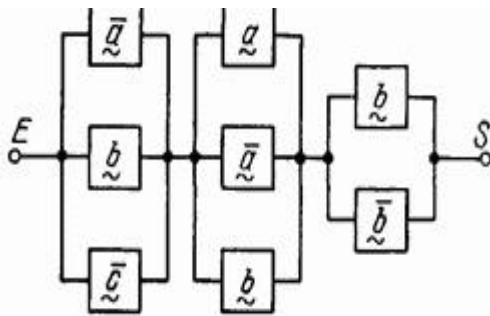
$f(a,b,c) = (a \wedge \bar{b} \wedge c) \vee (\bar{a} \wedge \vee \bar{b}) \vee (b)$	(38.20)
--	---------

які

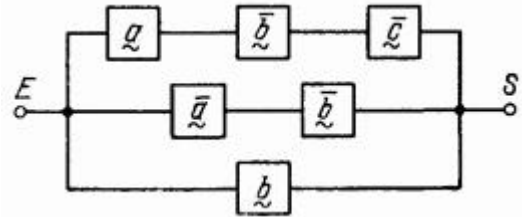
й відповідає послідовно-паралельна мережа, зображена на **рис. 38.14**.

Використовуючи відповідне розкладання, можна показати, що (38.19) та (38.20) дійсно представляють одну і ту ж функцію.

*Зауваження.* Ми знаємо<sup>\*\*</sup>), що будь-яка мережа контактних ланцюгів можна виконати з різних технічних елементів (діодів, мостів, транзисторів, інтегральних схем і т. п.).



**Рис. 38.13**



**Рис. 38.14**

Всі наведені теоретичні розгляди, що стосуються технологічних реалізацій функцій нечіткої логіки при відповідному виборі операторів, можна адаптувати до використання більш різноманітних технічних засобів. Але є небезпека, що технічна реалізація нечітких логік виявиться занадто дорогою (тут на відміну від бінарної техніки потрібен точний контроль потенціалів). Однак міркування про обмежені можливості техніки навряд чи залишаться справедливими навіть у недалекому майбутньому.

#### 41. Теорія нечітких підмножин і теорія структурних функцій

Між теорією нечітких змінних, як вона визначена в § 32 і наступних параграфах, і теорією структурних функцій, що вивчаються в теорії надійності системи, можна встановити деякі цікаві зв'язки. Спочатку нагадаємо основні поняття теорії структурних<sup>\*\*</sup>) функцій.

**Структурні функції.** Розглянемо змінні  $a, b, \dots \in \{0, 1\}$ .

Для цих бінарних змінних будемо використовувати тільки наступні операції<sup>\*\*\*</sup>):

a · b – звичайне множення	(41.1)
---------------------------	--------

$ab = a + b - a \cdot b, \text{ где } (+) \text{ означає звичайне додавання,}$ $a (-) \text{ звичайне віднімання.}$	<b>(41.2)</b>
---	---------------

Введемо функції цих змінних, для побудови яких використовуються тільки операції

$\hat{+}$  та  $\cdot$ .

Але спочатку розглянемо загальні властивості змінних  $a, b, \dots \in [0, 1]$  та операцій  $\cdot$  та  $\hat{+}$ .

$\left\{ \begin{array}{l} a * b = b * a, \\ a + b = b + a \end{array} \right\}$	комутативність	<b>(41.3)</b> <b>(41.4)</b>
$\left\{ \begin{array}{l} a \cdot (b \cdot c) = (a \cdot b) \cdot c, \\ a + (b + c) = (a + b) + c \end{array} \right\}$	асоціативність	<b>(41.5)</b> <b>(41.6)</b>
$\left\{ \begin{array}{l} a \cdot a = a \\ a + a = a \end{array} \right\}$	ідемпотентність	<b>(41.7)</b> <b>(41.8)</b>
$\left\{ \begin{array}{l} a \cdot (b + c) = a \cdot b + a \cdot c, \\ a + (b \cdot c) = (a + b) \cdot (a + c), \end{array} \right\}$	дистрибутивність	<b>(41.9)</b> <b>(41.10)</b>
$a \cdot 0 = 0,$		<b>(41.11)</b>
$a + 0 = a,$		<b>(41.12)</b>
$a \cdot 1 = a,$		<b>(41.13)</b>
$a + 1 = 1.$		<b>(41.14)</b>
Обозначимо через $\varphi$ структурну функцію змінних $\varphi(a, b, \dots)$ .		<b>(41.15)</b>

Наприклад,

$\varphi(a, b, c) = a + ab + bc$	<b>(41.16)</b>
----------------------------------	----------------

$\varphi$  є структурна функція.

Нагадаємо дві властивості поглинання, які дозволяють спрощувати структурні функції:

$a \cdot (a+b) = a,$	<b>(41.17)</b>
$a+ab = a.$	<b>(41.18)</b>

Обидві ці властивості виводяться з (41.3) и (41.14).

Використовуючи поняття максимального одночлена, будь-яку функцію  $\varphi(a,b,\dots)$  можна виразити в поліноміальній формі відносно  $\cdot$  та  $\hat{+}$ .

Наприклад, функція

$\varphi(a, b, c, d) = \overline{a+cb}+bd$	<b>(41.19)</b>
--	----------------

утворена трьома максимальними одночленами і не може бути спрощена далі.

Функція ж $\varphi(a, b, c, d) = a+b+bd+cd$ ☒ допускає спрощення - її можна звести до виду	<b>(41.20)</b>
---	----------------

$\varphi(a, b, c, d) = a+b+cd$ ☒	<b>(41.21)</b>
----------------------------------	----------------

Поліноміальна форма, що містить тільки максимальні одночлени, буде називатися *наведеною* або *канонічною*.

Будемо говорити, що дві структурні функції рівні або тотожні, якщо вони зводяться до однієї і тієї ж поліноміальної форми відносно отвору  $\cdot$  або суми  $\hat{+}$ .

Звичайно, будь-яку канонічну форму відносно  $\cdot$  можна перетворити в канонічну відносно  $\hat{+}$  і навпаки.

З кожною структурною функцією можна пов'язати представлення у вигляді мережі, в якій послідовне розташування елементів відповідає операції  $\cdot$ , а паралельне – операції  $\hat{+}$ .

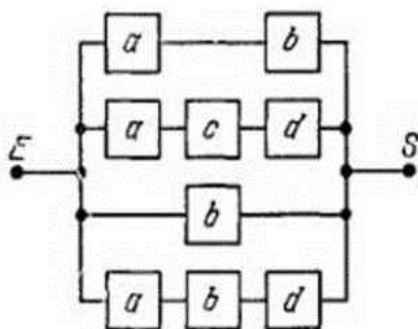


Рис. 41.1

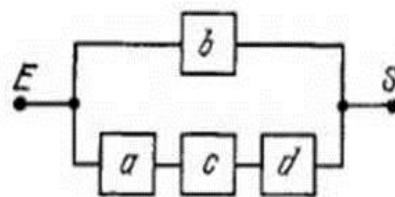


Рис. 41.2



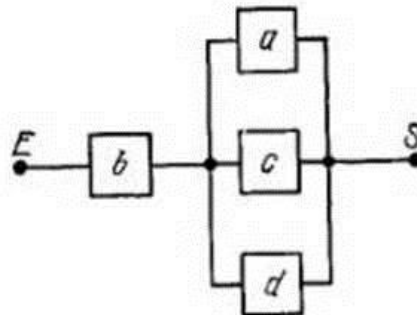


Рис. 41.4

Приклад. Розглянемо структурну функцію, відповідну мережі на рис. 41.1:

$\varphi(a, b, c, d) = ab \hat{+} acd \hat{+} b \hat{+} abd.$	(41.22)
---	---------

$b \hat{+} abd = b,$	(41.23)
----------------------	---------

то перше скорочення дає

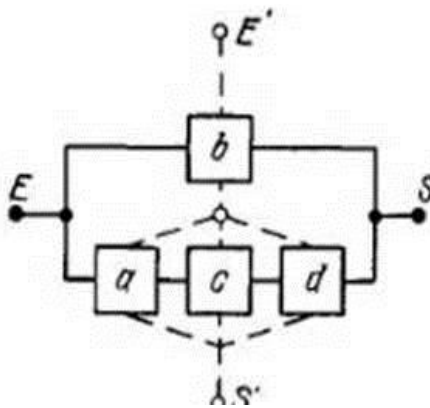
$\varphi(a, b, c, d) = ab \hat{+} acd \hat{+} b.$	(41.24)
---	---------

І так як

$ab \hat{+} b = b,$	(41.25)
---------------------	---------

то друге скорочення приведе до

$\varphi(a, b, c, d) = b \hat{+} acd.$	(41.26)
--	---------



Звичайно, обидва скорочення можна було б виконати відразу. Отже,

$\varphi_1(a, b, c, d) = b+acd.$	<b>(41.27)</b>
----------------------------------	----------------

- канонічна форма функції  $\varphi$ . Відповідна їй мережа представлена на **рис. 41.2**.

Використовуючи двоякісну мережу, зображену на **рис. 41.3**, отримуємо двоїсту канонічну форму, відповідну трьом паралельним маршрутам, що йдуть з  $E$  та  $S$  (див. **рис. 41.4**):

$\varphi_i(a, b, c, d) = ba+bc+bd,$	<b>(41.28)</b>
-------------------------------------	----------------

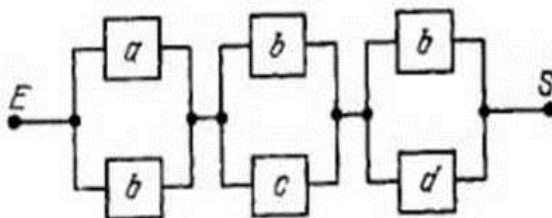
Взаємозамінюючи операції  $\cdot$  та  $\hat{+}$ , приходимо до

$\varphi_2(a, b, c, d) = (a+b)\cdot(b+c)\cdot(b+d).$	<b>(41.29)</b>
--	----------------

Мережа, відповідна цій другій канонічній формі функції  $\varphi$ , представлена на **рис. 41.5**.

**Перехід к поверхням.** Щоб уникнути занадто абстрактного викладу розглянемо конкретний приклад.

Припустимо, що змінні  $a, b, \dots \in \{0, 1\}$ , позначають стан компонентів  $A, B, \dots \in E$ , де  $E$  – множина компонентів. Цю множину будемо називати словами *складне обладнання, просте обладнання або система*.



**Рис. 41.5**

Будемо вважати, що компонент

$X$ – функціональний, якщо $x = 1$ , $X$ – не функціональний, якщо $x = 0$	<b>(41.30)</b>
---	----------------

де  $x \in \{0, 1\}$  – бінарна змінна, пов'язана з компонентом  $X$ .

У цьому випадку  $\varphi(a, b, \dots)$  представляє таку бінарну функцію, приймаючу значення  $\{0, 1\}$ , що система

$E$ – функціональна, якщо $\varphi = 1$ , $E$ – не функціональна, якщо $\varphi = 0$ ,	<b>(41.32)*</b>
---	-----------------

та  $\varphi(a, b, \dots)$  виражає залежність  $E$  від своїх компонентів.

Нехай  $p_x$  є ймовірність того, що компонент  $X$  – функціональний, та  $h_E$  – ймовірність того, що система  $E$  – функціональна.

Підрахуємо ймовірність  $h_E$  як функцію ймовірностей  $p_a, p_b, \dots$

Для того щоб показати, як провести відповідні обчислення, потрібно згадати дві ідемпотентні формули (41.7) и (41.8):

$a \cdot a = a$	(41.33)
$a + a = a \hat{+}$	(41.34)

оскільки, якщо обчислення проводяться відносно звичайного складання, то

$a + a = 2a \hat{+}$	(41.35)
----------------------	---------

і очевидно, що  $\{0,1\}$  не буде областю значень для цієї суми.

У теорії ймовірностей, яка застосовується в теорії надійності розглянутого тут класу систем, вважається, що якщо  $p_x$  – ймовірність того, що  $X$  – функціональний компонент, тоді  $1 - p_x$  є ймовірність того, що він не функціональний.

Розглянемо систему  $E = \{A, B\}$ . Для ймовірності функціонування цієї системи отримаємо

$h_E = p_a \cdot p_b$ , що відповідає $\varphi(a,b) = ab$	(41.36)
$1 - h_E = (1 - p_a) \cdot (1 - p_b)$ , що відповідає $\varphi(a,b) = a + b$ ,	(41.37)

останній вираз можна записати у вигляді

$h_E = 1 - (1 - p_a) \cdot (1 - p_b) = p_a + p_b - p_a p_b$ ,	(41.38)
---	---------

що відповідає  $\varphi(a,b) = a + b - ab$ .

Ми бачимо, що існує ізоморфізм \*) між функціями  $h_E$  та  $\varphi$ .

Однак закон дистрибутивності, справедливий на множині  $\{a, b, \dots\}$  змінних, що приймають значення з  $\{0,1\}$ , щодо операцій  $\cdot$  та  $\hat{+}$ , перестає бути справедливим для ймовірностей  $p_a, p_b, \dots$

Але дистрибутивність відновлюється, якщо розглядати звичайні операції  $+$  і  $\cdot$ . Для переходу від функцій  $\varphi$  к функціям  $h_E$  потрібно у  $\varphi$  замінити оператори  $\hat{+}$  на  $+$ , що може призводити до використання -;

потім можна перейти від  $\varphi$  до  $h_E$ , замінюючи  $x$  на  $p_x$  і не забуваючи, де це необхідно, застосовувати властивість ідемпотентності (41.7). Таким чином, ми припускаємо встановлювати ймовірність функціонування (так звану надійність) системи, для якої структурна функція виражається у вигляді (41.22), тобто такої системи  $E$ , що

$E$  функціонує, якщо  $A$  и  $B$  – функціональні,  
 $E$  функціонує, якщо або/та  $A, C$  та  $D$  – функціональні,  
 або/та  $B$  – функціональні,  
 або/та  $A, B$  и  $D$  – функціональні.

Очевидно, що

$\varphi(a, b, c, d) = ab + acd + b + abd = b + acd$	(41.39)
--	---------

(см. (41.27)) чи знову $\varphi(a, b, c, d) = b + acd - abcd$ ,	(41.40)
--	---------

$h_E(p_a, p_b, p_c, p_d) = p_a + p_a p_c p_d - p_a p_b p_c p_d$	<b>(41.41)</b>
---	----------------

Для канонічної форми (41.29) отримуємо

$\begin{aligned} \varphi(a, b, c, d) &= (a+b) \cdot (b+c) \cdot (b+d) \\ &= (ab+ac+b^2+bc) \cdot (b+d) = (ac+b) \cdot (b+d) \\ &= abc+acd+b^2+bd = b+acd. \end{aligned}$	<b>(41.42)</b>
--	----------------

**Показник якості функціонування системи.** У деяких завданнях, пов'язаних з оцінкою функціонування системи, враховують не тільки той факт, працює чи не працює система, а й рівень якості її роботи.

наприклад:

- працює відмінно,
- працює дуже добре,
- працює досить добре,
- працює досить погано,
- не працює.

Припустимо тепер, що кожному компоненту  $X_1$  системи або елементу множини

$$E = (X_1, X_2, \dots, X_n)$$

ми ставимо відповідно нечітку змінну

$$x_1 = \mu_A(X_1) \in [0, 1],$$

де  $A$  описує стан системи  $E$ , яке залежить від характеристики стану кожного свого компонента. У цьому випадку  $A$  дійсно є нечітка підмножина.

Якщо допустити, що рівень системи задається функцією

$$(x_1, x_2, \dots, x_n) = x_1 \wedge x_2 \wedge \dots \wedge x_n,$$

коли функціонування системи можна описати послідовною мережею і

$$(x_1, x_2, \dots, x_n) = x_1 \vee x_2 \vee \dots \vee x_n,$$

коли функціонування системи описується паралельною мережею, то доведеться звернутися до різних понять з теорії нечітких підмножин, виключаючи питання, пов'язані з поняттям доповнення, які не мають прямого відношення до задачі оцінки якості роботи системи.

Властивості змінних  $x_i$ ,  $i=1, 2, \dots, n$ , збігаються з тими, які були сформульовані в (32.12) – (32.23), а властивості (32.24) - (32.26) втрачають своє значення.

Відзначимо, що властивості поглинання

$$a \vee (a \wedge b) = a,$$

$$a \wedge (a \vee b) = a,$$

залишаються справедливими, що дозволяє скорочувати формули і дає можливість ввести як поняття максимального одночлена (відносно  $\wedge$  або  $\vee$ ), так і поняття наведеної поліноміальної форми.

Функції, подібні  $\varphi$ , називатимуться показниками якості функціонування системи.

**Приклад.** Розглянемо рисунок (41.6). За схемою легко визначити структурну функцію

$$\varphi(a,b,c) = [(a \wedge b \wedge c) \vee (a \wedge c) \vee b] \wedge a$$

Застосовуючи правило поглинання поліному в квадратних дужках, отримуємо

$$\varphi(a,b,c) = (a \wedge b) \vee (a \wedge c).$$

Функцію  $\varphi$  можна вивчати за допомогою таблиці значень, як це ми робили в § 32 для функції  $\varphi$ , але не включаючи в таблицю доповнення змінних. Розглянемо випадки, коли є

Одна змінна  $a$  :

$a$ ,

Дві змінні  $a$  и  $b$ :

$a \leq b$ ,

$b \leq a$ ,

$a \leq b \leq c$ ,

$a \leq c \leq b$ ,

три змінні  $a$ ,  $b$  та  $c$ :

$b \leq a \leq c$ ,

$b \leq c \leq a$ ,

$c \leq a \leq b$ ,

$c \leq b \leq a$ , і т.д. для чотирьох, п'яти, ... змінних.

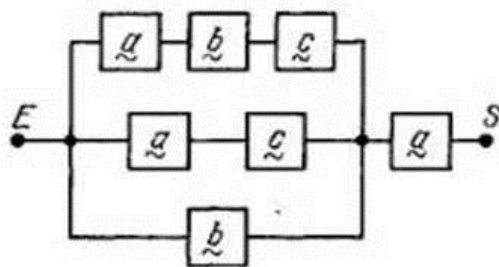


Рис. 41.6

Якщо  $\varphi$  - показник якості функціонування системи, що залежить від  $n$  змінних, то таблиця буде мати  $n!$  рядків: кожен рядок може приймати  $n$  значень, усього є

$$N = n^n$$

різних функцій. серед цих  $n^n$  функцій тільки невелика їх число може бути представлено в канонічній \*) формі (відносно  $\wedge$  або  $\vee$ ) і, отже, представлено надійною мережею.

Знову повертаючись до прикладу на рис. 41.6. отримуємо таблицю, яка наведена на рис. 41.7.

Розглянемо систему  $E$ , що складається з двох компонентів  $A$  і  $B$ .

Припустимо, що  $a$  характеризує якість роботи компонента  $A$ ,  $a$   $b$  - якість роботи  $B$ ,  $a$ ,  $b \in \{0,1\}$ . Припустимо також, що в мережі ці два компоненти з'єднані паралельно. тоді маємо

$\varphi(a,b) = a \vee b.$	<b>(41.64)</b>
----------------------------	----------------

Нарешті, припустимо, що  $a = b$ , тоді

$\varphi(a, b) = \varphi(a) = \varphi(b) = a$	<b>(41.65)</b>
---	----------------

Звідси слідує, що надмірність не змінює якості роботи системи.

Тепер розглянемо надійність тієї ж системи **E**. Якщо **a i b** - не випадкові змінні, **a, b ∈ {0,1}**, тоді маємо

$\varphi(a, b) = a+b = a+b-ab.$	<b>(41.66)</b>
---------------------------------	----------------

Звідси отримаємо

$h(p_a, p_b) = p_a + p_b - p_a p_b.$	<b>(41.67)</b>
--------------------------------------	----------------

*Припустимо, що  $p_a = p_b$ : тоді маєм*

$h(p_a, p_b) = 2p_a - p_a^2 \geq p_a$	<b>(41.68)</b>
$h(p_a, p_b) > p_a$ если $p_a \neq 0$ и $p_a \neq 1$	<b>(41.69)</b>

Таким чином, надмірність підвищує надійність, але не рівень функціонування.

Ці два поняття - «рівень функціонування» і «надійність» - не повинні змішуватися. Перша пов'язана з теорією нечітких підмножин, а друге - з теорією ймовірностей.

Отже, якщо кожен з двох компонентів працює досить добре, то їх паралельне з'єднання працює настільки ж добре, але не краще, зате надійність системи підвищується. Цей приклад добре ілюструє всю різницю, що існує між двома поняттями.

Про показники надійності та якості роботи системи можна сказати, що це монотонні показники. В результаті паралельного підключення мережі **R'** до мережі **R** не погіршується ні якість роботи системи, ні її надійність. У результаті послідовного з'єднання цих мереж ні надійність системи, ні якість її роботи не поліпшується.

Відзначимо, що намічена вище теорія функцій якості допускає узагальнення, в якому змінні **a, b, ...** що характеризують якість роботи компонентів системи, приймають значення не в **{0,1}**, а в довільній упорядкованій множині, як це і зроблено в гол. **V**, в якій розглядається важливе узагальнення теорії Заде.

Поняття показника якості може стати предметом різних визначень, що виникають в теорії таксономії \*).

### 1.1.3. Ступінь нечіткості

Залежно від інформованості та досвіду осіб, що рішення, подання, якими вони користуються при побудові нечітких множин, можуть володіти різним ступенем розпливчастості, в результаті чого виникає питання про градаціях нечіткості самих нечітких множин.

Формалізація заходи нечіткості має важливе практичне і теоретичне значення. Цьому питанню присвячено значну кількість робіт, в яких, за

винятком, розглядаються різні варіанти заходів нечіткості, що задовольняють вимогам, які введені вперше авторами роботи .

Нехай  $A$  - нечітка множина на  $X$  і  $d(A)$  - міра нечіткості  $A$ . Тоді повинні виконуватися такі вимоги:

- 1)  $d(A)=0$ , якщо  $A$  чітка множина;
- 2)  $d(A)$  досягає максимуму при  $\mu_A(X) = \frac{1}{2}$  для усіх  $x \in X$ ;
- 3)  $d(A) \geq d(A^*)$ , якщо 
$$\begin{cases} \mu_{A^*}(x) \geq \mu_A(x), & \mu_A(x) > \frac{1}{2}, \\ \mu_{A^*}(x) \leq \mu_A(x), & \mu_A(x) < \frac{1}{2}. \end{cases}$$

Цим вимогам задовольняє міра, аналогічна міру кількості ентропії Шеннона, якщо замість ймовірностей в ній використовувати функції приналежності.

Властивостям 1 ÷ 3 задовольняє досить широкий клас функцій. Аналізуючи недоліки такого роду конструктивних підходів до побудови заходів нечіткості, Р. Єгер ввів визначення, що впливає з формулювань операцій над нечіткими множинами. Він виходив з того , що єдиним корінним відзнакою алгебри нечітких множин від звичайної булевої алгебри є непустота перетину  $A$  і  $\bar{A}$  , тобто для нечітких множин

$$A \cap \bar{A} = A \cap (1 - A) = B, \tag{1.23}$$

де  $B$  – не порожня множина.

Очевидно, що чим ближче  $A$  до  $\bar{A}$ , тим більше  $B$ , і тим сильніше  $A$  відрізняється від чіткої множини. На підставі цього Р. Єгер запропонував для опису заходи чіткості нечіткої множини  $A$  вираз

$$D_p(A, \bar{A}) = \frac{1}{n} \left[ \sum_{i=1}^n |\mu_A(x_i) - \mu_{\bar{A}}(x_i)|^p \right]^{1/p}, \tag{1.24}$$

де  $p = 1, 2, \dots, \infty$  - ступінь нечіткості. На основі (1.24) конструюється вираз для визначення міри нечіткості

$$d_p(A, \bar{A}) = 1 - D_p(A, \bar{A}) / n^{1/p}.$$

Визначення (1.25) задовольняє всім сформульованим вище вимогам до міри нечіткості.

Для більш гнучкого використання ступеня нечіткості в роботі [341] введені спеціальні операції. Ступенем нечіткої множини  $A$  називається множина  $A^\varepsilon = \{\mu_A^\varepsilon(x_i)/x_i\}$ , де  $\varepsilon$  - деяке число. При  $\varepsilon = 2$  операція зводиться до зведення в квадрат - в операцію концентрації  $CON(A) = A^2$ , а при  $\varepsilon = 0.5$  - в операцію розтягнення  $DIL(A) = A^{0.5}$ . Оскільки ступінь приналежності до нечіткої множини  $\mu_A(x)$  - величина позитивна, що не перевершує одиниці, операція  $CON$  знижує ступінь нечіткості опису, в той час як  $DIL$  підвищує. Розгляд операції модифікування нечіткості відіграє важливу роль при формуванні згортки приватних критеріїв якості.

У конкретних прикладних проблемах можуть зустрітися ситуації, коли функції приналежності нечітких множин самі є нечіткими множинами. Така ситуація типова при завданні безлічі  $A$  на лінгвістичному рівні. Нехай, наприклад  $X = \{K1, K2, K3\}$  клієнти банку. Вони можуть оцінюватися як особи різного ступеня надійності з точки зору своєчасності повернення ними кредитів. Тоді нечітка множина «надійних клієнтів» можна уявити, наприклад, в такій формі:

$A = \{\text{«не дуже надійний клієнт»} / K1, \text{«надійний клієнт»} / K2, \text{«виключно надійний клієнт»} / K3\}$ .

У цій конструкції лінгвістичні визначення є стандартними функції приналежності, які за допомогою сучасних засобів теорії нечітких множин можуть бути замінені відповідними функціями приналежності, мають конкретну математичну форму. У результаті виходять математичні конструкції, що володіють фактично подвійний нечіткістю, оскільки їх функції приналежності самі є нечіткими множинами. Нечіткі множини з нечіткими функціями належності називають нечіткими множинами типу 2.



Нехай  $A$  - нечітка множина типу 2 на  $X$ . Тоді для кожного  $x \in X$  функція приналежності  $\mu_A(x)$  є нечітким безліччю в  $Y$  з функцією приналежності  $f_x(y)$ , тобто

$$\mu_A(x) = \int f_x(y) / y, \quad (1.26)$$

або для дискретного випадку

$$\mu_A(x) = \left\{ \frac{f_x(y_i)}{y_i} \right\}, \quad i=1, \dots, n. \quad (1.27)$$

$$\mu_{\bar{A}}(x) = \left\{ \frac{f_x(y_i)}{1 - y_i} \right\}, \quad i=1, \dots, n. \quad (1.28)$$

Якщо  $A$  і  $B$  - нечіткі множини типу 2 і

$$\mu_A(x) = \left\{ \frac{f_x(y_i)}{y_i} \right\}, \quad \mu_B(x) = \left\{ \frac{g_x(z_j)}{z_j} \right\}, \quad (1.29)$$

де  $i=1, \dots, n, j=1, \dots, m$ ,

то функція приналежності множини  $D = A \cup B$  має вигляд

$$\mu_D(x) = \left\{ \frac{f_x(y_i) \wedge g_x(z_j)}{(y_i \vee z_j)} \right\}, \quad i=1, \dots, n, j=1, \dots, m,$$

У тому випадку, якщо елементи  $x$  є нечіткими підмножинами-якого безлічі  $z$ , для дискретного випадку маємо

$$A = \left\{ \frac{\mu_A(x_i)}{x_i} \right\}, \quad x_i = \left\{ \frac{h_i(z_i)}{z_i} \right\}, \quad (1.31)$$

$$A = \left\{ \frac{\max_i [\mu_A(x_i) h_i(z_i)]}{z_i} \right\}, \quad i=1, \dots, n, j=1, \dots, m$$

Слідуючи Л Заде [344], на відміну від нечітких множин (1.26), (1.27) типу 2, отриману математичну конструкцію (1.31) будемо називати нечіткими множинами 2-го рівня. Таке уточнення термінології необхідно, оскільки в додатках обидва розглянутих виду подвійний нечіткості часто зустрічаються одночасно, що

призводить іноді до необгрунтованих операціями відомості одного до іншого і, відповідно, неадекватним висновків за результатами аналізу конкретних ситуацій.

З останнього виразу (1.31) випливає, що ступінь приналежності  $z_j$  до  $A$  визначається функцією

$$\mu_A(z_i) = \max_i [\mu_A(x_i)h_i(z_i)] \quad i = 1, \dots, n. \quad (1.32)$$

Відзначимо, що вирази ( 1.29 ) - ( 1.32 ) можна представити і в більш загальній формі, використовуючи замість операцій мінімуму, максимуму і множення, описані вище  $t$  – норми або  $s$  - норми. [ 238 ]. Однак у цій книзі для операцій з нечіткими множинами типу 2 та рівня 2 будуть використовуватися саме вирази ( 1.29 ) - ( 1.32 ), як ті, які найбільш відповідають особливостям вирішуваних завдань. Основи теорія нечітких множин типу 2 описані в роботах [ 47, 325, 264 ], де введені основні операції з ними. В роботі [ 325 ] відзначено, що, користуючись принципом узагальнення Л. А. Заде, можна побудувати нечіткі множини 3, 4, ..., N і т.д. типів.

#### 1.1.4. Нечіткі і можливі випадкові події

При вирішенні прикладних задач можуть зустрітися ситуації, коли одночасно присутні обидва види невизначеності - нечіткість і випадковість. У практичних додатках теорія нечітких множин добре працює спільно з методами теорії ймовірностей і математичної статистики. В роботі [47] введено важливе для практики поняття математичного очікування функції за наявності обох типів невизначеності.

Нехай  $A$ -нечітке підмножина множини  $X = \{x_i\}$ ,  $i = 1, \dots, n$ , зі ступенем приналежності  $\mu_A(x)$  і нехай  $P_i$  - ймовірність події  $x_i$ . Тоді математичне сподівання деякої функції  $f$ , визначеної на  $X$ , для дискретного випадку визначається

конструкцією:

$$C_p(f, A) = \sum_{i=1}^n f(x_i) \mu_A(x_i) P_i. \quad (1.33)$$

Запропоноване визначення (1.33) забезпечує органічний зв'язок теорії нечітких множин з традиційними підходами, заснованими на теоретико-ймовірнісній інтерпретації невизначеностей, і дозволяє більш детально описувати існуючі невизначеності у вихідних даних, причому, при необхідності в поєднанні з методами традиційної статистики.

В даний час прийнято виділяти чотири базових типи ймовірностей . Так, на основі аналізу відносин між суб'єктом , знанням і об'єктом визначені наступні види ймовірностей : класична , частотна , логічна і суб'єктивна . При цьому показано , що кожна з них задовольняє вихідним аксіомам теорії ймовірностей А. Н. Колмогорова.

Проте в загальному випадку на практиці постійно зустрічаються ситуації , коли жодна з цих інтерпретацій ймовірності не забезпечує адекватного опису прогнозу можливості тієї чи іншої події . Наочний приклад наведено в, де аналізується ситуація очікування автобуса групою людей. Автобус прибуває повний і кількість пасажирів , які можуть до нього потрапити , залежить від рішення водія. На цьому прикладі чітко проявляється різниця між можливістю ( ступенем легкості ) того , що в автобус увійдуть 1 , 2 , 3 , ... , і т.д. пасажирів і ймовірністю того , що водій їх візьме.

Підстави формалізованої теорії можливостей закладені Л.А.Заде в роботах [ 340 , 342 , 339 ] , де на відміну від імовірності можливість пропонується розглядати як ступінь легкості , з якою розглянута змінна може приймати певні значення. В якості основи теорії можливостей використаний математичний апарат теорії нечітких множин , аналогічно тому , як основою теорії ймовірності є теорія міри .

У роботі [ 339 ] розподіл можливостей визначено таким чином: нехай  $X$  - змінна, що приймає значення з безлічі  $U = \{ u_1 , u_2 , \dots \}$  . Тоді розподіл можливості  $X$  приймає значення з  $U$  і описується функцією  $\pi_x : U \rightarrow [ 0,1 ]$  , де  $\pi_x (X) = \text{Poss} : [X = u ]$  . Позначення  $\text{Poss}$  використовується для відображення конкретного значення можливості при фіксованому  $u \in U$ .

У розглянутому вище прикладі очікування автобуса  $X$  - кількість пасажирів, які мають можливість увійти в автобус;  $u$  - елементи множини  $U = \{ 1,2,3 \dots \}$ . Тоді  $\pi_x (2)$  - можливість, що в автобус увійдуть 2 людини,  $\pi_x (3)$  - 3 людини і т.д.

Перейдемо до розширення теорії можливостей для випадку оцінки можливості (легкості) реалізації множин. Нехай  $A$  - підмножина в  $U$ , тоді  $\pi_A (A) = \text{Poss} [u \in A]$ . У теорії можливостей постулюється, що

$$\pi_A (A) = \sup_{x \in A} [\pi_X (X)]. \tag{1.34}$$

Якщо  $A$  нечітка підмножина в  $U$ , то

$$\pi_{\hat{A}} (\hat{A}) = \sup_{x \in \hat{A}} [\pi_X (X) \wedge \mu_{\hat{A}} (X)],$$

де  $\mu_{\hat{A}} (X)$  - функція приналежності підмножини  $\hat{A}$  .

Доведено , що у випадках наявності лише непрямой інформації про розподілі можливостей  $\pi_x (X)$  , тобто , коли можна вказати лише , що  $\pi_x (X_1)$  більш можливо , ніж  $\pi_x (X_2)$  , без кількісної оцінки цього відносини ( linear ordering [ 204 ]), виразу ( 1.27 ) , ( 1.28 ) є єдино допустимими . У тих випадках , коли відносини між можливостями можна охарактеризувати конкретними числами ( cardinal ordering ), розподіл можливостей вироджується в імовірнісний розподіл [ 339 ] . Тоді з  $\pi_x (X)$  слід оперувати як із звичайною щільністю ймовірності . Звідси

автор

роботи [ 318 ] робить висновок , що теорію можливостей можна розглядати як узагальнення теорії ймовірностей.

Поява формалізованої теорії можливостей породило масу питань методологічного плану, зокрема , проблему єдиної трактування самої категорії

можливості [ 212 , 290 ] . Послідовний аналіз поняття можливості з загально філософських позицій з урахуванням результатів теорії проведено в роботі [ 86 ] .

Засновник теорії можливостей Л.А. Заде тісно пов'язував її з лінгвістичної

інтерпретацією нечітких множин. Водночас , як зазначається в [ 322 ] , ідея опису

істинності лінгвістичних змінних , сформульованих на природній мові з допомогою неоднозначною логіки , була висловлена ще Л.Вітгенштейном [ 317 ] .

Однак , суворі методи оперування з змінними лінгвістичної природи з'явилися лише з розвитком теорії нечітких множин. Лінгвістична інтерпретація

відіграє велику роль в теорії нечітких множин та теорії можливостей , дозволяючи

оперувати з об'єктами якісної природи , в тому числі з висловлюваннями на природній мові.

Головним достоїнством теорії можливостей з точки зору завдань оптимізації та

ідентифікації є наявність формалізованого математичного апарату , що дозволяє працювати в умовах невизначеності при відсутності інформації , достатньою для застосування теоретико - імовірнісних методів .

Проведений аналіз дозволяє зробити висновок про те , що теорія нечітких множин

знаходиться на стадії інтенсивного розвитку , не є концептуально замкнутої , допускає неоднозначність ряду математичних операцій всередині себе.

Неповнота існуючої теорії нечітких множин породжує появу цілого ряду альтернативних варіантів теорії [ 317 , 231 , 269 , 224 , 315 , 314 , 276 , 311 , 267 ] , вільних від тих чи інших недоліків аксіоматизації , запропонованої Л.А.Заде .

У ряді випадків [ 317 , 231 , 269 ] запропоновані варіанти виглядають вельми перспективними , розширюють концепцію нечіткості і вимагають більш детального вивчення та випробувань на практиці. В інших випадках , як визнає автор [ 311 ] , нові теорії нечітких множин мають лише академічний інтерес , будучи в значній ступеня практично неконструктивними .

Підводячи підсумок , слід підкреслити , що в теорії прийняття рішень , багатокритеріальної оптимізації , автоматичному управлінні і в багатьох інших прикладних задачах в даний час широко використовується варіант теорії нечітких множин , запропонований Л.А.Заде , що добре зарекомендував себе на практиці.

### **Нечіткість, можливість, ймовірність: методологічні проблеми взаємозв'язку і розмежування**

В даний час більшість дослідників схильні розглядати нечіткість і випадковість , як два якісно різних види невизначеності. З одного боку , нечіткість стосується величин і відносин , межі яких неточно визначені , тобто коли їх не можна адекватно визначити ( описати ) , використовуючи поняття звичайної множини, оскільки перехід від приналежності до неналежності безлічі вже не має стрибкоподібного характеру . З іншого боку , випадковість стосується ситуацій , в яких подія точно визначено, а невизначеним є ймовірність його настання.

Однак , легко уявити собі , що в щоденному житті зустрічається багато ситуацій , в яких нечіткість і випадковість виступають разом. Наприклад , якщо ми запитуємо , яка ймовірність того , що завтра буде гарна погода, або про те, яка ймовірність того , що в майбутньому році буде висока інфляція , то в обох випадках маємо неточно визначені ( нечіткі ) події , відповідно гарну погоду і високу

інфляцію. Очевидно, що для того, щоб можна було формально визначити ймовірності таких неточно певних подій, необхідно ввести поняття нечіткого події і ймовірності нечіткого події.

Перший підхід до вирішення цієї проблеми запропонував Л. Заде [343]. Його відправною точкою є поняття нечіткого події, яка є нечітка множина  $A$  в  $X = \{x\} = \{x_1, \dots, x_n\}$  з відповідною функцією приналежності. При цьому передбачається, що ймовірності (чітких) елементарних подій  $x_1, \dots, x_n \in X$  відомі і рівні, відповідно,  $p(x_1), \dots, p(x_n) \in [0, 1]$ , причому  $p(x_1) + \dots + p(x_n) = 1$ .

Серед найбільш важливих понять і властивостей, що описують нечіткі події, наведемо наступні:

Нечіткі події  $A$  і  $B$  в  $X$  незалежні тоді і тільки тоді, коли

$$p(AB) = p(A)p(B)$$

.

Умовну ймовірність нечіткого події  $A$  в  $X$  за умови настання нечіткого події  $B$  в  $X$  будемо позначати  $p(A|B)$  і визначимо як

$$p(A|B) = \frac{p(AB)}{p(B)}, p(B) > 0$$

Якщо нечіткі події  $A$  і  $B$  незалежні, то

$$p(A|B) = p(A).$$

Відзначимо, що обидва наведених поняття аналогічні їх чітким аналогам.

Чітка ймовірність нечіткого події  $A$  в  $X = \{x\} = \{x_1, \dots, x_n\}$  позначається  $p(A)$  і визначається як

$$p(A) = \sum_{i=1}^n \mu_A(x_i) p(x_i) \tag{1.36}$$

тобто як очікуване значення функції приналежності нечіткій множині  $A$ , з функцією

приналежності  $\mu_A$  (x).

Приклад. Нехай  $X = \{1, 2, \dots, 5\}$ ,  $p(x_1) = 0.1$ ,  $p(x_2) = 0.1$ ,  $p(x_3) = 0.1$ ,  $p(x_4) = 0.3$ ,  $p(x_5) = 0.4$  і  $A = 0.1 / 2 + 0.5 / 3 + 0.7 / 4 + 0.9 / 5$ .

Ноді

$$p(A) = 0.1 \times 0.1 + 0.1 \times 0.5 + 0.3 \times 0.7 + 0.4 \times 0.9 = 0.73$$

Головними властивостями вищеописаної ймовірності настання нечіткого події в сенсі Л. Заде [337] є наступні:

$$p(\emptyset) = 0$$

$$p(\neg A) = 1 - p(A)$$

$$p(A + B) = p(A) + p(B) - p(A \cap B)$$

$$p\left(\sum_{i=1}^r A_i\right) = \sum_{i=1}^r p(A_i) - \sum_{j=1}^r \sum_{k=1, k < j}^r p(A_j \cap A_k) + \sum_{j=1}^r \sum_{k=1, k < j}^r \sum_{l=1, l < k}^r p(A_j \cap A_k \cap A_l) + \dots + (-1)^{r+1} p(A_1 \cap A_2 \cap \dots \cap A_r)$$

Зауважимо, що в представленому підході Л. Заде, подія є нечітким, проте його ймовірність - чітка, і вона представлена дійсним числом з інтервалу  $[0, 1]$ . У певному сенсі це є протиріччям, і тому пізніше були запропоновані інші визначення нечіткої ймовірності нечіткої події. Далі наведені деякі з них.

Р. Єгоровим [320] нечітка подія визначена як нечітке множина  $A$  в  $X = \{x\} = \{x_1, \dots, x_n\}$  з функцією приналежності, вимірної в сенсі булевої логіки. при цьому передбачається, що ймовірності елементарних подій  $p(x_1), \dots, p(x_n) \in [0, 1]$  відомі, причому  $p(x_1) + \dots + p(x_n) = 1$ .

Нечітка ймовірність нечіткого події  $A$  в  $X = \{x\} = \{x_1, \dots, x_n\}$  позначається

$P(A)$  і визначається як наступне нечітке безліч в  $[0, 1]$



$$P(A) = \sum_{\alpha \in (0,1]} \alpha / p(A_\alpha) \quad (1.37)$$

Вираз (1.37), очевидно, є результатом застосування принципу розширення. У термінах функції приналежності визначення (1.37) має вигляд:

$$\mu_{P(A)}[p(A_\alpha)] = \alpha \quad \text{для кожного } \alpha \in (0,1],$$

де  $A_\alpha \in \alpha$ -рівень нечіткої множини  $A$ .

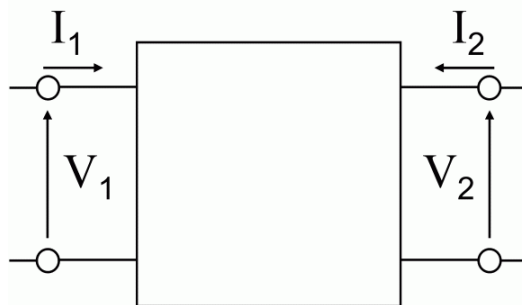
Приклад. Нехай  $X = \{1, 2, 3, 4\}$ ,  $p(1) = 0.1$ ,  $p(2) = 0.3$ ,  $p(3) = 0.5$ ,  $p(4) = 0.1$ , і  $A = 0.2 / 1 + 0.5 / 2 + 0.8 / 3 + 1/4$ .  
Тоді  $\alpha \in \{0.3, 0.5, 0.8, 1\}$  і  
 $A_1 = \{4\}$   $A_{0.8} = \{3, 4\}$   $A_{0.5} = \{2, 3, 4\}$   $A_{0.3} = \{1, 2, 3, 4\}$ .  
Або  $p(A_1) = 0.1$   $p(A_{0.8}) = 0.6$   $p(A_{0.5}) = 0.9$   $p(A_{0.3}) = 1$ .

Слід зазначити, що визначення Р. Егера також не дозволяє всіх суперечностей і проблем, так як воно не виконує вищенаведених властивостей ймовірності, яким задовольняє визначення Л. Заде. Відзначимо, що на практиці підхід Егера вже знайшов своє додаток в нечіткому управлінні на основі нечіткої логіки. Запропоновані та інші визначення нечіткого події, що спираються на те, що нечіткість і випадковість, будучи якісно різними видами невизначеності, не є взаємовиключними, а навпаки, взаємопов'язані і доповнюють один одного при аналізі одних і тих же подій.

## Розділ 7. ПРИКЛАДНІ АСПЕКТИ

### Чотириполіусник, як приклад моделі чорного ящика

Чотириполіусник - різновид багатополіусників, що має чотири точки підключення. При аналізі електричних ланцюгів дуже часто буває зручним виділити фрагмент ланцюга, що має дві пари затискачів. Оскільки електричні (електронні) ланцюги дуже часто пов'язані з передачею енергії або обробкою і перетворенням інформації, одну пару затискачів зазвичай називають «вхідними», а другу - «вихідними». На вхідні затискачі подається вихідний сигнал, з вихідних знімається перетворений. Такими чотириполіусниками є, наприклад, трансформатори, підсилювачі, фільтри, стабілізатори напруги, телефонні лінії, лінії електропередачі і т. д. Однак математична теорія чотириполіусників не передбачає ніяких зумовлених потоків енергії/інформації в ланцюгах, тому назви «вхідні» і «вихідні» є даниною традиції і з цією обмовкою будуть використовуватись далі.



Стан вхідних і вихідних затискачів визначаються чотирма параметрами: напругою і струмом у вхідних ( $U_1, I_1$ ) і вихідних ( $U_2, I_2$ ) ланцюгах. У цій системі параметрів лінійний чотириполіусник описується системою з двох лінійних рівнянь, причому два з чотирьох параметрів стану є вихідними, а два інші - обумовленими. Для нелінійних чотириполіусників залежність може носити більш складний характер. Наприклад, вихідні параметри через вхідні можна виразити системою:

$$\begin{cases} U_2 = b_{11}U_1 + b_{12}I_1 \\ I_2 = b_{21}U_1 + b_{22}I_1 \end{cases}; \begin{pmatrix} U_2 \\ I_2 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} U_1 \\ I_1 \end{pmatrix}$$

Оскільки чотириполіусник має чотири параметри стану, очевидно, що маєсья шість систем рівнянь, що виражають різні пари параметрів через два інших. Коефіцієнти цих шести систем рівнянь отримали традиційне найменування А-, В-, G-,

H-, Y- і Z-параметрів. Системи рівнянь і еквівалентні схеми чотириполусників при використанні кожного типу параметрів показані в таблиці нижче

	Система рівнянь	Еквівалентна схема	Вимірювання параметрів
<b>G</b>	$\begin{pmatrix} I_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} U_1 \\ I_2 \end{pmatrix}$		$g_{11} = \frac{I_1}{U_1} \Big _{I_2=0}; g_{12} = \frac{I_1}{I_2} \Big _{U_1=0}$ $g_{21} = \frac{U_2}{U_1} \Big _{I_2=0}; g_{22} = \frac{U_2}{I_2} \Big _{U_1=0}$
<b>H</b>	$\begin{pmatrix} U_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} I_1 \\ U_2 \end{pmatrix}$		$h_{11} = \frac{U_1}{I_1} \Big _{U_2=0}; h_{12} = \frac{U_1}{U_2} \Big _{I_1=0}$ $h_{21} = \frac{I_2}{I_1} \Big _{U_2=0}; h_{22} = \frac{I_2}{U_2} \Big _{I_1=0}$
<b>Y</b>	$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$		$y_{11} = \frac{I_1}{U_1} \Big _{U_2=0}; y_{12} = \frac{I_1}{U_2} \Big _{U_1=0}$ $y_{21} = \frac{I_2}{U_1} \Big _{U_2=0}; y_{22} = \frac{I_2}{U_2} \Big _{U_1=0}$
<b>Z</b>	$\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}$		$z_{11} = \frac{U_1}{I_1} \Big _{I_2=0}; z_{12} = \frac{U_1}{I_2} \Big _{I_1=0}$ $z_{21} = \frac{U_2}{I_1} \Big _{I_2=0}; z_{22} = \frac{U_2}{I_2} \Big _{I_1=0}$
<b>A</b>	$\begin{pmatrix} U_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} U_2 \\ I_2 \end{pmatrix}$		$a_{11} = \frac{U_1}{U_2} \Big _{I_2=0}; a_{12} = \frac{U_1}{I_2} \Big _{U_2=0}$ $a_{21} = \frac{I_1}{U_2} \Big _{I_2=0}; a_{22} = \frac{I_1}{I_2} \Big _{U_2=0}$
<b>B</b>	$\begin{pmatrix} U_2 \\ I_2 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} U_1 \\ I_1 \end{pmatrix}$		$b_{11} = \frac{U_2}{U_1} \Big _{I_1=0}; b_{12} = \frac{U_2}{I_1} \Big _{U_1=0}$ $b_{21} = \frac{I_2}{U_1} \Big _{I_1=0}; b_{22} = \frac{I_2}{I_1} \Big _{U_1=0}$

Як правило, моделі () доцільно використовувати для лінійних систем. При моделюванні високочастотних елементів в практику впроваджують удосконалені моделі, наприклад чотирохполюсники з X-параметрами [Root D.E., Horn J., Betts L. et al. X-параметры: новый принцип измерений, моделирования и разработки нелинейных ВЧ- и СВЧ-компонентов. – Контрольно-измерительные приборы и системы, 2009, №2, с.20.]

### «Фрактальні антени». Принципово новий тип "ламаних" антен

Перші публікації з електродинаміки фрактальних структур відносяться до 80-х років минулого століття. У публікаціях з історії фрактальних антен зазвичай згадується робота вчених Університету штату Пенсільванія Я.Кіма і Д.Джаггарда (Y.Kim and DLJaggard). Першість у теоретичних дослідженнях можливості застосування фрактальних форм для формування многополосних за частотою антен приписують вченому Технологічного університету Каталонії К.Пуенте (C.Puente). Початок же практичного застосування фрактальних антен в 1995 році поклав, як прийнято вважати, американський інженер Натан Коен (N.Cohen). Щоб обійти заборону бостонських влади встановлювати на будинках зовнішні антени, він замаскував антену своєї радіостанції під декоративну фігуру, виконану на основі фрактальної ламаної, описаної шведським математиком Хельге фон Кохом (Helge von Koch) в 1904 році.

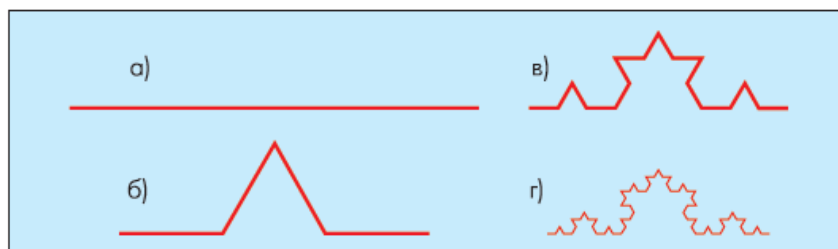
Опубліковані Коеном результати досліджень характеристик новою антеною конструкції привернули увагу фахівців. Завдяки зусиллям багатьох дослідників сьогодні теорія фрактальних антен перетворилася на самостійний, досить розвинений апарат синтезу та аналізу ЕМА.

Фрактальні антени - відносно новий клас електрично малих антен (ЕМА), принципово відрізняється своєю геометрією від відомих рішень. По суті, традиційна еволюція антен базувалася на евклідовій геометрії, що оперує об'єктами цілочислової розмірності (лінія, коло, еліпс, параболоїд і т. д.). Головна відмінність фрактальних геометричних форм - їх дробова розмірність, що зовні проявляється в рекурсивному повторенні в зростаючому або зменшуваному масштабах вихідних детермінованих або випадкових шаблонів. Фрактальні технології набули поширення при формуванні засобів фільтрації сигналів, синтезі тривимірних комп'ютерних моделей природних ландшафтів, стисненні зображень. Цілком природно, що фрактальна "мода" не обійшла стороною і теорію антен. Тим більше, що прообразом сучасних фрактальних технологій в антенній техніці

з'явилися запропоновані в середині 60-х років минулого століття логоперіодичні і спіральні конструкції. Правда, в строгому математичному математичному сенсі такі конструкції на момент розробки не мали відношення до фрактальної геометрії, будучи, по суті, лише фракталами першого роду [1]. Зараз дослідники, в основному методом проб і помилок, намагаються використовувати відомі в геометрії фрактали в антенних рішеннях. В результаті імітаційного моделювання та експериментів встановлено, що фрактальні антени дозволяють отримати практично той же коефіцієнт посилення, що й звичайні, але при менших габаритах, що важливо для мобільних додатків. Розглянемо результати, отримані в галузі створення фрактальних антен самих різних типів.

### **Заповнюючі простір криві. Випромінювачі на основі фрактала Коха**

Першою конструкцією фрактальної антени з найбільш повно вивченими електромагнітними і спрямованими властивостями стала антена на основі префрактальної кривої Коха. При побудові лінії Коха вихідний відрізок довжиною  $z$ , іменованій ініціатором фрактала, ділиться на три рівні частини. Центральна ділянка замінюють рівностороннім трикутником зі стороною  $z / 3$ . У результаті утворюється ламана, що складається з чотирьох ланок завдовжки  $z / 3$  кожен (рис.1). Цей процес повторюється для кожного окремого сегмента ламаній лінії: у другій ітерації на відрізках  $z / 3$  будуються трикутники зі сторонами  $z / 9$ , на них - трикутники зі сторонами  $z/27$  (третя ітерація) і т.д. Гранична крива і є крива Коха.



**рис.1. Побудова кривої Коха: а) перша, б) друга, в) третя і г) четверта ітерації**

Кожен крок синтезу збільшує довжину результуючої кривої у відповідності з виразом [2]:

$$L = z \left(\frac{4}{3}\right)^n \quad (1)$$

де  $n$  - число ітерацій,  $z$  - висота утворює шаблону (довжина вихідного відрізка).

Цей ефект мініатюризації антен є істотним лише при п'яти-шести перших ітераціях фрактала.

Строго кажучи, в антенних рішеннях використовуються не справжні фрактали, а лише кілька перших їх ітераційних форм, які отримали в

геометрії назву кривих, що заповнюють простір (Space-Filling Curves, SFC) [8, 9] або площину (Plane-Filling Curves, PFC) [10]. Рідше використовується термін "префрактали". Всі ці поняття стосовно до антенних конструкціям можуть вживатися як синоніми. Така історично сформована термінологія теорії фрактальних антен, хоча вона і не відповідає прийнятим математичним визначенням.

SFC можуть застосовуватися як шаблони для виготовлення монополів і плечей диполів, формування топології друкованих антен, частотно-селективних поверхонь (Frequency Selection Surfaces, FSS) або обичайок дзеркальних рефлекторів, побудови контурів рамкових антен і профілів апертури рупорів, а також фрезерування пазів в щілинних антенах. В англійській літературі відповідні антени нерідко називають "space-filling antenna" (SFA) (антени, що заповнюють простір).

У разі дротяних антен самоперетинів SFC допускається тільки в початковому (або кінцевому) пункті [9]. Інакше кажучи, фрактальна лінія може мати вигляд замкнутого контуру, але жодна з її частин не може бути замкнутим фрагментом. Відсутність точок самоконтакта в SFC-об'єктах дозволяє говорити про них як про "самоунікаючих" кривих [11]. Звідси, до речі, походить ще одна назва цих ламаних ліній - FASS-криві (space-Filling self-Avoidance Simplicity Similarity - самоунікаючі криві подібних сегментів, що заповнюють простір) [10].

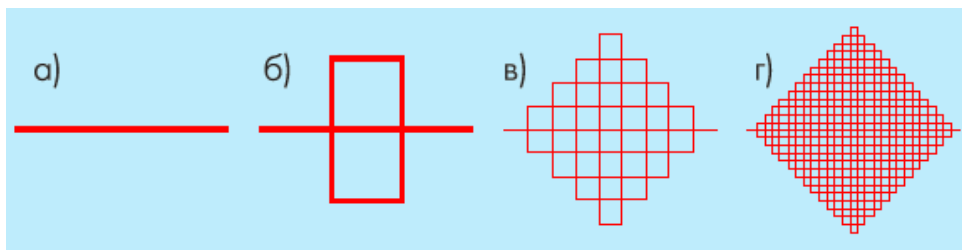
Існує й інше обмеження всіх типів фрактальних антен [9]: сегменти використовуваних в них SFC-ліній повинні бути коротше однієї десятої робочої довжини хвилі антени у вільному просторі. При цьому бажано, щоб загальна кількість пов'язаних SFC-сегментів в антенних топологіях перевищувало 10.

Експериментальні дані, отримані фахівцями компанії CushCraft: для кривої Коха, чотирьох ітерацій меандра і спіральної антени, дозволяють зіставити електричні властивості антени Коха з іншими випромінювачами з періодичною структурою [12]. Всі зіставлені випромінювачі володіли багаточастотними властивостями, що проявилось в наявності періодичних резонансів на графіках імпедансов. Однак для багатодіапазонні додатків найбільше придатний фрактал Коха, у якого із зростанням частоти пікові значення реактивних і активних опорів зменшуються, тоді як у меандра і спіралі вони зростають.

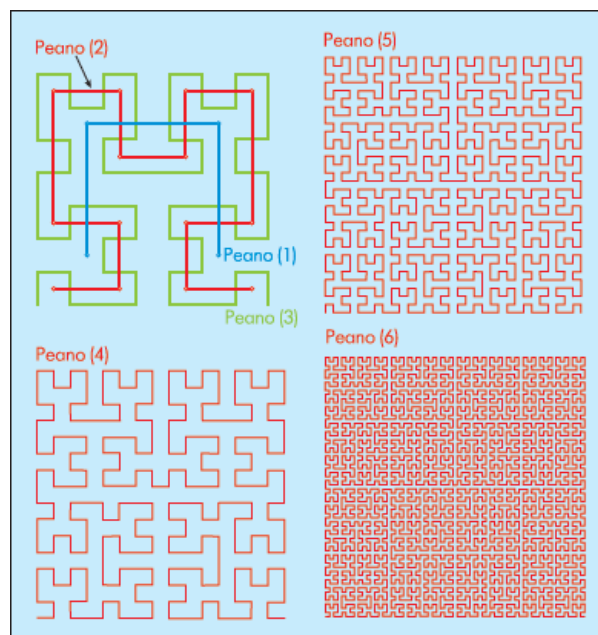
Загалом слід зазначити, що теоретично уявити механізм взаємодії фрактальної прийомної антени і падаючих на неї електромагнітних хвиль складно через відсутність аналітичного опису хвильових процесів в провіднику зі складною топологією. У такій ситуації основні параметри фрактальних антен доцільно визначати шляхом математичного моделювання. Чисельне дослідження електромагнітних процесів, що протікають у фрактальних антенах і при їх взаємодії з

предметами навколишнього середовища, присвячено досить багато робіт. Їх детальний огляд і аналіз виходить за рамки цієї статті. Загальний недолік всіх відомих публікацій за результатами досліджень фрактальних антен - відсутність вказівок на статистичну обробку результатів експериментів. Зокрема, в них не наводяться відомості про недовірчі інтервали для вимірних параметрів, що не дозволяє судити про точність отриманих у результаті емпіричних співвідношень. В цілому ж, статистична теорія фрактальних антен при розрахунку їх чисельними методами поки ще чекає своїх розробників.

Приклад побудови першої самоподобної фрактальної кривої продемонстрував в 1890 році італійський математик Джузеппе Пеано (РСАП) [8, 13, 14]. Запропонована ним лінія в межі повністю заповнює квадрат, оббігаючи всі його точки (рис.2). Надалі були знайдені й інші подібні об'єкти, що одержали по імені першовідкривача їх сімейства узагальнююча назва "криві Пеано". Правда, внаслідок чисто аналітичного опису кривої, запропонованого Пеано [13], виникла деяка плутанина в класифікації БРС-ліній. Насправді найменування "криві Пеано" варто було б давати лише оригінальним кривим, побудова яких відповідає аналітиці, опублікованій Пеано (рис.3) [13].



**рис.2. Ітерації кривої Пеано: а) вихідна лінія, б) перше, в) друге і г) третій ітерації**

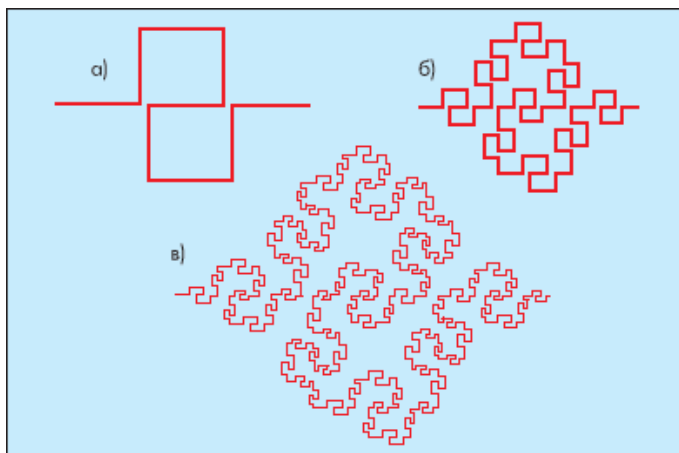


**рис.3. Ітерації ламаної, запропонованої Гильбертом в 1891 році [8]. нерідко трактується як рекурсивна крива Пеано [15]**

Тому для конкретизації розглянутих об'єктів антеною техніки при описі тієї чи іншої форми фрактальної антени слід, по можливості, згадувати та імена авторів, що запропонували відповідну модифікацію БРС. Це тим більше важливо, що згідно з підрахунками [16], число відомих різновидів БРС наближається до трьох сотень, причому ця цифра не є граничною.

Слід зазначити, що крива Пеано (див. рис.2) в початковому вигляді цілком придатна для виготовлення щілин в стінках хвилеводу, друкованих та інших апертурних фрактальних антен, але не прийнятна для побудови дротяної антени, оскільки має стикаються ділянки. Тому фахівцями компанії Fractus була запропонована її модифікація [9], що отримала назву "Peanodec" (рис. 4).

Представлена на рис.1 антена за фракталу Коха - лише один з варіантів, реалізований при використанні одностороннього ініціюючого трикутника рекурсії, тобто кут  $\theta$  при його підставі (indentation angle або "кут поглиблення") дорівнює  $60^\circ$ .



**рис.4. Варіант модифікації кривої Пеано ("Peanodec"):  
а) першого, б) другого в) третьої ітерації [9]**

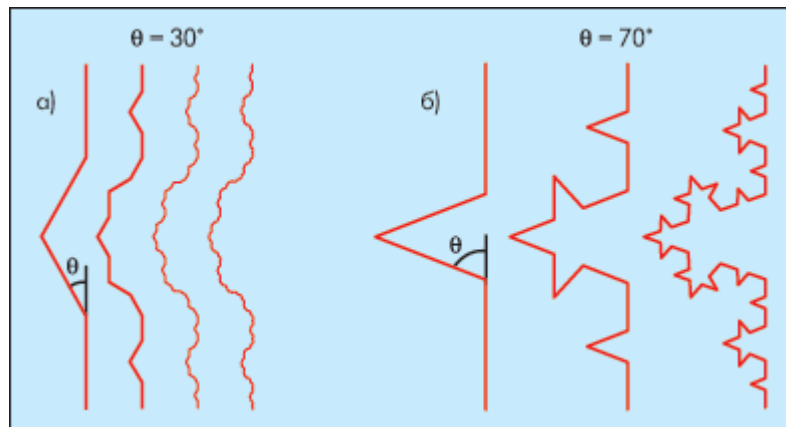
Такий варіант фрактала Коха прийнято називати стандартним. Цілком природно задатися питанням, чи можна використовувати модифікації фрактала з іншими значеннями цього кута. Ствердну і ґрунтовну відповідь на дане питання міститься в роботі вченого Пенсільванського університету К.Дж.Віной (KJVinoy) [10]. Виною запропонував розглядати кут при основі ініціюючого трикутника як параметр, що характеризує антенну конструкцію. Змінюючи цей кут, можна одержувати аналогічні рекурсивні криві різної розмірності (рис. 5). Криві зберігають властивість самоподібності, однак результуюча довжина лінії може бути різною, що впливає на



характеристики антени. Виною першим досліджував кореляцію між властивостями антени і розмірністю узагальненого фрактала Коха  $D$ , яка визначається в загальному випадку залежністю

$$D = \frac{\log 4}{\log (2(1 + \cos\theta))}$$

Було показано, що в міру збільшення кута  $\theta$  розмірність фрактала також збільшується, і при  $\theta \rightarrow 90^\circ$  наближається до 2. Слід зазначити, що використане в теорії фрактальних антен поняття розмірності суперечить поняттям, прийнятим в геометрії, де ця міра застосовна тільки до нескінченно рекурсивних об'єктів.



**рис.5. Побудова кривої Коха з кутом  $\theta$  а)  $30^\circ$  і б)  $70^\circ$  при основі трикутника в генераторі фрактала.**

Із збільшенням розмірності нелінійно зростає і загальна довжина ламаної лінії, що визначається співвідношенням [17]:

$$L_{n,\theta} = \left( \frac{2}{1 + \cos\theta} \right)^n L_0 \quad (2)$$

де  $L_0$  - довжина лінійного диполя, відстань між кінцями якого та ж, що і у ламаної Коха,  $n$  - номер ітерації. Перехід від  $\theta = 60^\circ$  до  $\theta = 80^\circ$  на шостій ітерації дозволяє збільшити загальну довжину префрактала більш ніж у чотири рази.

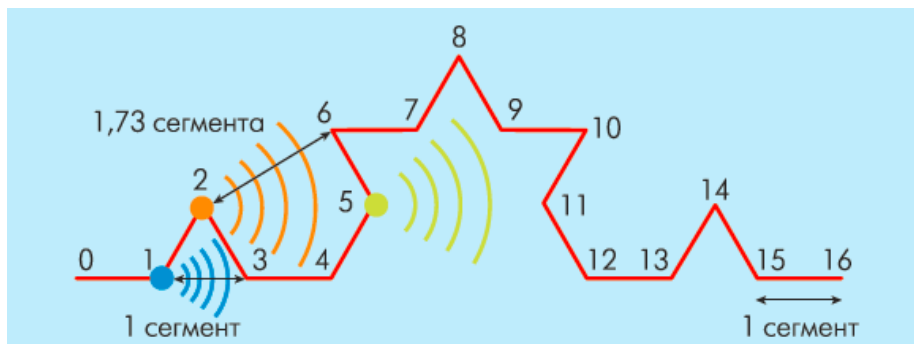
Як і слід було очікувати, між рекурсивною розмірністю і такими властивостями антени, як первинна резонансна частота, внутрішній опір на резонансі і багатодіапазонній характеристиці, існує прямий зв'язок. На основі комп'ютерних розрахунків Виною [10] отримав залежність першої резонансної частоти диполя Коха  $f_k$  від розмірності префрактала  $D$ , номера ітерації  $n$  і резонансної частоти прямолінійного диполя  $f_D$  тієї ж висоти, що й ламана Коха (по крайніх точках):

$$f_k = f_D \cdot \left[ 1 - \exp\left(\frac{n-1}{n}\right) \frac{\ln D}{D} \right] \quad (3)$$

При розрахунках він використовував модель диполя з радіусом провідника 0,1 мм, сегментованого на прямолінійні відрізки довжиною 0,5 мм. Загальна висота диполя (інтервал між його кінцями) дорівнювала 10 см. З розрахунків Виною випливає, що для реалізуються на практиці розглянутих параметрів диполя Коха резонансну частоту можна знизити до 30% від значення резонансної частоти 10-см прямолінійного диполя [10].

Незважаючи на величезну роботу, виконану Виною, залишається неясним, чи збережеться в силі залежність (3) при інших значеннях діаметра провідника. У всякому разі, Виною стверджує, що нормування резонансної частоти диполя Коха щодо частоти прямолінійного диполя дозволяє нівелювати вплив радіуса провідника, яке, як він сподівається, однаково для обох диполів. Однак не виключено, що при ітераціях більшого порядку через взаємного впливу близько розташованих сегментів фрактала можуть виникнути паразитні електродинамічні ефекти, що звужують область застосування виразу (3). Це підтверджує і робота вчених Політехнічного університету Каталонії [17], які пояснили уповільнення ефекту зниження резонансної частоти із збільшенням числа ітерацій префрактала "просочуванням" електромагнітних хвиль між близько розташованими зламами лінії. За їхньою гіпотезою, поверхнева хвиля сегментів префрактала, розміри яких багато менше довжини хвилі, "відривається" від конфігурації провідника антени (рис.6) і досягає точки підключення фідера швидше, ніж якби вона поширювалася, суворо дотримуючись геометрії електропровідного матеріалу.

Одна з серйозних проблем, що обмежує практичне використання стандартного диполя Коха, - зменшення його внутрішнього опору зі збільшенням числа ітерацій. Згідно Виною, цей недолік можна подолати, варіюючи значення кута  $\theta$ .



**рис.6.Ефект "просочування" електромагнітної хвилі**

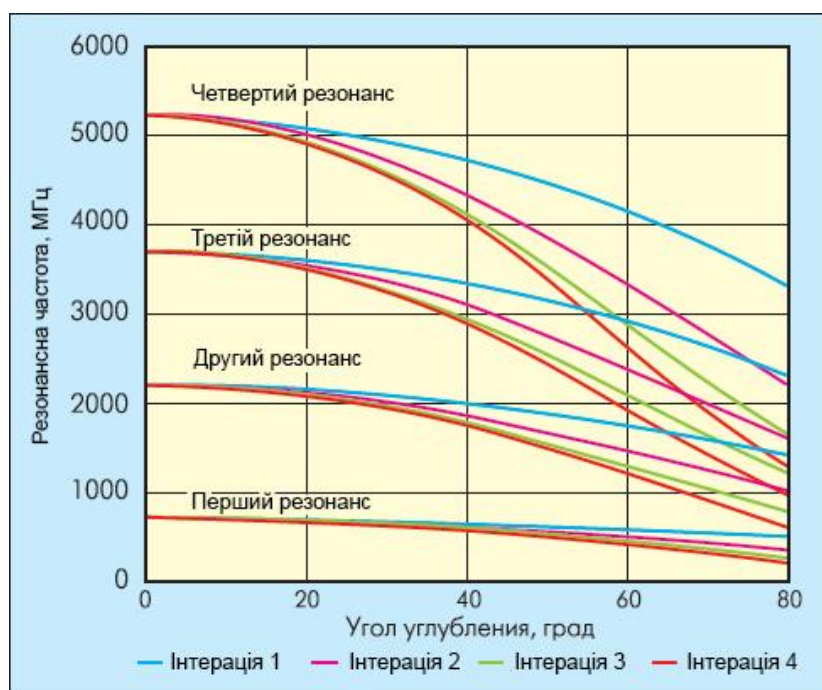
У загальному випадку для внутрішнього опору диполя Коха на першому резонансній частоті справедливо наближене співвідношення, наведене в [10]:

$$R_k = R_0 \cdot \left[ 1 - (1 + 0,9 \ln n) \frac{\ln D}{D} \right]^2 \quad (4)$$

де  $R_0$ - внутрішній опір лінійного диполя ( $D = 1$ ), яке в розглянутому випадку дорівнює 72 Ом.

Вирази (3) і (4) можна використовувати для визначення геометричних параметрів антени з необхідними значеннями резонансної частоти і внутрішнього опору.

Багатосмугові властивості диполя Коха також чутливі до значення кута  $\theta$ . Із збільшенням  $\theta$  номінали резонансних частот зближуються, а, отже, зростає їх число в заданому спектральному діапазоні (рис.7). При цьому, чим вище номер ітерації, тим сильніше це зближення.



**Рис.7. Ефект звуження інтервалу між резонансними частотами**

В Університеті штату Пенсільванія був досліджений ще один важливий аспект диполя Коха - вплив несиметричності його заживлення на ступінь наближення внутрішнього опору антени до 50 Ом [10]. У лінійних диполях точка харчування часто розташовується асиметрично [18]. Цей же підхід може бути використаний і для фрактальної антени у вигляді кривої Коха, внутрішній опір якої менше нормативних значень. Так, у третій ітерації внутрішній опір

стандартного диполя Коха ( $\theta = 60^\circ$ ) без урахування втрат при підключенні фідера в центрі становить 28 Ом. При фідера до одного з кінців антени можна отримати опір 50 Ом.

Всі розглянуті досі конфігурації ламаної Коха синтезувалися рекурсивно. Однак, згідно Виною, якщо порушити це правило, зокрема, задаючи різні кути  $\theta$  на кожній новій ітерації, властивості антени можна змінювати з більшою гнучкістю [10]. Для збереження подібності доцільно вибрати регулярну схему зміни кута  $\theta$ . Наприклад, змінювати його за лінійним законом  $\theta_n = \theta_{n-1} - \Delta\theta \cdot n$ , де  $n$  - номер ітерації,

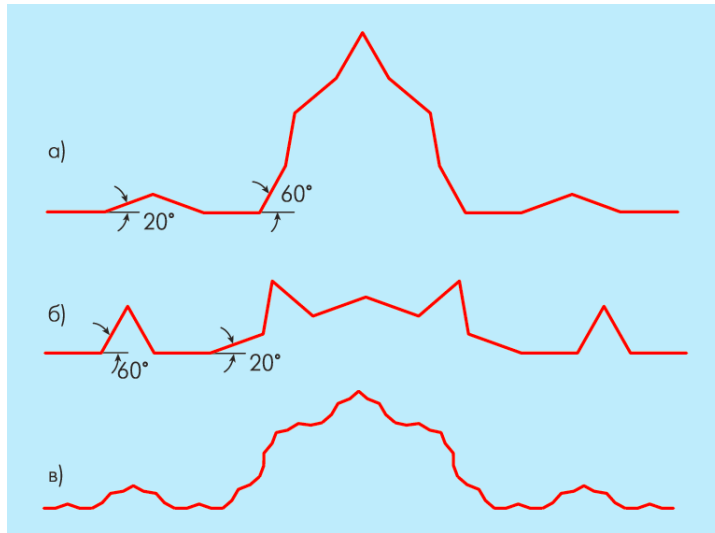
$\Delta\theta$  - приріст кута в підставі трикутника. Варіантом такого принципу побудови ламаної є наступна послідовність кутів:  $\theta_1 = 20^\circ$  для першого ітерації,  $\theta_2 = 10^\circ$  для другого і т.д. Конфігурація вібратора в цьому випадку не буде строго рекурсивної, проте всі його сегменти, синтезовані в одній ітерації, будуть мати однакові розміри і форму. Тому геометрія такої гібридної ламаної сприймається як самоподібна. При малому числі ітерацій поряд з негативним приростом  $\Delta\theta$  може застосовуватися квадратична або інша нелінійна зміна кута  $\theta_n$ .

Розглянутий підхід дозволяє задавати розподіл резонансних частот антени і значення її внутрішнього опору. Однак перестановка порядку зміни значень кутів  $\theta$  в ітераціях не дає еквівалентного результату. Як зазначено в роботах [10, 19], для однієї і тієї ж висоти ламаної лінії різні комбінації однакових кутів, наприклад  $\theta_1 = 20^\circ$ ,  $\theta_2 = 60^\circ$  і  $\theta_{+1} = 60^\circ$ ,  $\theta_2 = 20^\circ$  (рис .8), дають однакову розгорнуту довжину префракталов. Але, проти очікування, повний збіг параметрів не забезпечує тотожності резонансних частот та ідентичність многополосних властивостей антен. Причина - зміна внутрішнього опору сегментів ламаної, тобто ключову роль грає конфігурація провідника, а не його розміри.

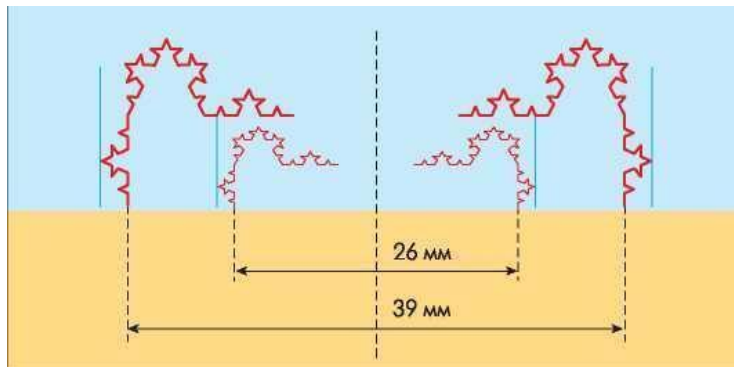
Перспективне застосування антен з топологією Коха - МІМО-системи зв'язку (системи зв'язку з багатьма входами і виходами).

Для мініатюризації антенних решіток абонентських терміналів в таких засобах комунікації фахівці Лабораторії електромагнетизму Університету Патрас (Греція) [20] запропонували фрактальну подобию перевернутої

L-антени (ІЛА). Суть ідеї зводиться до вигину вібратора Коха на  $90^\circ$  у точці, що ділить його на сегменти з співвідношенням довжин 2:1. Для мобільних засобів зв'язку з частотою несучої  $\sim 2,4$  Гц габарити такої антени в друкованому вигляді складають 12,33 x 10, 16 мм ( $\sim \lambda/10$  x  $\lambda/12$ ), смуга пропускання -  $\sim 20\%$  і ККД - 93% [20].



**рис.8. Узагальнені префрактали Коха другої ітерації з негативним приростом  $\Delta\theta$  (а), позитивним приростом  $\Delta\theta$  (б) і третьої ітерації з негативним приростом  $\Delta\theta = 40^\circ, 30^\circ, 20^\circ$  (в)**



**рис.9. Приклад двухдиапазонної (2,45 і 5,25 ГГц) антеною решітки**

Діаграма спрямованості по азимуту майже рівномірна, коефіцієнт посилення в перерахунку до входу фідера складає  $\sim 3,4$  дБ. Правда, як зазначено в статті [20], робота таких друкованих елементів у складі решітки (рис.9) супроводжується зниженням їх ККД в порівнянні з одиничним елементом [21]. Так, на частоті 2,4 ГГц ККД зігнутого на  $90^\circ$  монополя Коха знижується з 93 до 72%, а на частоті 5,2 ГГц - з 90 до 80%. Дещо краще йде справа з взаємним впливом антен високочастотної смуги: на частоті 5,25 ГГц розв'язка між елементами, що утворюють центральну пару антен, становить 10 дБ. Що стосується взаємного впливу в парі сусідніх різнодіапазонних елементів, то залежно від частоти сигналу розв'язка змінюється від 11 дБ (на 2,45 ГГц) до 15 дБ (на частоті 5,25 ГГц). Причина погіршення ефективності роботи антен - взаємний вплив друкованих елементів.

Таким чином, можливість вибору безлічі різноманітних параметрів антеною системи на основі ламаної Коха дозволяє при проектуванні задовольняти різні вимоги, пропоновані до значення внутрішнього опору і розподілу резонансних частот. Однак, оскільки взаємозалежність рекурсивної розмірності і характеристик антени може бути отримана тільки

для певної геометрії, справедливість розглянутих властивостей для інших рекурсивних конфігурацій потребує додаткового дослідження.

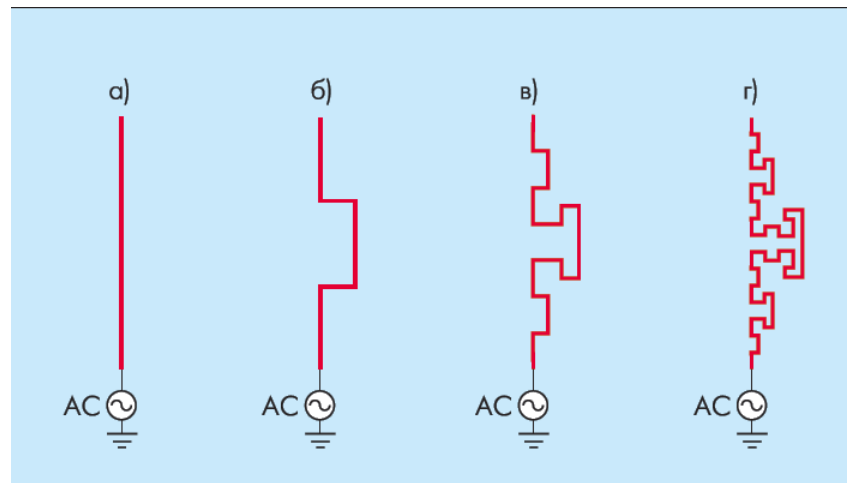
### **Диполі на основі фракталів Мінковського, Гільберта та їх модифікацій**

Весьма близький за своїми властивостями до антени на основі фрактала Коха - диполь, сформований за законом ламаної Германа Мінковського. При побудові цієї антени замість системи трикутників на прямий формуються меандри убуваючих розмірів. В Університеті Патраса запропонована модифікація ламаної Мінковського, заснована не на квадратному, а на прямокутному шаблоні. Така модифікація названа авторами фрактальної прямокутної ламаної ( Fractal Rectangular Curve , FRC (рис.10). При синтезі цієї ламаної використаний досить нетрадиційний для фракталів метод. Своєю появою FRC забов'язана так званою кривою квадратів ( Squares Curve ), ініціатором якої служить квадрат. На першому кроці рекурсії його вершини стають центрами чотирьох квадратів зі сторонами вдвічі меншої довжини (рис.11). Далі , в вершинах цих квадратів знову розташовуються центри ще менших чотирикутників , і процес рекурсивного синтезу повторюється. Примітно, що на кожному новому кроці периметр одержуваної фігури збільшується в два рази. Саме одну з чотирьох сторін цього геометричного об'єкта вчені Університету Патраса і запропонували використовувати як самостійну фрактальну конструкцію для побудови антен (див. рис. 10). Узагальнення полягало в переході від ініціюючого квадрата до довільного прямокутника .

Антени на основі FRC - ламаної мають ряд переваг перед вібраторами Коха і Мінковського. Наприклад, згідно, при друкованому виконанні монополя у вигляді FRC -3 на резонансну частоту 2,45 ГГц займає область, обмежену прямокутником площею 53,72 мм<sup>2</sup> зі сторонами 15,8 і 3,4 мм (  $\approx \lambda / 8 \times \lambda/36$  ). Для тієї ж частоти висота звичайного лінійного вібратора дорівнює 23,5 мм. Габарити монополя Коха вписуються в прямокутник площею 125 мм<sup>2</sup> , а монополя Мінковського - 64 мм<sup>2</sup>. При цьому смуга пропускання FRC - 3 - антени становить 770 МГц, а вібраторів Коха і Мінковського- 490 і 280 МГц, відповідно. Інша перевага FRC - 3 - антени в друкованому виконанні - збільшене (до 84 Ом) внутрішній опір проти 50 Ом для друкованих монополів Коха і Мінковського. ККД монополя з геометрією FRC - 3 дорівнює  $\sim 94$  %. Перераховані фактори, на думку авторів, роблять FRC - 3 - антени привабливими для реалізації на їх основі прийомних ґрат МІМО -терміналів в ноутбуках і кишенькових комп'ютерах.

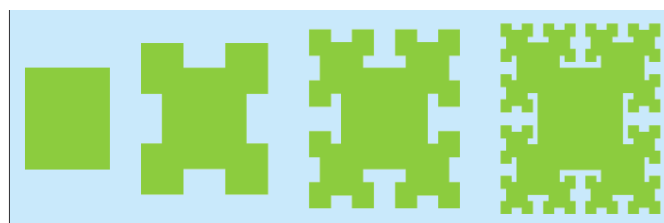
В антенних додатках зустрічається, що найменше, три модифікації ще однієї топології SFC-кривою, приписуваною Мінковському, шаблоном якої служить " біполярний стрибок " (рис.12).

Дві з них відрізняються порядком заповнення найбільш протяжних прямолінійних сегментів шаблону на наступних ітераціях . При цьому у разі строгого проходження

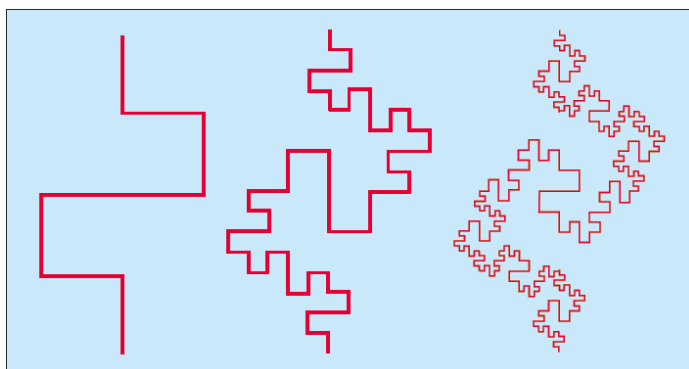


**рис.10. Формування фрактальної прямокутної ламаної антени: вихідний вібратор (а) і перші три ітерації монополя - FRC-1 (б), FRC-2 (в), FRC-3 (г)**

канонам фрактальної геометрії, ламана формується з різними по висоті "біполярними стрибками" (див. рис.12). Це закладає передумови для розширення смуги прийому антени. В альтернативному варіанті, запропонованому дослідниками Університету Патраса, протяжні сегменти в центрі "стрибка" на наступній ітерації замінюються парою ідентичних сегментів, рівних за розмахом іншим "скачком" вздовж ламаної (рис.13).



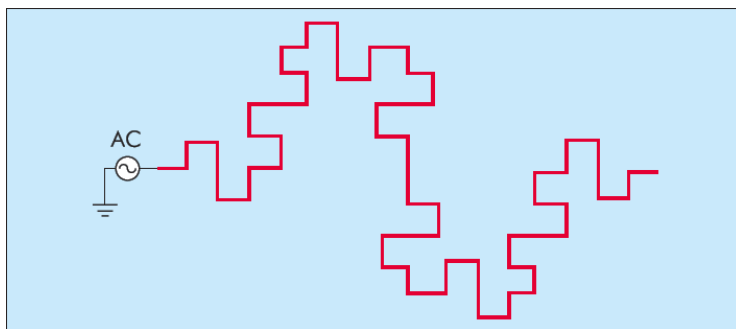
**рис.11.Ітерації прямокутної рекурсії, узагальнюючої Squares Curve**



**рис.12. Застосування шаблону виду "біполярний стрибок" при формуванні топології SFC-кривої Мінковського**

Ще одна модифікація полягає у відході від строгих правил побудови фракталів на користь комбінації " біполярних стрибків " Мінковського і модульованого ними меандра. В роботі (21) на основі антен такої комбінованої геометрії запропоновано реалізувати на РСМСІА- карті чотирьохелементну антенну решітку поперечного прийому для роботи в системі МІМО на частотах  $\sim 2,4$  ГГц (рис. 14). Як показали дослідження, проведені в Університеті Патраса (21), при такому конструктивному вирішенні взаємний вплив антен на деяких частотах більше, ніж у варіанту на основі елементів перевернутої F-антени (IFA), розміщених на чотирьох сторонах РСМСІА -модуля (мал.15).

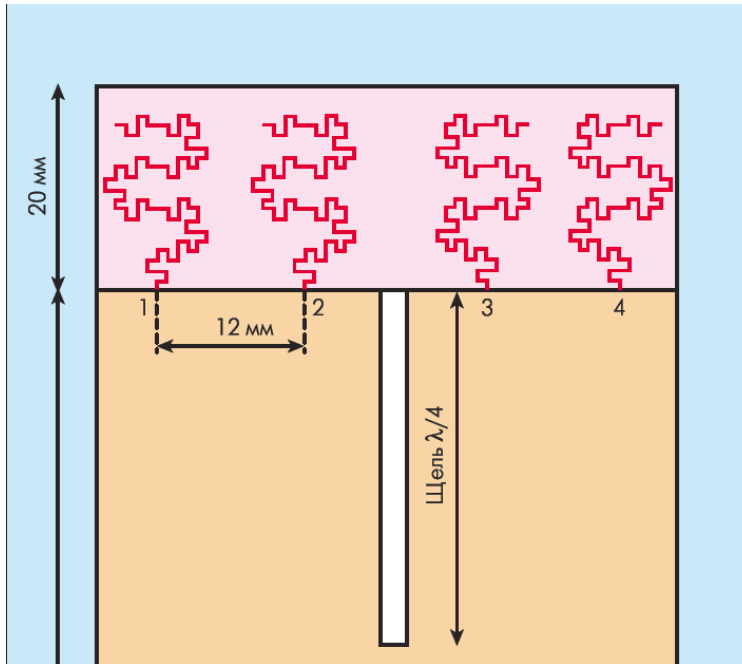
Але порівняно з IFA - ґратами висота антенних елементів у вигляді модифікованих " біполярних стрибків "



**рис.13. Модифікація "стрибків" Мінковського**

Мінковського менше. До того ж вони більш ширококугові. Існують і варіанти спільного застосування в одній РСМСІА-карті IFA і антен Мінковського [24]. Пильна увага, що приділяється вібраторам Мінковського [21, 24, 25], свідчить про те, що цей тип антен, можливо, незабаром можна буде віднести до розряду найбільш досліджених.





**рис.14. Чотирьохелементна антенна решітка МІМО-терміналу [21]**

За формою утворюючого елемента досить схожа на фрактал Мінковського згадувана ламана Гільберта (див. рис.3). Вперше мініатюрна антена, заснована на такій рекурсивній геометрії, була представлена вченими Політехнічного університету Каталонії [26]. Особливість найпростішого варіанта її побудови полягає в тому, що вихідний квадрат на першій ітерації розбивається на чотири симетричних підквадрата, центри яких з'єднуються П – подібною ламаною лінією. Процес декомпозиції підквадратів на нові квадранти повторюється, і П- сегменти, утворені при об'єднанні центрів квадрантів кривої, з'єднуються (див. рис.3). Розміри квадранта на n-ій ітерації, за умови одиничної довжини сторін вихідної фігури, визначаються як  $2^{-n} \times 2^{-n}$ , а довжина  $L$  - монополя Гільберта збільшується за кожної ітерації  $n$  згідно рівнянню [ 27,28 ]

$$L(n) = \frac{4^n - 1}{2^n - 1} h = (2^n + 1)h = (2^{2n} - 1)d, \quad (5)$$

Де  $h$  – повна висота монополя (довжина сторони описанного навколо нього квадрата),  $d$  – протяжність мінімального прямолінійного сегмента. Для п'ятої ітерації  $L(5) = 33h$ . Звідси напрашується висновок, що резонансна частота такого монополя Гільберта може бути в 33 рази менше основної резонансної частоти прямолінійного вібратора висотою  $h$ . Однак насправді таке зниження номіналу резонансних частот на практиці неможливо через наявність взаємного впливу близько розташованих сегментів

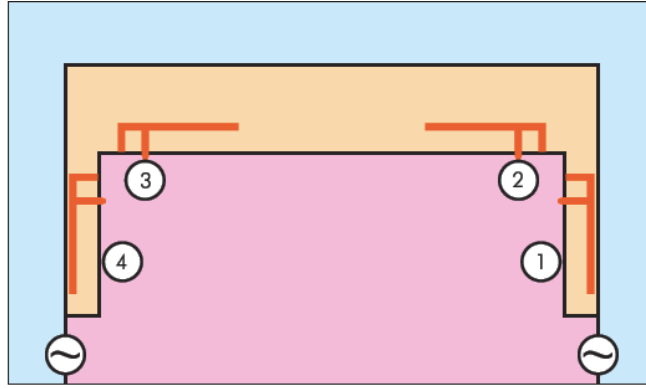


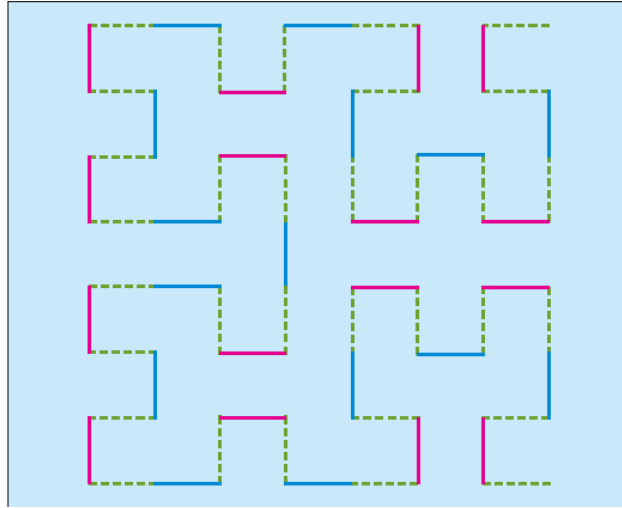
рис.15. Чотирьохелементна IFA-решітка МІМО-терміналу [21]

фрактальної антени, що приводить до укорочення її електричної довжини. Особливо сильна взаємодія виникає між зламами провідника, в результаті якого шлях поширення електромагнітних хвиль від одного кінця монополя до іншого стає більш коротким. Проте, навіть з урахуванням ефекту взаємного впливу резонансна частота може знижуватися на порядок і більше. В останні роки антени на основі кривої Гільберта, завдяки їх простоті і технологічності, лідирують по числу присвячених їм публікацій, потіснивши досить довго утримувавших пальму першості антени на базі префрактала Коха.

Диполі на основі ламаної Гільберта відносяться до розряду тих нечисленних фрактальних антен, для яких на сьогоднішній день отримані аналітичні співвідношення за розрахунком резонансних частот. У згадуваній вже роботі Виною [ 10 ] для виведення таких співвідношень запропоновано структурувати ламану на ключові, з погляду електродинаміки, сегменти. Зокрема в лінії Гільберта слід розрізняти паралельні провідники, короткозамикаючі перемички і сполучні сегменти ( рис.16) .

За допомогою методики Виною декілька перших резонансних частот можуть бути розраховані мінімізацією вираження

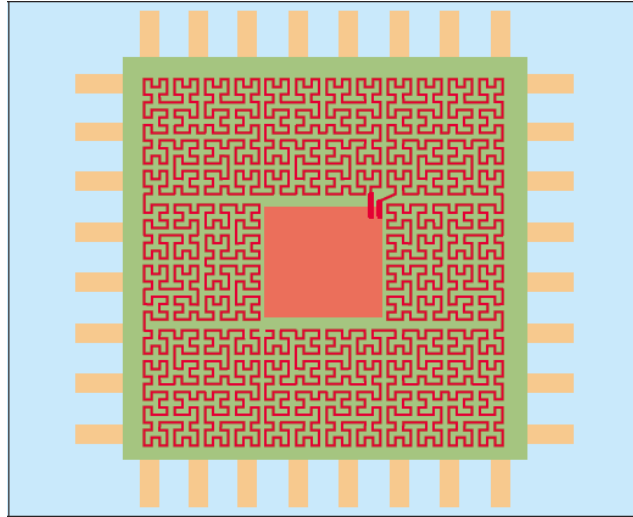
$$\left| m \frac{\eta}{4\pi f_r} \left[ \lg \left( \frac{2d}{b} \right) \right] \cdot \operatorname{tg} \left( \frac{\pi f_r d}{c} \right) + \right. \\ \left. + \frac{\mu_0}{2\pi} s \left[ \lg \left( \frac{4s}{b} \right) - 1 \right] - \frac{\mu_0}{\pi} \frac{rc}{4f_r} \left[ \lg \left( \frac{2rc}{bf_r} \right) - 1 \right] \right| \rightarrow \min, \quad (6)$$



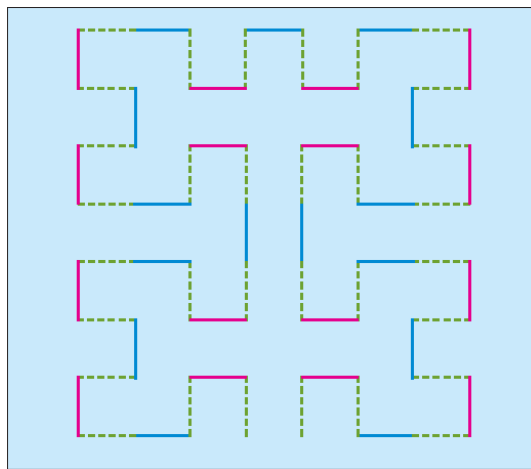
**рис.16. Сегментація ламаної Гільберта на електродинамічні значущі фрагменти**

де  $r \geq$  - порядковий номер резонансної частоти,  $f_r$  – значення  $r$ -ої резонансної частоти. Формула уточнена автором статті по відношенню до опублікованого в [10] висловом в результаті аналізу люб'язно наданого Виною тексту комп'ютерної програми, яка використовувалася для чисельного розрахунку резонансних частот. Формула ( 6 ) застосовна для відокремленого диполя. Вона не враховує високочастотні ефекти і, отже, не може дати точні результати для вищих резонансів.

Дослідження антен на основі фрактала Гільберта примітні тим, що вони породили перші серйозні сумніви в абсолютній перевазі фрактальних систем над евклідовими. Вчені Політехнічного університету Каталонії[ 28 ] показали, що монополь Гільберта по ефективності програє антені, виконаній у вигляді серпантинної лінії або меандру. Геометрія останніх забезпечує кращий ККД, більший внутрішній опір і меншу добротність антени, а отже, більш широку смугу пропускання. Однак, як з'ясувалося, відносно низький внутрішній опір притаманне лише варіанту монополя, запитанному по краю фрактала. Зовсім інший результат спостерігається, якщо змінити точку підключення фідера, змістивши її в глибину вздовж ламаної Гільберта. Як показали дослідники Університету штату Пенсільванія, при оптимальному виборі місця розташування контактної площадки фідерного з'єднання можна забезпечити приблизно 50 - Ом активний внутрішній опір антени на першій ( фундаментальної )резонансній частоті,



**рис.17. Варіант виконання фрактальної антени в складі системи на кристалі (SoC) [29]**

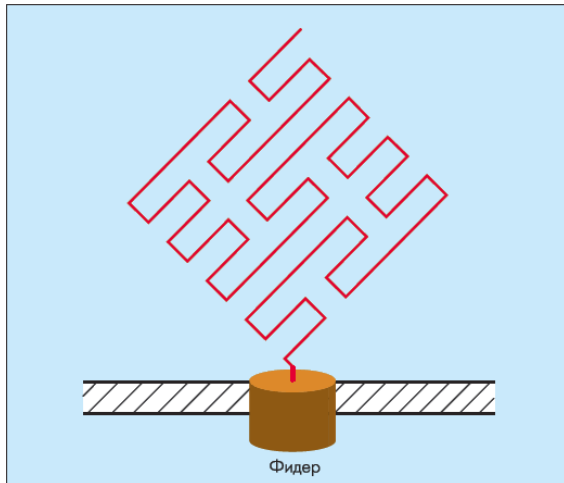


**рис.18. Ломана Мура**

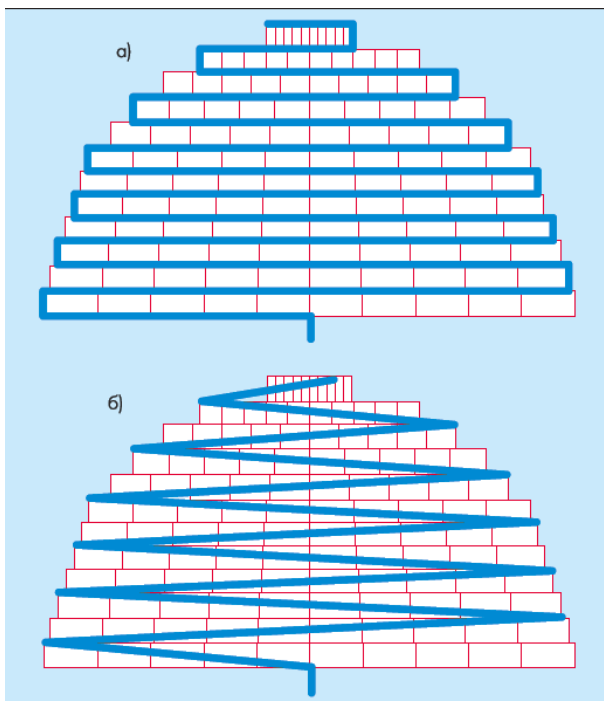
незалежно від номера інтеграції (27). Але платою за таку сталість значення внутрішнього опору антени є звуження смуги пропускання на низьких частотах (для перших ітерацій фрактала) порівняно зі смугою пропускання для ітерацій більшого порядку .

Антени на базі ламаної Гільберта рекомендується використовувати при створенні системи на кристалі (SoC) [29]. У цьому випадку антенний модуль і радіотехнічні пристрої можуть бути розміщені в одному корпусі ( рис.17 ), а більша фізична довжина лінії дозволяє диполям працювати на довжинах хвиль, незрівнянно більших, ніж габарити SoC. На основі цієї та інших SFC-кривих можуть бути виконані плоскі спіральні антени для пристроїв радіочастотної ідентифікації (RFID-мітки). Це дозволяє істотно зменшити габарити RFID-міток (див. рис.17).

Серед інших фрактальних антен, перш за все, слід згадати SFA на базі варіанта кривої Гильберта (див. Рис 16),



**рис.19. Варіант просторової орієнтації монополя Пеано [9]**



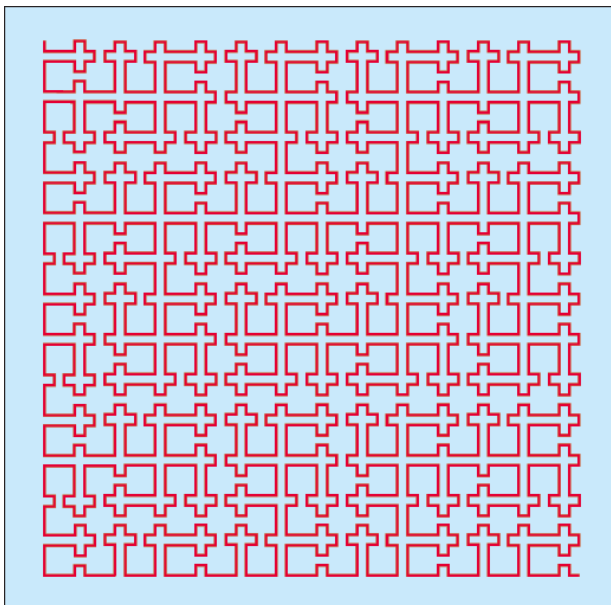
**рис.20. Монополі, виконані у вигляді серпантинних ліній евклідової геометрії: меандровий (а) і зигзагоподібний (б)**

який в 1900 році запропонував американський математик Мур ( Е.Н.Моore ) (рис.18) [30]. Вчені Політехнічного інституту Вустера ( США ) зіставили властивості цієї антени з друкованою антеною по рекурсії Гільберта [31]. В результаті моделювання зроблено висновок про досить близьких значеннях резонансних частот обох типів антен, реалізованих однаковим числом ітерацій, і про подібність залежностей, характеризують зміну їх внутрішніх імпедансів. Для наближеного розрахунку резонансних частот антени, отриманої шляхом сегментації провідника за законом ламаної Мура

(див. рис.18), слід використовувати те ж співвідношення ( 6 ), що і для антени Гільберта.

Збіг аналітичних співвідношень для розрахунку резонансних властивостей антен Мура і Гільберта - свідчення інваріантності аналітичного методу до особливостей геометрії антени. У цьому сенсі чисельне моделювання, наприклад в пакеті NEC, займає більш вигідні позиції, оскільки воно більш чутливо до відмінностей в топології фрактальних антен.

Серйозний недолік багатьох електрично малих резонансних антен, в тому числі і антени Гільберта, - відносно великі рівні крос- поляризації. Найменші крос- поляризаційні компоненти в наведених токах характерні для конструкції антени по фракталу Пеано(мал. 19). Зокрема, рівень придушення крос- поляризаційної перешкоди другої ітерації SFA - конструкції антени Пеано, вписаною в квадрат  $70 \times 70$  мм, перевищує 60 дБ в обох площинах поляризації. А рівень крос-поляризації антени Гільберта третього порядку майже тієї ж електричної площі і з аналогічною резонансною частотою становить близько -9,5 дБ [32].



**рис.21. Ломаная Осгуда [36]**

Така істотна різниця в крос- поляризаційних характеристиках двох типів антен- здебільшого наслідок дзеркальної симетрії Гільберт -геометрії уздовж однієї з осей, тоді як крива Пеано володіє діагональною центральною симетрією. Тому при прийомі сигналів діагональ SFA - антени Пеано зазвичай орієнтують паралельно вектору електричної напруженості падаючих електро магнітних хвиль (див. рис.19 ).

Основна відмінність лінії Пеано від кривих Гільберта і Мура - розбивка вихідного одиничного квадрата не на чотири, а на дев'ять частин з розмірами  $3^{-n} \times 3^{-n}$  кожна, де  $n$  - номер ітерації. При цьому довжина сполучних сегментів ламаної лінії Пеано дорівнює  $3^{-2n}$ . Для довільної висоти описаного квадрата  $h$  протяженність мінімального прямолінійного сегмента лінії  $n$ -ої ітерації може бути розрахована за формулою [ 32 ] :

$$d = \frac{h}{3^n - 1}, \quad (7)$$

де  $h$  - повна висота монополя.

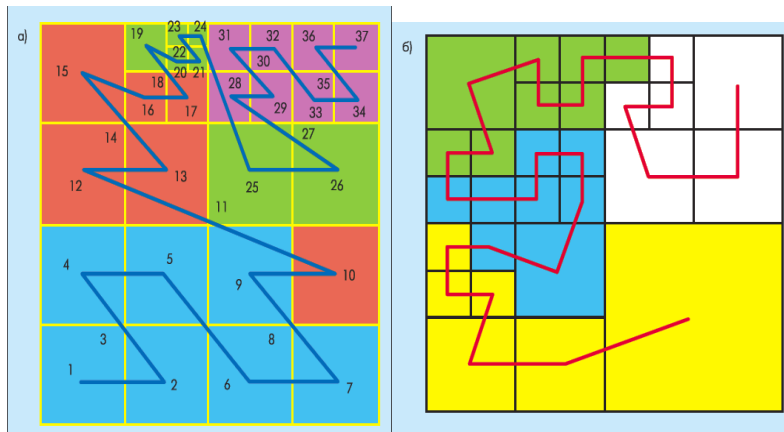
Повная довжина ламаної (див. рис.19) на  $n$ -ій ітерації, вписаною в квадрат зі стороною  $h$ , визначається залежністю [32-33]

$$L(n) = (3^{2n} - 1)d = (3^n + 1)h. \quad (8)$$

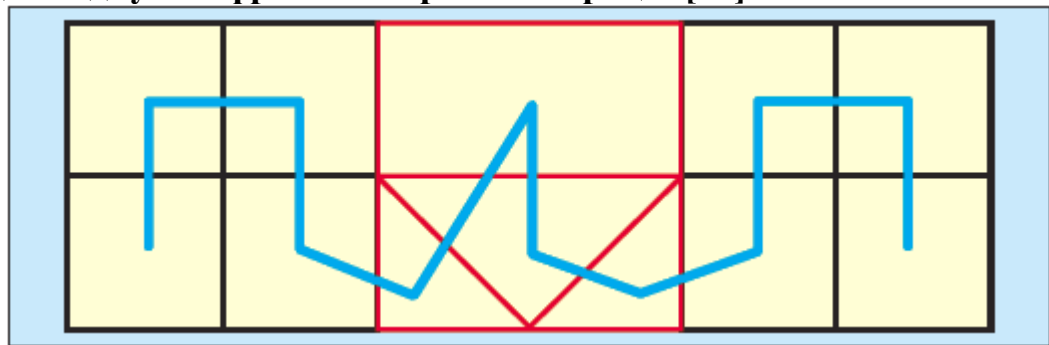
Неординарні поляризаційні властивості дозволяють рекомендувати антену Пеано для формування сигналів подвійної поляризації з метою підвищення пропускної здатності каналів передачі засобів зв'язку. Комбінацію пари таких монопольних антен, орієнтованих в ортогональних площинах, можна розглядати в якості компактної альтернативи турнікетною антені. На базі цього прикладу можна зробити висновок, що для досягнення низького рівня крос- поляризаційної перешкоди доцільно використовувати фрактальні антени з центральносиметричною структурою.

С точки зору ККД, розглянутий тип антени програє серпантинним вібраторам евклідової геометрії, варіанти яких показані на рис.20. Хоча цей програш при близьких електричних розмірах і не перевищує десятка відсотків, сам по собі цей факт позбавляє фрактальні антени ореолу переваги і винятковості, який їм спочатку приписувався.

Втім, серпантинні конструкції, подібно всім відомим фрактальним антенам, також не дозволяють досягти розглянутого автором в [ 33 ] фундаментальної межі добротності. Цей факт змушує з одного боку шукати нові антенні рішення, а з іншого - переглянути зазначені межі. Як наголошується дослідниками Каліфорнійського університету в Лос - Анджелесі [ 34 ], для наближення характеристик антени до граничних значень слід збільшувати діаметр її провідника, а також застосовувати в її топології тривимірні версії префракталов.



**рис.22. Модифіковані фрактали Лебега (а) і Гільберта (б), що поєднують фрагменти різних ітерацій [37]**



**рис.23. Комбінувальна рекурсія**

В Політехнічному університеті Каталонії був проведений порівняльний аналіз уточненої фундаментальної межі і оцінок добротності реальних фрактальних ЕММ [35]. При цьому для антен на основі префракталов Коха, Пеано, Гільберта та інших отримано досить гарний збіг кордону добротності з результатами моделювання. Відносно монополів, навантажених серпантинної (меандрової) лінією (MLLM, Meander Line Loaded Monopoles), по всій видимості, мають більш рівномірний розподіл струму, зазначена межа при  $ka < 0,8$  проходить трохи вище експериментальних значень.

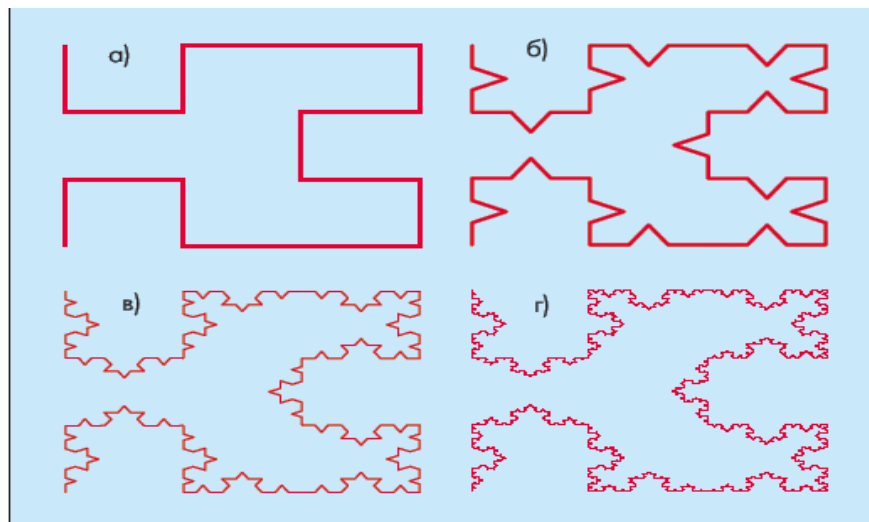
Досить цікавою для антенних додатків є SFC-конструкція Осгуда (рис.21) [36]. Лінії Осгуда мало відомі широкому колу дослідників, тому аналіз електродинамічних властивостей відповідних їм антен ще не проводився.

Не вивчені і електродинамічні властивості антен на основі змішаних ітерацій префракталов, окремі SFC - ділянки яких містять фрагменти різних рівнів рекурсії (рис.22) [37]. На цій основі можуть будуватися адаптивно Реконфігуровані антени, в яких за допомогою МЕМС - технології можна вибудовувати в єдину криву різні за формою ділянки вихідного фрактала вищого порядку. Про можливість подібної реконфігурації фрактальних антен свідчить опис патенту [38] і робота [10], де розглянуто спрямовані властивості монополя Гільберта при різних варіантах короткого замикання його фрагментів. За допомогою МЕМС - ключів в такій антені можна зміщувати



резонансну частоту і напрямок максимуму діаграми спрямованості. У рамках такого підходу можливі адитивні комбінації декількох різних геометричних фрагментів у межах вихідної фігури. Наприклад, в Університеті Британської Колумбії ( Канада ) на цій основі запропонована комбінована рекурсивна лінія, окремі фрагменти якої відповідають кривій Гільберта( рис.23 ) [39].

Слід зазначити, що комбінація декількох фракталів в одній лінії може бути виконана не тільки адитивним способом, а й мультиплікативно. Вперше електричні властивості фрактальних монополів, отриманих в результаті мультиплікативної комбінації кривих Коха і Гільберта( рис.24 ), досліджував Виной [10]. Як впливає з [10], додаткова " модуляція " сегментів плоского префрактала Гільберта за законом ламаної лінії Коха дозволяє, в кінцевому рахунку, збільшити загальну довжину лінії більш ніж у чотири рази. Однак набагато результативніше не замінювати прямолінійні сегменти ламаними Коха до другого ітерації включно, а підвищувати порядок рекурсії Гільберта, якщо це можливо.



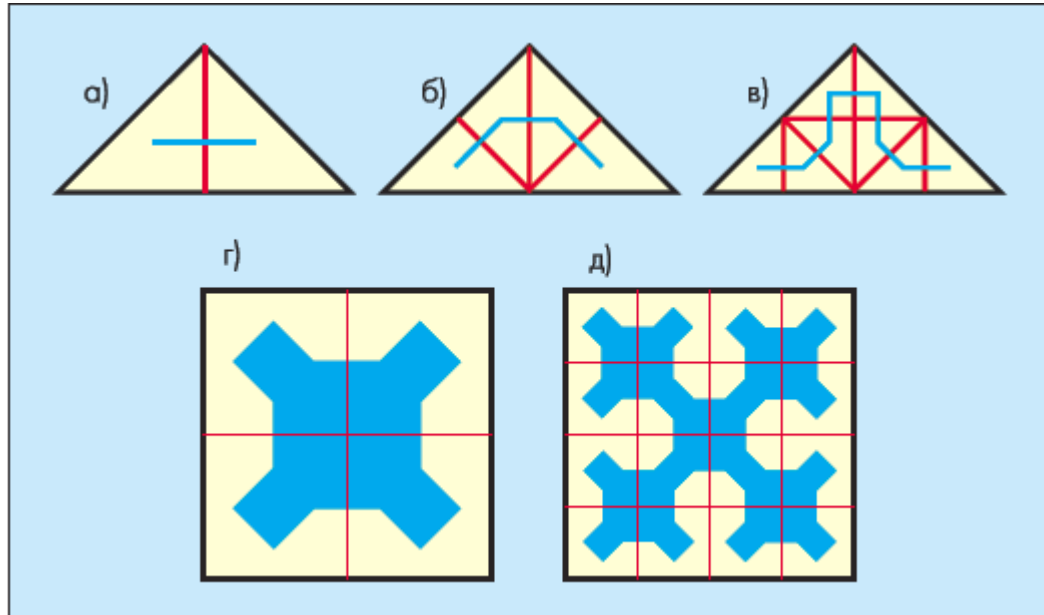
**рис.24. Симбіоз кривої Гільберта і Коха: а) базова геометрія кривої Гільберта; б) перша; в) друга; г) третя ітерації мультиплікативної комбінації кривих Коха і Гільберта [10]**

В цілому, гібридні антенні рішення, мультиплікативно поєднують кілька різних фракталів, заслуговують подальшого дослідження з метою розширення числа можливих рекурсивних комбінацій.

На завершення розгляду використовуваного в антенній техніці сімейства найпростіших фрактальних ліній, орієнтованих на заповнення квадратної або прямокутної площадки, слід зупинитися на ламаних Серпінського ( Sierpinski). Варіант SFC-конструкції Серпінського передбачає ітеріровану фрагментацію вихідного квадрата не тільки по горизонтальним і вертикальним лініях, а й по діагоналях [40]. Суть цієї ідеї найпростіше пояснити на основі трикутного фрагмента квадрата ( рис.25 ). Поступовий поділ трикутників шляхом розбиття навпіл однієї зі сторін, дозволяє отримати кілька варіантів рекурсивних кривих, що з'єднують центри

трикутних фрагментів. В роботі [ 41 ] викладено результати дослідження вченими Політехнічного університету Каталонії властивостей декількох ітерацій друкованої антени у вигляді стрілоподібної петлі Серпінського ( рис.26 ).

Розглянутий перелік фрактальних об'єктів, використовуваних в якості геометричного прототипу антени конструкції, може бути доповнений плоскими фігурами у вигляді фрактальних дерев, серветок Коха, Мінковського, Серпінського і т.п., а також їх об'ємними версіями [42].



**рис.25. Варіанти побудови ламаних Серпінського: а, б, в - технологія побудови стрілоподібної ламаної на нульовій, першій, другій ітераціях; г, д - перша і друга ітерації стрілообразної петлі**

Для виготовлення тривимірних конструкцій можуть використовуватися фаббер-технології тривимірного прототипування, що дозволяють прецизійно сформувати мініатюрні антени будь-якої складності [43].

Серед відомих прикладів серійних моделей фрактальних антен слід згадати продукцію іспанської компанії Fractus. Їй належить більше 140 національних і міжнародних патентів в цій області. Частина з них зареєстрована і в Російській Федерації (наприклад , патент № 2263378 ). Серед останніх розробок компанії – фрактальна антена Micro Reach Xtend <sup>TM</sup> на діапазон 2,4 ГГц (для систем стандартів Wi-Fi, Bluetooth і Zigbee). Її габарити складають всього  $3,7 \times 2$  мм, що відповідає розмірам рисового зерна. Інший напрям розробок Fractus – технологія створення " антени в корпусі", що отримала співзвучну назву AiP ( Antenna - in - Package). При цьому реалізуються ідеї, аналогічні представленим на рис.17, коли вся радіосистема, включаючи антену, виконується в одній мікросхемі .

На закінчення слід зазначити, що зусиллями ентузіастів і компаній типу Fractus антени на основі фрактальних композицій поступово перестають бути екзотикою. Зрозуміло, вони не вирішують всі проблеми мініатюризації антенних пристроїв. Але безсумнівно це напрямок створення ЕМА досить ефективно і займає цілком авторитетні позиції в сучасній техніці.

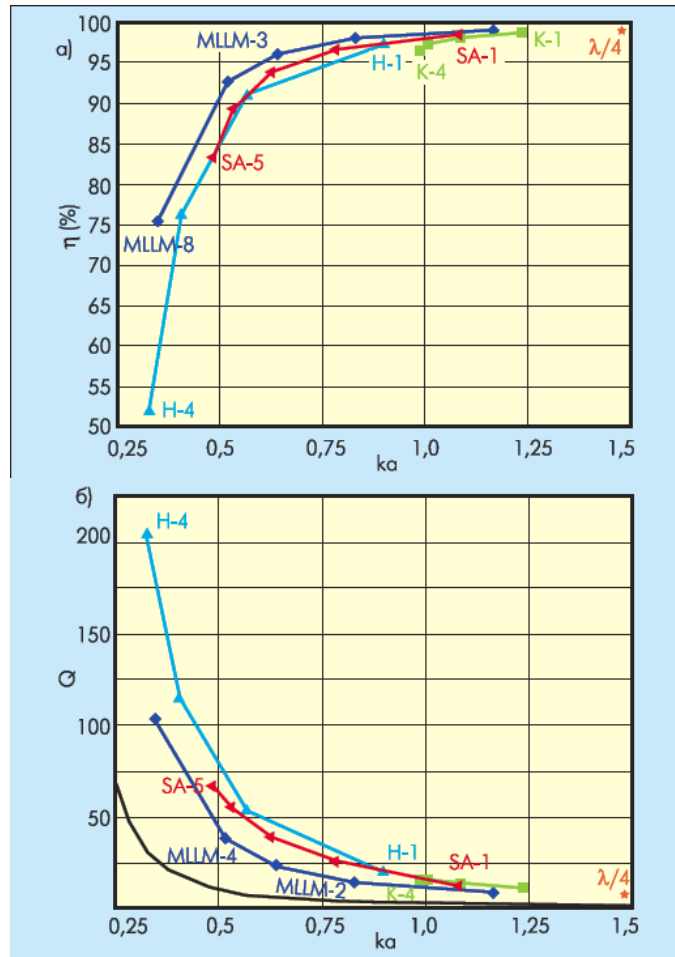


рис.26. Розрахункові значення КПД (а) і добротності (б) антен. MLLM - монополь, навантажений серпантинною лінією, К-п і Н-п - монополі Коха і Гільберта п-й ітерації, SA-п - антена на основі п-й ітерації стрілоподібній ламаній Ссерпінського. Нижня суцільна лінія в позиції (б) відповідає фундаментальній межі добротності  $Q_1 = 1 / (k^3 \cdot a^3) + 1 / (ka)$  [33]

