

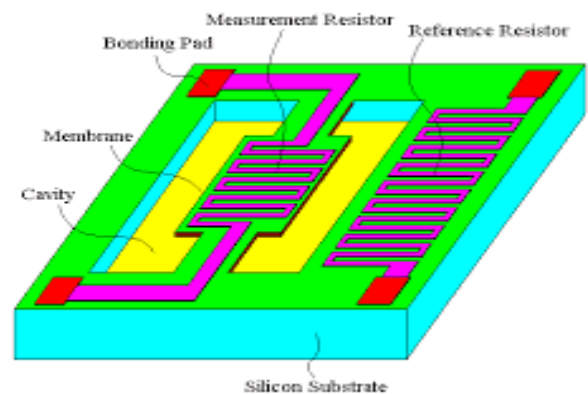
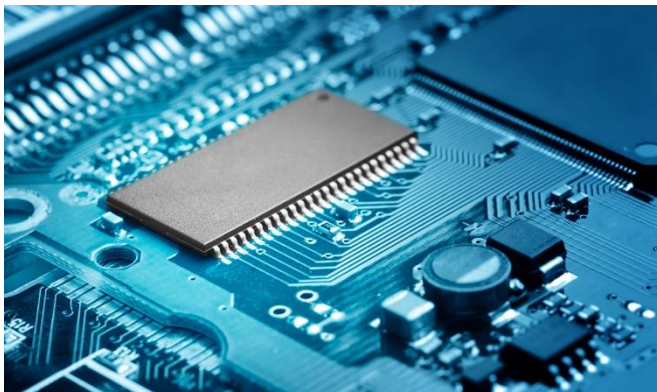
Національний університет "Львівська політехніка"

Харківський національний університет радіоелектроніки

Петро Кособуцький, Михайло Лобур, Світлана Чумаченко

Статистичне проектування мікроелектромеханічних систем

Навчальний посібник



№530785-TEMPUS-1-2012-1-PL-TEMPUS-JPCR



Tempus

Навчальний посібник **"Статистичне проектування мікроелектромеханічних систем"** створено для допомоги вищим навчальним закладам України впровадити нову магістерську навчальну програму "Проектування мікросистем".

Посібник **"Статистичне проектування мікроелектромеханічних систем"** створено при підтримці Європейського Союзу за Спільним Європейським Проектом "Curricula Development for New Specialization: Master of Engineering in Microsystems Design" (MastMST), ідентифікаційний номер 530785-TEMPUS-1-2012-1-PL-TEMPUS-JPCR.

Координатор проекту проф. Збігнєв Лісік, Технічний університет м.Лодзь, Польща.

Учасники проекту:

- Національний університет "Львівська політехніка", м. Львів, Україна, координатор проф. Михайло Лобур.
- Київський Національний університет ім. Тараса Шевченка, м. Київ, Україна, координатор проф. Валерій Скришевський.
- Харківський національний університет радіоелектроніки, м. Харків, Україна, координатор проф. Володимир Хаханов.
- Донецький національний технічний університет, м. Красноармійськ, Україна, координатор проф. Володимир Святний.
- Технічний університет м.Ільменау, Німеччина, координатор проф. Іво Рангелов.
- Ліонський Національний інститут прикладних наук, Франція, координатор проф. Александра Апотолук
- Університет Павії, м.Павії, Італія, координатор проф. Паоло Ді Барба

Посібник схвалено редакційним комітетом (проф. Паоло Ді Барба (Університет Павії) - співголова, проф. Александра Апостолук(Ліонський Національний інститут прикладних наук) – співголова, члени: проф. Збігнєв Лісік(Технічний університет м.Лодзь), д-р Яцек Подгурські (Технічний університет м.Лодзь), Д-р Януш Возний (Технічний університет м.Лодзь), Д-р Валентин Ішук (Технічний університет м.Ільменау), Д-р Марія-Евеліна Могначі (Університет Павії), Д-р Роберто Галді (Університет Павії)) 6 травня 2016, м. Павія, Італія

Автори висловлюють глибоку вдячність керівництву вищеназваних університетів за всебічну підтримку Проекту.

Textbook "**Statistical design of microelectromechanical systems**" developed to help higher education institutions in Ukraine to introduce new master's educational program "Designing microsystems".

Textbook "**Statistical design of microelectromechanical systems**" was created with the support of the European Union within the Joint European Project "Curricula Development for New Specialization: Master of Engineering in Microsystems Design" (MastMST), identification number 530785-TEMPUS-1-2012-1-PL-TEMPUS-JPCR.

Project Coordinator prof. Zbigniew Lisik, Lodz University of Technology, Lodz, Poland.

Учасники проекту:

- Lviv Politechnical National University, Lviv, Ukraine ,
Coordinator prof. Mykhailo Lobur.
- Taras Shevchenko National University of Kyiv, Ukraine,
Coordinator prof. Valeriy Skryshevsky.
- Kharkiv National University of Radioelectronics, Ukraine ,
Coordinator prof. Vladimir Hahanov.
- Donetsk National Technical University, Krasnoarmiysk,
Coordinator prof. Volodymyr Sviatny.
- Ilmenau University of Technology, Germany,
Coordinator prof. Ivo Rangelow.
- Lyon Institute of Applied Sciences, France,
Coordinator prof. Alexandra Apostoluk.
- University of Pavia, Italy,
Coordinator prof. Paolo Di Barba.

The Handbook was approved by Editorial Committee (prof. Paolo Di Barba (University of Pavia) - Co-Chair, prof. Alexandra Apostoluk (Lyon Institute of Applied Sciences) – Co-Chair, members: prof. Zbigniew Lisik (Lodz University of Technology), Dr Jacek Podgorski (Lodz University of Technology), Dr Janusz Wozny (Lodz University of Technology), Dr Valentyn Ishchuk (Ilmenau University of Technology), Dr Maria Evelina Mognaschi (University of Pavia), Dr Roberto Galdi (University of Pavia) May 6, 2016, Pavia, Italy.

The authors express their deep gratitude to the aforementioned universities for full support of the project.

Національний університет "Львівська політехніка"
Харківський національний університет радіоелектроніки
Петро Кособуцький, Михайло Лобур, Світлана Чумаченко
Статистичне проектування мікроелектромеханічних систем
Навчальний посібник

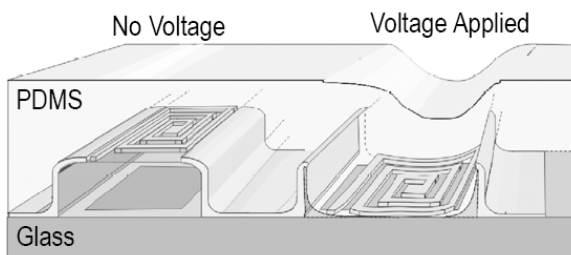
Роботу виконано за проектом Curricula Development for New Specialization: Master of Engineering in Microsystems Design / MastMST, ідентифікаційний номер 530785-TEMPUS-1-2012-1-PL-TEMPUS-JPCR. Програма фінансування ЄК: Міжрегіональна програма Європейського сусідства і партнерства. Цільова група: студенти, випускники, викладачі та адміністрація університетів, керівники промислових підприємств, міністерство освіти і науки України.

Основною метою проекту є створення умов в українських технічних університетах для наскрізного 3-рівневого навчання за спеціальністю "Проектування та інженерія мікросистем" відповідно до регіональних потреб ринку праці.

У навчальному посібнику розглянуто особливості автоматизованого проектування мікроелектромеханічних систем з врахуванням статистичних параметрів. Коротко розглянуто моделі елементів і пристроїв MEMS. Найбільшу увагу приділено статистичним методам проектування і моделювання MEMS. Приведено практичні результати статистичного проектування акселерометра

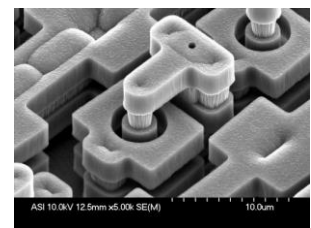
Запропонований посібник буде корисним для студентів, аспірантів та науковців, які спеціалізуються в галузі автоматизованого проектування та математичного моделювання мікроелектромеханічних систем, деякі із них зображені нижче.

MEMS: Micro Electro Mechanical Systems

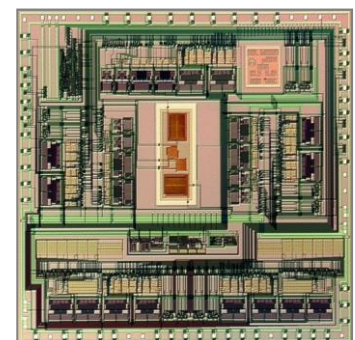


Electrostatic Actuator for chronic drug dosing

<http://www.wimserc.org> [http:// www.memx.com](http://www.memx.com)
<http://mems.sandia.gov/about/electro-mechanical.html>



Mechanic actuator



Accelerometer

ЗМІСТ

Передмова

1. Базові методи та моделі САПР мікросистемної техніки	3
1.1. Вступ	3
1.2. Основні задачі та цілі автоматизації проектування мікросистемної техніки.	4
1.3. Базові засади САПР проектування мікросистемної техніки та інтелектуальних і вбудованих систем на їх основі	13
2. Базові процеси і моделі елементів мікросистемної техніки	22
2.1. Базові технології мікро- і наносистемної техніки	22
2.2. Функціональні матеріали мікро- і наносистемної техніки	31
2.3. Базові фізичні моделі теплових сенсорів	50
2.4. Мікросенсиори на бар'єрних переходах	58
2.5. Магнітоелектричні мікросенсиори	63
2.6. Сенсори деформацій	71
2.7. Пасивні сенсори фізичних полів	78
2.8. Сенсори на квантових ефектах	79
2.9. Фізична механіка макро- і мікросенсорів руху	82
2.10. Динаміка процесів в сенсорах руху осциляторного типу	95
2.11. Мікроакселерометри, мікрогіроскопи і мікроприводи механічного руху	106
3. Моделі та методи статистичного аналізу мікросистем	134
3.1. Вступні поняття	134
3.2. Числові характеристики та закони розподілу дискретних випадкових величин	136
3.3. Основні моделі та закони розподілу дискретних випадкових величин	146
3.4. Основні моделі, числові характеристики та закони розподілу неперервних випадкових величин	150
3.5. Моделювання розподілів випадкових величин та критерії перевірки згоди між дослідними і теоретичними розподілами	161
3.6. Статистичний аналіз флуктуацій в мікросистемах сенсори деформацій	179
3.7. Метод статистичного моделювання Монте-Карло	195
4. Методи статистичної оптимізації мікросистем	200
4.1. Елементи математичних основ оптимізації мікросистем	200
4.2. Умовна мінімізація за наявності обмежень	205
4.3. Оптимізація методами покоординатного і градієнтного спуску	206
4.4. Метод генетичних алгоритмів	208
4.5. Оптимізація багатокритеріальної задачі та крива компромісів Парето	211
4.6. Оптимізація пружних мікросистем кантилеверного типу	213
4.7. Оптимізація пружних мембран	219
4.8 Структурна оптимізація мікросистем методом змінної густини	220
Список літератури	222

1. Базові методи та моделі САПР мікросистемної техніки

1.1. Вступ

Проектування - процес складання опису, необхідного для створення в заданих умовах ще не існуючого об'єкта, на основі первинного опису цього об'єкта і (або) алгоритму його функціонування. Проектування включає в себе комплекс робіт з вишукування, дослідження, розрахунками та конструювання. Проектування - це складний специфічний вид творчої діяльності людини, заснований на глибоких наукових знаннях і творчому пошуку, використанні накопиченого досвіду і навичок у певній сфері, не позбавлений, однак, необхідність виконання трудомістких рутинних робіт.

Під автоматизацією проектування (АП) розуміється такий спосіб виконання процесу розробки проекту, коли проектні процедури здійснюються розроблювачем виробу при тісній взаємодії з ЕОМ. АП передбачає систематичне використання засобів обчислювальної техніки (ОТ) при раціональному розподілі функцій між проектувальником і ЕОМ і обґрунтованому виборі методів машинного рішення задач. Раціональний розподіл функцій між людиною і ЕОМ передбачає, що людина повинна в основному вирішувати завдання творчого характеру, а ЕОМ - завдання, що допускають формалізований опис у вигляді алгоритму, що дозволяє досягти більшої ефективності в порівнянні з традиційним ручним способом.

Суттєва перевага машинних методів проектування полягає в можливості проводити на ЕОМ експерименти на математичних моделях (ММ) об'єктів проектування, відмовившись від дорогого фізичного моделювання або скоротивши його.

Мета АП - підвищення якості, зниження матеріальних витрат, скорочення термінів проектування і ліквідація тенденції до зростання числа інженерно-технічних працівників, зайнятих проектуванням, підвищення продуктивності їхньої праці. АП виникла на базі досягнень конкретних технічних дисциплін, обчислювальної математики і ОТ.

1.2. Основні задачі та цілі автоматизації проектування мікросистемної техніки.

Проектування елементів мікросистемної техніки (МСТ), в якому всі проектні розв'язки або їх частина отримується із застосуванням ЕОМ, називаються автоматизованими (системи автоматизованих проектувальних робіт (САПР)) (в англ. мовному представленні CAD (Computer Aided Design)). Тут ручне моделювання переважно використовується для нескладних елементів чи конструкцій.

Основою САПР є ММ, як множина математичних об'єктів і відношень між ними, які адекватно відображають конкретні властивості МСТ. Якщо в проектуванні використовують різні рівні абстрагування, то такий підхід називається багато масштабним моделюванням. В залежності від послідовності процедур і етапів розрізняють низхідний, висхідний і змішані етапи проектування. Послідовність розв'язування задач від нижніх рівнів до верхніх характерна для висхідного проектування. Зворотний напрям – це низхідне проектування. Для проектування мікроелектромеханічних і наноелектромеханічних систем (МЕМС і НЕМС) потрібні інтегровані системи, що включають в себе: - проектування на компонентному рівні; - проектування на міждисциплінарному рівні; - проектування електронних компонентів; - проектування з врахуванням забезпечення надійного функціонування МСТ.

Однак математична постановка для більшості проектних процедур неочевидна, а їх подальша алгоритмічна реалізація існуючими математичними методами часто є незадовільною. Тому формалізація завдань, вибір і розробка ММ, методів і алгоритмів виконання проектних процедур значною мірою визначають зміст теорії САПР.

Необхідна умова реалізації алгоритмізованих проектних процедур - наявність відповідних засобів ОТ. Особливістю САПР є існування завдань, можливості вирішення яких без прийняття спрощень знаходяться за межами можливостей ОТ. Тому в великих системах САПР використовують ЕОМ граничної продуктивності. Пристосування завдань до можливостей наявної ВТ відбувається, по-перше, на основі спеціальних прийомів поділу

процесу проектування на ряд ієрархічних рівнів і аспектів, по-друге, завдяки збереженню за людиною в САПР тих функцій, які не можуть бути виконані формальними методами з прийнятними витратами часу і засобів. В результаті процес САПР зводиться до необхідності вирішення кінцевої послідовності завдань прийнятною складністю в режимі взаємодії людини і ЕОМ.

Необхідність взаємодії людини і ЕОМ і специфіка проектних завдань породжують багато додаткових вимог до технічних засобів (ТЗ). Крім пристроїв програмної обробки даних, необхідні спеціальні пристрої оперативного обміну інформацією, документування та архіву проектних рішень, зберігання інформації у вигляді бази знань, що відбиває накопичений досвід проектування. Питання організації спільного функціонування ТС в складі САПР утворюють важливий розділ теорії САПР. Ці питання стосуються не тільки апаратних засобів, але навіть в більшій мірі коштів програмного забезпечення (ПО) САПР.

Таким чином, автоматизація проектування мікроелектромеханічних систем включає всебі: **1.**Методологію САПР; **2.**МО, що об'єднує ММ, методи і алгоритми для виконання різних проектних процедур; **3.**Питання комплексування ТС і розробки спеціалізованої апаратури для САПР; **4.** Розробку і використання програмно-інформаційного забезпечення банків даних, пакетів прикладних програм (ППП), операційних систем ЕОМ; **5.** Технологію виготовлення мікроелектромеханічних систем.

Виділення АП як самостійного науково-технічного напрямку пов'язано з тим, що постановка і методи вирішення проектних завдань при автоматизованому і неавтоматизованому проектуванні істотно розрізняються. Предметом АП є формалізація проектних процедур, структурування і типізація процесів проектування, постановки, моделі, методи та алгоритми розв'язання проектних задач, способи побудови ТС, створення мов, опису програм, банків даних, а також питання їх об'єднання в єдину проектує систему.

Застосування ЕОМ для вирішення інженерних задач почалося відразу ж після появи перших ЕОМ. Однак це застосування вимагало від користувача трудомісткою підготовки завдань до вирішення, що полягає в математичному формулюванні завдання, виборі чисельного методу, розробці алгоритму і його

записи на одному з мов програмування. АПР відрізняється від подібного використання ЕОМ перш за все тим, що майже всі з перерахованих операцій автоматизовані і виконуються на ЕОМ за допомогою заздалегідь розробленого ПО, розрахованого на багаторазове застосування при вирішенні певного класу проектних завдань. Від користувача потрібно лише описати вихідні дані завдання на проблемно-орієнтованій мові і бути готовим до оцінки результатів і прийняття рішень за отриманими від ЕОМ відомостями.

У комплекс засобів АП, поряд з технічним, математичним і іншими видами забезпечення, входить ПО. Важливо відзначити, що програми в САПР розробляються не інженер-користувачами, а фахівцями з САПР і предметних областей. Програми розробляються раз, а застосовуються багаторазово в різних ситуаціях, що виникають при проектуванні багатьох об'єктів. Незважаючи на це, інженеру-користувачеві необхідно знати методи і алгоритми, реалізовані в програмах САПР, що допоможе уникнути помилок у формулюванні завдань, виборі вихідних даних, інтерпретації результатів і отримати їх з найменшими витратами загального і машинного часу.

Основна функція САПР - здійснення АПР на всіх або окремих стадіях проектування об'єктів та їх складових частин на основі застосування математичних та інших моделей, автоматизованих проектних процедур і засобів ОТ. Функціонування САПР повинно забезпечувати отримання проектних рішень, виконаних в заданій формі. Проектне рішення - опис об'єкта проектування або його складової частини, достатню для розгляду і ухвалення висновку про закінчення проектування або шляхах його продовження. Результатом проектування в САПР є задоволення заданим вимогам сукупність закінчених проектних рішень, необхідних для створення об'єкта проектування.

Сьогодні найбільш перспективним напрямком автоматизованого проектування є проектування мікроелектромеханічних систем, адже ріст виробництва таких систем складає 13% в рік (рис 1.2.1)

Прогноз ринку MEMC на 2011-2017 у вартісному вираженні

(Джерело: звіт MEMC промисловості, липень 2012, Yole Development)

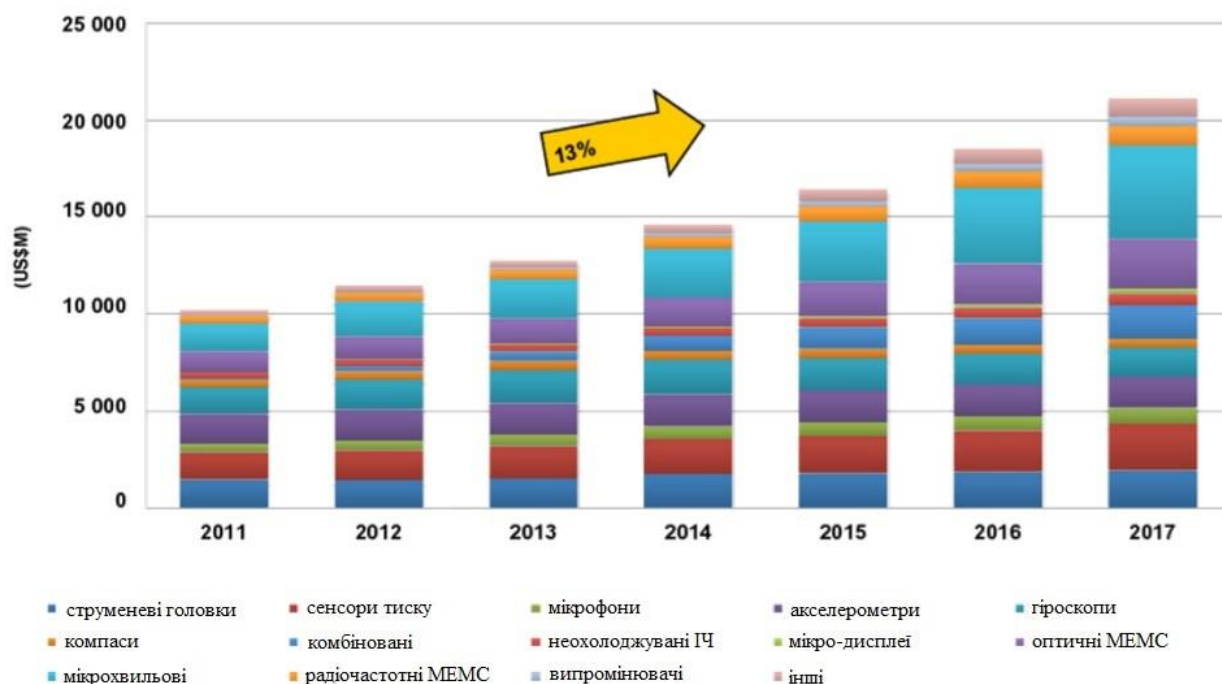


Рис.1.2.1 – Прогноз ринку MEMC

Проектування MEMC і мікросистем відрізняється від традиційного інженерного проектування тим, що необхідно здійснювати проектні рішення в різних предметних областях в єдиному технологічному циклі.

На рис.1.2.2 приведена структурна схема проектування MEMC. Слід зауважити, що для типового процесу проектування MEMC важливою складовою є моделювання технологічного процесу їх виготовлення (рис.1.2.3). При такому підході важливо врахувати статистичні характеристики технологічного процесу виготовлення MEMC і їх вплив на область працездатності елементів MEMC і пристроїв в цілому, це дозволить визначити найбільш впливові параметри.

На сучасному етапі розвитку технологічних процесів виготовлення пристроїв вбудованих систем велике значення відводиться статистичному аналізу параметрів і визначенню процента виходу працездатних пристроїв на етапі проектування. Все це пов'язано в повній мірі груповою технологією і

неможливістю здійснення корекції параметрів елементів пристроїв на етапі експлуатації.

Типова структурна схема проектування МЕМС

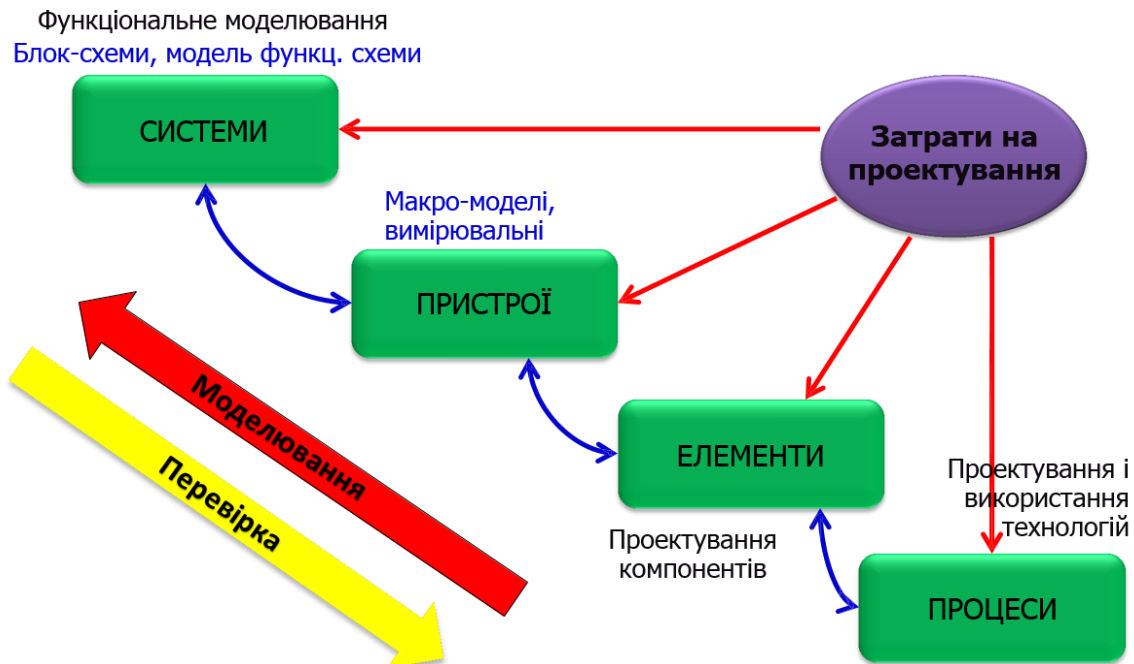


Рис.1.2.2

Наявність випадкових розподілів параметрів окремих елементів і кореляційних зв'язків [1.2] між ними, зумовлених особливістю виготовлення, суттєво впливає на величину і характер відхилень вихідних параметрів від заданих. У зв'язку з цим під час проектування вбудованих систем необхідно враховувати статистичний розподіл параметрів елементів, які визначаються технологією їх виготовлення.

Одноваріантний аналіз дозволяє отримати інформацію про стан і поведінку проектного пристрою в одній точці простору внутрішніх X параметрів і зовнішніх Q параметрів. Очевидно, що для оцінки властивостей проектного пристрою цього недостатньо. Потрібно виконувати багатоваріантний аналіз, тобто дослідити поведінку пристрою в ряді точок згаданого простору, яке для стислості надалі будемо називати простором аргументів.

Найчастіше багатоваріантний аналіз в САПР здійснюється в інтерактивному режимі, коли розробник неодноразово змінює в математичній моделі ті чи інші параметри з множин X або Q , виконує одноваріантний аналіз і фіксує значення вихідних параметрів. Подібний багатоваріантний аналіз дозволяє

оцінити область працездатності, ступінь виконання умов працездатності, а отже, ступінь виконання ТЗ на проектування.



Рис.1.2.3

Основою САПР MEMC і НЕМC є математична модель, яка формує множину математичних об'єктів і відношень між ними, адекватно відображаючи їх конкретні властивості. Причому для проектування MEMC і НЕМC часто використовують різні рівні абстракцій. Такий підхід називається багато масштабним моделюванням. В залежності від послідовності процедур та етапів проектування, розрізняють низхідний (обернена послідовність проектування), визхідний (від нижчого рівня до вищого) та змішаний етапи проектування. В цілому САПР MEMC і НЕМC базується на ітераційному підході, коли наближення до кінцевої мети здійснюється шляхом багатократного виконання одної і тої ж послідовності процедур. Ітерації можуть охоплювати як декілька операцій, так і декілька етапів [1.22].

Основні концепції САПР MEMC і НЕМC методом міждисциплінарного підходу. Математичними моделями на міждисциплінарному рівні є диференційні та інтегральні рівняння, що описують поля фізичних величин, тобто на цьому рівні використовуються моделі з розподіленими параметрами, базовими серед яких є просторові декартові координати та час. До таких відносять рівняння

математичної фізики з крайовими умовами, які включають також початкові умови, що характеризують просторовий розподіл залежних змінних в початковий момент, і граничні умови, що задають значення цих змінних на границях розглядуваної області [1.22].

Такі задачі переважно розв'язуються чисельними методами шляхом дискретизації незалежних змінних у вузлах (сіткові методи). Класичними чисельними методами є метод скінчених різниць; метод скінчених елементів та метод граничних елементів. Перший із цих методів базується на реалізації різницевого оператора, який відповідає вихідному диференційному рівнянню. Другий – базується на розбитті досліджуваного тіла на ряд під областей простішої геометричної форми (скінчені елементи). В основу моделі покладений варіаційний принцип механіки – принцип мінімуму потенціальної енергії. Альтернативним розглянутим методам є інтегральний метод.

МЕМС і НЕМС, що інтегрують для свого функціонування різні хімічні і фізичні явища (електричні, механічні, гідравлічні, теплові, тощо), відносяться до класу фізичних неоднорідних. Фізична неоднорідність МЕМС і НЕМС суттєво ускладнює процедуру моделювання, а, отже, і синтезу, тому розробка ефективних методів і програмних засобів моделювання є фундаментальною науковою задачею.

З цією метою розроблені спеціалізовані системи моделювання і проектування МЕМС – CoventorWare, EM3DS (Coventors, Inc.) IntelliSuite (IntelliScience Corp.), MEMCAD ((Microcosm Technologies Inc.) MEMS Pro (Tanner Research Inc.) NODAS (Carnegie Mellons University MEMS LAB). При моделюванні на мікрорівні використовуються універсальні програми ANSYS (ANSYS, Inc.) NASTRAN (MSC. Software Corp.) Abaqus (Dassault Systems), COMSOL Multiphysics (раніше Femlab, Comsol Inc.), які дозволяють розв'язувати задачі механіки деформованого тіла і механіки конструкції, механіки рідини і газу, теплопередачі і теплообміну, електродинаміки і акустики, тощо.

Для моделювання МЕМС на макрорівні використовуються програмні засоби: 1.Схемотехнічного моделювання, де модель неелектричної частини МЕМС будується на базі методу аналогій; 2.Загальноматематичного призначення

(SUGAR (University of California, Beerkly), що має в собі бібліотеку MEMC для системи Matlab. Перераховані вище програмні засоби дозволяють дослідити функціональні моделі неоднорідних MEMC на макрорівні, але не орієнтовані на розв'язування задач моделювання приладів та пристроїв MEMC в рамках САПР наскрізного проектування, що зумовлено реалізованими в них методами моделювання.

В нановимірних об'єктах змінюється традиційне співвідношення між силами, вимагається врахування сил в'язкості, поверхневого натягу, сухого тертя, електростатичні і магнітні сили. Відповідно для аналізу фізичних процесів вимагається моделювання мультіфізичних середовищ, використання методів і професійних мов різних предметних областей. В зв'язку з цим, часто йдуть шляхом використання САПР і методу фізичних аналогій, незважаючи на те, що даний підхід трудомісткий.

Розроблений фірмою IntelliScience Corp САПР IntelliSuite має широку сферу застосування – від проектування НЕМС та створення моделей процесів, до розробки технологічного маршруту його сегментів різного призначення. Пакет програм має декілька розділів: IntelliFab, AnisE, Mematerial, IntelliMask. Догий час цей комплекс був стандартним інструментом розробки НЕМС широкого призначення.

І на завершення відзначимо, що використання MEMC і НЕМС приводить до масштабного ефекту, який проявляється, з одної сторони, в зростанні ролі тертя і погіршення тепловідведення в мікропристроях, з іншої – в підвищенні міцності деталей. Масштабний ефект в MEMC може бути зумовлений близькістю характерних розмірів структурних складових матеріалу деталей, самих деталей. Так модуль зчеплення надмініатюрних зубчатих передач стає спів вимірним з розмірами надмолекулярної структури самих матеріалів, що приводить до зростання ролі поверхневих сил тертя і адгезії в порівнянні з об'ємними (наприклад, інерційними) силами і, природно, до зниження надійності самих деталей. Тому практично всі САПР доповнюються модулями для проведення статистичних досліджень впливу дестабілізуючих факторів на надійність MEMC і НЕМС.

Інтегровані пакети автоматизованого проектування MEMC і НЕМC та САПР ANSYS. В даний час розроблені різні пакети автоматизованого проектування MEMC і НЕМC, що дозволяють виконати проектування на всіх основних етапах. Одним із таких пакетів є пакет програм CoventorWave, який включає в себе модулі: Architect (для візуалізації MEMC і НЕМC, дослідження перехідних процесів в них та модальний аналіз, наявна бібліотека базових елементів); Desinger; Analyzer; Memulator; Integrator, Etch3D. Модуль моделювання дозволяє провести розрахунок робочих точок по постійному струму, провести аналіз перехідних процесів, частотний аналіз та аналіз передвальних функцій по постійному струму та чутливість вузлів. Наявний генератор топології дозволяє провести генерацію топології розробленої MEMC і НЕМC.

Розроблений інтегрований пакет автоматизованого проектування MEMS Pro із модулів MEMS Modeler; L-Edit і Desing Rule Cheker. Програма MEMS Modeler дозволяє отримати поведінські моделі. В цілому САПР MEMS Pro - це інтегрований пакет проектування MEMC і НЕМC, який призначений для проектування: механізмів; інерційних блоків; електромеханічних структур; п'єзоелементів підтримує формат GDSII, CIF та допускає обмін моделями програмним комплексом ANSYS. Розроблений також інтегрований пакет САПР MEMC і НЕМC IntelliSuite фірми IntelliSense Software (США). Дані САПР підтримують маршрут проектування: мікропристроїв (акселерометрів, гіроскопів, запам'ятовуючих пристроїв, тощо); мікродвигуни, зубчаті передачі, тощо; електромагнітні MEMC; НВЧ MEMC; актюатори (теплові, механічні, електростатичні, п'єзоелектричні, тощо). Для моделювання використовується комбінація методу скінченних елементів і методу граничних елементів.

Підвищення функціональності вбудованих систем вимагає застосування спеціалізованих обчислювальних систем, яким є ANSYS, що дозволяє проводити розрахунки в напрямках електромеханічного, електротепломеханічного, високо- і низькочастотного електромагнітного та акустомеханічного та п'єзоелектричного аналізів. Розрахунок вбудованих систем з використанням ANSYS зводиться до трьох етапів: 1. Розташування пристроїв, їх структура розробляються

спеціалізованими програмними пакетами Vulcain і MEMSPro, де моделі пристроїв створюються прогнаною L-Edit. 2. Розроблена модель імпортується в розрахункове середовище ANSYS, формулюються граничні умови, модельований об'єкт розбивається на скінченно-елементну сітку з подальшим розв'язуванням задачі. 3. По завершенню процедури розрахунку, розробник має можливість повторно ввести розроблені пакетом параметри знову в обчислювальне середовище вручну, якщо обчислення проводяться в Vulcain чи за допомогою програми T-Spice, якщо обчислення проводяться в MEMSPro.

1.3. Базові засади САПР проектування інтелектуальних і вбудованих систем на їх основі [1.13-1.29].

Мікроелектроніка сьогодні розвивається надзвичайно швидкими темпами. Розробляються нові технології і зразу ж з'являються нові інструментальні засоби для їх проектування. Один з швидко прогресуючих напрямків в галузі мікроелектроніки на сьогодні – вбудовані системи (ВС).

Основною особливістю вбудованих систем є те, що вони суміщають в собі електронну та механічну частини. Одним з основних компонентів вбудованих систем є датчі (мікросенсиори), модуль обробки інформації (мікропроцесор) та виконуючі механізми (актюатори).

Якщо сучасний етап розвитку твердотіЛЬНОЇ мікроелектроніки дає змогу вирішити всі проблеми, які виникають під час розробки та виготовлення модуля мікропроцесорної обробки інформації, то цілком інша ситуація спостерігається під час проектування та виготовлення сенсорів і актюаторів. Передусім це визначається тим, що всі складові вбудованої системи доцільно виготовляти в одному технологічному процесі. По-друге, сенсори та актюатори можуть містити складові, які виконують значні механічні переміщення в часі. По-третє фізичний принцип роботи цих пристроїв базується на стику наукових галузей: механіки, електрики, біології, мікроелектроніки, оптики тощо.

Тому процес проектування вбудованих систем має свої особливості, на відміну від технології проектування інтегральних схем (IC), яка на сьогодні, досить добре відпрацьована і використовує сучасні програмні засоби, що

дозволяють швидко та ефективно організовувати проектування та виготовлення нових інтегральних пристроїв.

Під час проектування вбудованих систем використовується класичне низхідне проектування, тобто проектування “зверху – вниз”. Цей підхід передбачає, що проектування ВС розбивається на розв’язання окремих підзадач. Для прикладу, на проектування функціонально довершених модулів вбудованої системи, а саме: сенсорів, мікропроцесорного модуля, виконуючих пристроїв тощо. Подальша декомпозиція проектування передбачає розбиття задач проектування сенсорів та актюаторів на задачі проектування елементів. Це є задачі, пов’язані з проектуванням статора чи ротора мікродвигуна, пружини чи джерела живлення електростатичного актюатора. Наступний крок декомпозиції передбачає розбиття задач проектування компонентів ВС на задачі проектування елементів вбудованих систем. Тому структурна схема проектування вбудованих систем, яка показана на рис. 1.3.1, визначає чотири рівні проектування. Перший рівень називається системним, який призначений для розробки технічного завдання на ВС, другий – структурно-функціональним, третій – компонентний рівень проектування вбудованої системи, а четвертий – елементний.

Цей підхід до проектування вбудованих систем має низку переваг, а саме: з простішими об’єктами проектування зручніше працювати, побудувати математичну модель даного об’єкта, провести моделювання його роботи, виконати верифікацію та тестування результатів проектування тощо. Структурна схема проектування та виготовлення вбудованих систем наведена на рис.1.3.2.

Якісно оцінюючи складність систем, розрізняють великі системи, вбудовані системи, підсистеми (субсистеми), комплекси, пристрої. При цьому чіткої границі в проведенні класифікації не існує. В одному випадку комплекс може бути системою, а в іншому – пристроєм залежно від його функціонального призначення. Це свідчить про таке: система – поняття відносне і може використовуватися у всіх випадках проектування, коли для прийняття ефективного рішення необхідно враховувати взаємодію багатьох факторів і обмежень.

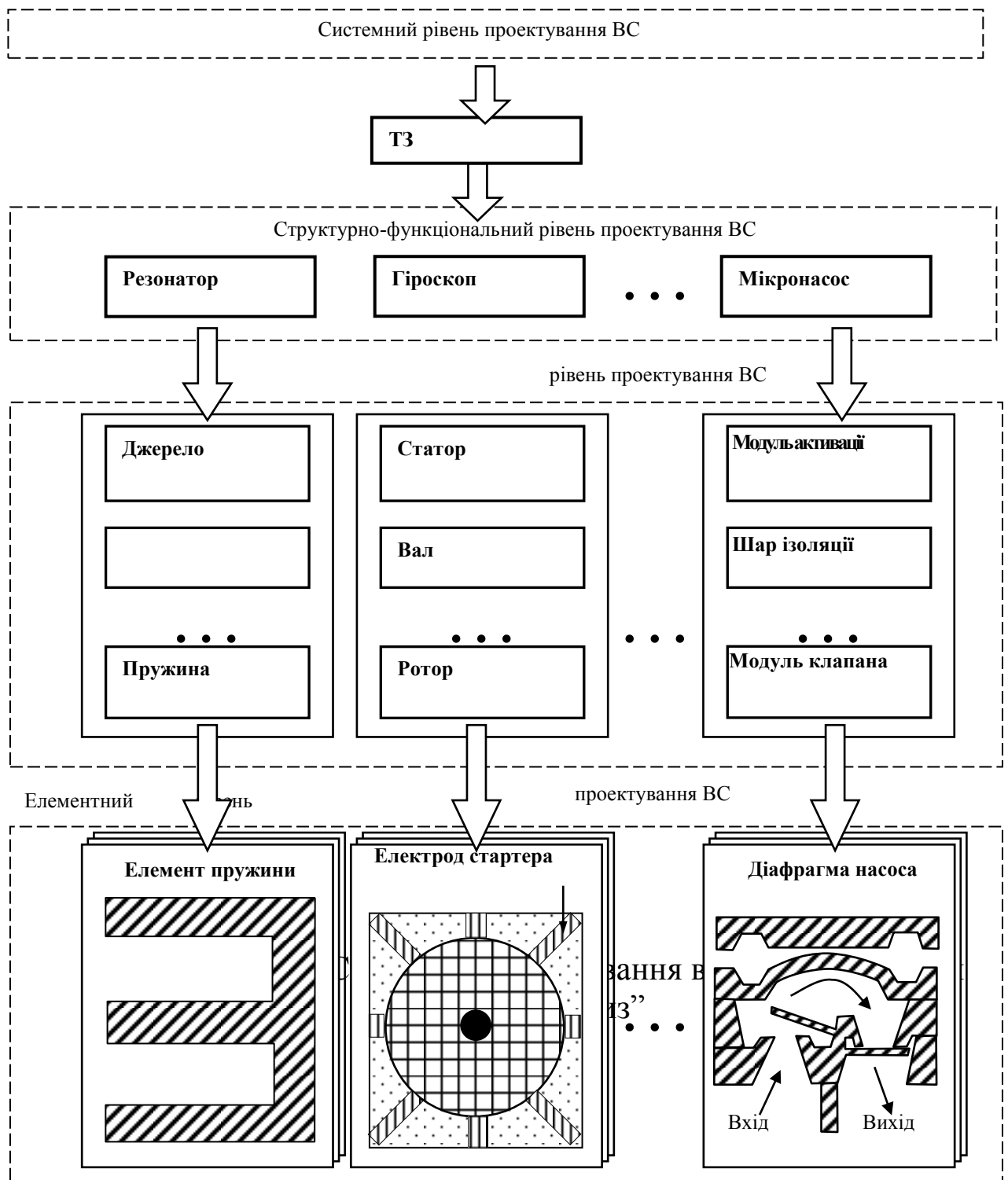


Рис.1.3.1

Системи автоматизованого проектування (САПР) вбудованих систем на основі MEMS і NEMS охоплює всі основні питання автоматизації їх проектування. Методи проектування NEMS суттєво відрізняються від методів проектування традиційних електронних систем. Це пов'язано з тим, що кількість компонент суттєво зростає, а типові розміри елементів наближаються до нанометрів, що вимагає використання не лише потужних обчислювальних

система, але й складних систем проектування. Тенденції до підвищення їх функціональної складності при одночасному зменшенні розмірів створюваних сьогодні мікро- і наносистеми мають – це не лише технічна вимога, але й економічна.



Рис.1.3.2

Більшість інструментальних засобів САПР для проектування вбудованих систем, на першому етапі його розвитку, використовувались для розв’язання окремих підзадач проектування, зокрема: тривимірний аналіз (OYSTER, EMCAD), визначення різних характеристик сенсорів (SENSIM, CAEMEMS) чи актюаторів. Вони не були завершеними САПР повного циклу проектування вбудованих систем. З кінця 90-х років минулого століття ситуація на ринку програмного забезпечення проектування вбудованих систем суттєво змінилась – з’явилися пакети, які підтримують процес наскрізного проектування. Це такі, як система проектування MEMC MEMSCAP фірми Mentor Graphics, Tanner Tools

MEMS фірми Tanner Research, Sc-Smash фірми Dolphin та MEMCAD [1.21] Масачусетського технологічного інституту (США).

САПР інтелектуальних систем. В 1950 р. британський математик А. Тьюрінг надрукував в журнали «Mind» свою працю «Обчислювальні машини і інтелект», в якій описав тест для перевірки програми на інтелектуальність. Він запропонував дослідника і програму розташувати в різних кімнатах і до тих пір, доки дослідник не визначить, хто за стіною – людина чи програма, вважати поведінку програми розумною. Це був один з перших запропонованих тестів для перевірки програми на можливість генерування інтелектуального розв’язку задачі. В 1956 р. на семінарі в Дортмундському коледжі (США) був запропонований до вжитку термін «штучний інтелект» (AI – Artificial Intelligence). В подальшому розвинувся цілий науковий напрям інтелектуальної діяльності людини – штучний інтелект, інтегрований в ланцюг – система автоматизованого виробництва, робото технічний комплекс і електронно-обчислювальна система. По способу адаптації, такі інтелектуальні комплекси стали поділяти на самонастроюючі, самонавчальні та самоорганізуючі. Інтелектуальні технологічні системи виробництва мають в собі змінну, настроювану модель і реальну виконавчу систему з об’єктом керування. Сама інтелектуальна технологічна система складається із декількох елементів, кожний із яких є, за свою суттю, окремим науковим напрямком: інтелектуальне програмування і інтелектуальний інтерфейс; система адаптивного пошуку логічного висновку і мовні засоби; система подання знань і формування понять; система логічного аналізу та розпізнавання; система штучного зору і розпізнавання. Проектування складних інтелектуальних систем вимагає розробку інтелектуальних САПР. Аналіз САПР проектування сенсорної мережі [1.24] свідчить, що створення інтелектуальних САПР є актуальним завданням і може слугувати ефективним засобом проектування складних систем. Інтелектуальна САПР будувалася як сукупність кластерів, які формуються класичними алгоритмами кластеризації, як алгоритм k – середніх чи алгоритми ієрархічної кластеризації або алгоритми кластеризації методами теорії графів. Показано, що

потреба створення інтелектуальних САПР є особливо відчутною у випадку, коли має місце неповнота і неточність інформації.

САПР «інтелектуального будинку». Розумний будинок чи офіс – це автоматизована система керування, що призначена для контролю та керування різноманітними функціональними діями, закладеними в житловий чи офісний комплекс дистанційно у відповідності до алгоритмів використовуваного програмного забезпечування. Розглядувана система динамічна, оскільки підлягає активному перепрограмуванню функціональних дій, в залежності від вимог, що формуються у вигляді вхідних завдань системи, забезпечуючи при цьому режим узгодженої роботи різних ділянок системи [1.25]. Для проектування «інтелектуального будинку» єдиного САПР не існує. Так для створення проектної документації найбільш часто використовується програма програма AutoCAD. На аналогічному рівні використовується програма КОМПАС компанії АКСОН чи пакет інтегрованого САПР – Solid Works.

Нижче в таблиці наведений порівняльний аналіз декількох САПР.

Опис	AutoCAD	Solid Works	PCAD	OrCAD
Робота з планами поверхів будинку	+	+	-	-
Можливість спільної роботи з іншими напрямками проектування	+	+	-	-
Наявність бібліотеки шаблонів	+	+	+	+
Можливість розширення функцій пакету за рахунок підключення сторонніх модулів	+	+	+	+
Генерація креслень і проектної документації на фінальному етапі, який не вимагає доведення іншими програмами	+	+	-/+	+
Можливість експорту документації в сторонньому форматі	+	+	+	+
Можливість виконання всієї проектної документації без використання сторонніх програм	+	+	-	+
Можливість оптимізації проектного вирішення засобами пакету	-	-	+	+
Можливість моделювання поведінки системи	-	-	-	-

Як випливає із таблиці, не можна надати переваги жодній із САПР, оскільки жоден із них не розроблявся і не адаптовувався з метою цілісного проектування «інтелектуального будинку».

2. Базові процеси і моделі елементів мікросистемної техніки

2.1. Базові технології мікро- і наносистемної техніки

Переважаю MEMC виготовляються груповою технологією, в якій одночасно опрацьовується велика кількість елементів. Її спрощений маршрут поданий на рис.2.1.1. Кожний виріб MEMC проходить три стадії: розробка, що включає в себе аналітичне математичне моделювання за допомогою пакетів прикладних програм ANSYS, Solid Works і комп'ютерне моделювання з використанням програмних продуктів PTC, Mentor Graphics; виготовлення – кристалльне продукування мікроелементів і мікросхем, збірка модулів систем, тестування і випробовування та температурне і кліматичне випробування та вихідний контроль. Виготовлення МСТ в основному здійснюється двома об'ємною і поверхневою кремнієвими технологіями. Об'ємна – це процес створення електронних приладів на основі анізотропного травлення кремнію в масці із оксиду або нітриду кремнію, а та золота і хрому. Для одержання об'ємних 3D-структур, використовують глибоке сухе анізотропне травлення, як реактивно-іонне газовою плазмою або здійснюють монтаж типу кремній-кремній чи кремній-скло.

На поверхні кремнієвої підкладки орієнтацією (100) створюється легований шар p -типу провідності, на поверхню якого наноситься епітаксій на плівка кремнію. В подальшому шляхом окиснення створюються діелектричні шари SiO_2 , після чого метод літографії і травлення формується топологічний рисунок. Завершується технологічний процес видаленням решток матеріалу шляхом травлення. Недоліком об'ємної мікрообробки є залежність геометрії мікроструктури від кристалічної структури підкладки.

Тому для формування MEMC-структури використовують або глибоке анізотропне сухе травлення, або з'єднують між собою дві підкладки типу кремній-кремній чи кремній-скло, на кожній із яких виготовлена MEMC-структура і мікросхема, що дає можливість створювати складні тривимірні MEMC-елементи.

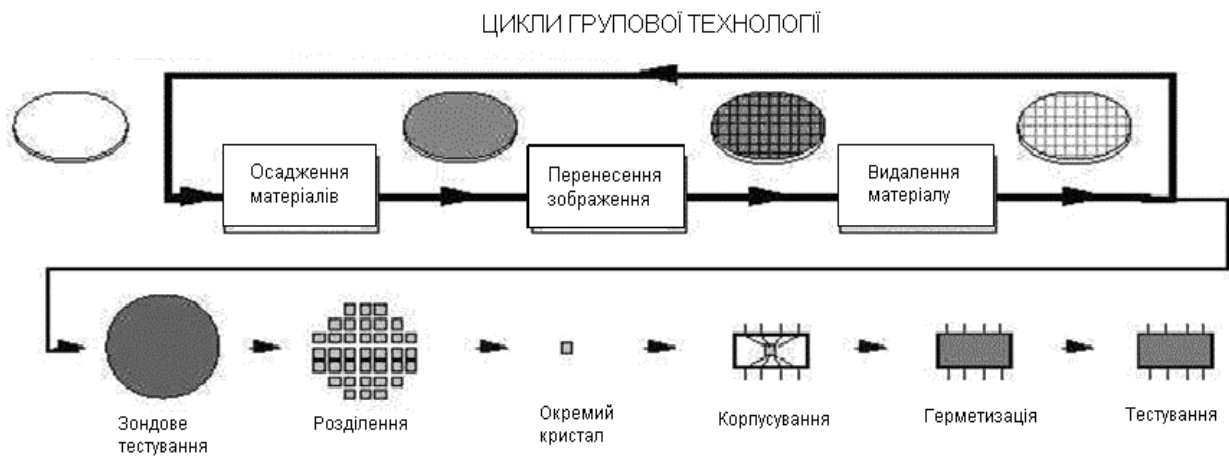


Рис.2.1.1 [2.1]

В технології поверхневої обробки, структури створюються на поверхні кремнієвої підкладки, на яку наносяться плівки полі кремнію чи оксиду кремнію. Технологічна лінійка етапів кремнієвої поверхневої обробки зображена на рис.2.1.2,а. Після створення структури, рештки шарів видаляються шляхом хімічного травлення. Підкладка в поверхневій технології відіграє роль механічної основи. Використання структур типу полі кремній-оксид кремнію дозволяє мінімізувати механічні напруги в шарах. Серед MEMC -технологій поширення набули:

1.EFAB (аббревіатура від слів *Electrochemical FABrication*)-технологія. Вона ґрунтується на гальванічному осадженні металів на ізолюючі поверхні, з подальшим травленням структур в ізоляційному матеріалі, що дозволяє створювати, в тому числі тривимірні мікроструктури. Основними розробниками цієї технології є *Information Science Institute (ISI)* та *University of Southern California*. EFAB –технологія не вимагає надчистих приміщень, повністю автоматизована та має меншу кількість технологічних етапів.

2.LIGA (аббревіатура від слів *Litographie* – літографія та *Galvanoformung*-гальванообробка) - технологія. Ця об'ємна технологія і розроблена фахівцями німецького Центру ядерних досліджень (Карлсруе) з використанням жорсткого випромінювання синхротрона, прецизійного лиття полімерів заданої форми та гальванічного осадження металів на мікроповерхнях. Випромінювання синхротрона має малу кутову розбіжність і спроможне проникати в глибину полімеру на декілька міліметрів, що дозволило створювати об'ємні структури.

Лінійка **LIGA**-технології наведена на рис.2.1.2,б. Для її реалізації використовується рентгенове випромінювання від синхротрону. Глибока його проникність дозволяє створювати структури в товщі полімерного матеріалу декількох міліметрів. LIGA-технологія дозволяє використовувати інші матеріали, як кераміку, метали, полімери, скло і на їх створювати тривимірні MEMC-структури

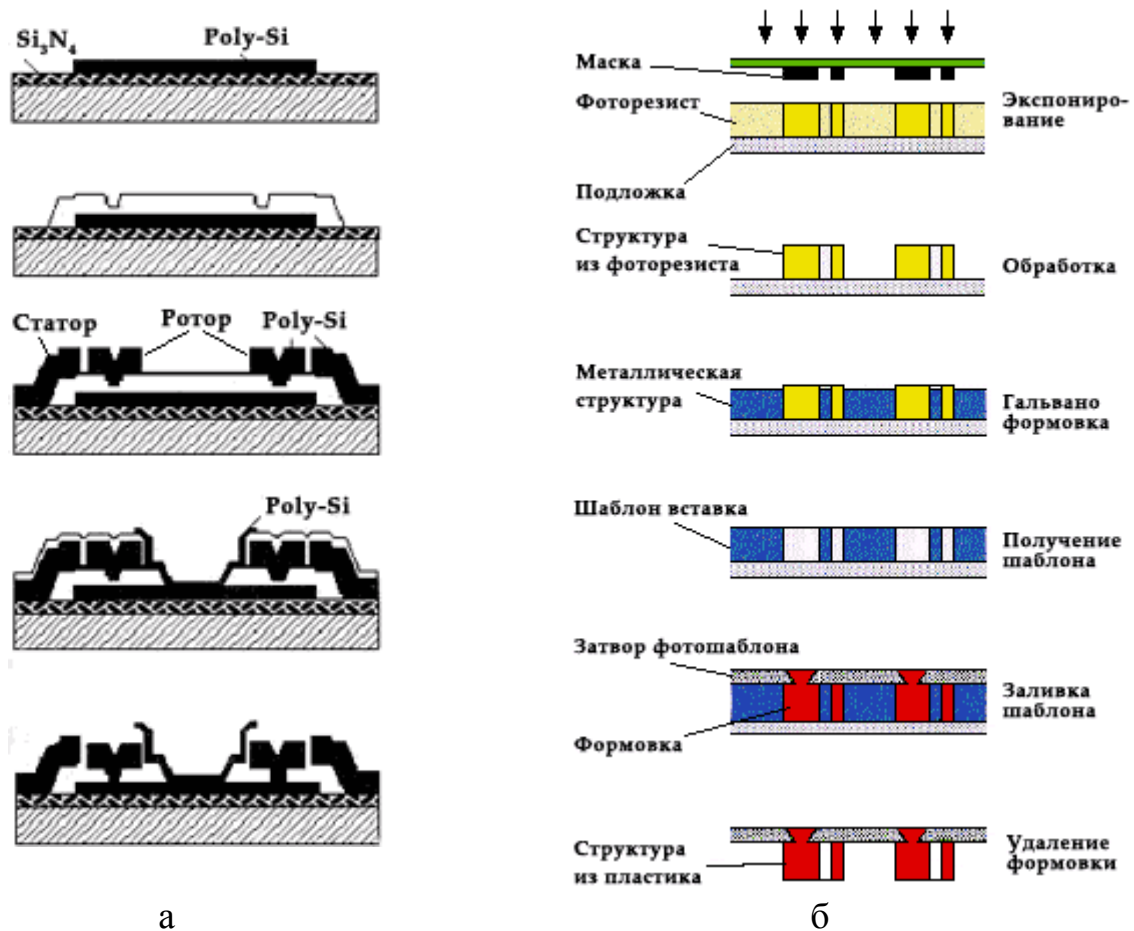


Рис.2.1.2 [2.2].

3.SUMMiT (аббревіатура від слів *Sandia Ultra-planar Multi-level MEMS Technology*)-технологія. Це поверхнева технологія виготовлення (М,Н,О)ЕМС і ґрунтується на формуванні чотирьохшарової полікристалічної кремнієвої механічної структури, де перший нерухомий шар (кремнієва підкладка) утворює механічну і електричну основу для решта трьох рухомих шарів. SUMMiT технологія легко інтегрується в кремнієву ІС. Механічні структури MEMC систем створюються методами тонкоплівкової фотолітографії та хімічного травлення. Розроблена п'ятишарова SUMMiT-V технологія, спроможна створювати 24-битні механооптичні замки.

4. MUMPs (Multi User MEMS Process) –технологія. Це тришарова поверхнева технологія полікремнієвого процесу виготовлення MEMC. Вона складається із створення ізоляційного шару Si_3N_4 , шару полікремнію, на поверхні якого формуються рухомі частини із двох інших структурних шарів полікремнію. За потреби наносяться шари для виготовлення масок фотолітографії для створення сенсорів і актюаторних елементів та шар металізації. Недоліком цього процесу є неможливість виготовлення елементів MEMC із іншого матеріалу окрім полікремнію.

Елементи нанотехнологій

Класифікація нанотехнологій наведена на рис.2.1.3 [2.3]. Базовими, на яких вони ґрунтуються, є фізичні, хімічні і біологічні процеси, які реалізуються за принципами «згори-вниз» та «знизу-вгору».

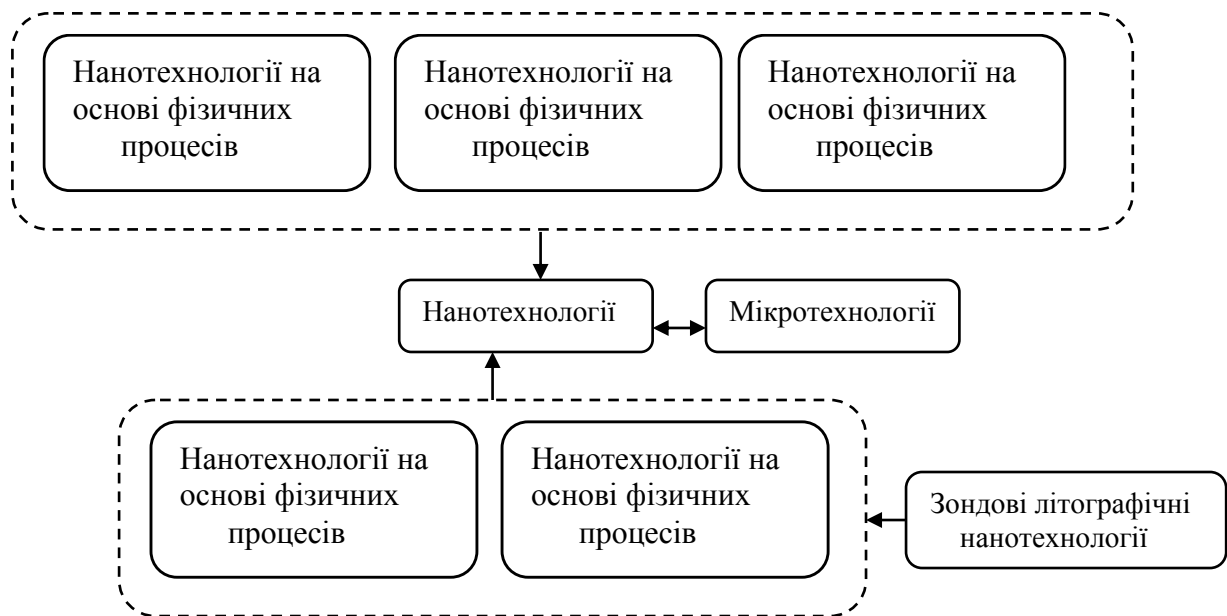


Рис.2.1.3

Атомна інженерія наноструктур. Нанотехнологічні можливості скануючої тунельної мікроскопії можна використовувати для керованої локальної модифікації поверхні з атомарною роздільною здатністю. Цьому сприяють екстремальні умови, в яких перебувають поверхні зонда і зразка, а саме високі електричні поля до 10^9 В/м та здатність пропускати великі струми до 10^7 А/см^2 , що дозволяє здійснювати локальне окислення і осадження на рівні маніпулювання одиницями атомами. На відміну від інших технологічних процесів, процес атомної інженерії поверхні представляється можливим вивчати в режимі

тенульної спектроскопії. Причому переміщення атомів вздовж поверхні можна здійснювати як за рахунок дифузії електричним полем між зондом і поверхнею зразка, так і паралельного ковзання силами взаємодії атомів зонду з взаємодіючими атомами. Величина цих сил визначається потенціалом міжатомної взаємодії. В обох випадках зв'язки між пересуваючими атомами і підкладкою не розриваються. Адсорбований атом завжди перебуває в потенціальній ямі. Енергія, яка необхідна для пересування атома, відповідає енергетичному бар'єру для дифузії по поверхні.

В атомній інженерії можуть мати місце так звані перпендикулярні процеси, тобто переносом атомів між поверхнями зонда і зразка. В цьому випадку ефективність процесу залежить від відстані зонда до поверхні зразка і в залежності від ефективності дії електричного поля маніпуляція атомами може бути здійснена за допомогою контактного переносу, польового випаровування і електроміграції. У першому випадку зонд у відсутності електричного поля наближають до поверхні зразка настільки, щоб від скотактував з атомом, потім його з цим атомом відводять від поверхні і, при потребі, переносять в іншу ділянку. Під час польового випаровування атом з поверхні досліджуваного зразка переноситься на поверхню зонда електричним полем. В електроміграційному процесі потік носіїв струму захоплює атоми внаслідок передачі їм частки кінетичної енергії.

Нанолітографія. Зменшення довжини світлової хвилі, використання Х-променів та електронного променя дозволяє створювати літографічні рисунки з розмірами елементів менше 100 нм. Особливо перспективною в цьому плані є електронно-променева літографія, яка дозволяє досягати розмірів сфокусованого п'ятна розміром 1-1.5 нм. Таким чином створюються маски з розмірами елементів в діапазоні 10-1000 нм. Визначальною характеристикою літографічного процесу з точки зору організації виробничого процесу є забезпечення продуктивності експонування не менше $1 \text{ см}^2/\text{с}$. Нагадаємо, що можливості оптичної літографії задовольняють 10-100 $\text{см}^2/\text{с}$, однак хвильова природа світла накладає так звані дифракційні обмеження. Джерела короткохвильового синхротронного випромінювання надто громіздки, тому перспективні в цьому плані плазмові

джерела лазерного випромінювання, рентгенівська та електронно-променева літографія, яка дозволяє досягти роздільної здатності до 5 нм.

Застосування нанолітографії на основі АСМ і СТМ дозволяє отримувати релаксаційні квантові діоди і транзистори, робота яких ґрунтується на ефектах розмірного квантування і релаксації нерівноважних електронів. Для цього розміри приладів повинні бути менші деякого критичного значення – довжини розмірного квантування [2.4] $L = \sqrt{3\hbar^2 / 2m^*k_B T}$, де m^* - ефективна маса електрона в об'ємі кристалу.

Напрофілювання. Суть цього перспективного процесу створення наноструктур полягає в тому, що оксидні плівки SiO_2 товщиною менше 1 нм опромінюються гостро сфокусованим електронним променем при кімнатній температурі. Оксид в опромінених ділянках розкладається і випаровується при подальшій термообробці у вакуумі при температурі 720-750°C. Такий процес багато в чому нагадує звичайне електрохімічне анодне окислення, коли на зонд подається від'ємний потенціал. Вологість оточуючої атмосфери відіграє роль електроліту. Внаслідок капілярного ефекту і сильного електричного поля вода конденсується на кінчику вістря. Там молекули води дисоціюють :

$\text{H}_2\text{O} \leftrightarrow \text{H}^+ + \text{OH}^-$. Електричне розділяє ці іони, скеровуючи OH^- до поверхні зразка, в результаті чого вона окислюється.

Нанодрук. Використовуються два методи: класичний чорнильний друк рисунків та механічне притискання штампу, поверхня якого покривається фарбою. Ці методи позбавлені таких недоліків як дифракція та розсіяння хвиль.

Саморегульовані нанопроцеси. Мікросистеми можна поділити на ізольовані і відкриті. В ізольованих процеси відбуваються у відповідності законів термодинаміки і рано чи пізно, вони прямують до стану термодинамічної рівноваги. Нагадаємо, що в цьому стані мікросистеми «забувають» свою передісторію, відповідна інформація збережена лише в законах збереження, як сталість повної енергії, імпульсу, маси системи і тощо.

Процес самоорганізації є властивістю відкритої системи. Вважається, що джерелом впорядкування відкритих систем є їх нерівноважність [2.5].

Саморганізація може виникнути із флуктуацій, має пороговий характер і зв'язана з формуванням структури більш складної, ніж початкова. Процес самоорганізації представляється як перехід від хаотичного руху, хаотичного стану через інтенсифікації флуктуацій до нового порядку. Причому, це зовсім не новий статичний порядок, а динамічний, який підтримується підведенням до системи енергії та маси з зовні. В самоорганізованих системах проявляються колективні ефекти, як наприклад, метод хімічної нанообробки поверхневих наноструктур. Саморегулювання відноситься до одного з найбільш організованих процесів в природі, в результаті якого система приходить до найбільш стійкого стану. В нанотехнології також реалізовані процеси самоорганізації наноструктур, причому розроблені два підходи – самовпорядкування і самозбірка.

Самовпорядкування – це процес адсорбції і специфічного розташування молекул на твердій поверхні. Основним процесом в цьому є хемосорбція, яка супроводжується високоенергетичними реакціями між адсорбентом та адсорбованою поверхнею, тоді як взаємодія між самими молекулами залишається слабкою. Такі плівки мають надзвичайно малу густоту дефектів, досить стабільні, мають значну механічну міцність, тому їх використовують в якості трафарета в літографічному процесі. В якості груп, прикріплюючи весь молекулярний блок до поверхні підкладки, частіше всього використовують сілани

$RSiX_3$ ($R = CH_3, C_2H_5, \dots$), для того щоб прикріпитися до гідроксильних груп OH .

Використовують інші амінові групи NH_2 , які здатні приєднувати до себе галогени (хлор, йод та інш.). Поверхні, покриті алкільними групами, інертні та гідрофобні. За хімічною активністю вони нагадують парафін.

Самозбирання або самоорганізація – це певне розташування взаємодіючих атомів в твердому тілі. Такий процес ефективний для само зборки квантових точок та квантових шнурів. Під час само збирання атомна структура намагається прийняти конфігурацію, що відповідає її потенціальної енергії. До таких процесів в твердих тілах відноситься процес спонтанної кристалізації. Під час кристалізації, на відміну від аморфного стану, атоми чи молекули формують впорядковану кристалічну структуру.

Таблиця 2.1.1 [2.7]

Параметр	Фізичне обмеження
Мінімальна величина одного елемента – 0.3 нм	Статистичні флуктуації легування підкладки, роздільність фоторезиста, космічні промені і радіоактивність, скінчена ширина $p-n$ -переходу
Товщина підзатворного діелектрика – 2.3 нм	Тунельні струми крізь діелектрик
Мінімальна напруга живлення – 0.025 В	Тепловий потенціал $k_B T / q$
Мінімальна густина струму – 10^{-6} А/см ²	Дискретність заряду електрона, флуктуація вбудованого заряду
Мінімальна потужність – 10^{-12} Вт/елемент для частоти $\nu = 1$ кГц	Шуми, теплова енергія, діелектрична стала
Гранична швидкодія – 0.03 нс	Швидкість світла
Максимальна напруга живлення	Пробій під затворного діелектрика, змикання областей витоку і стоку
Максимальна густина струму	Електроміграція, спад напруги на паразитних опорах контактів
Максимальне легування підкладки	Тунельний пробій $p-n$ -переходу стоку
Максимальна потужність	Теплопровідність підкладки і компонентів схеми
Кількість елементів на кристал - 10^9	Сукупність всіх перелічених раніше обмежень

Фізичні і структурні обмеження кремнієвої технології [2.6]. Довжина хвилі де Бройля для електрона в зоні провідності складає порядку 5 нм. Товщина під затворного діелектрика із SiO_2 за технологічними можливостями не може бути меншою ніж 1 нм в два молекулярних шари, оскільки за менших значень посилюється тунелювання електронів. Її технологічна межа оцінюється в 2 нм. Мінімальна довжина затвору визначається підсиленням транзистора і тунельним струмом стік-підкладка. Якщо концентрація легуючої домішки в підкладці висока, то тунельний струм досягає величини струму каналу. Якщо концентрація легуючої домішки в підкладці низька, то ОПЗ витоку і стоку перекриваються, напруга стік-витік відкриває $p-n$ - перехід витіку і крізь підкладку тече струм, обмежений ОПЗ. Підсилення транзистора при цьому зменшується. Тому довжина затвору складає $14 \div 18$ нм. Ємність затвору

мінімального за розмірами 10×10 нм транзистора порядку $2 \cdot 10^{-18}$ фарад. Тому для робочої напруги $0.8V$ і пороговій напрузі $0.2V$ в каналі не може перебувати більше 6 електронів.

Швидкодія транзистора визначається ємністю затвору і крутизною ВАХ, тому його значення обмежене дробовістю характеру струму із-за квантованості заряду електрона. Крім цього при зменшенні розмірів актуальними стають ефекти радіаційного пошкодження космічними променями і розпадом радіоактивних атомів в матеріалах. Не можна також зменшувати напругу живлення нижче $0.5V$, оскільки контактна напруга не масштабується. Обмеженою швидкістю світла є швидкість поширення сигналу, яка крім цього обмежується реактивністю контура $R-L-C$ транзисторної структури.

Для МДП-транзисторів фізичні обмеження мікромініатюризації згруповані у вигляді табл.2.1.1. Для надтонких МДП структур із товщиною діелектрика ~ 1 нм можливе використання лише SiO_2 , який забезпечує низьку густину поверхневих станів на межі поділу. Для діелектриків із товщиною $1,5$ нм і більше можливе використання двохшарових структур, в яких нижній шар SiO_2 , а верхній Si_3N_4, Al_2O_3 або інший діелектрик із більшим значенням діелектричної проникності, як $SrTiO_3$ із $\varepsilon \approx 1500$.

Для транзисторів із мінімальними розмірами (~ 10 нм) неможливо створювати статистично однорідний профіль легування підкладки, але, зменшити струм, обмежений ОПЗ використовуючи ретроградний профіль легування і зниження температури до $77K$ (рідкий азот). Для створення контактів витік і стік необхідні. Для транзисторів із мінімальними розмірами (~ 10 нм) неможливо створювати статистично однорідний профіль легування підкладки, але, зменшити струм, обмежений ОПЗ використовуючи ретроградний профіль легування і зниження температури до $77K$ (рідкий азот). Для створення контактів витік і стік необхідні будуть нові процеси, що дозволяють змінювати нерівноважний рівень легування Si вище межі розчинності.

Пропорційне масштабування вимагає зменшення розмірів не лише транзисторів, але й провідників. Питомий опір металів визначається довжиною

вільного пробігу електронів, яка залежить від температури і густини дефектів. При зменшенні розмірів зменшується механічна міцність провідників із-за появи її острівкової структури і її опір прямує до безмежності. Так для мідної плівки критична товщина складає 30нм.

Обмеження мають літографічні розміри із-за оптичних властивостей матеріалів. Так для короткохвильового випромінювання з довжиною хвилі 157нм (лазери на F_2) використовується оптика на основі CaF_2 , як матеріалу із найбільшою шириною забороненої зони і мінімальний розмір малюнка складає 60÷70нм. При одержанні розмірів елементів менших 60нм треба використовувати електронну літографію і неорганічні електрорезистори. Розміри органічних молекул в резистах складають десятки нанометрів і отримати розміри менше одної молекули неможливо. Крім цього, чутливість відомих неорганічних резистів значно нижча, ніж органічних.

2.2. Функціональні матеріали мікро- і наносистемної техніки^{*)}

В технології МінСТ використовуються різні групи матеріалів, з фізичними і оптичними властивостями [2.8].

1. Конструкційні (композиційні) матеріали. Цей клас матеріалів виконують функції несучих конструкцій механічних рухомих і струмонесучих частин, опор та змазок. До таких матеріалів належать скло, пластик, монокристалічний кремній, полікристалічний кремній, поруватий кремній, діоксид та нітрид кремнію, поліамід, вольфрам, нікель, мідь, золото та вуглець. Для забезпечення тривалого функціонування МЕМС в агресивних середовищах, перспективною є композиція «карбід кремнію на нітриді алюмінію». Карбід кремнію – напівпровідник, а нітрид алюмінію – діелектрик з п'єзоелектричними властивостями. Крім цього, обидва матеріали оптично активні, в тому числі в УФ ділянці спектру, мають високу теплопровідність та температуру Дебая, тобто термічно, хімічно та радіаційно стійкі.

Кремній. Кремній – це базовий матеріал МЕМС. Його елементарна комірка має структуру алмазу (рис.2.2.1,а), в якій відсутня щільна упаковка, але є

^{*)} Значення параметрів подані на сайтах: <http://users.omskreg.ru/~kolosov/atlas/3D-crystals/indexr.htm> ; www.matweb.com;

площини, упаковані щільніше, ніж інші. Атоми займають вузли двох гранецентрованих комірок, вставлених одна в одну і зсунуті вздовж діагоналі куба на $1/4$ частину її довжини l , тому в елементарній комірці алмазу 8 атомів, а її період дорівнює $a = 4l / \sqrt{3}$ і координаційний багатогранник у вигляді тетраедра утворюють найближчі атоми першої координаційної сфери із числом 4.

Кремній має унікальні фізичні властивості, завдяки яким став широко використовуватись для виготовлення МЕМС. Завдяки алмазній структурі кристалічної ґратки, кремній характеризується анізотропністю швидкості травлення від орієнтації в кристалі. Це пояснюється різною щільністю упаковки атомів в різних напрямках. Найбільша щільність упаковки в площині (111), найменша – (100), тому найменшу швидкість травлення має площина (111). Розроблена САПР ACES (Anisotropic Crystalline Etching Simulation) для моделювання тривимірних процесів мікропрофілювання в напівпровідникових матеріалах, яка включає в себе анізотропне і ізотропне рідке та реактивне травлення. Моделювання ґрунтується на статичній клітковій моделі (Cellular Automata Model) зв'язку атомів, де враховується їх розташування в кристалічній комірці.

Матеріали маскуючих покриттів в електронних технологіях. В кремнієвій технології для ізоляції переважно використовують діоксин кремнію (SiO_2), нітрид кремнію (Si_3N_4), фоторезист та плівки деяких металів ($\text{Al}, \text{Ag}, \text{Cu}, \text{Cr}$). Низьку швидкість травлення в KOH мають шари кремнію, леговані бором до концентрації $7 \cdot 10^{19} \text{ см}^{-3}$ (p-Si). Найбільш часто шари SiO_2 формуються шляхом термічного окислення Si в парах води або сухого кисню при температурі $900 - 1200^\circ \text{C}$. Моделювання пасивації SiO_2 здійснюється за допомогою програмного пакету ACES.

Кварц. Кристалічний кварц SiO_2 (кремнезем, безводний двооксид кремнію) - найбільш поширена сполука на Землі. Прозорі кристали кварцу – цінна технічна сировина. Кварц відноситься до числа найбільш твердих мінералів

(твердість 7 за десятибальною шкалою, густина - 2,65 г/см³). Розрізняють чотири температурні модифікації кварцу, з яких основне призначення в техніці знаходить низькотемпературна модифікація - α -кварц. Комірка SiO_2 побудована з кремнієкисневих тетраедрів $[\text{SiO}_2]$ з атомами кремнію в центрі (рис.2.2.1,б). В тетраедрі кожний іон Si має позитивний заряд $+4e$, що пов'язаний з чотирима іонами кисню, кожний з яких має від'ємний заряд $-2e$, тобто кожний іон кисню зв'язує між собою два іони Si .

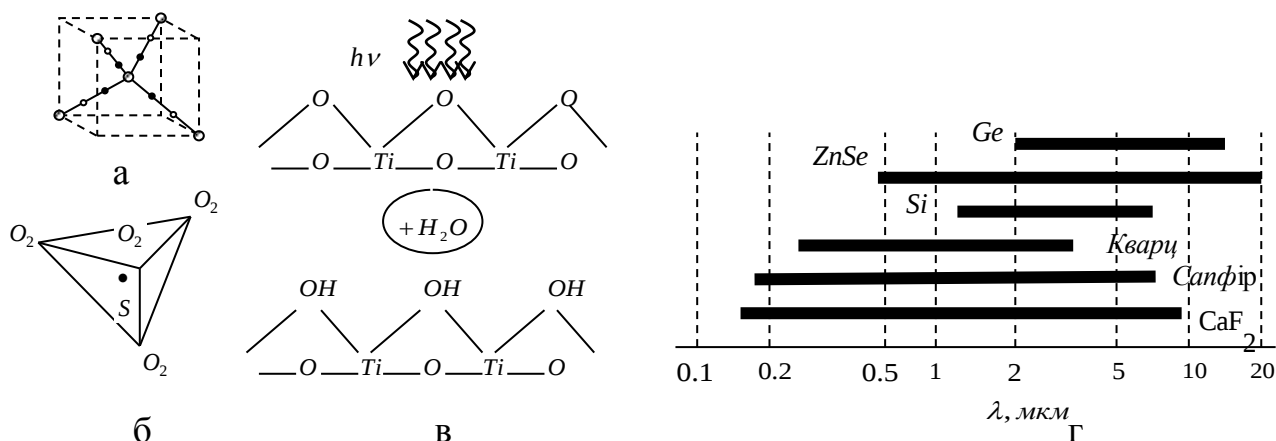


Рис.2.2.1*)

При температурі близько 573⁰С у результаті фазового перетворення виконується перебудова тригональної кристалічної структури кварцу (група симетрії 3:2) в гексагональну (група симетрії 6:2) – перетворення α - кварцу в β -кварц, який має зовсім інші п'єзоелектричні постійні та інші фізичні властивості. При температурі більше 1710⁰С кварц переходить в рідкий стан (плавиться). Частіше всього зустрічається α -кварц таких кольорів: безколірний – гірський криштал; димчатий – топаз; чорний – моріон; золотисто-жовтий – цитрин; бузковий - аметист. Нагрівання кристалів кварцу до температури 350⁰С призводить до втрати його кольору. Сьогодні широко використовують синтетичні кристали кварцу, які практично не поступаються за своєю якістю природним.

Діоксид титану. Японський вчений А.Фудзисима відкрив надзвичайну цікаву властивість діоксиду титану – здатність взаємодіяти з ультрафіолетовим випромінюванням. Поглинаючи ультрафіолет, діоксиду титану конвертує його таким чином, що утворюються вільні радикали, здатний як фотокаталізатор

*) www.tehnoluch.com

генерувати водень із води за допомогою освітлення електродів (рис.2.2.1,в). В термінах окисно-відновлювальних потенціалів, реакційна здатність електронів і дірок на поверхні характеризується потенціалами -0.1 eV і $+3\text{ eV}$ відносно нормального водневого електрода, тому електрон здатний реагувати з киснем, утворюючи окисні радикали O і OH . Matsunaga відкрив бактероцитну властивість TiO_2 . Діоксид титану (TiO_2) – неорганічний фотокаталізатор сполук NO_x . Випромінювання ними світла довжинами хвиль $360 - 380\text{ нм}$ супроводжується переходом електронів із зв'язаного стану в зону провідності з подальшою реакцією утворення вільних радикалів на поверхні – окисненням NO_x з кінцевим утворенням соляної кислоти. Порошок діоксиду титану інертний, тому його можна додавати в різні конструкційні матеріали і генерувати в них фотокаталітичні властивості поверхні.

2. Інтелектуальні матеріали. Друга група матеріалів – це так звані «інтелектуальні матеріали», які в (М,Н,О)ЕМС виконують функції джерел руху, механізмів для передачі руху, сенсорних та активних середовищ, геометричної та об'ємної пам'яті за рахунок електростатичних, магнітних, п'єзоелектричних та електромеханічних явищ та сил. До таких матеріалів перш за все відносяться сплави титану і нікелю, оксид цинку, кварц, пьезокераміка. В даний час особлива увага приділяється використанню в (М,Н,О)ЕМС -технологіях нового класу наноматеріалів.

3. Оптичні матеріали. Окрім традиційних фізичних (механічні, електричні, теплові, тощо) властивостей, для оптичних матеріалів важливими характеристиками є показник заломлення і область прозорості (рис.2.2.1,г). Значення фізичних параметрів основних оптичних матеріалів наведені в [2.9].

4. Матеріали рідинних мікропотоків систем. Класичні мікропотоків системи складаються з двох базових деталей: підкладка із мережею мікроканалів, та кришка зі скла або пластику для їх герметизації, яку приклеюють. Для підкладки використовуються скло, полімери, кремній та полімерні матеріали. Полімери як PMMA і PDMS, низьковартності та технологічні з точки зору створення в них мікро- і наноканалів. Однак, використання кремнію дозволяє використати добре відпрацьовану кремнієву технологію виготовлення ІС та МСТ.

5. Наноматеріали [2.10]. Наносистема – це сукупність нуль-вимірних ($0D$) квантових точок, одновимірних ($1D$) квантових дротин, двохвимірних ($2D$) наноплівки та тривимірних ($3D$) – просторових нанокристалітів чи їх сукупності. Обов’язковим компонентом наносистем є навколишнє середовище, з яким вона обмінюється енергією і речовиною. Тому важливою ознакою наносистеми є залежність властивості від маси наночастинок.

В нанотехнології використовують два фундаментальні підходи: - від меншого до більшого (“*bottom-up*” approach), коли наноматеріали збираються самостійно за принципами хімії і молекулярного розпізнавання; - від більшого до меншого (“*top-down*” approach), коли закономірність конструювання нанооб’єктів здійснюється на рівні атомів. Включення нанорозмірних елементів в електромеханічну систему формує НЕМС.

Однак створення нового класу функціональних пристроїв типу НЕМС вимагає застосування нанотехнологічних процесів синтезу і вимірювання параметрів. Базовим в цьому є подавно відомий в квантовій механіці ефект переборення ПБ за умови, коли на відповідне існує заборона класичної фізики. Великий масив експериментальної, теоретичної і технологічної інформації дозволяє узагальнювати її у вигляді стандартів. Вже розроблені базові терміни та класифікації, які у стислому вигляді подані в таблицях 2.2.1 і 2.2.2.

Якщо за критерій взяти «розмір», то наноматеріали за їх кількістю можна розбити на $1D$ (дискретно-компонентні та багатоконпонентні наношари і наноплівки, нанокомпозити), $2D$ (нановолокна, нанотрубки, нанострижні, нанодроти) і $3D$ (наночастинки, в тому числі нанокапсульовані, квантові точки фулерени, алмазоїди, нанооніони) розмірні категорії. Сучасне розуміння будови наноматеріалів відображає класифікаційна система «нанодерево».

Наноматеріали – це матеріали, структуровані на рівні молекулярних розмірів або близьких до них. Структура може бути більше чи менше регулярною або випадковою. Поверхні із випадковою наноструктурою одержуються опроміненням іншими частинками, плазмовим травленням та іншими іонно-фотонними технологіями. Наноструктури можуть складатись за так званою технологією «самозбирання», яка широко розповсюджена в живій природі. Так,

структура всіх живих тканин визначається їх самозбиранням із клітин; структура клітин – само збиранням із окремих молекул. Технологія само збирання молекулярних компонент використовується як спосіб побудови періодичних структур для нанoeлектронних пристроїв. Нанопроцеси не підлягають опису законами класичної фізики, а квантової механіки.

Таблиця 2.2.1 [2.11]

Терміни	Визначення
<i>Нанотехнологія</i>	Застосування наукового знання для впливу на речовину з розмірами у наношкالی та її контролю, щоб на відміну від окремих атомів і молекул чи масивних матеріалів, використовувати розмірно- й структуровано-залежні властивості та явища.
<i>Наноматеріал</i>	Матеріал із принаймні одним зовнішнім розміром у наношкالی або матеріал, який має внутрішню чи поверхневу нановимірну структуру.
<i>Наноструктура</i>	Структурна композиція із взаємозалежних складових частин, у якій хоча б одна є нанорозмірною областю
<i>Наноструктурований матеріал</i>	Матеріал, що має внутрішню або поверхневу наноструктуру
<i>Промисловий наноматеріал</i>	Наноматеріал із заданими властивостями та складом, спеціально вироблений із комерційною метою.
<i>Промислове нановиробництво</i>	Цілеспрямований синтез, формування й контроль наноматеріалів із комерційною метою
<i>Нановластивість</i>	Характеристика нанооб'єкта або нанорозмірної області

Щоб зрозуміти суть фізичних особливостей наносвіту, та пошуку аналогій їм у макросвіті, розглянемо добре вивчений у квантовій фізиці ефект переборення мікрочастинкою ПБ, коли її повна енергія менша висоти самого бар'єра. Європейська комісія затвердила рекомендації щодо офіційного визначення і використання терміну «наноматеріал» (http://dt.ua/SCIENCE/evropeyskiy_soyuz_viznachivsyia_z_nanomaterialami-90481.html): *Наноматеріал — це матеріал природного або штучного походження, що містить частинки, понад 50% яких мають лінійний розмір у межах 1—100 нанометрів.* В таблиці 2.2.2. наведена класифікація наноструктур за нанобазисом [2.12]. Нагадаємо, що 1 нанометр (нм) — це довжина ланцюжка з 5—10 атомів.

Наведемо перелік основних фізичних явищ, якими супроводжується перенесення досліджень в нанодіапазон: **1.** Атомна дифузія крізь межуючи поверхні породжує ефективний механізм переносу речовини при відносно більш низьких температурах у порівнянні із звичним процесом, що дозволяє значно збільшити чутливість азотових сенсорів чи знизити робочу температуру в топ

ливних елементах з твердим оксидом. 2. Якщо розмір кристалу значно менший середньої довжини вільного пробігу електрона, із-за розсіяння на межі розділу зменшуватиметься електронна провідність і температурний коефіцієнт. 3. Завдяки поверхневим ефектам і малому розміру частинок, змінюється фотонний спектр із проявом ефектів локалізації. 4. Зміна ширини забороненої зони із зменшенням розмірів частинок призводить до зменшення довжини хвилі випромінювання. 5. Зменшення розмірів квантових точок до відповідних з декількома електронами сприяло появі нового напрямку – спінтроники. 6. Зменшення до нанорозмірів суттєво змінює трибологічні властивості речовини, завдяки яким суттєво зменшується тертя.

Таблиця 2.2.2

«Класичні» твердотільні нановимірні структури (НВС)	Синтетичні НВС	Нанорозмірні біоструктури
Наночастинки	Нанополімери	Біомолекулярні комплекси
Нанотрубки	Синтетичні нановолокна	Модифіковані віруси
Металеві, н/п, діелектричні плівки	Синтетичні плівки	Органічні наноструктури
Квазіодновимірні провідники	Наноколоїди	
Квазінульвимірні металеві, н/п. діелектричні об'єкти	Нанокристали (каучук, тефлон, тощо)	
Нанокристали тощо		

Квантовро-розмірні матеріали

Зменшення розмірів приладних структур з одної сторони призводило до порушення стабільності роботи класичних приладів, а з іншої – давало можливість все більше зрозуміти значення квантових властивостей електрона. Нагадаємо, що з позицій квантової механіки електрон можна уявити як хвилю, тому на підставі квантових уявлень поширення електрона в середовищі нанорозмірів може супроводжуватись принципово новими ефектами – **розмірним квантуванням електронних станів**, що характерні для хвильових процесів, як інтерференція, дифракція та тунелювання.

Розмірне квантування електронних станів

Зменшення розмірів приладних структур з одної сторони призводило до порушення стабільності роботи класичних приладів, а з іншої – давало можливість все більше зрозуміти значення квантових властивостей електрона.

Нагадаємо, що з позицій квантової механіки електрон можна уявити як хвилю, тому на підставі квантових уявлень поширення електрона в середовищі нанорозмірів може супроводжуватись принципово новими ефектами – **розмірним квантуванням електронних станів**, що характерні для хвильових процесів, як інтерференція, дифракція та тунелювання.

Суть розмірного квантування електронних станів схематично ілюстрована на рис.2.2.2,а. Тут фізичний об'єкт зображений у формі паралелепіпеда таких малих розмірів, що в його ширину вкладається декілька електронних хвиль. Класичним аналогом такого ефекту може бути коливання струни, що закріплена з обох кінців. Як відомо з класичної фізики, в цьому випадку збуджуються стоячі хвилі довжиною $\lambda_n = \frac{2b}{n}$, де $n = 1, 2, 3, \dots$. Якщо цю механічну аналогію формально перенести на рух електрона в обмеженому просторі (електронну хвилю), то аналогічне проявлятиметься завдяки розміного квантування електронного стану. Це означає, що для електрона, що рухається вздовж виміру b , дозволеними будуть лише квантові хвильові числа із значеннями $k_n = \frac{2\pi}{\lambda_n}$, тоді як вздовж третього виміру c електрони можуть рухатись із довільними значеннями k . Нагадаємо відомий в квантовій механіці принцип невизначеності. Згідно з ним, якщо квантова частинка зазнає обмежень в русі в якомусь напрямку виміром a , то в цьому напрямку зростатиме її імпульс на величину $\frac{\hbar}{a}$, що, в свою чергу, веде до зростання енергії частинки на величину $\delta W = \frac{\hbar^2 k^2}{2m^*} = \left(\frac{\hbar^2}{2m^*} \right) \left(\frac{\pi}{a} \right)^2$. Як доводиться в квантовій механіці і було обґрунтовано вище, наявність штучної періодичності, що накладається на атомну структуру кристалу з точністю майже до одного атома, суттєво змінює енергетичний стан електрона. Наприклад, в об'ємі напівпровідника *GaAs* n -типу провідності (рис.2.2.3а) електрони можуть займати довільні енергетичні рівні, оскільки в ньому густина станів $g(E)$, як функція від енергії W , неперервна.

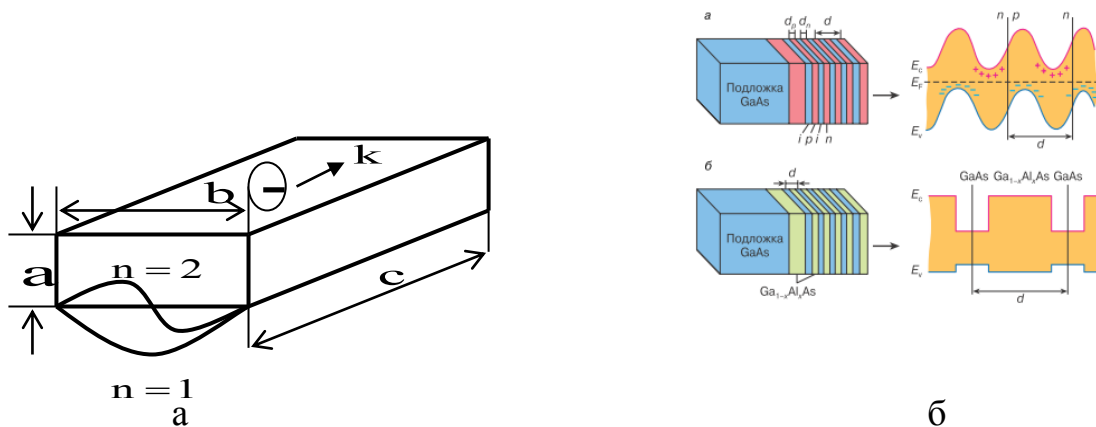


Рис.2.2.2

Квантова яма. Квантовий характер руху електрона проявлятиметься, якщо хоч один із лінійних розмірів структури є порядку довжини хвилі де Бройля. В *GaAs* при кімнатній температурі довжина хвилі де Бройля для електрона порядку $200 \text{ \AA}$. Тому у найпростішому випадку досить застосувати одновимірне обмеження для руху електрона. Ним може бути квантова яма у вигляді плівки *GaAs* шириною $200 \text{ \AA}$ порядку довжини хвилі електрона (темна полоса на рис.2.4.3б), яка з обох сторін оточена більш широкозонними шарами із *GaAlAs*. Останні для нього відіграють роль потенціальних бар'єрів. Тому внаслідок розмірного квантування у напрямку, що перпендикулярний до площин надграток, функція густини станів вже не плавна, а ступінчата і енергетичний спектр електрона у даному напрямку квантується. Отже, характер руху електрона між шарами визначається тунелюванням його через бар'єри, а їх інженерна динаміка визначатиме квантові властивості відповідного електронного пристрою. Основні явища, характерні для квантових яв, це розмірне квантування електронного спектру, квантовий (цілозначний і дробовий) ефект Холла. Молекулярно-пучкова епітаксія дає можливість формувати гетеро структури з квантовими ямами, в яких досягаються досить високі значення рухливості і, отже, швидкодій електронних приладів. Пристрої на структурах із квантовими ямами в наногетероструктурах використовуються для приладів НВЧ-діапазону, гетеро лазерів і світло діодів з широким спектром випромінювання, модуляторів і приймачів, тощо.

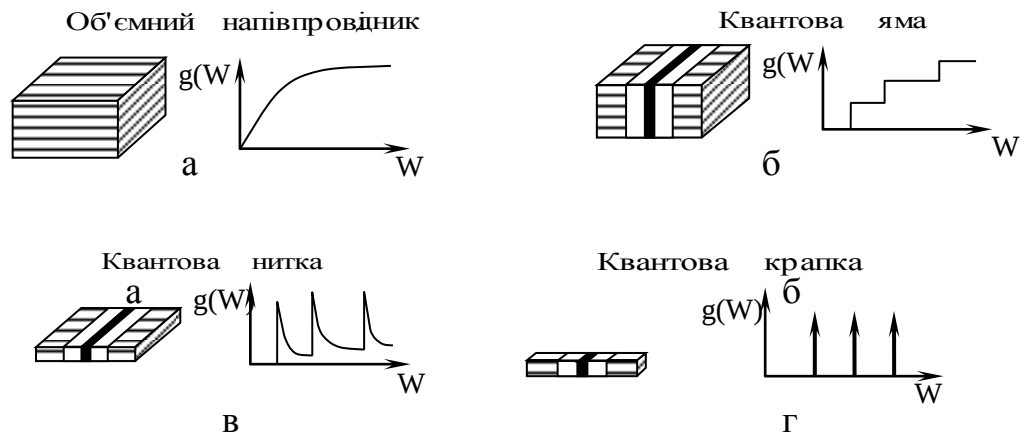


Рис.2.2.3

Квантова нитка. Якщо тепер до руху електрона ввести ще одне просторове обмеження, то одержимо випадок двовимірного обмеження руху електрона. Така структура називається **квантовою ниткою** (рис.2.2.3,в). В них рух носія струму кантується в двох вимірах. Квантові нитки формуються із квантових ям шляхом створення потенціального бар'єру за допомогою двох затворів. До основних властивостей, характерних для квантових ниток належать, квантова провідність, суттєво скорельований перенос носіїв електронів.

Квантова точка. Випадок тривимірного квантування руху електрона реалізується в **квантових крапках** (рис.2.2.3,г). Окрема квантова точка представляє собою матеріальний об'єкт, сформований із нановимірних острівків одного напівпровідника з меншою шириною забороненої зони в матриці іншого, з більшою шириною забороненої зони. Із-за різниці в значеннях ширини забороненої зони носії заряду виявляються локалізованими в межах одного острівка. Електрони в ньому поведуть себе як окреми стоячі хвилі, тобто так, як в окремому атомі. Електронну хмару, що утворюється при цьому, називають «штучним атомом». Штучний атом, на відміну від природного, не має ядра. Якщо із великої кількості «штучних атомів» сформувати об'ємну структуру по типу кристалічної ґратки напівпровідника, то новий матеріал буде мати інші властивості. Так, сформований таким чином напівпровідник, матиме діелектричні чи металічні властивості.

Надґратки. Структурами з штучним квантуванням електронного спектру є твердотільні **композиційні надґратки** (або **гетероструктури**) двх типів:

надгратки з почерговим легуванням (рис.2.2.2,б), в яких періодичність параметрів створюється введенням незначної кількості атомів у вузькі ділянки, та надгратки із змінним складом. В надгратках **першого типу** розриви в зоні провідності та у валентній зоні мають протилежні знаки. В надгратках **другого типу** модуляції краю зони провідності та валентної зони створюються із-за періодичного чергування широкозонного та вузькозонного напівпровідників. Композиційні надгратки першого типу теж умовно поділяють ще на два класи. До першого класу належать такі, в яких чергуються шарі з різними ширинами заборонених зон. Розриви зон на гетерограницях є потенціальними бар'єрами для електронів та дірок і створюють надгратковий періодичний потенціал в зоні провідності та валентній зоні. До таких структур належать надгратки із $GaAs$ та $Al_{1-x}Ga_xAs$.

Матеріали фотоніки. ФК прийнято називати середовища, в яких діелектрична проникність періодично змінюється в просторі з періодом, що допускає брегівську дифракцію світла. ФК – це надграткова система із періодом, що значно перевищує сталу гратки кристалу (рис.2.2.4,а,б). ФК. Щоб здійснити чисельне моделювання ФК, треба ввести структуру двохвимірної чи тривимірної моделей. Такими надгратками можна розглядати звичайну кубічну чи гексагональну, гранецентровану, тощо гратки, вузлами яких переважно розглядають сфери чи куби з діелектричною проникністю значно більшою, ніж діелектрична проникність решта області ФК. Щоб такі структури могли керувати оптичним променем, треба щоб кількість вузлів із зміненими показниками заломлень була не менша декількох десятків.

Для моделювання поширення ЕМХ переважно використовується найбільш поширений чисельний метод для розв'язування рівнянь Максвелла- метод FDTD (Finite-Difference Time Doma [2.13; 2.14]. Як бачимо із рис.1.3.10,в, якщо для вільного фотона закон дисперсії неперервний у вигляді параболи, то для фотона, що поширюється в періодичній структурі ФК, розподіл функції густини станів має локальний характер. Для оптичного діапазону період ФК складає порядку довжини світлової хвилі. За аналогією із звичайними кристалами, в яких є заборонені зони для енергій електронів, твердотільні структури, завдяки періодичній зміні показника заломлення, в ФК також можна побудувати

аналогічні, так звані «фотонні зони» в яких показник заломлення змінюється періодично, називаються фотонними.

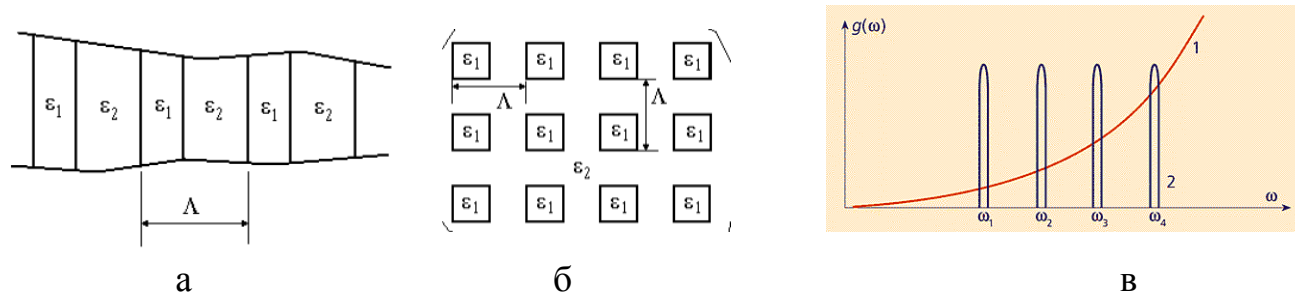


Рис.2.2.4

Відповідність між електронними і фотонами зонами зображена на рис.2.2.5,а. Тут m^* - ефективна маса електрона, а n^* - ефективний показник заломлення, які характеризують вплив кристалічної ґратки на електрон і фотон, відповідно. Як переконаємось нижче, на межах забороненої зони і в самій забороненій зоні власні функції кристалу – стоячі хвилі, так що хвилі з такими енергіями поширюватися не можуть. Всередині забороненої зони, хвильові числа уявні, тому для їх значень ефективні значення маси і показника заломлення від'ємні і формується область відбиття. Однак між дисперсіями фотонів і електронів є дві принципові відмінності: **1.** Лінійний закон дисперсії вільних фотонів замість параболічного для вільних електронів. **2.** Фотонна зона Бріллюена значно менша електронної, так що період фотонного кристалу в 10^3 разів більший сталої ґратки.

Із рис.2.2.5,б випливає, що прогрес в мікроелектроніці пов'язаний з безперервним підвищенням робочих частот і зменшенням розмірів мікроелементів. Однак реалізація наноприладів з розмірами елементів <100 нм не виводить робочі частоти за межі гігагерцового діапазону, що значно нижче терагерцових частот в фотоніці. З іншої сторони, розміри фотонних елементів обмежені значенням довжини світлової хвилі λ , тобто складає одиниці-десятки-сотні мікрон, що затруднює інтеграцію електронно-фотонних систем.

Скло – найбільш поширений матеріал фотоніки для виготовлення ВОЛЗ, який використовується для передачі оптичного сигналі, на вході і виході яких використовуються фото- і світлоелементи н/п сполук $A_3B_5, A_2B_6, \dots$ Н/п матеріали

можуть одночасно використовуватись як матеріали для виготовлення як електронних, так і фотонних компонент [2.16].

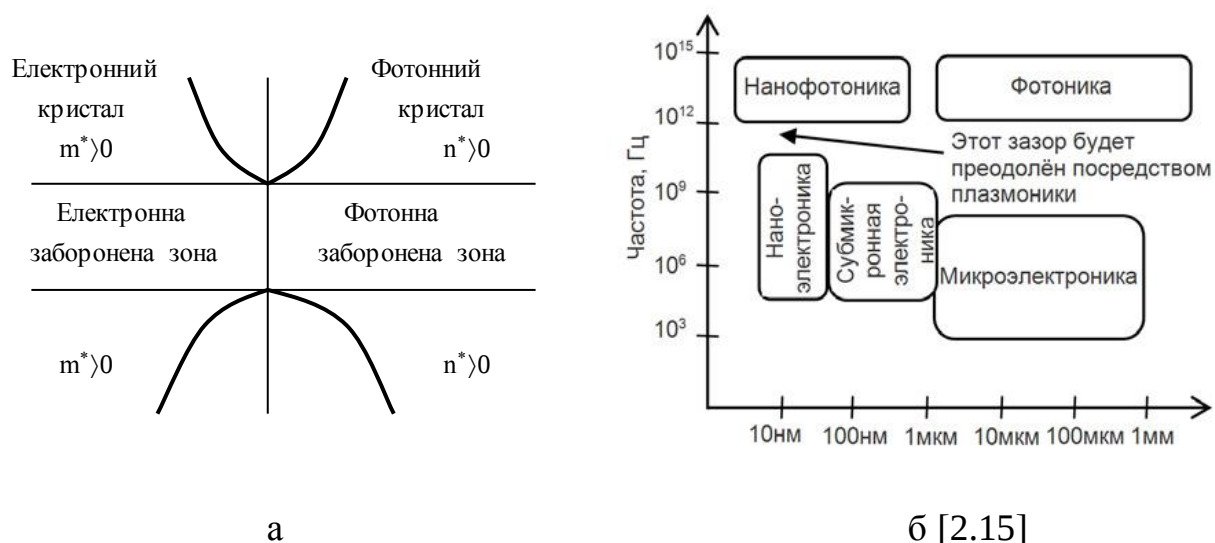


Рис.2.2.5

Нанокристали для фотонної оптоелектроніки. Наноструктурування значно підвищує ефективність люмінесценції кремнію, що дозволяє інтегрувати формування нанокристалів із розмірами менше 10 нм в планарну кремнієву КМДП технологією імплантації їх в діелектричну матрицю $\text{SiO}_2, \text{Si}_3\text{N}_4$. В цьому випадку реалізується просторове обмеження і, відповідно, квантово-вимірний ефект, що підвищує ймовірність між зонної випромінювальної рекомбінації. Такі структури формуються шляхом окислення нанокремнію ($nc-Si$) в $\text{SiO}_2, \text{Si}_3\text{N}_4$. Оскільки в них зростає показник заломлення, то такі структури проявляють хвилеводні властивості, де оптичні резонатори формуються шляхом нанесення багатошарового покриття або створення фотонного кристалу. Хоч самі по собі нанокристали Si не світять в ІЧ області, що використовується в телекомунікації, однак легування їх іонами ербію робить їх світними на довжині хвилі порядку 1.5 мкм, яка добре узгоджується із областю прозорості оптичного волокна. Крім цього $nc-Si$ перспективні для сонячної енергетики, в тому числі шляхом імплантації нанокристалів в органічні н/п [2.17].

Вуглецеві наноструктури. Наноматеріали представляють значний інтерес з точки зору використання їх в якості інструментальних матеріалів. Це пов'язано з тим, що наноструктурування різко збільшує механічну міцність. Перспективними в цьому плані є порошкоподібні матеріали на основі

наноалмазів, оксиду алюмінію, оксиду цирконію, карбидів тугоплавких матеріалів. Перспективним є напрям наноструктурних функціональних покриттів поверхонь, що значно підвищує їх ізносостійкість. Велике значення має розробка нового покоління біосумісних матеріалів, перш за все створення покриттів в стоматології чи імплантатів в ортопедії.

Стислий розгляд фізичних основ наноструктур був би незавершеним, не згадавши про недавно відкриті нові форми існування вуглецю також у вигляді наноструктур^{*)}. Добре відомі дві алотропні форми вуглецю: сажа, алмаз та графіт^{**)}. Кристаліти графіту складаються з двох паралельних двовимірних площин – графенів. Завдяки ковалентним зв'язкам, атоми вуглецю об'єднуються в молекули фуллерену. Молекула фуллерену подібна до багатогранника, побудованого з 12 п'ятикутників і 20 шестикутників, що відтворює форму футбольного м'яча і має розмір порядку 0.7_{нм} (рис.2.2.6,а). Тут для порівняння, приведені інші кристалічні та аморфні утворення з атомів вуглецю, як алмаз, графіт та аморфний вуглець.

Для одержання фуллеренів використовують електричну дугу, яку запалюють у спеціальній камері між двома графітовими електродами в атмосфері гелію. Атоми вуглецю, що випаровуються з електродів, досить довго «блукають» у цій атмосфері, багатократно зустрічаються один з одним, об'єднуються і таким чином утворюють молекулу фуллерену. За своїм зовнішнім виглядом осад на стінках камери нічим не відрізняється від сажі. Згодом було синтезовано багато інших вуглецевих молекул такого типу, як C₆₀, але серед них вона є найбільш симетричною і стійкою. Її відкриття поклало шлях до створення принципово нових матеріалів, перш за все вуглецевих нанотрубок.

Молекула C₆₀ складається з 60 атомів вуглецю, sp²- орбіталі яких представляють собою замкнений випуклий кластер з п'ятикутними і

^{*)} Відкриті молекули C₆₀ отримали назву фуллеренів на честь американського архітектора Б.Фуллера, що побудував геодезичний купол з п'яти- та шестикутників

^{**)} В 1967 році в СРСР синтезували карбін з лінійних молекул з sp-хімічними зв'язками. В 1973 р. Бочвар і Гальперн, а трохи раніше Осава квантово-механічними розрахунками передбачили стабільні сферичні молекул вуглецю завдяки замкнутій атомній оболонці та ароматичності. В 1985 р Смоллі, Керл і Крото відкрили нову форму – молекулу **фуллерену** C₆₀, а в 1990 році Кречмер і Хаффман розробили метод синтезу фуллеренів випаровуванням графіту в електричній дузі в атмосфері гелію.

шестикутними гранями, має 60 вершин, 12 пентагонів і 20 гексагонів. Пентагони формуються одинарними, а гексагони – подвійними С-С зв’язками. В звичних умовах фуллерени утворюють молекулярні кристали (фуллерити) з Ван-дер-Ваальсовим зв’язком.

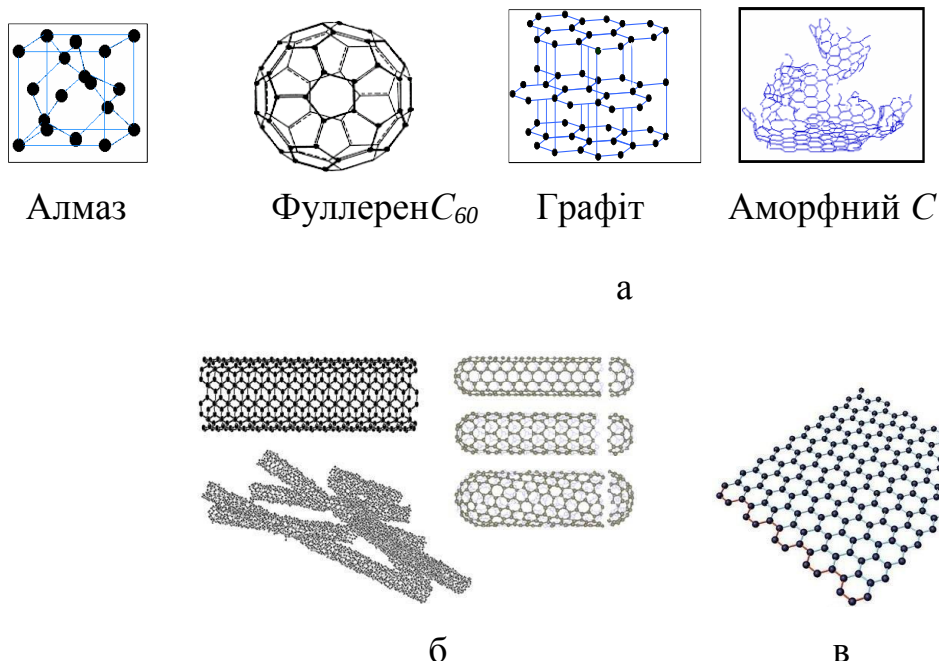


Рис.2.2.6

Вуглецеві нанотрубки*)** – це своєрідні багатоатомні вуглецеві циліндричні молекули діаметром у долі – одиниці нанометрів і завдовжки до декількох мікрометрів. Вуглецева нанотрубка – це згорнутий вуглецевий листок одноатомної площини. Її стінка має таку структуру, як атомні шарі вуглецю в графіті (рис.2.2.6,б). Однак виявилось, що їх міцність на розрив майже два порядки перевищує міцність сталі, незважаючи на те, що густина одношарових вуглецевих нанотрубок майже в 5 разів менша густини сталі. Вже запропоновані аналогічні структури з напівпровідникових матеріалів.

Графітові вуглецеві нанотрубки одержують такими методами: **1.** Лазерним випаровуванням з метало-графітових електродів; **2.** Електродуговим випаровування графіта в присутності каталізаторів; **3.** Каталітичним піролізом вуглеводнів; **4.** Диспропорціюванням окислу вуглецю в металевих каталізаторах; **5.** Електролізом розплавлених солей на графітових електродах.

***)) Технологія одержання нанотрубок вуглецю шляхом електролізу в розплаві була розроблена в Києві Делімарським ще в 1964 р. Вуглецеву нанотрубку діаметром 1.4 нм повторно синтезував Іідзіма в 1991 р

Перспективним застосуванням нанотрубок є хімічні джерела струму, акумулятори водню, оптичні фільтри. Вуглецеві нановолокна дозволяють створювати міцні композити, літієві батарейки, а виготовлені з напівпровідника – польові емітери, зонди тунельних мікроскопів, “квантові дротини”. Їх електронні властивості залежать від напрямку, в якому згорнутий графеновий листок. Графен – це двохвимірний сотовий сітка товщиною в один атом (рис. 2.2.5,в), що утворена атомами вуглецю, зв’язані між собою sp^2 - електронними станами. Ефективна маса електрона графена прямує до нуля, тому ці електрони рухливі, яка в сотні разів перевищує відповідне в кремнії.

Графен дозволяє створювати транзистори достатньо малих розмірів, декілька нанометрів і в них менш актуальні частинкові властивості електрона, а переважаючими стають хвильові властивості. Провідні електронні стани в елементарній комірни графена мають визначену просторову орієнтацію. Графен проводить електричний струм лише в тих напрямках, в яких розташовані так звані конуси густини електронів в k – просторі. Тому одностінні вуглецеві трубки, згорнуті навколо вісі, перпендикулярної $C-C$ – зв’язкам, ають напівпровідникові властивості ($W_g = 0.4 \div 0.7 \text{ eV}$), згорнуті перпендикулярно $C-C$ – зв’язкам – металічні. Одностінні нанотрубки поводять себе як одновимірні структури.

Нанопоруваті системи

Відомо [2.16; 2.34], що властивості поруватих систем суттєво відрізняються від об’ємних матеріалів чи окремих наночастинок, що формують композити. Серед великого спектру різних типів пористих систем, особливе місце займають поруваті напівпровідники і діелектрики, що сформовані шляхом видалення частки матеріалу з об’єму. При цьому виникають нанорозмірів пори та частинки, які призводять до модифікації властивостей матеріалів. Цьому можна виділити такі причини: **1.** Квантоворозмірні ефекти актуальні для нанооб’єктів, розміри яких спів вимірні з довжиною хвилі де Бройля. **2.** Формування поруватості значно збільшує ефективну площу поверхні, в зв’язку з чим суттєвими стають поверхневі ефекти, пов’язані з появою нових електронних і фононних станів. **3.** Утворення нанооб’єктів призводить до виникнення значних локальних полів в

нанокомозитних середовищах, величина яких визначається розмірами, формою та орієнтацією пор.

Поруватий кремній. Поруватий кремній (*porous Si* або *por-Si*) – це наноструктурований напівпровідниковий матеріал. Вперше поруватий кремній був створений Улром в 1956 р. під час анодного (електрохімічного) травлення пластин у водно-спиртовому розчині плавикової кислоти *HF*.

Схема електрохімічного травлення пластини *Si* в електроліті плавикової кислоти показана на рис.2.2.7, а і б. Тут приведені схеми експериментальної установки для формування шарів поруватого кремнію з вертикальною (а) та горизонтальною (б) електролітичними комірками, де 1–кремнієвий зразок (анод), 2–платиновий контрелектрод (катод), 3–фторопластова кювета з електролітом, 4–джерело стабілізованого струму, 5–металічний контакт.

Для ініціювання електрохімічного розчинення на пластину кремнію подають додатний (анодний) відносно платинового електрода потенціал. Процес електрохімічної обробки проводять при певній густині анодного струму, що дозволяє отримувати необхідну поруватість, товщину поруватого шару і забезпечити їх відтворюваність. Саме за цих умов об'ємний кристалічний *Si* може перетворюється на ниткоподібну матрицю. Нитки, стовчики чи дротинки в них зв'язані між собою водневими зв'язками, тому поруватий кремній легко оксидується у звичайній атмосфері. В ідеальному випадку *por-Si* складається з мікрокристалітів та пор. В залежності від методу його виготовлення, *por-Si* може містити в собі включення аморфного кремнію, кремнієвих оксидів і навіть кремнійорганічних сполук, що появляються внаслідок його обробки в органічних сумішах. Щоб досягнути максимальної однорідності властивостей *por-Si*, застосовують перемішування електроліту.

В макропоруватому їх радіус не перевищує 2 нм, в мезапоруватому – менше 50 нм, а при більших – макропоруватий. Ідеалізована модель *por-Si* показана на рис.2.6.7, в. Бачимо, що їх можна моделювати як двовимірні у вигляді системи паралельних каналів, заповнених повітрям, що перпендикулярні до поверхні

кремнію. В таких структурах поглинання світла є на 2 порядки більшим, ніж для монокристалічного кремнію.

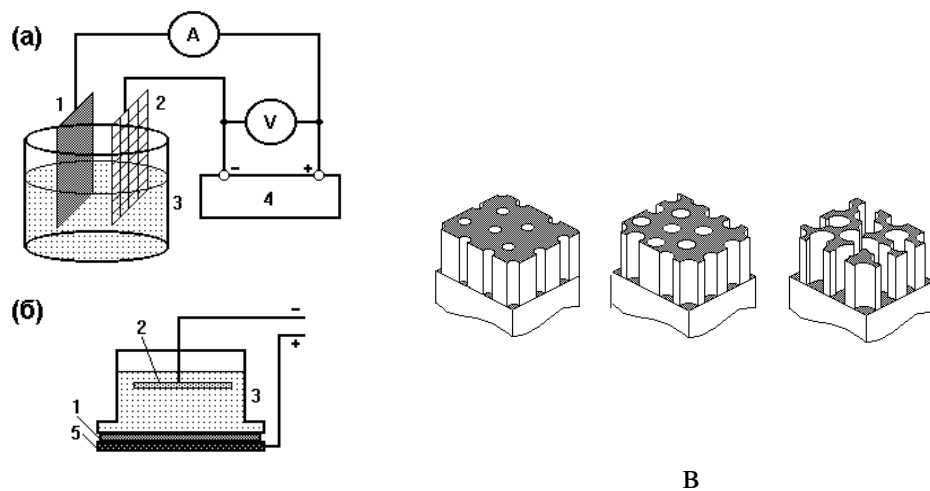


Рис.2.2.7

Кремній між порами зберігає свою вихідну кристалічну структуру. Самі пори та їх переплітання формують різні наноструктури у вигляді нанокластерів і наношнурів (стрижнів). Причому всі ці наноструктури проявляють свої властивості лише в ансамблі, статистичні властивості якого визначають фізичні властивості *por-Si*. Унікальні властивості наноструктурованого кремнію, перш за все зумовлені значною питомою густиною поверхні $\approx 10^3 \text{ м}^2 / \text{грам}$, здатністю світитись у видимій області спектру, в тому числі під час адсорбції озону та генерація синглетного кисню за рахунок того, що порування збільшує його ширину забороненої зони на 100-500 *meV*. Ці фізико-хімічні властивості дозволяють використовувати його для газових сенсорів, вибухових речовин, як матрицю для створення нанокомпозитів з фаз аморфного кремнію і нанокристалів кремнію.

Поруватий *Si* також напівпровідник, з дуже малою концентрацією вільних носіїв і високим питомим опором порядку $10^8 \text{ Ом} \cdot \text{см}$. В залежності від умов травлення, його електричний опір може змінюватися в дуже широких межах ($10^2 \dots 10^{11} \text{ Ом} \cdot \text{см}$). Теплопровідність *por-Si* більш як на порядок нижча, ніж в монокристалічного.

Тривалий час шари поруватого *Si* використовувалися для виготовлення ізоляційних шарів в електронних приладах за технологією *кремній-на-ізоляторі*.

Після відкриття в поруватому кремнії інтенсивної люмінесценції при кімнатній температурі, його стали інтенсивно досліджувати з точки зору застосування в оптоелектроніці та сенсоріці, технологія яких сумісна традиційною кремнієвою ІС. Внутрішня поверхня пор свіжо виготовленого нанопоруватого кремнію, що піддана пасивації воднем у вигляді комплексів $Si-H_n, n=1\div3$, забезпечує захист від окиснення. Водень нейтралізує електронні дефекти (типу обірваних зв'язків), мінімізуючи безвипромінювальні втрати енергії у збуджених нанокристалах. Разом з величезною питомою поверхнею, наноструктурований Si перспективний для хімічних сенсорів. Важливою практичною властивістю нанопоруватого кремнію є його біосумісність з організмом людини, що дозволяє застосовувати його як фотосенсибілізатор у терапії раку.

Для нанесення на поверхню поруватого кремнію спряженого полімерних сполук використовується метод електрохімічної полімеризації. Для цього електрохімічна комірка, у якій робочим електродом служить пластинка поруватого кремнію, закріплюється у металевому тримачі, де допоміжним електродом є платинова пластинка. Для контролю електрохімічного потенціалу робочого електроду використовують електрод порівняння з хлорсрібла $Ag/AgCl$, потенціал якого прижається за нульовий.

Окиснений поруватий кремній. Найбільш ефективним способом окиснення плівок $por-Si$ є їх термічний прогрів. Окиснення на повітрі при температурі $950^{\circ}C$ протягом 2.5 год. Призводить до формування прозорих плівок окислів у видимій області спектру. Рентгенівська дифракція окислених плівок $por-Si$ свідчить, що вони мають аморфну структуру.

Поруватий фосфід галію. Шари поруватого фосфіду галію також формуються шляхом електрохімічного травлення в матеріалі електронного типу з утворенням пор розмірами від 0.05 до 1 μm . Було встановлено, що коли піддавати електрохімічному травленню поверхню орієнтації (111), то пори формуються вздовж нормалі до поверхні. З оптичної точки зору, такі шари проявляють анізотропію показника заломлення в оптично додатних кристалах.

Поруватий оксид алюмінію. Аналогічним методом електрохімічного травлення формуються порувати шари оксиду алюмінію на поверхні фольги із

алюмінію. Поруватий шар складається з набору гексагональних оксидних комірок, по центру яких проходить пустий канал. Розміри пор залежать від температури травлення та вибору травника. Наноструктуровані матеріали стали основою для НЕМС, які стали подальшим кроком переносу електромеханічних систем в наносвіт. Ці об'єкти розвивають наносили під дією електричного чи магнітного поля, в тому числі електромагнітної хвилі, та при прикладанні зовнішнього фізичного зусилля відгукуються електромагнітним сигналом. НЕМС включають в себе наночутливий сенсор, трансдюсер та процесор обробки інформації. Разом із актюатором, вони дозволяють створювати замкнуту функціональну систему на нанорівні, здатних працювати в дистанційному режимі і за допомогою електромагнітних хвиль об'єднуватись з зовнішніми процесорними ситемами – створювати саморганізовані сенсорні мережи.

2.3. Базові фізичні моделі теплових сенсорів

В якості сенсорних елементів використовують елементи ІС, як резистори, конденсатори, діоди і транзистори, так і мікромеханічні структури, як мембрани, консолі. Класифікація сенсорів провадиться за різними ознаками: **1.** За технологією виготовлення – гібридно-плівкових, твердотільно біполярних або МДН-структур, тощо. **2.** За вихідним сигналом – ємнісні, резистивні, індуктивні, тощо. **3.** За вимірюваним параметром – сенсори температури, тиску, магнітного поля, тощо. Сутність перетворення сенсором фізичного сигналу в інформаційний відображена в таблиці 2.3.1.

В залежності від фізичної природи первинного сигналу, інтелектуальні сенсори поділяються на термоелектричні; піроелектричні; п'єзоелектричні; електромагнітні на явищі електромагнітної індукції ; фотоелектричні та холлівські.

Таблиця 2.3.1

Ефект або вплив	Перетворення	Сутність
Піроелектричний ефект	Температура – електрика	Виникнення електричних зарядів на гранях кристалу під час його нагрівання
Термоелектричний ефект	Теплова енергія – електрони	Емісія електронів нагрітим металом
Електротермічний ефект Пельт'є	Електрика – тепла енергія	Поглинання (генерація) теплової енергії при протіканні через біметалеві контакти електричного струму
Електротермічний ефект Томсона	Температура і електрика – тепла енергія	Поглинання (генерація) теплової енергії між контактами, які нагріті до різних температур
Теплопровідність	Теплова енергія – зміна властивостей	Перехід теплоти від нагрітого до холоднішого кінця
Теплове випромінювання	Теплова енергія – інфрачервоні промені	Теплове випромінювання нагрітим тілом
Ефект Зеебека	Температура – електрика	Виникнення ЕРС в колі з біметалевими контактами при різних температурах
Фотогальванічний ефект	Світло – електрика	Виникнення ЕРС під час опромінення світлом <i>p-n</i> переходу
Фотопровідність	Світло – електроопір	Зміна електроопору <i>n/p</i> , опроміненого світлом
Ефект Зеемана	Світло, магнетизм – спектр	Розщеплення спектральних ліній в магнітному полі
Ефект Рамана	Світло – світло	Виникнення світла із зміненими частотами при проходженні його крізь кристалічне середовище
Ефект Поккельса	Світло і електрика – світло	Розщеплення променя світла в кристалі під дією електричного поля
Ефект Керра	Світло і електрика – світло	Розщеплення променя світла в кристалі під дією механічної напруги
Ефект Фарадея	Світло і магнетизм – світло	Повертання площини поляризації парамагнетиком в магнітному полі
Ефект Холла	Магнетизм і електрика – електрика	Виникнення різниці потенціалів на протилежних поверхнях провідника із електричним струмом у магнітному полі
Ефект Доплера	Звук, світло – частота	Зміна частоти звукової хвилі, що генерується рухомом джерелом
Магнітоопір	Магнетизм і електрика	Зміна електричного опору твердого тіла в магнітному полі
Магнітострикція	Магнетизм – деформація	Деформація феромагнітного тіла в магнітному полі
П'єоелектричний ефект	Тиск – електрика	Виникнення різниці потенціалів на гранях деформованого сегнетоелектрика

За загальним принципом, їх можна поділити на механічні, акустичні, електричні, електромагнітні, електрохімічні, оптичні, як це показано на рис.2.3.1.

Окрему групу складають біологічні сенсори (біосенсори). Оптичні сенсори, в свою чергу, також поділяються на спектрофотометричні, люмінесцентні, сенсори на поверхнево-плазменному резонансі, інтерферометричні та поляризаційні. Останнім часом значна увага приділяється розробці квантових сенсорів.

Основним елементом сенсора є давач. За означенням – це електронний пристрій, що перетворює параметр дії зовнішнього середовища в напругу, або струм на виході сенсорного пристрою. Класифікують сенсорні давачі за двома ознаками: **1.** За тим параметром зовнішнього середовища, який перетворюється у напругу або струм. За цією ознакою давачі поділяються на сенсори тиску, сили, положення (по координаті або по куту), переміщення, швидкості (лінійної або кутової), прискорення (лінійного або кутового), вібрації, температури, газового складу середовища, обертів, деформації, вологості тощо. **2.** За фізичним принципом функціонування, за яким давачі поділяють на ємнісні, п'єзоелектричні, диференційно-трансформаторні, термопарні, тензорезистивні, потенціометричні, струмовихрові, термісторні, на ефекті Холла, п'єзорезистивні, оптоволоконні, магніторезистивні, на ПАХ, індукційні, тощо. До перспективних відносяться давачі на квантово-розмірних ефектах (ФК, наноб'єкти, тощо).

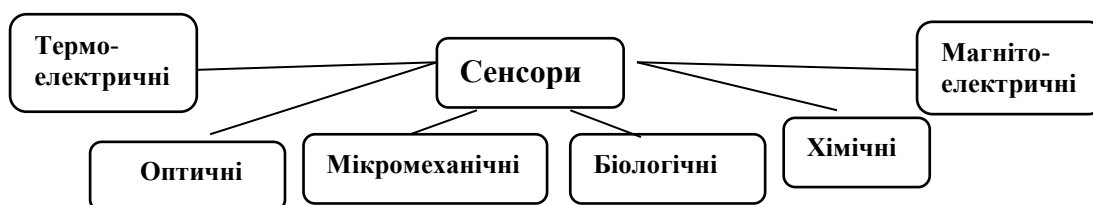


Рис.2.3.1

Сенсори поділяють на активні і пасивні. Принцип дії активного сенсора ґрунтується на фізичних ефектах, перерахованих в таблиці 2.3.2 [2.18]. Активні сенсори забезпечують перетворення вимірюваної величини в електричну форму енергії. Оскільки в активних сенсорах сигнал генерується внаслідок зміни електричного опору ємності чи індуктивності давача, то в їх коло треба включати електричне джерело збудження. Тому такі сенсори (ємнісний, резистивний або індуктивний) ще називають параметричним. В пасивних сенсорах деякі параметри вихідного імпедансу можуть змінюватись під дією вимірюваної величини, який визначається геометричними розмірами і фізичними

параметрами середовища. Базові фізичні процеси пасивних сенсорів перелічені в таблиці 2.3.3. [2.14]. Імпеданс пасивного давача і його зміну вимірюють за допомогою електричної схеми із джерелом живлення.

Таблиця 2.3.2

Вимірювана величина	Базовий ефект	Вихідна величина
Температура	Термoeфект	Електрична напруга
Потік оптичного випромінювання	Піроелектричний ефект, зовнішній і внутрішній (в н/п із $p-n$ -переходами) фотоefекти, фотоелектромагнітний ефект	Заряд, Електричний струм, Напруга, Напруга
Сила, тиск, прискорення	Пізоелектричний ефект	Заряд
Швидкість	Електромагнітна індукція	Напруга
Зміщення	Ефект Холла	Напруга

Таблиця 2.3.3

Вимірювана величина	Електричний параметр, який змінюється під дією вимірюваної величини	Матеріал сенсора
Температура	Опір	Метали (платина, нікель, мідь), н/п
Наднизькі температури	Діелектрична проникність	Скло, кераміка
Потік оптичного випромінювання	Опір	Н/п
Деформація	Опір	Сплави нікелю, легований Si
Зміщення	Магнітна проникність, опір	Феромагнітні сплави, магніторезистивні матеріали: вісмут, антимонід індію
Вологість	Опір, діелектрична проникність	Хлористий літій, оксид алюмінію, полімери
Рівень	Діелектрична проникність	Рідкі ізоляційні матеріали

Окрему нішу в (М,Н,О)ЕМС займають «інтелектуальні сенсори» (smart sensor) [2.19]. Вони є основним елементом інформаційно-вимірювальної ланки інтелектуальної системи, пасивно реагують на подразник із зовні, самі спеціальним чином діють на об'єкт дослідження, сприймаючи та аналізуючи зміни, що виникають при цьому, здійснюючи статистичну обробку результатів вимірювань. Для цього, пасивний сенсор об'єднують з трандюсером в єдиний функціональний елемент, управління яким здійснюється мікро- чи нанопроцесором.

В сенсорах теплового випромінювання падаюча енергія поглинається атомами, молекулами або вільними електронами, в результаті чого змінюється їх коливний стан, внаслідок чого виникає в поглинальному середовищі градієнт

температури. Щоб його виміряти, використовують ряд фізичних ефектів, як: піроелектричний ефект – залежність поляризації сегнетоелектричних кристалів від температури; болометричний ефект – зміна опору (провідності) при зміні граткової температури матеріалу фотоприймача; ефект Зеебека – залежність термоЕРС від значення градієнта температури; ефект деформації і розширення тіл при нагріванні; зміна провідності електронного газу шляхом його нагрівання (болометри на гарячих вільних носіях). В більшості випадках зміна електричного сигналу ТП зв'язана з тепловими процесами, тому його чутливість S_p практично не залежить від довжини хвилі падаючого випромінювання, тоді як для квантового фотонного приймача вона має різку межу спаду зі сторони довгих хвиль.

В цілому теплові мікросенсиори поділяють на три групи [2.20]: 1.Сенсиори, що використовують зміни фізичних властивостей, зумовлених зміною температури (*sensor using temperature measurement*). До цієї групи віднесені теплові приймачі випромінювання та електротеплові перетворювачі. 2.Сенсиори, що використовують теплове збудження (*sensor using thermal actuation*). До цієї групи відносяться теплові мікро сенсиори для вимірювання потоків теплового випромінювання, вакуумні мікро сенсиори, резонансні сенсиори положення і концентрації газів, в яких тепловий сигнал використовується для керування резонансною частотою. 3.Сенсиори для дослідження теплофізичних властивостей матеріалів (*sensor for thermal material properties*). В цю групу включені мікросенсиори для дослідження розподілу температури по поверхні, теплопровідності, теплоємності та термоЕРС. Існує й інша класифікація теплових мікросенсиорів: сенсиори теплової потужності (*thermal power sensor*) і сенсиори теплової провідності (*thermal conductor sensor*), в основу якої покладена аналогія між тепловими і електричними явищами в мікросенсиорах за технологією UFPA .

За класифікацією [2.15], теплові МКС поділяють на сенсиори, які використовують температурні зміни (*sensor using temperature measurements*); сенсиори, які використовують теплове збудження (*sensors using thermal actuation*) і сенсиори для дослідження теплофізичних властивостей матеріалів (*sensors for thermal material properties*), тоді як за класифікацією [2.21], на підставі

електротеплової аналогії $\Delta T = P / G_T$ (P - теплова потужність, що генерується в сенсорі, G_T - теплова провідність між сенсором і оточуючим середовищем), теплові МКС поділяють на сенсори теплової потужності (thermal power sensors) і сенсори теплової провідності (thermal conductance sensors).

Порогові характеристики теплових сенсорів. 1.Поріг чутливості: $S = \frac{\sqrt{\langle \varepsilon^2 \rangle}}{V}$, де $S = \frac{V}{P}$ - вольт (параметр вихідного сигналу)-ватна (характеристика вхідного сигналу) чутливість приймача. За теоремою Найквіста $\langle \varepsilon^2 \rangle = 4k_B T R_{ш} \cdot \Delta f_{e.ш.}$, де $R_{ш}$ -шумовий опір, $\Delta f_{e.ш.}$ - еквівалентна шумова смуга, що вибирається так, щоб спектральна густина шумів залишалась незмінною. 2.Порогова чутливість. За Джонсоном $D = 1/S$. 3.Поріг чутливості за енергією: $\Delta = \frac{U_{\max}}{W \sqrt{\langle \varepsilon^2 \rangle}}$, де W - енергія

вхідного сигналу. Поріг чутливості приймача теплового випромінювання обмежується в основному трьома видами шумів, фундаментальними за своєю природою, як теплові або джонсонівські шуми, що зумовлені флуктуацією носіїв заряду в електричному опорі, незалежних від матеріалу приймача; радіційні або фотонні шуми, зумовлені флуктуацією кількості фотонів, що випромінюється джерелом випромінювання в даний момент часу та температурні шуми, що пов'язані з флуктуаціями температури із-за статистичної природи теплообміну між чутливим елементом і оточуючим середовищем.

На рис.2.3.2 наведена базова структура теплових мікросистем у вигляді резистора, сформованого на поверхні тонкої мембрани. Головною особливістю таких систем є малий об'єм нагрівного елемента, їх теплова ізоляція та мала теплова інерційність, що дозволяє SiC-резистор розташувати на тонкій мембрані. Діелектрична плівка AlN ізолює SiC-резистор від підкладки.

Термометри опору або терморезистори). У терморезисторах використовується відоме явище зміни електричного опору провідника чи напівпровідника зі зміною температури. Воно спричинене зміною інтенсивності теплових коливань іонів із зміною температури, що призводить до залежності від температури електричного опору. Ця залежність описується ТКО $\alpha_R = \frac{1}{\rho_R} \frac{d\rho}{dT}$ і

визначає відносний приріст опору при зміні температури на один градус. Вимірність і одиниця α_R дорівнюють: $\dim \alpha_R = \Theta^{-1}$, $[\alpha_R] = 1K^{-1}$. Група датчиків, принцип роботи яких базується на цьому принципі, називаються термісторами або терморезисторами. Вони виготовляються із сплавів металевих оксидів хрому, кобальту, заліза, марганцю, нікелю, тощо. Графік температурної залежності опору терморезистора нелінійний $R_T \approx \exp\left(\frac{a}{T}\right)$, тому з ростом температури опір зменшується. Однак, може мати місце протилежний ефект.

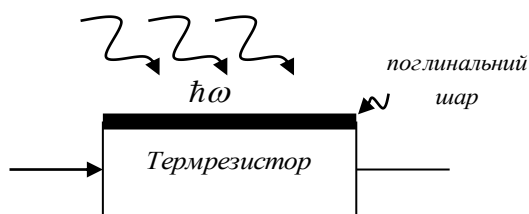


Рис.2.3.2

Електричний опір металів з температурою змінюється за законом $R(T) = R_0(1 + aT + bT^2 + cT^3)$. Для терморезисторів $R(T) = R_0 \exp(k \exp(T^{-1} - T_0^{-1}))$. В інтервалі малих змін температури цю залежність можна лінеаризувати як $R(T + \Delta T) = R(T)(1 + \alpha_R \Delta T)$. Отже, мінімальне значення вимірюваної температури буде дорівнювати $\Delta T_{\min} = \left(\frac{1}{\alpha_R}\right) \left(\frac{\Delta R}{R_0}\right)_{\min}$. Платиновий терморезистор в околі $0^\circ C$ має

значення $\Delta T_{\min} = 2.6 \cdot 10^{-4} \text{ } ^\circ C$, що забезпечує значення $\left(\frac{\Delta R}{R_0}\right)_{\min} = 10^{-6}$, тоді як

термістор цю точність забезпечує для $\Delta T_{\min} = 2.0 \cdot 10^{-5} \text{ } ^\circ C$. Деформаційна

чутливість терморезистора визначається як $S_T = \frac{\Delta R / R}{\Delta \ell / \ell}$, де для прямолінійного

дроту $R = \rho_R \frac{\ell}{S}$. При деформації відносний приріст електричного опору можна

оцінити як $\frac{\Delta R}{R} = \frac{\Delta \rho_R}{\rho_R} + \frac{\Delta \ell}{\ell} - \frac{\Delta S}{S} \cong 2 \frac{\Delta \ell}{\ell} = 2\varepsilon$, де використано наближення

$\frac{\Delta \rho_R}{\rho_R} \cong \frac{\Delta V}{V} = \frac{\Delta S}{S} + \varepsilon$, яке справджується не для усіх матеріалів.

Коефіцієнт α_R залежить від температури, тому залежність $\rho(T)$ нелінійна. Однак, якщо за значення α_R взяти середнє значення в певному інтервалі

температур, то температурну залежність $\rho(T)$ можна подати у вигляді $\rho_R = \rho_0(1 + \alpha_R T)$, ρ_0 – це значення питомого опору при кімнатній температурі $T = 273\text{K}$. Ця формула справджується для електрично однорідного провідника. Якщо провідник електрично неоднорідний вздовж довжини, то опір встановлюється електричний опір його елементарної ділянки $dR = \rho_R \frac{d\ell}{S}$, а потім обчислюється сумарний електричний опір як інтеграл $R = \int dR$, де $dR = \rho_0 \frac{d\ell}{S} (1 + \alpha_R T)$. Якщо допустити, що в провіднику температура по довжині змінюється так, що $\frac{\Delta T}{\ell} = \frac{dT}{d\ell}$, то $d\ell = \frac{\ell}{\Delta T} dT$ – зміна температурі провідника на довжині $d\ell$ і тоді $dR = \rho_0 \frac{\ell}{S \Delta T} (1 + \alpha_R T) dT$, звідки після інтегрування одержуємо вираз $R = \rho_0 \frac{\ell}{S} \left(1 + \frac{1}{2} \alpha_R (T_1 + T_2) \right)$, який покладений в основу принципу дії термометра опору. Нелінійні характеристики терморезисторів можуть бути лінізовані, якщо паралельно чи послідовно йому увімкнути опрі з сталим значенням опору R_ℓ (лінеаризуючий опір), величина якого визначається із умови квазілінійної залежності вимірюваної напруги U_m від температури T в околі вибраної точки T_j . Ця квазілінійна залежність насправді відповідає точці перегину кривої $U_m(T)$ при $T = T_j$, що виражається математичною умовою $\left. \frac{dU_m}{dt^2} \right|_{T=T_j} = 0$. Цій умові задовольняє двополіусник із опорів сенсора $R_s(T)$ і додатково R_ℓ .

Теплові сенсори консольного типу. Конструктивно вони виглядають так. Консоль (мікробалка), наприклад, із Si/SiO_2 , в якій один кінець закріплений, а інший вільно коливається (в англійській літературі – кантелівер) характеризується певним значенням резонансної частоти коливань. Якщо її поверхня адсорбує молекули, то її маса змінюється і змінюється частота резонансних коливань кантелівера, що покладено в основу побудови сенсора. Чутливість консольного сенсора консольного пропорційна зміні густини маси

$S = \frac{\partial \nu_0}{\partial \rho_{sub}} \frac{\partial \rho_c}{\partial c}$ і сягає значення до 0.67 нг/см^2 . Чутливість калориметричного порядку 10^{-5} К . Нагрівання протилежних поверхонь кантелівера до різних температур призводить до його вигину величиною $\Delta x = \frac{k_\sigma T}{\mu}$, що призводить до різниці сил поверхневого натягу $\Delta F = \frac{4\mu\ell}{3(1-\nu_\sigma)ah} \Delta x$. Тода за формулою Стоні*) відхилення кантелівера товщиною b дорівнює: $\Delta z = \frac{3(1-\nu_\sigma)\ell^2}{E_\sigma b^2} \sigma$.

2.4. Мікросенсори на бар'єрних переходах

В переважній більшості матеріалом чутливого елемента є напівпровідник. В ньому значення деяких параметрів залежать від рівня легування (рівня Фермі) та від поверхневого вигину зон (поверхневого потенціалу). На ступінь прояву повехневих ефектів (положення рівня Фермі) впливає парціальний тиск вимірюваного активного газу. В датчиках поверхневого (дифузійного) типу, газ, що аналізується, взаємодіє з поверхнею (хемісорбується), формуючи при цьому збіднену область просторового заряду, змінюючи поверхневий заряд і поверхневий потенціал.

Омічність контактів. Дослідження властивостей н/п матеріалів чи конструювання електронних приладів завжди супроводжується вирішенням питання про роль контактів. До останніх відносять не лише область, що включає в себе ділянку прямого контакту між металом та н/п, але й область просторового заряду, яка прилягає до неї. Ширина цієї області залежить від висоти та форми ПБ на границі розділу, який в свою чергу визначається різницею робіт виходу електрона з металу та н/п, ступенем заповнення поверхневих рівнів, наявністю в приконтактній області шарів із іншими характеристиками. Важливою характеристикою контакту є їх омічність. Під омічним контактом розуміють контакт металевого електрода з н/п із лінійною ВАХ. Для нього спад напруги в залежності від сили струму в режимі низького рівня інжекції неосновних носіїв струму лінійно змінюється, а отже справджуватиметься закон Ома. Якщо ця

*) Stoney G.G. Proc. R. Soc., Lond. Ser. A 82 (1909)

умова не задовільняється, то збільшуватиметься сила зворотного струму. Розглядуваний контакт не дає випрямлення сили струму, який протікає через нього, тому відносять його до невиконаних. Крім цього омичний контакт повинен мати ще й малий опір. Для створення омичного контакту найчастіше використовують свинець, олово, золото, алюміній.

Напівпровідникові сенсори [2.22]. Широкого використання набули н/п сенсори, в яких відбувається взаємодія газів з сенсорною структурою, яка перетворює цю взаємодію в електричний, оптичний чи іншого виду сигнал. Це так званий "електронний ніс", що складається з двох блоків – блоку хімічного сенсора (чутливого шару, нанесеного на поверхню підкладки) і блоку математичної (*pattern recognition method (PRM)*) – алгоритм розпізнавання образів. Блок-схема "електронного носа", що пояснює принцип його дії, приведена на рис.2.4.1,а,б.

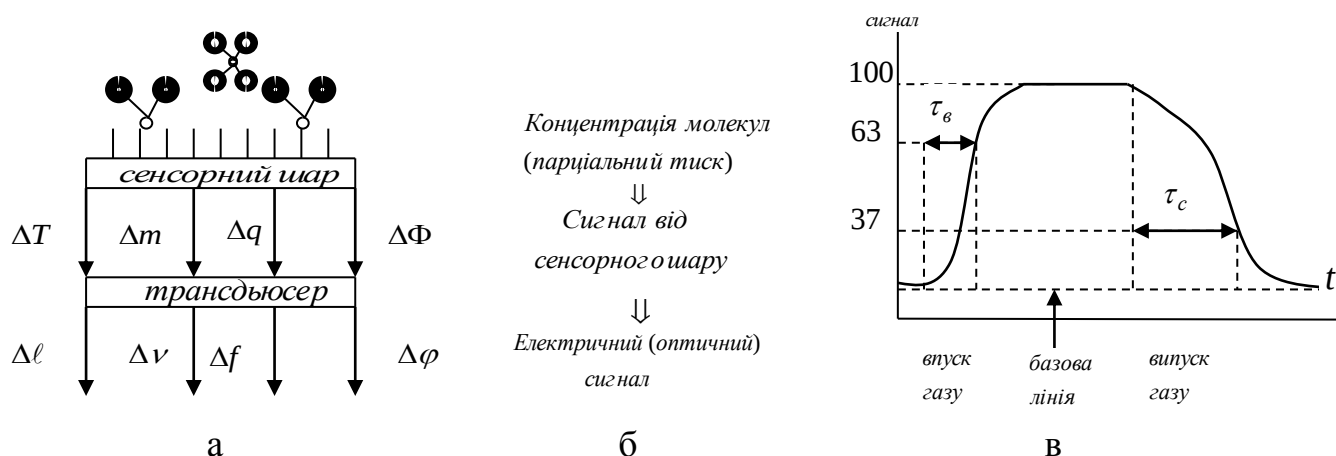


Рис.2.4.1

В цілому, сенсори створюють на тих матеріалах, у яких фізичні властивості змінюються залежно від концентрації певних хімічних сполук. Результатом такої взаємодії може бути зміна температури ΔT , маси Δm , заряду Δq та світлового потоку $\Delta \Phi$. Ці міни параметрів сенсора передаються трансдюсеру, який перетворює зміну фізичних параметрів на електричний (провідність, ємність, потенціал тощо) або на оптичний (показник заломлення, інтенсивність випромінювання тощо) сигнали. Однак розроблені інтегральні сенсори, в яких функції окремих сенсора і трансдюсера поєднані в одному пристрої, як це має

місце в сенсорі на водень на основі ПТ. В ньому паладієвий електрод затвора відіграє роль чутливого елемента, який нейтральні молекули H_2 атоми водню, що накопичуються на межі метал-діелектрик, як диполі, а зміна заряду й потенціалу перетворюється на зміну електричного струму витік-стік у транзисторі, що й вимірюється.

На рис.2.4.1,в, наведений принцип паспортизації сенсорів, де τ_e - час відповіді сенсора, протягом якого вхідний сигнал сенсора досягає 63% від максимального значення при зміні концентрації (парціального тиску) вимірюваного газу від нуля до цієї концентрації (парціального тиску). Часто вводять інше значення часу відповіді – 90% від максимального рівня сигналу.

Важливим параметром сенсора є час його відповіді τ_c , за який вихідний сигнал зменшується до 37% (в e разів), чи до 10% від максимального рівня при зміні концентрації чи парціального тиску вимірюваного газу до нуля. Якщо діапазон вимірювання сенсором фізичної величини $\Delta x = x_{\max} - x_{\min}$, що відповідає проміжку реєстраційної шкали Δy і залежність вихідного сигналу від вхідного описується функцією $y = f(x)$, то відношення $S(x_a) = \left. \frac{dy}{dx} \right|_{x=x_a}$.

Бар'єрні сенсори типу метал-напівпровідник. Широкого застосування в сенсоріці набули бар'єрні контакти метал-н/п. Такі прилади працюють на зміні ВАХ за умови адсорбції газів (вимірюється зміна струму $\Delta I / I$ при $V = const$). Якщо контакт утворює ідеальний без поверхневих (s -) станів н/п, наприклад, електронного типу, то з нього в метал будуть перетікати електрони доти, доки не вирівняються їх рівні Фермі і встановиться термодинамічна рівновага між їх електронними системами.

Для металу робота виходу рівна $e\varphi_m$, для н/п - $e\chi + W_F$. Тому за умови термодинамічних рівноваги, контактна різниця потенціалів буде дорівнювати $W_{кон} = e\varphi_m - e\chi + W_F$. Отже, зі зменшенням відстані між контактом, в металі зростає концентрація електронів, а приконтактна область н/п заряджатиметься позитивно, що призведе до вигину енергетичних зон на величину $e\varphi_{виг,n} = e(\varphi_m - \chi)$ в електронному н/п і $e\varphi_{виг,p} = W_g - e(\varphi_m - \chi)$ в н/п p -типу. Якщо концентрація

s-станів, наприклад електронного типу значна, то у відсутності контакту з металом (термодинамічної рівноваги), s-стани заповнені до рівня Фермі, то під час контакту з металом вони приймають на себе процес встановлення термодинамічної рівноваги, тому за їх наявності величина ОПЗ в н/п при зближенні контактних поверхонь залишається незмінною. Тоді для реального контакту треба ввести деякі поправочні параметри і вираз $e\varphi_0 = 2.3k_B T(pH^* - pH)$ можна переписати як $e\varphi_{\text{виг},n} = Ae(\varphi_m - \chi) + B$. В загальному діодний струм при напрузі v за величиною рівний $J = J_0 \left[\exp\left(\frac{eV}{\kappa \cdot k_B T}\right) - 1 \right]$, де κ - деякий поправочний коефіцієнт, що враховує не ідеальність діода. Якщо діодний струм має надбар'єрну природу, наприклад, зумовлений термоелектронною емісією, то струм насичення прямує до значення $J_0 = AT^2 \exp\left(-\frac{e\varphi_{\text{виг},n}}{k_B T}\right)$.

Сенсори на p-n-переходах. Сенсори на основі p-n-переходах мають переваги над сенсорами на основі полікристалічних плівок та діодів Шотткі. Перш за все, p-n-переходи на основі широкозонних напівпровідників мають високий ПБ для носіїв заряду, відносно високу чутливість при кімнатній температурі та селективність до компонентів газової суміші. Чутливість p-n-переходів на основі A_3B_5 до парів аміаку, води та етанолу зростає, якщо поверхня піддається легуванню атомами сірки, як донорами в даних сполуках. В зв'язку з високим рівнем технологічного розвитку кремнієвої технології, особливу увагу в газових сенсорах привертають також кремнієві p-n-переходи. Встановлено [2.23], що тривала обробка їх в парах аміаку, суттєво покращує струмові характеристики виготовлених на їх основі сенсорів.

Фотоопори. Поширеним фотосенсором є фотоопір, електричний опір якого змінюється при освітленні його поверхні. Фотосенсором є фотодіод. Він складається із сполучених між собою двох н/п n- і p-типу провідностей. Залежність фотоструму від прикладеної напруги U та освітленості Φ має вигляд

$$I = I_0 \left[\exp\left(\frac{eU}{k_B T}\right) - 1 \right] - \eta e \Phi, \text{ де } \eta - \text{квантовий вихід, } I_0 - \text{темновий струм.}$$

Щоб забезпечити малу інерційність фотовідгуку, треба зменшити ємність ФД. Це досягається тим, що поблизу p -області аноду формують розширену область, збіднену носіями струму, майже з ізолюючими властивостями (i -область) (рис.2.4.2). Такий ФД називається $p-i-n$ ФД.

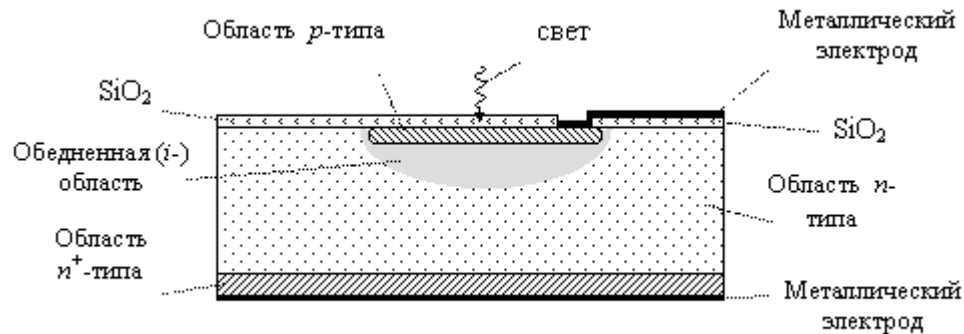


Рис.2.4.2

Щоб в процесі поглинання n/p фотона виникла пара носіїв струму (електрон + "дірка"), енергія фотона повинна бути більша ширини забороненої зони. Наприклад для кремнію вона дорівнює 1.12 eV. Таку енергію мають кванти світла з довжиною хвилі меншою 1.1 $\mu\text{м}$ – це червона межа фотоефекту для чистого кремнію. Оскільки світло довжиною хвилі, меншою 1.1 $\mu\text{м}$, суттєво поглинається кремнієм, тому p -область виготовляють відносно тонкою, щоб світло поширювалось до $p-n$ -переходу. Із-за вузької ширини забороненої зони основні n/p для ФД $\text{Si}, \text{PbS}, \text{PbSe}, \text{InSb}, \text{CdHgTe}$, темновий та дробовий струми в таких ФД за кімнатних температур досить значні, тому ТП вимагають охолодження.

Фотодіоди Шотткі. Для забезпечення чутливості кремнієвих ФД в ультрафіолетовій і фіолетовій областях використовують ФД Шотткі (рис.2.4.3).

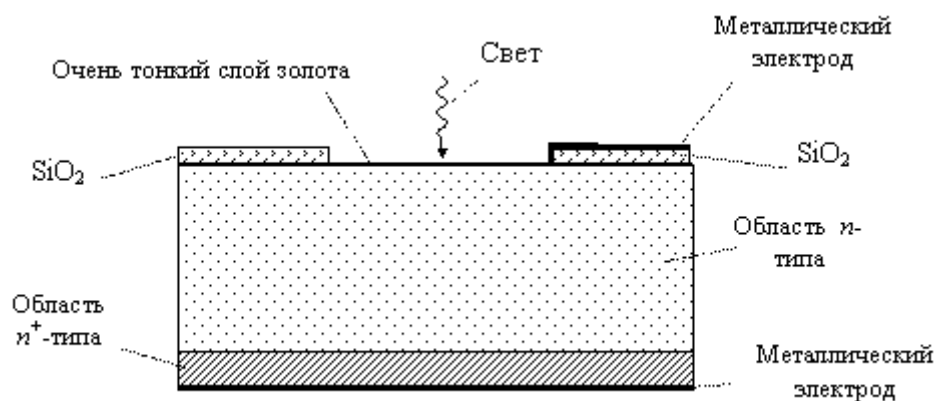


Рис.2.4.3

В них замість $p-n$ - переходу формують БШ, який виникає на межі метал-н/п. Для цього на фото чутливій області кремнію напиленням у вакуумі наносять настільки тонкий шар золота, прозорий для видимого і ультрафіолетового світла.

Фототранзистори. В фототранзисторі фоточутливий елемент має будову транзистора, завдяки чому досягається внутрішнє підсилення. Конструктивно фототранзистор виготовляють так. В підкладку із Ge з обох сторін впаюють індієві контакти, які відіграють роль колектора та емітера. Якщо фототранзистор увімкнути в електричне коло так, що вільна база B (сила струму бази $I_b = 0$), то під час освітлення в ній виникають вільні електрони та дірки.

Для бази, що має n - тип провідності (фототранзисторі $p-n-p$ типу), дірки є неосновними носіями зарядів, тому вони втягуються електричним полем колекторного переходу в колектор, збільшуючі силу колекторного струму. За цієї умови в базі залишилися основні носії заряду - електрони, які збільшують просторовий заряд і зменшують висоту потенціального бар'єру емітерного переходу. Це означає, що діркам стає значно легше переходити з емітера в базу, а потім в колектор, тобто сила колекторного струму через опір R зростатиме. Тому навіть незначне освітлення фототранзистора призводитиме до різкого зростання колекторного струму через опір R , що й визначає його високу фоточутливість.

2.5. Магнітоелектричні мікросенсори [2.24-2.27]

Магнітогальванічні сенсори Холла. Схема досліду *Холла* схематично зображена на рис.2.5.1,а. До провідника прямокутної форми вздовж вісі X прикладене електричне поле E_x , що генерує електричний струм густиною j_x . Вектор індукції магнітного поля спрямований перпендикулярно до поверхні зразка вздовж вісі Z . Зі сторони магнітного поля на рухомі носії струму діє сила Лоренца (рис.2.5.1,а), яка відхиляє електрони провідності вздовж вісі Y і призводить до просторового розділення зарядів.

Електричне поле, яке виникає при цьому називається холлівським. Таке розділення зарядів відбувається до тих пір, доки сила Лоренца не зрівноважиться

з електричною силою $e\mathcal{G}B = eE_{Hy}$. Оскільки $j = e\mathcal{G}n$, то $\mathcal{G} = \frac{j}{en}$ і напруженість

холлівського поля буде дорівнювати $E_{Hy} = \frac{1}{en} Bj = R_{Hy} B j_x$, де R_{Hy} стала Холла.

Холлівське поле спрямоване вздовж вісі Y спрямоване вздовж вісі Y проти додатного напрямку струму, тому стала R_{Hy} від'ємна і за знаком співпадає з знаком носія струму в металі. Це дозволяє за знаком сталої Холла зробити висновок про тип носіїв заряду.

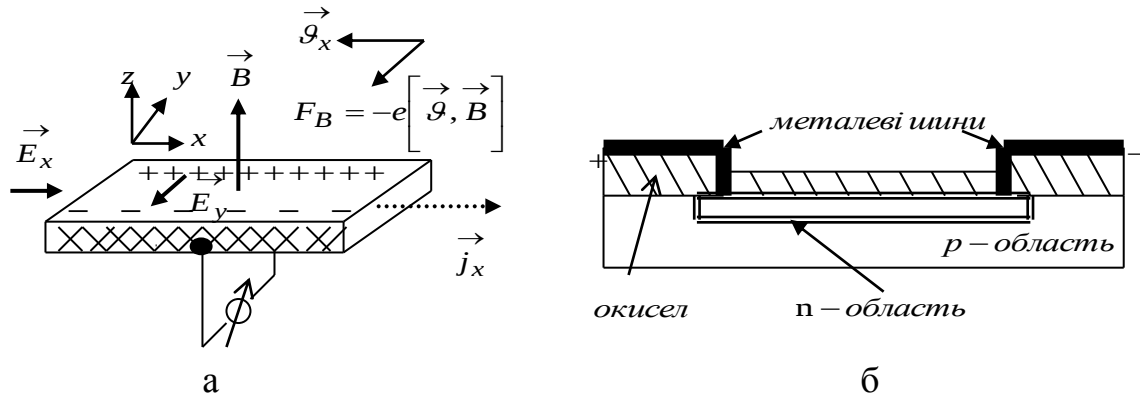


Рис. 2.5.1

У більшості металів стала Холла від'ємна. Однак є й такі, в яких вона додатна, що має квантову природу. Переважно вимірюють холлівську різницю потенціалів між зарядженими поверхнями зразка, відстань між якими дорівнює b .

Тоді $E_y = \frac{\Delta\phi_y}{b}$ і холлівська різниця потенціалів дорівнює $\Delta\phi_y = R_y \frac{BI}{b}$.

Конструкція холлівського сенсора наведена на рис.2.5.1,в. В кристалі Si p -типу провідності формують область n -типу провідності, потім на поверхні ізолюючого шару формують електричні контакти. Одну пару провідників використовують для пропускання крізь n -область електричного струму, а другу — для вимірювання холлівської різниці потенціалів.

Найбільш суттєвою характеристикою давача Холла, як вимірювача параметрів зовнішнього магнітного поля, є його лінійна залежність холлівської напруги від індукції магнітного поля. Сенсори виготовляють у вигляді тонкої плівки напівпровідника Ge (рухливість $3800 \text{ см}^2 / \text{В} \cdot \text{с}$, $GaAs$ ($8500 \text{ см}^2 / \text{В} \cdot \text{с}$) чи $InSb$ ($77000 \text{ см}^2 / \text{В} \cdot \text{с}$), які мають найбільшу рухливість носіїв заряду та амплітуду холлівської напруги. Однак мала амплітуда вихідного сигналу суттєво обмежує їх використання. Давачі Холла виготовляють або з тонких n/p пластин, або з

напилених плівок. Переважно вони мають квадратну форму з чотирма електродами, симетрично розташованими з усіх чотирьох сторін пластини.

Сенсори на спіновому транспорті електронів . Спінотроніка (від англ. spin transport electrons) – це новий напрям магнітонаноелектроніки, який ґрунтується на врахуванні квантових властивостей електрона – спіну, та дії спіну на рухливність електронів у феромагнітних матеріалах^{*)}. Спін електрона може приймати одне із двох значень: $+1/2$ і $-1/2$ (рис.2.5.2,а). Спін має вимірність $\text{Дж} \cdot \text{с}$, квантова ний пропорційно сталій Планка. Основу спінотроніки складають два основні ефекти, які є наслідками особливостей транспорту заряду, що контролюється спіном електрона в нановимірних структурах: гігантський магнітоопір і тунельний магнітоопір.

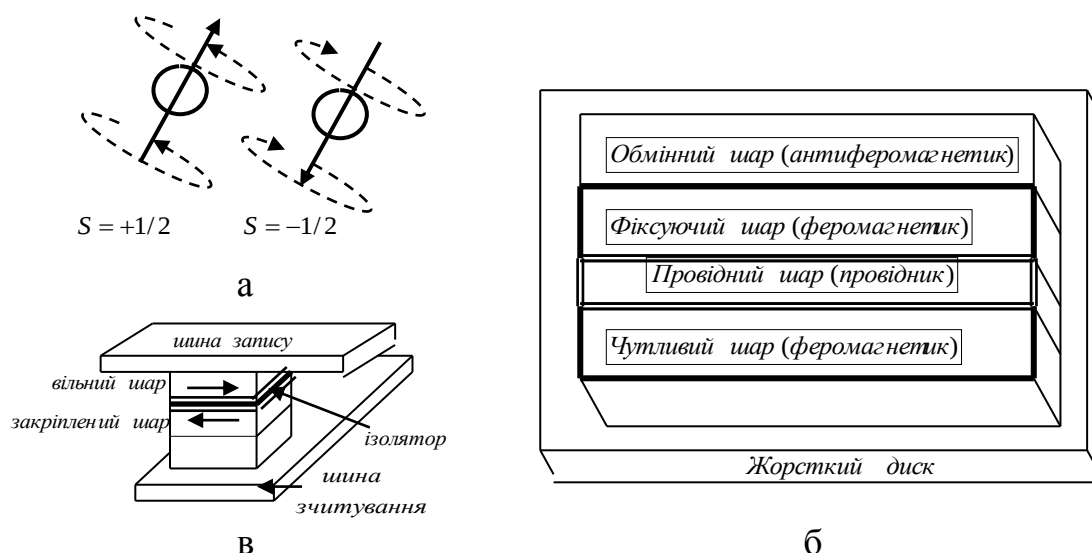


Рис.2.5.2

Явище гігантського магнітоопору. ГМО відкрили Ферт і Грюнберг в 1988 р. у багатошарових структурах Fe/Cr . Суть його полягає в тому, що електрони з різними спінами по-різному рухаються у магнітному полі і опір такої структури при антипаралельній орієнтації намагніченостей сусідніх феромагнітних шарів значно більший, ніж при паралельній орієнтації. Гігантське значення магнітоопору дало поштовх до створення магнітних жорстких дисків з новим типом запису і зчитування інформації із значно більшою густиною (з 1 до

^{*)} Модель спин-залежної провідності була запропонована ще в 1936 р. Моттом для пояснення електричного опору феромагнітних матеріалів поблизу температури Кюрі: А.Ферт. Походження, розвиток і перспективи спінотроніки. (Вибрані розділи Нобелівської лекції. Стокгольм, 8 грудня 2007 р. [Куницький Ю.А. та інші. Основи спінотроніки: матеріали, прилади та пристрої. Суми, СДУ, 2013]

600 Гбіт/дюйм² до 2007 р.). Це відкрило шлях до створення небувалої за ємності зовнішньої пам'яті (до 1Терабайта) для відеозапису.

Велике значення ефекту ГМО мало для створення магнітних сенсорів чи RAM-пам'яті, де основним матеріалом був магнетит Fe_3O_4 . Магнітна головка (рис.2.5.2,б) складається з чотирьох шарів. Верхній шар із антиферомагнетика називається обмінним (exchange) і призначений для того, щоб зафіксувати магнітне поле другого шару, який так і називається - фіксуєчим (pinned). Другий шар виготовляється із феромагнетика (сплави нікелю, заліза, кобальту). Магнітне поле фіксуєчого шару завжди спрямоване в одну сторону. Третій шар – провідний (conducting) – виготовляється із міді і призначений для розділення фіксуєчого і чутливого шарів. Четвертий чутливий шар (sensing) виготовлений із феромагнетика. На відміну від фіксуєчого шару, напрям магнітного поля в чутливому шарі визначається зовнішнім магнітним полем – полем комірки жорсткого диска, що має один біт інформації. В *MRAM* (Magnetic Random Access Memory) -пам'яті основним елементом є тунельний магнітний контакт, в якому два магнітні шари розділені тонким шаром ізолятора (рис.2.5.2,в).

Магнетит – це напівметал, що має максимально можливу поляризацію спінів електронів, майже 100% та відносно високу температуру Кюрі – порядку 858K, тому придатний для пристроїв, що працюють при кімнатній температурі. Плівки магнетиту використовуються для виготовлення сенсорів зчитуючи головок та елементів магніторезистивної пам'яті. Переважно епітаксійні плівки нарощуються на поверхні підкладок із кремнію, сапфіру чи арсеніду галію. Якщо наноструктура виготовлена у вигляді седвича $Fe_3O_4 - Y_2O_3 - Fe_3O_4$, то поляризацією спінів визначається опір тунельного переходу електронів між шарами феромагнетика і може регулюватися зовнішнім магнітним полем. Опір тунельного переходу різний, в залежності від того, паралельні чи антипаралельні спіни в шарах.

Магніторезистивні сенсори. Магніторезистивний ефект був відкритий Томсоном ще в 1856 р., як прояв дії сили Лоренца на носії струму зі сторони зовнішнього магнітного поля, що приводить ефект зміни електричного опору в магнітному полі (так званий магніторезистивний ефект або магнітоопір). В

магніторезисторах довжина перевищує ширину, тому ефектом Холла переважно нехтують. Це пов'язано з тим, що на повільніші електрони сильніше діє поле Холла, тоді як на більш швидкі – сила Лоренца. Це означає, що траєкторія руху носіїв, швидкість яких відрізняється від середньої, не буде співпадати з напрямком

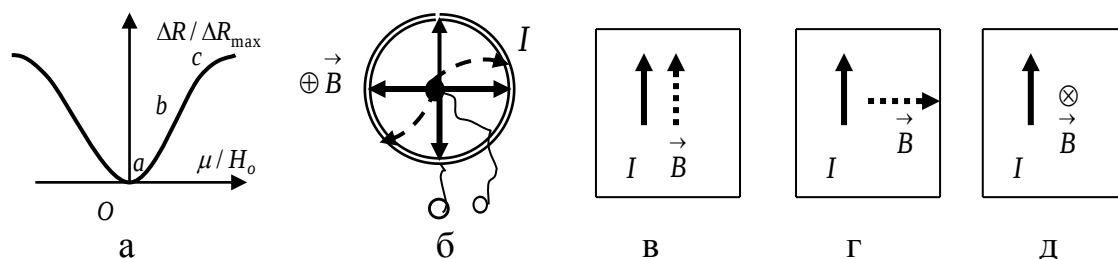


Рис.2.5.3

зовнішнього електричного поля. Таким чином зменшується складова швидкості в напрямку поля і відповідно електричний струм, тобто сила струму стає залежною від величини зовнішнього магнітного поля. Це основна причина появи магнітоопору у металах чи металевих сплавах, чи н/п. Тому оперують відносною зміною значення магнітоопору $\frac{\Delta \rho}{\rho} = \frac{R(H)}{R(H=0)} - 1$, яка при температурі $T = 300\text{K}$ складає до 0.1% у полях $H = 0.8\text{MA/м}$. Переважно із збільшення магнітного поля магнітоопір зростає. Значно більші зміни магнітоопору до 20-30% проявляються в манганітах лантану із структурою перовскіту.

Магніторезистивна характеристика показана на рис.2.5.3,а. Тут H_0 - ефективне поле анізотропії в металі, як сума розмагнічу вального поля анізотропії і поля анізотропії, що виникає під час елементарного процесу осадження матеріалу. За відсутності зовнішнього магнітного поля в елементі домена в точці *a* формується доменна намагніченість. Із збільшення напруженості поля магнітоопір зростає доти, доки на ділянці *b* (характеристика лінійна) елементи не обернуться на 45° щодо довжини. Подальше збільшення H призводить до насичення (точка *c*). Отже, магніторезистивний сенсор працює модуляції домена зовнішнім магнітним полем. Оскільки магніторезистивний ефект зростає із зменшенням відношення довжини до ширини сенсора, то для одержання високого внутрішнього опору за відсутності магнітного поля і суттєвої зміни

опору з ростом магнітного поля, кілька сенсорів сполучають між собою послідовно.

Наглядом моделлю проектування МРС необмежених розмірів є диск Кобрино (рис.2.5.3,б). В ньому один контакт нанесений в центрі диску, а другий на його бокову поверхню. Без магнітного поля струм спрямований вздовж радіусу. В магнітному полі, вектор індукції якого перпендикулярний поверхні диска, носії відхилятимуться від радіальних напрямків. Лінії струму видовжуються, однак виникнення холлівського поля не відбувається. Тому диск Корбино має максимальний ефект магнітоопору. Якщо диски Корбино виготовляються з н/п матеріалів, то омичні контакти формуються шляхом створення областей з високою провідністю. Такі сенсори створені на монокристалах $InSb - NiSb$. Як встановлено експериментально, в малих магнітних полях, зміна опору пропорційна квадрату магнітної індукції, тоді як в сильних – першому ступеню. Основними елементами структури датчика є два феромагнітні шари, виготовлені із сплаву Co, Ni, Fe і розділених прошарком немагнітного матеріалу Cu, Ag, Au . Фіксований шар із $FeNi, FeIr, NiO$ забезпечує обмінну взаємодію з феромагнітним шаром. МРС використовуються також для вимірювання напруженості магнітного поля в магнетометрах.

МРС складаються із двох чутливих елементів у вигляді меандри із електричним опором від 30 Ом до 1000 Ом. Їх вмикають у місткову схему, на виході якої вимірюється зміна електричної напруги, зумовленої зміною магнітоопору плівкового магнітрезистора із $Ni - Fe$ чи $Ni - Fe - Co$ під дією магнітного поля. Використовуються різні геометрії вимірювання магнітоопору, в залежності від взаємної орієнтації підкладки і напрямку протікання струму відносно напрямку вектора магнітної індукції (показано на рис.2.5.3: в- поздовжня; г- поперечна і д- перпендикулярна геометрії).

Окремий клас магнітосенсорів складають квантові магнітометри, принцип дії яких базується на вільній прецесії магнітних моментів ядер або електронів у зовнішньому магнітному полі частотою $\Omega = \gamma H$ (γ - гіромагнітне відношення) та інших квантових ефектах, як ядерний магнітний резонанс чи електронний парамагнітний резонанс. Чутливість квантових магнітометрів

визначається шириною резонансної лінії (порівняльна характеристика для деяких подана в таблиці 2.5.1.

На рис.2.5.3,в схематично показаний іверсний магнітострикційний сенсор на основі ефекту Віларі, де чутливим елементом є мікротрансформатор, що виготовлений із плівки магнітоеластичного сплаву $Ni_{0.45}Fe_{0.55}$ і має *O*-подібну форму [2.26; 2.27]. Чутливий до магнітного поля сенсор виготовляють у вигляді

Таблиця 2.5.1

Тип магнетометра	Ширина резонансної лінії, Г, нТл
Цезій	20
Гелій 3	2.74×10^{-5}
Гелій 4	70
Оверхаузен	4
Калій	0.1 – 10
Протон	15

кантелівера, на поверхню якого наноситься речовина із ефектом магнітострикції. Під дією магнітного поля балка вигинається, змінюючи резонансну частоту коливання. Ці сенсори здатні фіксувати магнітні поля до 10^{-5} Гаусс або 10^{-9} Тесла .

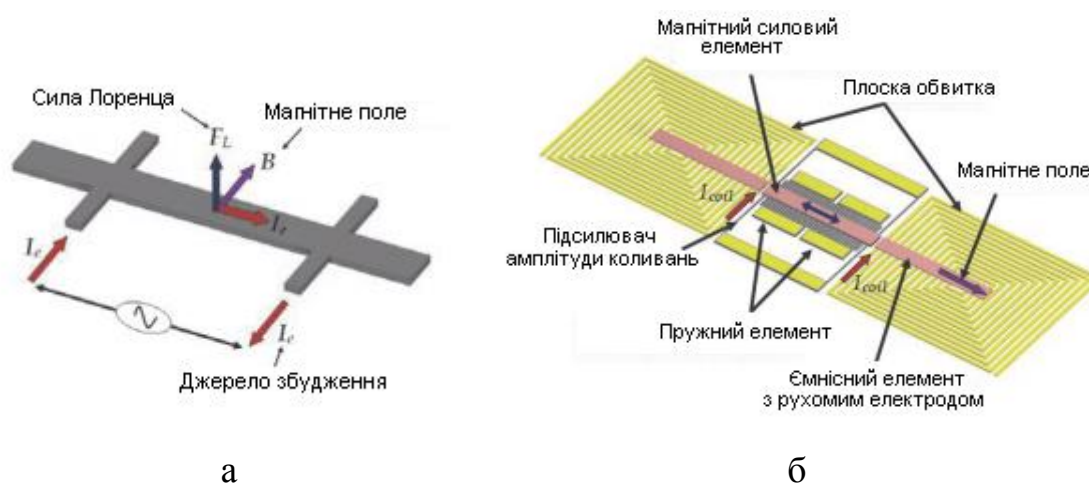


Рис.2.5.4 [2.26;2.27]

Магніторезонансні сенсори. Фізичний принцип їх конструювання зображений на рис.2.5.4. На рис.2.5.4,а, показана орієнтація вектора сили Лоренца зі сторони зовнішнього магнітного поля, що діє на провідник із струмом. Сполучення

магнітодинамічного контура з пружними елементами формує резонансні властивості коливної системи (фізичний аналог конденсатора, рухома пластина якого підвішена на пружині). На рис.2.5.4,б, показані розглядувані сенсори в

інтегральному виконанні. Якщо сенсор сполучений із джерелом змінного струму $I_0 = I_0 \sin \omega t$, то змінного характеру має дія сили Лоренца на кантелівер, коливання якого мають резонансний характер.

2.6. Сенсори деформацій

Тензорезистивні сенсори. Тиск є одним із основних контрольованих параметрів в системах керування багатьма процесами. Тому основною їх ланкою є первинні перетворювачі тиску, які працюють на різних принципах, одним з яких є тензорезистивний ефект. Тензорезистивним ефектом називається ефект зміни електричного питомого опору матеріалу під дією деформації. В напівпровідниках Si, Ge він пов'язаний із багатодолинною структурою енергетичних зон. За відсутності зовнішньої механічної дії, мінімуми зони провідності (долини) симетричні і концентрація електронів в них однакова. Під дією деформації в напрямку одної із кристалографічних вісей, мінімуми вздовж них перебуватимуть не в однакових умовах по відношенню деформації. Зміщення долин, яке при цьому виникає, призводить до перетікання електронів із одної долини в іншу, тобто деформація спричиняє появу анізотропії електропровідності. В полікристалах і металах тензоефект зумовлений іншими механізмами, як зміною енергії іонізації домішок, зміною рухливості носіїв заряду, зміною ширини забороненої зони. Треба відзначити, що явище тензоефекту справджується лише в анізотропних матеріалах.

Електричний опір робочого тіла тензометра у вигляді бруска визначається за формулою $R = \rho_R \frac{\ell}{ab}$, тому зміна опору дорівнює

$$dR = \frac{1}{ab} d\rho_R + \frac{\rho_R}{ab} d\ell - \frac{\rho_R \ell}{a^2 b} da - \frac{\rho_R \ell}{ab^2} db, \text{ звідки відносна зміна } \frac{dR}{R} = \frac{d\rho_R}{\rho_R} + \frac{d\ell}{\ell} - \frac{da}{a} - \frac{db}{b}.$$

Якщо позначити $\varepsilon_\ell = \frac{d\ell}{\ell}$, $\varepsilon_a = \frac{da}{a} = -\nu_\sigma \varepsilon_a$, $\varepsilon_b = \frac{db}{b} = -\nu_\sigma \varepsilon_b$, де ν_σ -модуль Пуассона,

то з формули $\frac{\Delta R}{R} = k_\sigma \varepsilon$ маємо, що тензометричний коефіцієнт (або фактор тиску)

дорівнює $k_\sigma = \frac{d\rho_R / \rho_R}{\varepsilon_\ell} + (1 + 2\nu_\sigma)$. Перший доданок описує внесок

п'єзореzystивного ефекту $\frac{d\rho_R / \rho_R}{\varepsilon_\ell}$, а другий – зумовлений зміною геометричних

розмірів $(1 + 2\nu_\sigma) \cong (1 \div 2)$. У сенсорах на металічній основі геометричний ефект є домінуючим, а в напівпровідникових сенсорах п'єзореzystивний ефект є приблизно в 50 разів більший, ніж геометричний. Фізичні принципи вимірювання тензометричних коефіцієнтів для різних кристалографічних орієнтацій монокристалічного кремнію показані на рис.2.6.1,а. Зв'язок п'єзометричних коефіцієнтів з кристалографічними напрямками (індекси Міллера) для кубічного кристалу показаний на рис.2.6.1,б.

Нижче приведена система рівнянь, що зв'язує між собою п'єзореzystивні і

$$\begin{aligned} E_1 &= \rho_R I_1 + \rho_R \pi_{11} \sigma_1 I_1 + \rho_R \pi_{12} (\sigma_2 + \sigma_3) I_1 + \rho_R \pi_{44} (I_2 \tau_3 + I_3 \tau_2), \\ \text{електричні параметри: } E_2 &= \rho_R I_2 + \rho_R \pi_{11} \sigma_2 I_2 + \rho_R \pi_{12} (\sigma_1 + \sigma_3) I_2 + \rho_R \pi_{44} (I_1 \tau_3 + I_3 \tau_1), \\ E_3 &= \rho_R I_3 + \rho_R \pi_{11} \sigma_3 I_3 + \rho_R \pi_{12} (\sigma_1 + \sigma_2) I_3 + \rho_R \pi_{44} (I_1 \tau_2 + I_2 \tau_1); \end{aligned}$$

$$\frac{1}{\rho_R} \begin{bmatrix} \Delta\rho_{R1} \\ \Delta\rho_{R2} \\ \Delta\rho_{R3} \\ \Delta\rho_{R4} \\ \Delta\rho_{R5} \\ \Delta\rho_{R6} \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & 0 & 0 & 0 \\ \pi_{12} & \pi_{11} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{12} & \pi_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \pi_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \pi_{44} \end{bmatrix} \cdot \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix}, \text{ значення яких приведені нижче.}$$

	ρ_R	π_{11}	π_{12}	π_{44}
	Ω, cm	$10^{-11} Pa^{-1}$	$10^{-11} Pa^{-1}$	$10^{-11} Pa^{-1}$
$n-Si$	11.7	-102.2	53.4	-13.6
$p-Si$	7.8	6.6	-1.1	138.1

На рис.2.6.1,в показаний принцип створення резистивного тензодатчика на поздовжньому (ℓ) і поперечному (t) ефектах. В цих випадках відносна зміна опору визначається як $\frac{\Delta R}{R} = \pi_l \sigma_l + \pi_t \sigma_t$ [26]. Значення коефіцієнтів для n - і $p-Si$ π -коефіцієнти дорівнюють:

$n-Si$	$\pi_l = -31.2 \cdot 10^{-11} Pa^{-1}$	$\pi_t = -17.6 \cdot 10^{-11} Pa^{-1}$
$p-Si$	$\pi_l = +71.8 \cdot 10^{-11} Pa^{-1}$	$\pi_t = -66.3 \cdot 10^{-11} Pa^{-1}$

Розглянемо приклад побудови сенсора тиску, побудованого на п'єзореzystивних елементах. Для того, щоб збільшити чутливість сенсорів, використовують поєднання кількох п'єзореzystорів, з'єднаних за класичною схемою у вигляді містка Сімпсона [2.35]). Якщо на тонку пластину сенсора діє тиск, то виникають

механічні напруження і максимальний прогин досягається посередині цієї пластини. П'єзореzystори, орієнтовані перпендикулярно до границь пластини, зазнають позитивного приросту опору, тому що присутній поздовжній ефект. Відповідно поперечний п'єзореzystивний ефект присутній в паралельно орієнтованих п'єзореzystорах і опір у них зменшується. Структурна схема сенсора тиску показана на рис.2.6.2,а, а місткова схема сплучення між собою окремих тензоелементів показана на рис.2.6.2,б. Температурна залежність у такому сенсорі практично відсутня, оскільки всі резистори при мостовому

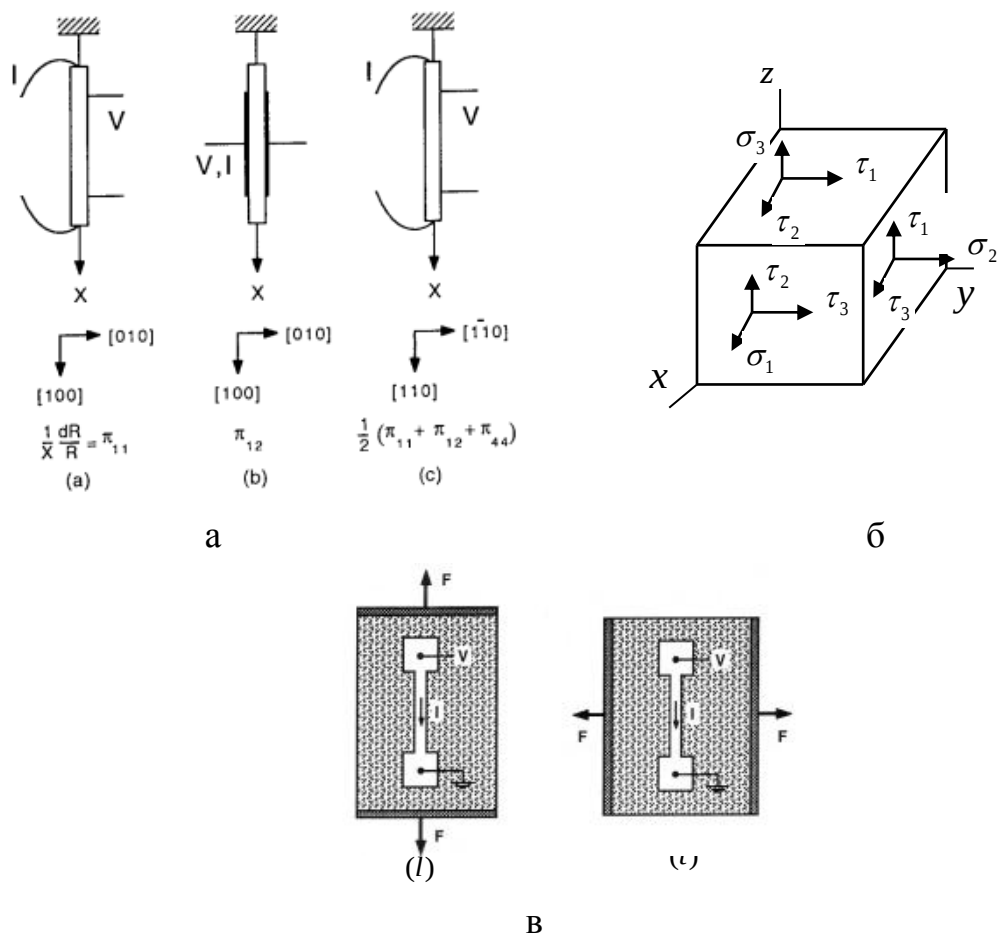


Рис.2.6.1

Нижче приведений зв'язок тензометричних коефіцієнтів з кристалографічними напрямками – індексами Міллера для кубічних кристалів (Si, Ge).

l – напрям	π_l	t – напрям	π_t
(100)	π_{11}	(010)	π_{12}
(001)	π_{11}	(110)	π_{12}
(111)	$\frac{1}{3}(\pi_{11} + 2\pi_{12} + 2\pi_{44})$	(110)	$\frac{1}{3}(\pi_{11} + 2\pi_{12} - \pi_{44})$

(110)	$\frac{1}{2}(\pi_{11} + \pi_{12} + \pi_{44})$	(111)	$\frac{1}{3}(\pi_{11} + 2\pi_{12} - \pi_{44})$
(110)	$\frac{1}{2}(\pi_{11} + \pi_{12} + \pi_{44})$	(001)	π_{12}
(110)	$\frac{1}{2}(\pi_{11} + \pi_{12} + \pi_{44})$	(110)	$\frac{1}{2}(\pi_{11} + \pi_{12} - \pi_{44})$

з'єднанні залежать від однієї температури в сенсорі. Проте внаслідок температурної залежності вбудованих резисторів змінюється чутливість вимірювання тиску. Тому, звичайно потрібно використовувати схему для температурної компенсації вихідного параметра. Для прикладу можна використати опорне джерело струму замість джерела напруги. У цьому випадку при мостовому з'єднанні резисторів, вихідна напруга, практично, не піддається впливу температури.

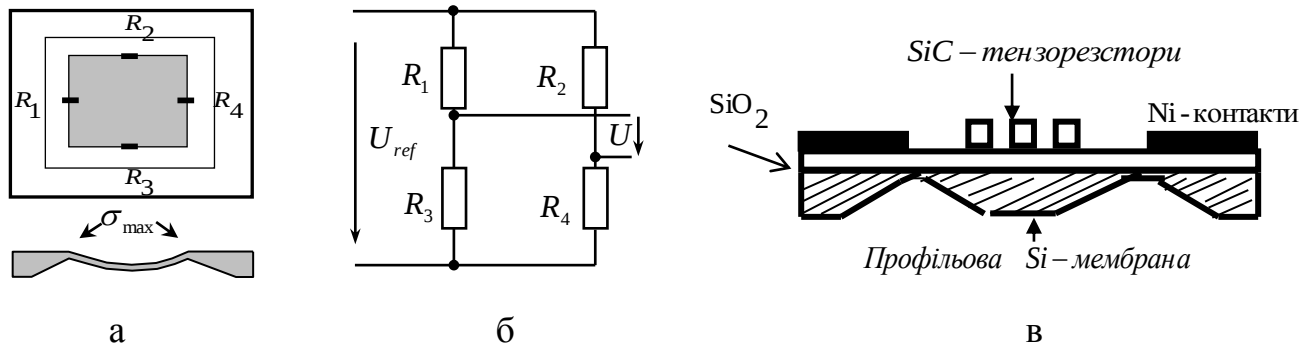


Рис.2.6.2

Поширені тензoeлементи з профільними мембранами вузлом чутливого елемента із кристалу $n-Si$, як пружного елемента. На лицевій поверхні нанесений плівка Si_3N_4 / SiO_2 , на якій сформовані резистори, терморезистор із SiC та металеві контакти (рис.2.6.2,в).

Частіше мембрану виготовляють із Si n -типу провідності (для електронного напівпровідника тензометричний коефіцієнт від'ємний, тобто електричний опір зменшується із збільшенням деформації) в площині (100), прямокутної форми, тензорезистори - p -типу провідності, що дозволяє застосувати об'ємну технологію мікромеханіки формування ІС опрацювання сигналу разом із тензорезистивною структурою на поверхні кремнієвої пластини. Молекулярний контакт тензорезисторів із мембраною дозволяє на зворотній поверхні формувати маски з подальшим анізотропним травленням.

Одним із методів контролю товщини мембрани є ІЧ інтерферометрія Фабрі-Перо, згідно з яким коефіцієнт пропускання за інтенсивністю обчислюється як $T = \frac{(1 - \sqrt{R})^2 \exp(-2\alpha_I h)}{1 - R^2 \exp(-2\alpha_I h)}$, де α_I - коефіцієнт загасання, що обчислюється як $\alpha_I = -\frac{1}{2\tau_I} \ln\left(\frac{T}{(1 - \sqrt{R})^2}\right)$. Якщо тензомембрана виготовлена у вигляді рухомого елемента, то її можна використати як рухомий електрод. Тоді ступінь деформації мембрани фіксується за величиною зміни електричної ємності відносно нерухомого електрода (ємнісний метод).

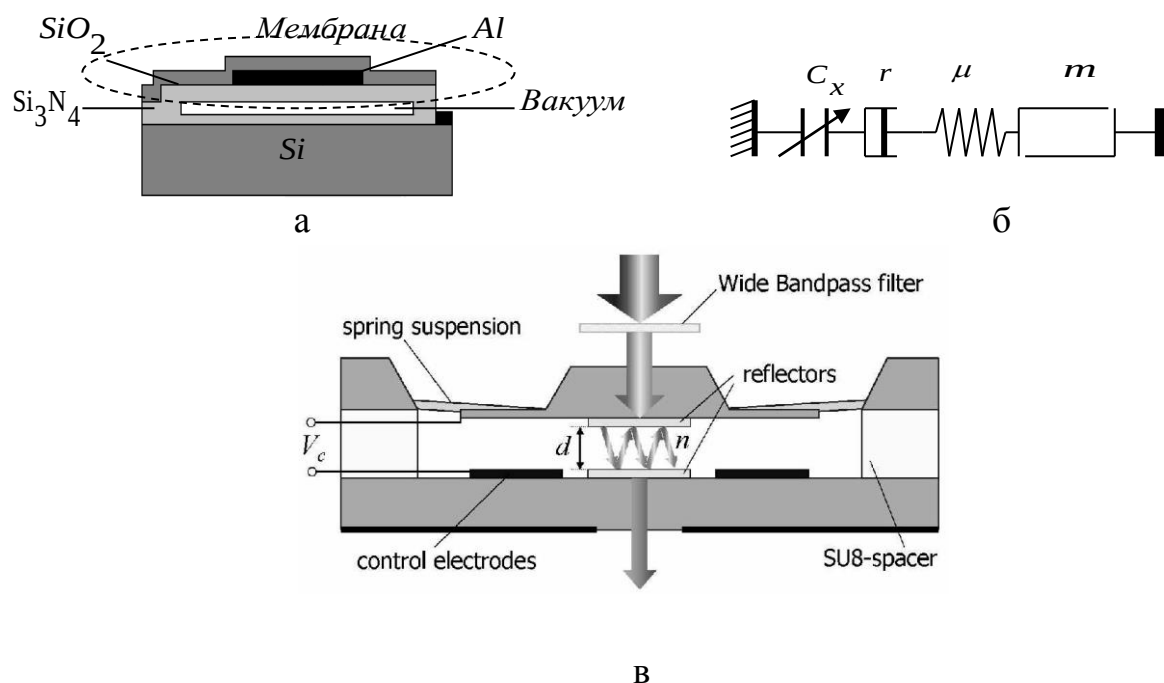


Рис.2.6.3

Ємнісні сенсори тиску. Фізична модель ємнісного сенсора наведена на рис.2.6.3,а, а на рис.2.6.3,б – його електромеханічна модель. Система рівнянь, що описує фізичний принцип роботи ємнісного сенсора має вигляд:

$$\begin{cases} dF = \frac{CU}{x} (dU - \frac{U}{x} dx) \\ dQ = -\frac{CU}{x} dx + CdU \end{cases}. \text{ Ці рівняння нелінійні, однак в диференційній формі}$$

правильно відображають взаємозалежності параметрів механічної і електричної частин. Записана система не враховує втрати енергії в процесі її перетворення. Так, для ємнісного сенсора з круглими електродами радіусом R_0 , за відсутності

зовнішнього тиску, ємність конденсатора дорівнює $C = \frac{\varepsilon_0 \pi R_0^2}{d}$, де d - початкова відстань між пластинами конденсатора. При деформації мембрани, відносна зміна ємності дорівнює [2.29]: $\frac{\Delta C}{C} = \frac{(1-\nu_\sigma^2)R_0^4}{16E_\sigma b d} p$, b - товщина мембрани, p - зовнішній тиск.

Оптичні сенсори. Поширеним оптичним детектором величини деформації і тиску є метод ІФП (рис.2.6.3,в). В ньому промінь світла від лазерного світлодіода скеровується на плоско паралельну структуру, що складається з скляної пластини (ІФП) зверху і діафрагми знизу. Два промені, що відбилися від двох дзеркал ІФП інтерферують між собою і результуюча інтенсивність фіксується ФП. В процесі деформації діафрагми, змінюються конструктивні параметри інтерферометра, що зумовлює зміну сигналу інтерференції.

Сенсори на ниткоподібних напівпровідникових кристалах. НК переважно кремнію і германію та їхніх твердих розчинів виявились найкращою елементною базою для створення сенсорів різних фізичних величин, як тиску і деформацій, механічних рухів, температури, параметрів атмосфери, тощо. Сенсори на їх основі характеризуються високою чутливістю до зовнішньої дії, можливістю роботи в динамічних режимах, компактні у виконанні, тому знайшли широке використання в науці і народному господарстві. За характером прояву фізичних властивостей, НК поділяють на групи за розмірами: діаметром 50 нм , в яких проявляються розмірні ефекти, мезоскопічні - $50\text{ нм} \div 0.5\text{ мкм}$ та макроскопічні, якщо $d > 1\text{ мкм}$.

Якісні НК одержуються шляхом газотранспортних реакцій. З точки зору застосування НК для тензосенсорів, здебільшого використовуються кристалографічні напрямки $\langle 111 \rangle$ для Si і $\langle 001 \rangle$ для Ge , що пов'язано з тим, що у цих матеріалах саме в таких напрямках проявляється максимально тензоефект. Крім цього, на відміну від масивних зразків, НК характерна структурна досконалість, що особливо робить їх перспективними для сенсорної електроніки. Комплексні дослідження авторів [2.30] засвідчили, що шляхом легування моделюється функціональна залежність між електричними і механічними

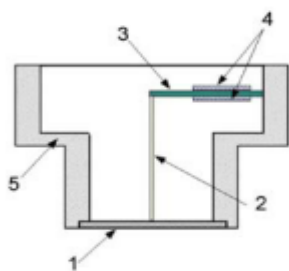
закономірностями, що розширює можливості інтеграції сенсорів на НК в МЕМС. При цьому важливо зазначити прояв гігантського пьезорезистивного ефекту при пониженні температури в області концентрацій, при яких формується перехід метал-діелектрик. Істотний вплив механічної деформації на електричні властивості функціональних матеріалів свідчить про необхідність врахування деформаційно-стримульованих ефектів під час моделювання технічних параметрів МЕМС.

Зміна траєкторії руху носіє заряду в матеріальному середовищі під дією магнітного поля призводить до появи так званого його магнітного опору. В загальному, має місце анізотропія магнітного опору, в тому числі він може досягати від'ємних значень, що пояснено у рамках теорії квантових поправок до провідності за рахунок слабкої локалізації дірок. Наближення до фазового переходу метал-діелектрик з діелектричного боку приводить до зміни механізму магнітоопору: якщо в глибокій діелектричній області – це сильна локалізація дірок унаслідок стискання хвильових функцій локалізованих дірок магнітним полем, то в області переходу із металевого боку переходу метал-діелектрик – це механізм слабкої локалізації дірок.

В цілому вплив температури і деформації в НК зменшує їх магнітний опір, причому ступінь прояву їх залежить від температури. Було встановлено, що магнітоопір залежить від ступеня легування, типу і рівня деформації. Розчин $Si_{1-x}Ge_x$ ефективний термоелектричний матеріал і широко використовується в термоелектричних перетворювачах. Як і є перспективними для застосування як чутливі елементи сенсорів теплових величин. Робота термоелектричних сенсорів ґрунтується на ефектах Зеебека та п'єзо-Зеебека. В різних температурних ділянках має місце різний характер прояву термоелектрики НК. При $T < 50K$ проявляється аномальний ефект пьезоопору, тому аномальність має місце і для термоелектричного ефекту. За температур $T > 50K$ спостерігається класичний характер п'єзо-термоЕРС, коли, аналогічно п'єзоопору, він зростає під час розтягу і зменшується за стиску. З точки зору сенсорних застосувань, цікавими є закономірності прояву термоЕРС за підвищених температур.

Важливим параметром з точки зору сенсорики є термоелектрична добротність. Виявилось також, що магнітне поле має істотний вплив на термоелектричні характеристики НК. Ця зміна незначна за низьких температур, коли істотний вплив на термоЕРС має деформація і проявляється неklasична зміна коефіцієнта Зеебека. В цілому, як засвідчив широкий спектр досліджень (з посиланнями на інші роботи можна ознайомитися в [2.107]), для створення ниткоподібних сенсорів теплових величин рекомендовано використовувати деформовані стиском кристали для інтервалу кріогенних температур та вільні зразки, леговані одночасно домішками бору та золота, для інтервалу підвищених температур. Термоелектричні параметри НК слабочутливі до магнітних полів, що робить перспективними їх для контролю параметрів теплових полів за присутності зовнішнього магнетизму.

Схематичний вигляд н/п НК-сенсора динамічних деформацій на поверхні елемента конструкції, показаний на рис.2.6.4,а. Тут прогин мембрани 1 під дією тиску передається через шток 2 на консольну балку 3 з закріпленими на ній тензорезисторами 4, викликаючи її прогин. Оскільки товщина тензорезисторів набагато менша від товщини балки, при її прогині вони реагують на одновісну деформацію розтягу (на верхній її стороні) і деформацію стиску (на нижній поверхні балки). Обидва тензорезистори утворюють напівміст, вихідний сигнал якого пропорційний тиску.



а



б

Характеристики розроблених високотемпературних сенсорів тиску

Параметри	Числове значення
Діапазон тисків	Від $0...10^3$ Па до $0...20$ МПа
Діапазон робочих температур	$-60...+350$ °С
Чутливість при 20 °С	$40 - 60$ мВ/В
Температурний коефіцієнт $U_{вих}$.	$-0,2\% \cdot \text{град}^{-1}$
Основна похибка	$0,1-0,5\%$
Резонансна частота	$4 - 40$ кГц

в

Рис.2.6.4 [2.107]

На рис.2.6.4,б, показана лінійка високотемпературних сенсорів тиску на основі напівпровідникового кремнію *p*-типу провідності з різними концентраціями бору з питомим опором 0.005 – 0.006 Ом·см. Характеристики розроблених сенсорів тиску приведені на рис.2.6.4,в. На основі кремнієвої технології, розроблені сенсори для вимірювання малих перепадів тиску ($\pm 6 \cdot 10^{-3}$ Па) на фоні високого статичного тиску для космічних потреб, які не зберігали робочі характеристики при низьких температурах 77-100 К.

2.7. Пасивні сенсори фізичних полів

Крім активних сенсорів, існує клас пасивних сенсорів. В пасивних сенсорах параметри вихідного сигналу (імпедансу) змінюються під дією вимірюваного сигналу. Імпеданс сенсора визначається його геометрією, розмірами та властивостями матеріалу - питомим електричним опором, магнітною та діелектричною проникностями. Зміна імпедансу може бути зумовлена дією вимірюваної величини на геометрію, як зміна форми при нагріванні електричним струмом, тощо. Однак, на відміну від активних датчиків, пасивні не вимагають додаткової витрати енергії для генерування інформаційного сигналу. Фізичні принципи перетворення величин і матеріали, що використовуються для побудови пасивних сенсорів, згруповані в таблицю 2.7.1.

П'єзоелектричні сенсори

Принцип дії ППС ґрунтується на п'єзоелектричному ефекті, який проявляється в кристалах і текстурах, що електризуються (поляризуються) під

Таблиця 2.7.1

Вимірювана величина	Фізична характеристика, що піддається зміні під дією вимірюваної величини	Використовуваний матеріал
Температура	Опір	Метали (платина, нікель, мідь), н/п
Наднизькі температури	Діелектрична проникність	Скло, кераміка
Світловий потік	Опір	Напівпровідники
Деформація	Опір	Сплави нікелю, легований кремній
Вологість	Опір, діелектрична проникність	Хлористий літій, оксид алюмінію, полімери
Переміщення	Магнітна проникність, опір	Феромагнітні сплави, вісмут,

		антимонід індію
Рівень	Діелектрична проникність	Рідкі ізоляційні матеріали

дією механічних напруг (прямий п'єзоефект) та деформуються в електричному полі (зворотний п'єзоефект).

Існує декілька технологічних способів поляризації середовища. Поширеною є теплова поляризація, що складається з таких етапів. **1.**Середовище п'єзоелемента, в якому диполі мають довільну орієнтацію, повільно нагрівають до температури, що не перевищує точку Кюрі. Під час нагріву диполі збуджуються і орієнтуються в потрібному напрямку. **2.**П'єзоелектрик вставляють в сильне електричне поле, що спроможне орієнтувати диполі в потрібному напрямку. **3.**Матеріал охолоджується при одночасній дії електричного поля. **4.**Після охолодження п'єзоелектрика до потрібної температури, електричне поле виключають і процес поляризації вважається завершеним. Після охолодження диполі підтримують наведену поляризацію. П'єзоефект знаочутливий, тобто знак деформації залежить від напрямку поля. П'єзоелектричні перетворювачі виготовляють з кварцу, сульфату літію, ніобіту і танталату літію. Широке застосування має п'єзоелектрична кераміка. Перед монокристалом вона має ту перевагу, що перетворювач можна виготовляти довільної форми і розмірів.

Поширеним кристалічним п'єзоелектрикою є кварц. В ньому під час стиску та розтягу виникають поляризаційні заряди елементарній комірці з атомів Si , спарених атомами O_2 , тобто в одну комірку входить три атоми кремнію та шість атомів кисню.

Недеформована кристалічна комірка електрично нейтральна. Деформація комірки супроводжується зміщенням зарядів і між протилежними поверхнями формується різниця потенціалів поляризаційних зарядів. Якщо сили, що зумовлюють поляризацію і напруженість поля не дуже великі, то можна записати

$$\begin{cases} D = \varepsilon_0 E + q \sigma_n; \\ \xi = q' E + \frac{\sigma_n}{E_\sigma}, \text{ де } \xi = \frac{\Delta \ell}{\ell} - \text{відносна зміна лінійних розмірів п'єзоелектрика при} \end{cases}$$

механічній нарузі σ_n , q – і q' - п'єзоелектрична стала прямого та оберненого ефектів.

2.8. Сенсори на квантових ефектах

Сенсори на тунельних діодах. Напівпровідниковий діод на основі виродженого напівпровідника, в якому тунельний ефект приводить до появи на вольт-амперній характеристиці при прямій напрузі ділянки відємної диференційної провідності, відносять до тунельного приладу. Для їх виготовлення використовується напівпровідник з дуже високою концентрацією домішки (10^{18} - 10^{20} см^{-3}). Наслідком високої концентрації домішки в областях, що прилягають до *p-n*-переходу, є мала товщина переходу (біля 10^{-2} мкм), тобто на два порядки менша, ніж в інших напівпровідникових діодах. Через тонкі потенціальні бар'єри відбувається тунелювання носіїв заряду. Другим наслідком високої концентрації носіїв є розщеплення домішкових рівнів з утворенням домішкових зон, які наближені до зони провідності в *n*-області і до валентної зони в *p*-області. Рівні Фермі при цьому вже розташовані в дозволених зонах.

Сенсори Есакі. Незважаючи на те, що аномальний характер ВАХ за умови протікання через нього електричного струму у напрямку виродженого *p-n* переходу спостерігався багатьма дослідниками, лише в 1958 році Есакі вдалося пояснити це явище як квантове тунелювання електронів. З тих пір тунельний діод або, як його ще називають діод Есакі, широко застосовується в практиці як швидкодіючий електронний прилад.

Явище тунелювання електронів можна одержати й між близько розташованими між собою напівпровідниками з різними типами провідності. Чим менші ефективні маси, у порівнянні із масами вільних носіїв, тим для більших відстаней між напівпровідниками можна одержати ефект тунелювання. Діод Есакі конструктивно представляє собою складну періодичну структуру (рис.2.8.1,а) нанорозмірів. В ділянці ВАХ з від'ємним диференціальним опором $R_{\text{диф}} = dU / dI < 0$, діод працюватиме як коливний контур. В іншій ділянці вольт – амперної характеристики опір діода додатний, тому електричні коливання Час тунелювання $\approx 10^{-13} \div 10^{-14} \text{ с}$ і в базі завдяки невеликим напругам неосновні носії заряду не накопичуються. Така мала інерція процесу тунелювання забезпечує використання тунельних діодів до частот сотень гігагерц. Такі властивості резонансно-тунельних діодів роблять їх перспективними для високошвидкісних

приладів терагерцового діапазону і пристроїв з часом перемикання порядку 10^{-12} с і менше.

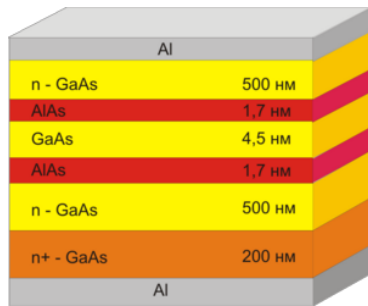
Сенсори на ефекті Джозефсона. Тунельний ефект лежить також в основі ефекту Джозефсона. Суть його полягає в протіканні надпровідного струму між двома надпровідниками крізь екстремально тонкий шар діелектрика. Тоді при низькій температурі між двома надпровідниками через діелектричний шар тече надпровідний електричний струм при відсутності прикладеної з зовні напруги. Щоб надпровідний струм мав помітну величину, товщина діелектричного шару повинна бути дуже малою: $d \approx 10^{-7}$ см. Зазначимо, що в ефекті Джозефсона тунелюють куперівські пари. Наноструктурована джозефсонівська електроніка перспективна для конструювання квантових комп'ютерів. Передбачається, що може бути створений новий тип комп'ютерної асоціативної пам'яті, розподіленої по всій структурі, подібно нейронним мережам живих організмів. Такі система буде спроможна розпізнавати образи, приймати рішення у багатофакторних ситуаціях в реальному часі. В даний час розроблені самі різноманітні квантові пристрої даного типу, в тому числі сенсорного типу, чутливість яких лежить поблизу фундаментального квантового переходу, що значно перевищує відповідну класичних пристроїв.

Відомо також, що коли до джозефсонівських контактів прикласти електричну сталу напругу U , то виникає генерація електромагнітного випромінювання на джозефсонівській частоті $\varpi = \frac{2eU}{\hbar}$. Це явище називається нестационарним ефектом Джозефсона. Генератор на основі ефекту Джозефсона випромінює випромінювання в терагерцовому діапазоні. Так, при напрузі $U = 1$ мВ частота $\varpi = 483.6$ ГГц. В класичному наближенні, фізичні властивості джозефсонівських контактів описується рівнянням Синус-Гордона, а їх шумові флуктуації - квантовою формулою Найквіста.

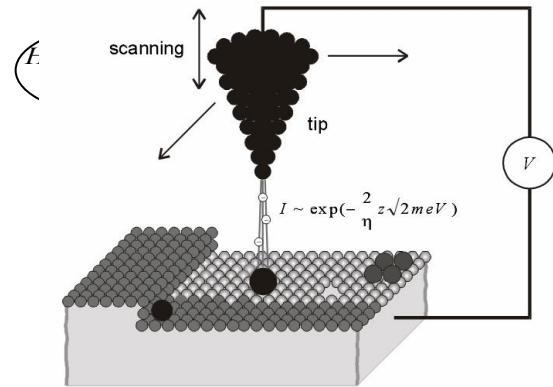
Сенсори кантилеверного типу для скануючих тунельних мікроскопів (*scanning tunneling microscopy* - STM). Тунельний ефект дослідно також був підтверджений явищем виривання електричним полем електронів із поверхні металу при низьких температурах (холодна емісія), що є прямим наслідком

нерівності Гейзенберга. Згідно з нею, якщо в довільний момент часу точно визначити форму потенціального бар'єру, як функцію просторової координати, то до безмежності зростає невизначеність у встановленні кінетичної енергії електрона в цьому бар'єрі, тобто існує певна ймовірність того, що

задовольняється нерівність $\frac{p^2}{2m_0} > W_p$.



а [2.31]



б

Рис.2.8.1

З поміж новітніх практичних застосувань тунельного ефекту особливої уваги заслуговує скануючий тунельний мікроскоп, в якому досягнуто збільшення 10^8 разів і був створений в 1981 р. Біннінгом і Рорером. Головна частина мікроскопа [2.32] – це тунельний сенсор з високим просторовим розділенням, який вимірює струм між всітрян і зразком. В основу принципу роботи мікроскопа покладена можливість позиціювати з високою точністю атомарно гостре вістря (зонда) поблизу поверхні зразка з точністю 0.01 нм . Вістря із вольфраму настільки гостре, що його вершину формує декілька атомів. Зауважимо, що крім тунельного сенсора, в зондовій спектроскопії використовується інший тип атомно-силовий сенсор, на основі якого створений так званий атомно-силовий мікроскоп. Зонд наближають до поверхні кристалу і підводять до них електричну напругу. Оскільки, напруженість електричного поля залежить від характеру нерівностей на поверхні, то пересуваючи вздовж неї вістря, матимемо пульсуючий тунельний струм.

2.9. Фізична механіка макро- і мікросенсорів руху

Вступні поняття. Фізична механіка – це синтез кінематики та динаміки. В кінематиці, використовують різні фізичні моделі, які є спрощеними копіями реальних фізичних тіл і об'єктів, матеріальна точка (*м.т.*), їх система, абсолютно тверде тіло, тощо, які займають певне положення, в тому числі змінне, в просторі в кожний момент чи в певному проміжку часу. Під проміжком часу розуміють перебіг часу між двома фізичними подіями, а моментом часу називають границю між двома суміжними проміжками часу. Його початок – це момент, з якого починається відлік часу.

В класичній механіці постулюється наявність системи відліку, відносно якої простір однорідний та ізотропний. Однорідним є також і час, що плине рівномірно та не залежать від фізичних процесів, які протікають у розглядуваній ділянці простору.

Положення тіла чи *м.т.* в просторі вибраної системи відліку описує геометричне місце послідовних положень кінців радіус – вектора $\vec{r}(t)$ (годограф вектора $\vec{r}(t)$), тому аналітичним параметром руху є радіус-вектор переміщення $\vec{dr} = dx\vec{h} + dy\vec{k} + dz\vec{l}$, де $\vec{h}, \vec{k}, \vec{l}$ – одиничні вектор (орти). Перша похідна по часу від радіус-вектора $\vec{g} = \frac{d\vec{r}}{dt}$ – за означенням вектор швидкості, а перша похідна по часу від вектора швидкості $\vec{a} = \frac{d\vec{g}}{dt}$ – вектор прискорення. Вибравши кількість руху $\vec{p}$ як міру механічного руху, Ньютон узагальнив поняття сили $\vec{F}$, як основну причину руху, тобто якщо імпульс є мірою руху тіла, то сила – мірою дії на нього зі сторони інших тіл. Тому відпала потреба в подальшому вводити спеціальні назви для швидкості зміни $\frac{d\vec{a}}{dt}$ і т.д.

Між кінематикою поступального та обертового рухами існує повна аналогія. Різниця лише полягає в тому, що вектори, що описують обертовий рух, аксіальні або вісьові, що визначаються через векторні добутки звичних векторів, які застосовують для опису поступального руху. Так, замість вектора елементарного

переміщення $\vec{dr}$, в обертовому русі користуються вектором малого кута повороту $\vec{d\phi}$, що входить у векторний добуток як: $\vec{dr} = [\vec{d\phi}, \vec{r}]$ (рис.2.9.1). Замість вектора лінійної швидкості $\vec{g}$, обертовий рух описують вектором кутової швидкості $\vec{\omega}$, що також входить у векторний добуток як: $\vec{g} = [\vec{\omega}, \vec{r}]$. Аналогія проявляється і між тангенціальним $\vec{a}_\tau$ та кутовим $\vec{\varepsilon}$ прискоренням: $\vec{a}_\tau = [\vec{\varepsilon}, \vec{r}]$. Вектори, що входять у ці

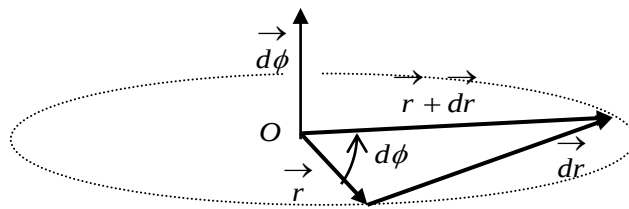


Рис.2.9.1

векторні добутки, задовольняють правилу “буравчика” з правою нарізкою, яке для векторного добутку трьох довільних векторів $\vec{a} = [\vec{b}, \vec{c}]$. Відзначимо також, що на відміну від векторів $\vec{r}, \vec{g}, \vec{a}$, аксіальні вектори $\vec{\phi}, \vec{\omega}, \vec{\varepsilon}$ не мають точок прикладання, тому відноситься до псевдовекторів. Якщо $\frac{d\omega}{dt} > 0$, то вектори $\vec{\omega}$ і

$\vec{\varepsilon}$ спрямовані в одну сторону, а при $\frac{d\omega}{dt} < 0$, вони взаємно протилежні. Трійка

векторів $\vec{g}, \vec{\omega}, \vec{R}$ утворює правоґвинтову систему. Для довільної точки, що не лежить на вісі обертання, в обертовому русі ці вектори пов’язані між собою векторним добутком $\vec{g} = [\vec{\omega}, \vec{R}]$, який називається формулою Ейлера. Коли рух *a.t.t.* складний, то, як перше наближення, його можна розглядати як суперпозицію поступального і обертового рухів. В цьому випадку для довільної

j -тої точки твердого тіла з координатою $\vec{r}_j$, вектор її миттєвої швидкості дорівнює $\vec{g}_j = \vec{g}_{j,c} + [\vec{\omega}, \vec{r}_j]$, де $\vec{g}_{j,c}$ - вектор швидкості поступального руху центру

мас, $[\vec{\omega}, \vec{r}_j]$ - вектор швидкості руху по колу j -тої точки твердого тіла з координатою $\vec{r}_j$.

Ротація швидкості. Нехай тверде тіло одночасно переміщається в просторі поступально з швидкістю руху центра мас $\vec{g}_C$ і обертається навколо вісі, що проходить через нерухому точку з кутовою швидкістю ω , то довільна точка тіла з координатою $\vec{r}$ обертається по колу навколо вісі з лінійною швидкістю $[\vec{\omega}, \vec{r}]$ і повна швидкість руху цієї точки дорівнює $\vec{g} = \vec{g}_C + [\vec{\omega}, \vec{r}]$. Візьмемо ротор від цього виразу.

Так як швидкість поступального руху всіх точок тіла однакова то $\text{rot } \vec{g}_C = 0$ і $\text{rot } \vec{g} = \text{rot } [\vec{\omega}, \vec{r}]$. Якщо вісь симетрії співпадає з віссю Z системи декартових координат, а площина XOY співпадає з площиною обертання точки з координатою $\vec{r}$, то $\omega_x = \omega_y = 0$, а $\omega_z = \omega$ і обертаюча точка характеризується двома координатами (x, y) . Тому, згідно з правилом векторного множення векторів маємо, що: $[\vec{\omega}, \vec{r}] = -\omega y \vec{h} + \omega x \vec{k}$. Підставивши декартові компоненти в ротор векторної функції $\text{rot } \vec{g} = \left(\frac{\partial g_y}{\partial x} - \frac{\partial g_x}{\partial y} \right) \vec{l}$ одержимо, що $\text{rot } [\vec{\omega}, \vec{r}] = 2\omega \vec{l}$, або $\text{rot } \vec{g} = 2\vec{\omega}$.

На відміну від полярного вектора, яким є швидкість поступального руху $\vec{g}$, вектор $\text{rot } \vec{g}$ є аксіальним вектором. Усі точки твердого тіла обертаються з однаковою кутової швидкістю, тому сталим є для нього ротор вектора швидкості його поступального руху (рис.2.9.2,а).

Метод динаміки. Імпульс $\vec{p} = m\vec{g}$ з проекціями $p_x = m g_x$, $p_y = m g_y$, $p_z = m g_z$ - це векторна міра руху *м.т.* Це поняття ввів Декарт і для системи *м.т.*

сформулював закон збереження^{*)}: $\vec{p} = \sum_i \vec{p}_i$. Оскільки $\frac{d\vec{g}}{dt} = \vec{a}$, то зміну

^{*)} Однак Декарт не розглядав імпульс як векторну величину, що привело його в подальшому до помилкових висновків. Це зробив де Меран, що дало йому можливість обґрунтувати закон збереження імпульсу і кінетичної енергії під час удару між тілами.

імпульсу з часом тіла можна виразити як $\vec{F} = \frac{d\vec{p}}{dt}$. Це основний закон динаміки або другий закон Ньютона в імпульсній формі. Він виражає іншу можливість появи сили за рахунок зміни маси тіла. Така сила називається реактивною, а рух тіла під її дією – реактивним.

Добуток $\vec{F} \cdot dt$ - це імпульс сили і за скінчений проміжок часу він дорівнює інтегралу $\int_{t_1}^{t_2} \vec{F} \cdot dt = \vec{p}_2 - \vec{p}_1$, який виражає суть теореми: *Вектор імпульсу сили за проміжок часу від t_1 до t_2 дорівнює зміні вектора імпульсу руху протягом цього ж проміжку часу.*

В динаміці обертового руху інерційні властивості тіла визначаються не лише його масою, але й характеру розподілу самої маси в ньому. Такою характеристикою є момент інерції, який для *м.т.* відносно нерухомого центру означений як $J = m(\vec{r} \cdot \vec{r}) = mr^2$. Якщо через нерухомий центр обертання провести вісь Z^* , навколо якої *м.т.* обертається по колу радіусом r , то момент інерції її відносно цієї вісі дорівнює: $J = mr^2$. В цьому записі вважається, що вісь перпендикулярна площині обертання *м.т.* по колу радіусом r . Так означений момент інерції називається вісний або полярний. Крім вісного моменту інерції, розрізняють відцентрові моменти інерцій. Для системи *м.т.*, вісний момент інерції тіла відносно нерухомої вісі Z обчислюється за принципом суперпозиції

$J_z = \sum_{i=1}^N m_i(\vec{r}_i \cdot \vec{r}_i) = \sum_{i=1}^N m_i r_i^2$, де m_i - маса i -ї точки тіла, що віддалена від вісі обертання на відстань r_i . Для абсолютно твердого тіла з неперервним розподілом маси, момент інерції відносно нерухомої вісі обчислюється як інтеграл $J_z = \int r^2 \cdot dm$, де r – радіус інерції.

Досвід свідчить, що одна і та ж сила, що прикладена до різних точок тіла, як механічної системи, викликає різні види рухів. Дія сили на механічну систему вичерпно характеризується трьома величинами: модулем сили, напрямком її дії та

лінією дії. Усі три величини об'єднуються в таке поняття, як момент сили^{*)} $\vec{M}_O$, модуль якого дорівнює $|\vec{M}_O| = Fr \sin \delta = Fl_F$, де $l_F = r \sin \delta$ - плече сили $\vec{F}$. Напрямок вектора $\vec{M}_O$ визначають за правилом “буравчика”. Лінія, вздовж якої спрямований вектор сили, називається лінією дії сили. Виходячи з означення векторного добутку приходимо до висновку, що вектор моменту вектора сили відносно точки можна записати як: $\vec{M}_O = [\vec{r}, \vec{F}]$ (рис.2.9.3,б). Важливою динамічною характеристикою обертання є вектор моменту імпульсу $\vec{L} : \vec{L}_O = [\vec{r}, \vec{p}]$ (рис.2.9.3,в). Модуль вектора $\vec{L}$ дорівнює $L_O = rp \sin \delta$, де δ – кут між векторами $\vec{r}$ і $\vec{p}$. Підкреслимо також, що частинка з визначеним імпульсом, може мати різні моменти імпульсів відносно різних точок обертання.

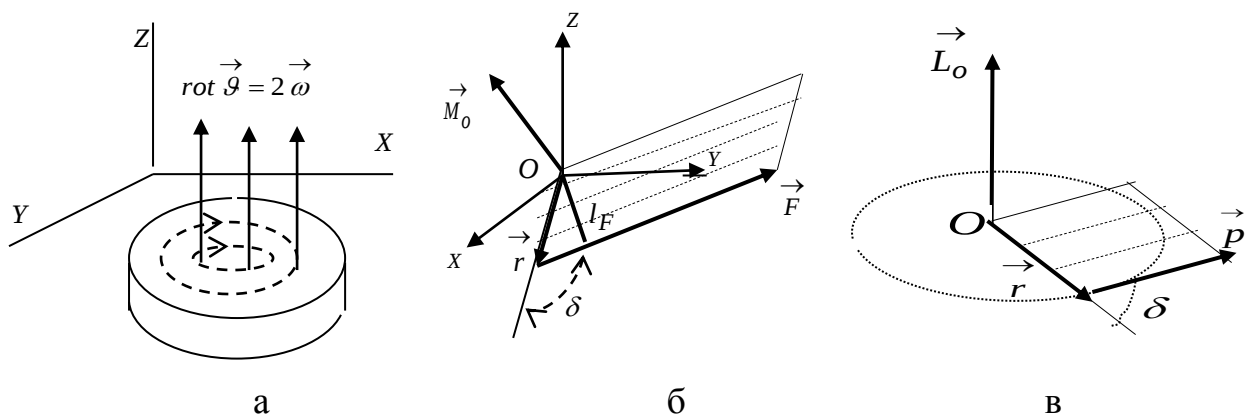


Рис.2.9.2

Треба зазначити, що моменти імпульсу відносно точки і вісі – це також різні поняття. Проекцію вектора $\vec{L}_O$ на довільну вісь чи лінію L_z , що проходить через фіксовану точку відліку, називають моментом імпульсу частинки відносно вісі: $L_z = [\vec{r}, \vec{p}]_z$. Спроектувавши рівняння моментів на довільну вісь Z одержимо рівняння моментів для обертання *a.m.m* навколо нерухомої вісі: $\frac{dL_z}{dt} = M_z$.

^{*)} Вісі обертання, які без спеціального закріплення зберігають свій напрям у просторі, називаються **вільними**. Такими є вісі обертання Землі, дзиги, тіла, що вільно рухається і обертається.

^{*)} Поняття “моменту сили” ввів Леонардо да Вінчі

Важливий висновок випливає з рівняння $\frac{dL_z}{dt} = M_z$. Якщо $M_z = 0$, то $L_z = \text{const}$, хоч

сам вектор моменту імпульсу може змінюватись. Наприклад, вектор $\vec{L}_O$ може обертатись (прецесіювати) навколо вісі під незмінним кутом δ (рис.2.9.3,а).

Якщо **м.т.** масою обертається по колу радіусом R , то проекція на вісь обертання її моменту імпульсу дорівнює $L_z = m g R = m R^2 \frac{g}{R} = J_z \frac{g}{R} = J_z \omega$.

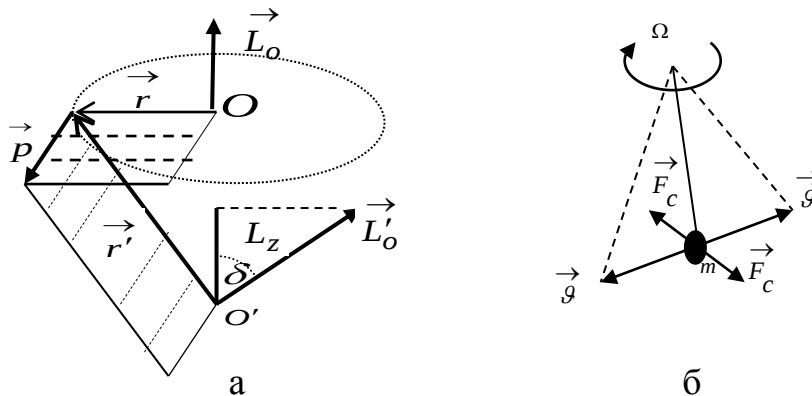


Рис.2.9.3

Сила Коріоліса. Виявляється, що коли тіло відносно системи відліку, яка обертається, додатково зазнає поступального переміщення із відносною швидкістю $\vec{g}_\partial$, то при цьому воно зазнає додаткової дії сили інерції Коріоліса

$\vec{F}_C = m \vec{a}_K$, де a_C - так зване прискорення Коріоліса. Вектор прискорення Коріоліса

визначається як векторний добуток $\vec{a}_C = 2[\vec{g}_\partial, \vec{\Omega}]$. Його діаграма виникнення показана на рис.2.9.3,б на прикладі плоских коливань маятника Фуко. Бачимо,

якщо вони здійснюються в площині вектора швидкості $\vec{g}$, то обертання такого маятника навколо вісі, що проходить крізь точку підвісу з кутовою швидкістю Ω , то сила Коріоліса діють в перпендикулярному напрямку до площини коливань. Це поворотне прискорення.

Сформулюємо основні закономірності коріолісового прискорення і, відповідних сил, пов'язаних з ним: **1.** Сила Коріоліса виникає тоді, коли тіло рухається відносно системи відліку, що перебуває в обертовому русі. Її вектор перпендикулярний до вектора кутової швидкості і завжди лежить у площині, що перпендикулярна до вісі обертання системи відліку; **2.** Вектор сили Коріоліса

перпендикулярний до вектора відносної швидкості тіла $\mathbf{v}_0$; **3.** На відміну від гравітаційної сили та сили пружності, сила Коріоліса діє саме на рухоме тіло, не змінює при цьому його механічну енергію, а отже не змінює модуль вектора швидкості поступального руху тіла, а лише його напрям^{*)}.

Елементи механіки руху за наявності дисипації

Закон тертя Амонтона – Кулона. Тертя поділяють на тертя спокою або статичне тертя, та тертя руху. Тертя спокою виникає при відносному спокою поверхонь тіл, що дотикаються між собою, тоді як тертя руху з'являється між дотичними поверхнями під час їх відносного руху. Переважно сила тертя спокою перевищує силу тертя руху.

Амонтон і Кулон дослідним шляхом встановили наближений закон тертя ковзання або тертя першого роду у вигляді залежності $F_{mp} = k R_n$. Це закон Амонтона – Кулона, де k - коефіцієнт тертя ковзання. Закон Амонтона – Кулона наближений і може використовуватись за умови поміркованих тисків поверхонь одна до одної^{*)}.

Якщо одне тіло котиться вздовж поверхні іншого, то виникає **тертя кочення**. Силу тертя кочення визначають за формулою $F_{T,k} = \mu_k \frac{N}{R}$, де μ_k - коефіцієнт тертя кочення, який має вимірність довжини, R - радіус поверхні тіла, що котиться, а N - сила реакції опори. Якщо сила, що діє на тіло перевищує граничне значення сили тертя спокою, то сила тертя спокою переходить в силу ковзання і тіло набуває прискореного руху. В загальному випадку, сила тертя ковзання залежить від відносної швидкості тіл, поверхні яких труться між собою. При нульовому значенні відносної швидкості, по абсолютній величині сила ковзання може приймати довільні значення в межах $\pm F_{mp}$ значення сили тертя кочення.

Сила лобового опору. Тертя поверхні тіла об рідину під час руху в ній, називається мокрим або рідким тертям, при якому відсутні сили тертя спокою.

^{*)}У фізиці аналогічний характер дії проявляє сила Лорентца, що діє на рухому в магнітному полі заряджену частинку

^{*)} Сухим тертям займалось багато вчених, починаючи з Леонардо да Вінчі. Саме завдяки їх працям, Кулон і Амонтон сформулювали емпіричні закони, які не завжди задовільно узгоджувались з експериментальними даними.

Під час руху тіла в рідкому чи газоподібному середовищі, зі сторони рідини чи газу виникає лобовий опір. Він зумовлений втратою енергії рухомим тілом на виконання роботи для розсування частинок рідини або газу, в середовищі якого воно рухається, та надання частинкам середовища деякої кінетичної енергії. Ньютон встановив, що для невеликих швидкостей руху тіла в рідині чи газі справджується закон $\vec{F}_{on} = -r \vec{g}$, де r - коефіцієнт опору. Вектор сили лобового опору спрямований протилежно до напрямку руху тіла в середовищі.

Сили, які перешкоджають рухові, називаються дисипативними. Вони призводять до незворотних втрат механічної енергії тілом. Характерною особливістю дисипативної сили $\vec{F}_{on} = -r \vec{g}$ є те, що, наприклад, амплітуда коливання фізичного маятника навколо положення рівноваги в середовищі з опором руху зменшується за експоненціальним законом.

Сила внутрішнього тертя Ньютона. Сила внутрішнього тертя в рідині чи газі виникає за наявності градієнта швидкості між шарами рідини (рис.2.9.5) і зумовлена обміном імпульсів між ними. Модуль сили внутрішнього тертя

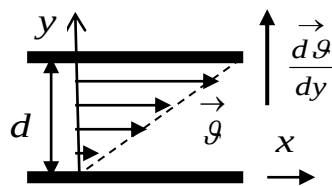


Рис.2.9.5

пропорційний модулю градієнта швидкості $\frac{d g}{d y}$, домноженому на площу самих

шарів S : $F = -\eta \frac{d g}{d y} S$. Знак мінус свідчить, що перенос імпульсу відбувається в

напрямку, протилежному до напрямку градієнта швидкості. Вираз $F = -\eta \frac{d g}{d y} S$ - це

закон Ньютона. Коефіцієнт пропорційності η називається коефіцієнтом внутрішнього тертя або динамічної в'язкості рідини чи газу.

Метод механіки деформованого твердого тіла. Твердим тілам властива міцність і жорсткість, тобто здатність чинити опір при дії зовнішніх сил не руйнуючись і суттєво не змінюючи геометричних розмірів. Під дією сил в

Так, ще в 1895 р. Пенлеве показав, що можливі випадки, коли закон Амонтона-Кулона протирічає основним законам динаміки.

елементах конструкцій виникають зміни розмірів та форми, тобто деформації. В опорі матеріалів розрізняють деформації таких чотирьох основних простих видів: деформація розтягу, як зміна початкової довжини ℓ , яка характеризується абсолютним $\Delta\ell$ та відносним ε видовження (укорочення); деформація зсуву (зрізу), під час якої зовнішні сили зміщують два близько розташованих паралельні плоскі перерізи один відносно одного. Деформація зсуву характеризується абсолютним ΔS та відносним γ зсувом поверхні; деформація кручення вала завдовжки l характеризується взаємним поворотом кінцевих перерізів на кут закручування; деформація згину характеризується лінійним переміщенням у точок вісі стрижня (балки) і кутами повороту перерізів θ своїх початкових положень.

Гіпотези механічних опору та напруження в матеріалах. Основними з них є: пружне тіло суцільне, однорідне і ізотропне; матеріал є однорідним і ізотропним; деформації розтягу пружні і задовольняють лінійному закону *Гука*^{*)}: $\vec{F} = -\mu \vec{x}$, де $\vec{x}$ - вектор деформації, μ - коефіцієнт жорсткості. Пружна сила відноситься до потенціальних сил; мірою інтенсивності внутрішніх сил під час деформації розтягу (або стиску) у даному перерізі тіла є механічне напруження. Кількісною її характеристикою є нормальна механічна напруга σ_n , яка чисельно дорівнює пружній силі F_n , що прикладена до одиниці площі S перерізу тіла, нормально до напрямку поверхні S : $\sigma_n = \frac{F_n}{S}$.

Отже, якщо в напрямку деформації тіло мало початковий розмір x_0 , то згідно з законом *Гука*, відносна деформація пружно деформованого тіла дорівнює відношенню $\varepsilon = \Delta x / x_0$. За помірних навантажень, відносна деформація прямо пропорційна механічній напрузі $\frac{\Delta x}{x_0} = \mu \frac{F_n}{S}$, де Δx - абсолютна зміна його розмірів.

Обернену величину $\kappa = \frac{1}{\mu}$ називають модулем пружності. При деформації поздовжнього розтягу, модуль пружності називають модулем *Юнга*. Модуль *Юнга* дорівнює $E_\sigma = \frac{\sigma_n}{\varepsilon}$. Тоді закон *Гука* набуває вигляду $\sigma_n = E_\sigma \varepsilon$.

Дотична деформація та модуль зсуву. У механіці деформованого тіла крім нормального напруження розглядають дотичне або тангенційне (зсувне) механічне напруження. Якщо в прямокутному бруску закріпити нижню грань, а до верхньої прикласти дотичну силу F_τ (рис.2.9.6), то він зазнаватиме деформації зсуву. Якщо ці деформації пружні, то закон *Гука* запишеться як: $\sigma_\tau = G\varphi$. Тут $\sigma_\tau = \frac{F_\tau}{S}$ – тангенційна механічна напруга, що виникає під час деформації зсуву в розрахунку на одиницю площі верхньої S , G – модуль зсуву.

Напруження у поперечних перерізах. Поздовжня сила N у поперечному перерізі стрижня являє собою рівнодійну нормальних напружень σ_n , що виникають у ньому при розтягу (стиску) $N = \int \sigma_n dS$, де S - площа поперечного перерізу стрижня; $\sigma_n dS$ - елементарна поздовжня сила на безмежно малій площині dS перерізу. Виявляється, що поперечні перерізи стрижня, плоскі перед прикладанням навантаження, залишаються плоскими також після нього (гіпотеза плоских перерізів Бернуллі). Отже, нормальні напруження σ_n розподіляються рівномірно на площі поперечного перерізу стрижня: $N = \sigma_n S$, звідки $\sigma_n = \frac{N}{S}$. Такий розрахунковий випадок називають простим розтягом (стиском). Простий розтяг або стиск справджується лише у перерізах, що віддалені від місць прикладання навантаження.

Напруження в похилих перерізах. Для визначення напружень у довільному похилому перерізі, що визначається кутом α (рис 2.9.6,а) між віссю бруса x і зовнішньою нормаллю n до перерізу, використаємо метод перерізів. Відкинувши верхню частину бруса, з рівняння проєкцій на нормальну вісь n , одержимо вираз для нормального напруження σ_α (інше позначення введене для похилого перерізу тому, що напрям $\vec{n} = \vec{n}(\alpha)$) у похилому перерізі $\sigma_\alpha = \sigma_1 \cos^2 \alpha$, де введене позначення $\sigma_1 = \frac{N}{A}$ – нормальне напруження у поперечному перерізі. Тоді дотичне напруження τ_α , яке діє в похилому перерізі, виразиться

^{*)} У 1676 році Гук сформулював свій закон так: «Яке подовження, така і сила»

як: $\tau_\alpha = \frac{\sigma_1}{2} \sin 2\alpha$, σ_α і τ_α вважаються додатними (рис. 2.9.6,а), а в граничних

випадках рівні: $\alpha = 0, \sigma_\alpha = \sigma_1, \tau_\alpha = 0$;
 $\alpha = 90^\circ, \sigma_\alpha = 0, \tau_\alpha = 0$. Отже, в перерізах, перпендикулярних до

осі стержня і паралельних осі X , немає дотичних напружень, а нормальні напруження набувають екстремальних значень. Площини, на яких немає дотичних напружень, називаються головними площинами, а нормальні напруження, що на них виникають, - головними напруженнями. Встановлено, що при розтягу або стиску матеріал стержня знаходиться в лінійному напруженому стані. У перерізі, що утворює з віссю стержня кут $\alpha=45^\circ$, виникає максимальне дотичне напруження, яке дорівнює $\tau_{\max} = \frac{\sigma_1}{2}$.

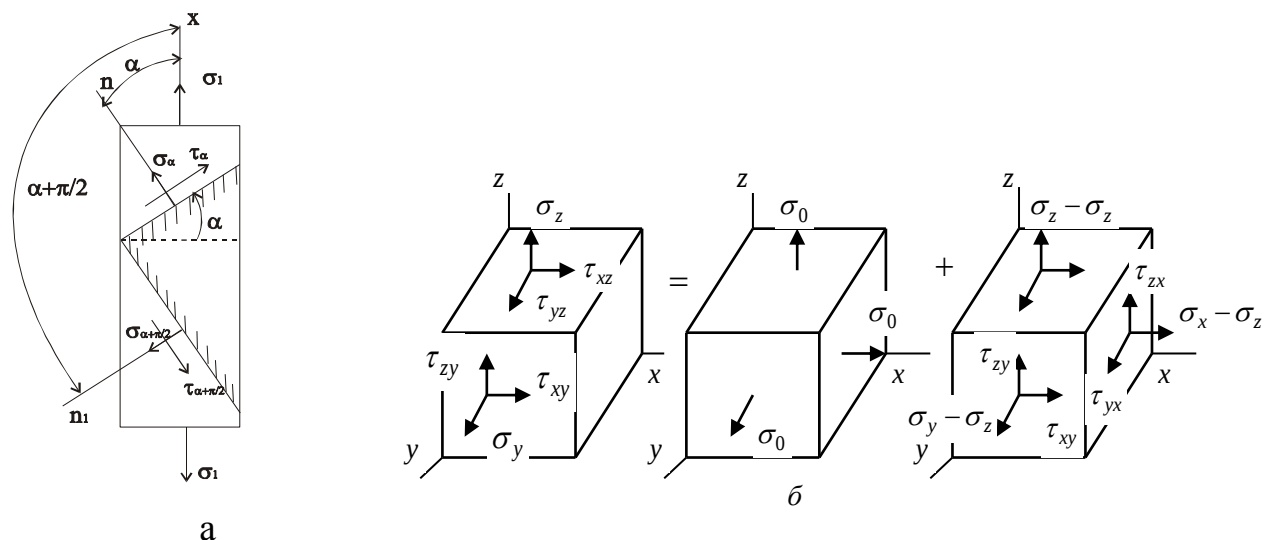


Рис.2.9.6

Тензор напруг. Відомо, що через довільну точку тіла можна провести безмежну кількість похилих площин. Щоб визначити напругу на довільній похилій площині, досить знати напруги в трьох взаємно перпендикулярних площинах вздовж напрямків декартових вісей, тому є дев'ять складових напруги, з яких тільки шість $\sigma_{x,y,z}$ і $\tau_{xy,yz,zx}$ різні і повністю визначають напружений стан тіла в точці. Напругу в кожній точці тіла описує тензор напруг, компонентами

якого є напруги, що утворюють матрицю: $T_\sigma = \begin{vmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z \end{vmatrix}$. В силу закону парності

дотичних напруг, тензор напруг симетричний.

Якщо в трьох декартових напрямках механічні напруги рівні між собою, то $\sigma_0 = \frac{1}{3}(\sigma_x + \sigma_y + \sigma_z)$ і тензор напруг набирає сферично-симетричного вигляду:

$$T_0 = \begin{vmatrix} \sigma_0 & 0 & 0 \\ 0 & \sigma_0 & 0 \\ 0 & 0 & \sigma_0 \end{vmatrix}. \text{ Ввівши одиничний тензор } T_\sigma^0 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}, \text{ сферичний тензор можна}$$

записати так $T_\sigma^0 = \sigma_0 T_1$ звідки загальний одержується новий тензор напруг

$$D_\sigma = \begin{vmatrix} \sigma_x - \sigma_0 & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_y - \sigma_0 & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_z - \sigma_0 \end{vmatrix}, \text{ який називається девіатором напруг. Тоді маємо}$$

рівняння $T_\sigma = \sigma_0 T_1 + D_\sigma$. Кульовий тензор виділяє з напруженого стану рівномірне всебічне розтягання (стиск), при якому змінюється тільки об'єм даного елемента без зміни його форми. Девіатор напруг характеризує стан зсуву, при якому змінюється форма елемента без зміни його об'єму. Кульовий тензор і девіатор показані на рис.2.9.6,б.

Зв'язок між тензорами напруг і деформацій. Узагальнений закон Гука

В механіці пружно-деформованого середовища в лінійному наближенні $\sigma = E_\sigma \varepsilon$ для поздовжньої пружності і $\tau = G\gamma$ для деформації зсуву обґрунтовується зв'язок $G = \frac{E_\sigma}{2(1+\nu_\sigma)}$, де τ - тангенційна напруга, G -модуль зсуву,

ν_σ -модуль Пуассона, що визначається як відношення відносного поперечного

видовження ε_T до відносного поздовжнього видовження ε_L : $\nu_\sigma = \frac{\varepsilon_T}{\varepsilon_L}$. Напруга σ_x

викликає вздовж ребра куба відносне видовження $\varepsilon'_x = \frac{\sigma_x}{E_\sigma}$, тоді як напруга σ_y

викликає вздовж іншого ребра куба відносне видовження $\varepsilon'_y = \frac{\sigma_y}{E_\sigma}$ та скорочення

ребра X на величину $\varepsilon''_x = -\nu_\sigma \varepsilon_y = -\nu_\sigma \frac{\sigma_y}{E_\sigma}$. Аналогічне скорочення викликає

напруга σ_z : $\varepsilon'''_x = -\nu_\sigma \varepsilon_z = -\nu_\sigma \frac{\sigma_z}{E_\sigma}$. Отже шість формул $\varepsilon_x = \frac{\sigma_x - \nu_\sigma(\sigma_y + \sigma_z)}{E_\sigma}$,

$\varepsilon_y = \frac{\sigma_y - \nu_\sigma(\sigma_x + \sigma_z)}{E_\sigma}$, $\varepsilon_z = \frac{\sigma_z - \nu_\sigma(\sigma_x + \sigma_y)}{E_\sigma}$ і зв'язок між кутовими деформаціями і

дотичними напруженнями для кожної координатних площин

$$\gamma_{xy} = \frac{\tau_{xy}}{G}, \gamma_{yz} = \frac{\tau_{yz}}{G}, \gamma_{zx} = \frac{\tau_{zx}}{G}$$

називаються узагальненим законом Гука.

2.10. Динаміка осциляторів інерції в сенсорах руху [2.33]

Осциляторні системи з одним ступенем вільності. Простою коливною системою з одним ступенем вільності (для характеристики положення досить одної координати) є пружинний маятник. В разі відсутності незворотних втрат енергії, для нього в горизонтальному положенні сила тяжіння не відіграє ніякої ролі, а у вертикальному лише зміщує положення стійкої рівноваги. Інші типи простих коливних систем зображені на рис.2.10.1 у вигляді комбінації з пружинних (а), крутильних маятників на одинарній (б) та подвійній (в) підвісках та закріпленої з обох кінців струни (г). Для них сила, що повертає систему в положення стійкої рівноваги, пружна і підлягає закону Гука.

Вільні коливання осцилятора в однорідному полі пружних сил

Коливання пружного осцилятора з зосередженими параметрами, для якого інерційні властивості осцилятора зосереджені в його інертній масі, а пружні – в силі зв'язку з положенням рівноваги, описується рівнянням $\frac{d^2 y}{dt^2} = -\omega_0^2 y$, розв'язок

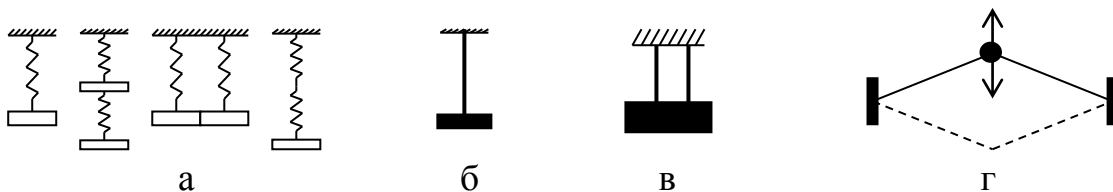


Рис.2.10.1

має вигляд: $y(t) = A \sin \omega_0 t + B \cos \omega_0 t$, де $\omega_0 = \sqrt{\frac{\mu}{m}}$, яке має вимірність частоти, яка називається циклічною частотою. Частота ω_0 і період коливань $T = \frac{2\pi}{\omega_0}$ не залежить від амплітуди, а лише від інертної маси m та коефіцієнта динамічної жорсткості μ , тому такий осцилятор має одну частоту коливань осцилятора ω_0 , тобто одну ступінь вільності і його ще називають одномодовим. Незалежність

періоду коливання від амплітуди називається ізохронністю. Ізохронність вільних коливань була відкрита *Галілеєм*^{*)}.

Загасні вільні коливання. На практиці важливе значення мають загасні коливання. Вплив сил опору або дисипативних сил на характер коливання найпростіше дослідити допустивши, що осцилятор рухається в рідині або в газі з не дуже великою швидкістю. Тоді можна покласти, що сила опору пропорційна швидкості руху осцилятора: $F = -\zeta \dot{y} = -2\gamma m \dot{y}$, де $\gamma = \frac{\zeta}{2m}$, ζ - коефіцієнт пропорційності, а γ - коефіцієнт загасання. Сила опору спрямована завжди проти напрямку руху, що відображено знаком мінус у формулі $F = -\zeta \dot{y} = -2\gamma m \dot{y}$, тому $\zeta > 0$. Сила опору завжди призводить до зменшення повної механічної енергії осцилятора, тому вона неконсервативна. За наявності опору рухові ДР коливань загасного осцилятора запишеться як: $\frac{d^2 y}{dt^2} = -2\gamma \frac{dy}{dt} - \omega_0^2 y$.

Оскільки $\gamma > 0$ і незначне, то коливання й надалі залишаються квазігармонічними (майже гармонічними), однак із зміненою циклічною частотою $\omega = \sqrt{\omega_0^2 - \gamma^2}$. Це зумовлено тим, що опір $F = -\zeta \dot{y} = -2\gamma m \dot{y}$ не змінює лінійності диференційного рівняння і його розв'язок можна виразити через добуток експоненти та гармонічної функції: $y(\gamma, t) = e^{-\gamma t} (A_1 \cos \omega t + A_2 \sin \omega t)$.

За початкових умов $t = 0, y = y_0, \dot{y}_0 = \frac{dy}{dt}|_{t=0} \neq 0$ сталі інтегрування $A_1 = y_0, A_2 = \frac{\dot{y}_0 + \gamma y_0}{\omega}$ і загальний розв'язок має вигляд $y(\gamma, t) = e^{-\gamma t} \left(y_0 \cos \omega t + \left[\frac{(\dot{y}_0 + \gamma y_0)}{\omega} \right] \sin \omega t \right)$. З рис.2.10.2 переконуємось, що коли сила опору коливному руху осцилятора пропорційна його швидкості, то обвідна загасання амплітуди коливань має експоненційний характер. Даний розв'язок має періодично змінний характер за умови $\omega_0 > \gamma$.

^{*)} Своє відкриття Галілей зробив в Пізанському соборі, спостерігаючи за коливаннями важкої люстри і порівнюючи їх з ударами власного пульсу

Насправді, умова періодичності для консервативного осцилятора $y(t+T)=y(t)$ і не виконується, однак задовольняється періодичність повторення нулів функції $F(y) = -2\gamma \frac{dy}{dt} - \omega_0^2 y$. Тому для загасаючого осцилятора із законом опору $F = -\zeta \dot{y} = -2\gamma m \dot{y}$ все одно можна ввести умовний період, як $T = \frac{2\pi}{\omega} > \frac{2\pi}{\omega_0}$.

Визначений таким чином період насправді не дорівнює тому періоду, що визначений як час, через який осцилятор періодично перебуватиме у положенні із тим же значенням амплітуди відхилення, та в одному і тому ж напрямку від положення рівноваги.

Декремент загасання. Загасні коливання з силою опору $F = -\zeta \dot{y} = -2\gamma m \dot{y}$ зручно характеризувати відношенням двох послідовних амплітуд відхилення осцилятора від рівноважного положення в одну і ту ж сторону: $D = \frac{y_{0,n}}{y_{0,n+1}} = \exp(\gamma T) = \exp(\pi \delta_y)$, де $\delta_y = \frac{2\gamma}{\omega}$. Величина D називається

декрементом загасання коливань. Взявши від нього натуральний логарифм, одержимо логарифмічний декремент загасання. Одиницею вимірювання логарифмічного декременту є непер. Якщо $\ln D = 1$, то кажуть, що в системі є загасання, яке дорівнює одному неперу. Виходячи з відношення амплітуд $\frac{y_{0,n}}{y_{0,n+1}}$,

визначають величину $\ln D = 20 \lg \frac{y_{0,n}}{y_{0,n+1}}$ дБ. Якщо $\ln D = 1$, то кажуть, що демпфування дорівнює 1 децибелу, тобто $1_{непер} = 8.68 \text{ дБ}$.

Вплив загасання на амплітуду коливань значно вагоміший, ніж на період. Дійсно, переписавши вираз для логарифмічного декременту у вигляді

$$\ln D = \gamma T = \gamma \frac{2\pi}{\sqrt{\omega_0^2 - \gamma^2}} = 2\pi \sqrt{\left(\frac{\omega_0}{\omega}\right)^2 - 1} \text{ маємо, що коли період загасання коливань}$$

відрізняється від відповідного для незагасаючих T_0 на $0.1T_0$, то $\ln D = 2.879$ і $\exp(\gamma T) = 17.8$, тоді як за період коливань амплітуда зменшується майже у 316 разів.

Час релаксації. Інформативним для характеристики загасаючих коливань є час τ , протягом якого амплітуда коливань зменшується в e (основа натурального

логарифму) разів. Цей час називається часом релаксації. Кількість коливань N_e за час τ дорівнює відношенню τ до періоду T загасаючих коливань $\frac{\tau}{T} = N_e$. Оскільки $\tau = \gamma^{-1}$, то $N_e = \gamma^{-1} T$.

Аперіодична динаміка. При зростанні загасання амплітуда коливань зменшується. Якщо задовольняється умова $\gamma = \omega_0$, то вони стають аперіодичні*) і їх напрям змінюється лише один раз. В цьому випадку розв'язком є плавна функція $y(t) = e^{-\gamma t}(At + B)$ (рис.2.10.2,а), де A і B – деякі сталі. Ця функція – це спадна експонента без періодичних осциляцій синуса чи косинуса. Така закономірність важлива в практичних застосуваннях в електровимірювальних приладах – у балістичних гальванометрах для повернення стрілки у стан рівноваги після акту вимірювання та швидкого встановлення стрілки на належну поділку шкали.

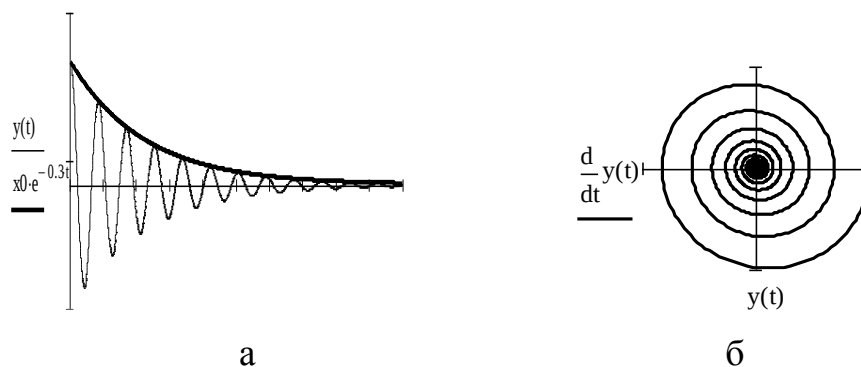


Рис.2.10.2

Фазовий портрет. Нагадаємо, що фазовим простором системи називається простір з вісями “координата–імпульс**). Тоді крива, що отримується з параметричних рівнянь
$$\begin{cases} p(t) = m \frac{dy}{dt}, \\ y(t), \end{cases}$$
 шляхом виключення параметра t , називається фазовим портретом процесу. В кожному осциляторі має координату y і швидкість $\dot{y}$, які на фазовому портреті відповідають одній точці, що називається зображеною. З часом значення величини координати зміщення

*) Аперіодичні – це такі системи, в яких через великі втрати енергії не можуть виникати власні коливання. Період аперіодичного коливання набуває безмежного значення.

**) Фазовий портрет будується і для інших фізичних величин, якщо вздовж одної вісі відкладати значення

у і швидкості $\dot{y}$ змінюються, тому геометричне місце зображуваних точок на площині утворюватиме фазову траєкторію.

Запишемо рівняння лінійних коливань $\frac{d^2 y}{dt^2} = -\omega_0^2 y$ у

вигляді $\left\{ \frac{d^2 y}{dt^2} + \gamma \frac{dy}{dt} \right\} + \left\{ \gamma \frac{dy}{dt} + \omega_0^2 y \right\} = 0$. Обидва доданки мають однаковий вигляд

диференційного рівняння. Перший із них описує фазовий портрет у координатах прискорення - швидкість, а другий – швидкість - координата. Фазовий портрет, що описується першим з них, приведений на рис.2.10.2,б. Це збіжна спіраль, оскільки завдяки наявності сили опору коливанням, що пропорційні швидкості руху осцилятора, амплітуда коливань з часом зменшується за експоненціальним законом.

Енергетичні характеристики осциляторних процесів. Домножимо рівняння руху на елементарне зміщення dy : $m \frac{d\dot{y}}{dt} dy + m\omega_0^2 y \cdot dy = 0$ і перетворимо одержану рівність до вигляду $m\dot{y} \cdot d\dot{y} + m\omega_0^2 y \cdot dy = 0$. Проінтегрувавши цю рівність одержимо так званий перший інтеграл рівняння руху осцилятора з одним ступенем вільності $\frac{1}{2} m\dot{y}^2 + \frac{1}{2} m\omega_0^2 y^2 = W_k + W_p = W_0$.

Дисипативна функція Релея. Введемо функцію $Q(\gamma, t)$, що описує зменшення повної механічної енергії осцилятором під час загасаючих коливань. Якщо в початковий момент осцилятор мав енергію W_0 , то в довільний момент вона дорівнює $W_0 - W(\gamma, t) = W_k(\gamma, t) + W_p(\gamma, t) = W_0 - Q(\gamma, t)$. Обчисливши роботу дисипативних сил та провівши елементарне інтегрування одержимо рівняння балансу енергій (перший інтеграл) загасаючого одновимірного осцилятора: $\frac{m\dot{y}^2(\gamma, t)}{2} + \frac{m\omega_0^2 y^2(\gamma, t)}{2} = W_0 - 2\gamma m \int_0^t \left[\frac{dy}{dt} \right]^2 dt = W_0 + Q(\gamma, t)$. Функція $Q(\gamma, t)$ - називається пасивною функцією Релея.

Встановимо її явний вигляд для дисипативної сили виду $F = -\zeta \dot{y} = -2\gamma m \dot{y}$. Якщо в початковий момент $t = 0$ осцилятор відхилений з положення рівноваги

$y|_{t=0} = y_0$ і перебуває в стані спокою $\vartheta|_{t=0} = 0$, то після інтегрування одержимо, що

$$Q(\gamma, t) = -m y_0^2 \omega_0^2 e^{-2\gamma t} \left\{ 1 + \left(\frac{\gamma^2}{\omega_0^2} \right) \cos 2\omega t + \frac{\gamma \omega}{\omega_0^2} \sin 2\omega t \right\}. \text{ Графік залежності пасивної функції}$$

Релея з часом приведений на рис.2.10.3,а одночасно з закономірностями зміни кінетичної і потенціальної енергії осцилятора. Бачимо, що функція Релея є їх обвідною. Причому дисипативні втрати енергії осцилятором корелюють із тою миттю коливного процесу, в якому енергія осцилятора запасена у вигляді кінетичної. Тому обвідна $Q(\gamma, t)$ не плавна, а має ступінчатий характер і її зміна найбільша в той момент, в який осцилятор проходить через стан рівноваги з найбільшою швидкістю.

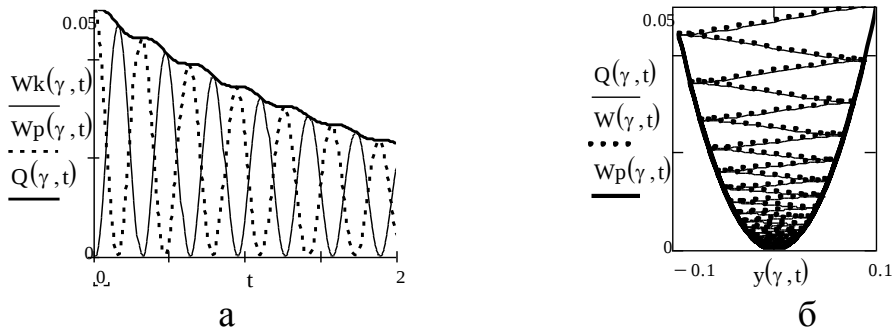


Рис.2.10.3

В гармонічному наближенні повна енергія осцилятора з часом зменшується за експоненційним законом $W(\gamma, t) = W_0 \exp(-2\gamma t)$, де $w_0 = \frac{1}{2} \mu y_0^2$. Тому, після інтегрування одержимо, що дисипативну функцію можна записати в такому вигляді $Q(\gamma, t) = \frac{1}{2} \zeta \vartheta^2$.

За своїм зовнішнім виглядом, формула $Q(\gamma, t) = \frac{1}{2} \zeta \vartheta^2$ формально нагадує вираз для кінетичної енергії, в якій замість інертної маси фігурує коефіцієнт опору. Тому кажуть, що при наявності опору $F = -\zeta \vartheta = -2\gamma m \vartheta$ параметр ζ відіграє роль своєрідної антиінерції. Вимірність антиінерції: $\dim \zeta = \dim M$, а одиниця вимірювання $[\zeta] = [m] = 1 \text{ кг}$.

Наочна модель, що ілюструє характер коливань за наявності дисипативних втрат енергії осцилятором, зображена на рис.2.10.3,б. Тут використана відома у

фізиці модель потенціальної ями (суцільна товста лінія), роль якої відіграє потенціальна енергія деформації пружного зв'язку. Завдяки наявності опору рухові осцилятор із часом скочується на дно ями, завершуючи там при цьому коливання. Енергетичною характеристикою загасаючих коливань осцилятора є час релаксації повної енергії τ_w . Це такий проміжок часу, протягом якого повна енергія зменшується в e разів. Крім цього, можна розглянути потужність енергетичних втрат $\frac{dW}{dt}$.

Добротність коливної системи. Добротність – кількісна характеристика коливної системи. Як випливає із $W(\gamma, t) = W_0 \exp(-2\gamma t)$, швидкість втрати енергії дорівнює $\frac{dW(\gamma, t)}{dt} = -2\gamma W(\gamma, t)$. Протягом періоду втрата енергії дорівнює

$$T \left| \frac{dW(\gamma, t)}{dt} \right| = 2\pi \frac{2\gamma}{\omega} W(\gamma, t), \quad \text{звідки} \quad \text{відношення} \quad \frac{\omega}{2\gamma} = 2\pi \frac{W(\gamma, t)}{T \left| \frac{dW(\gamma, t)}{dt} \right|}$$

домноженого на 2π відношення запасу енергії до енергії, що втрачається осцилятором за період. Тому величина $\Xi = \frac{1}{\delta_\gamma} = T\gamma$ обернено пропорційна

логарифмічному декременту і називається добротністю осцилятора. Вона є мірою відносної дисипації енергії і характеризує здатність осцилятора зберігати енергію під час коливань протягом періоду, яку накопичив на початку періоду. Оскільки

$$\Xi = 2\pi \frac{W(\gamma, t)}{\Delta W(\gamma, T)} = \frac{2\pi}{1 - \exp(-2\gamma T)}, \quad \text{то при незначному затуханні} \quad \Xi \approx \frac{\pi}{\delta_\gamma} = \pi N_e.$$

Отже, добротність пропорційна тій кількості коливань, протягом якої вони, як такі, зникнуть. Добротність, ще визначає відношення, яке свідчить у скільки разів сила

$$\text{пружності більша від сили тертя: } \Xi = \frac{\mu y}{\zeta dy/dt} = \frac{\mu}{\zeta \omega_0^2} = \frac{m\omega_0}{\zeta}.$$

Осцилятор зміщується з положення рівноваги на амплітуду y_0 . Тому, порівнявши енергію осцилятора $\frac{1}{2} m \omega_0 y_0^2$ з втратою енергії за чверть періоду коливання одержуємо, що відповідне відношення дорівнює добротності коливної системи.

Динаміка вимушених коливань. У цьому випадку рівняння руху неоднорідне і його розв'язок представляється у вигляді суми

$y(\Omega, t) = y_1(t) + y_2(\Omega, t)$, де перший доданок описує вільні коливання і функція $y_1(t) = A_1 \cos \omega_0 t + A_2 \sin \omega_0 t$ є розв'язком однорідного рівняння $\frac{d^2 y_1(t)}{dt^2} + \omega_0^2 y_1(t) = 0$, а другий - описує вимушені коливання осцилятора. Функція $y_2(t)$ є частковим розв'язком неоднорідного рівняння $\frac{d^2 y_2(\Omega, t)}{dt^2} + \omega_0^2 y_2(\Omega, t) = \frac{F_0}{m} \exp(i\Omega t)$, яке записане для випадку дії гармонічної сили $F(\Omega, t) = F_0 e^{i\Omega t}$ з циклічною частотою Ω та амплітудою F_0 у середовищі без опору.

Прави їх частини записані у комплексному вигляді, що спрощує алгоритм розв'язування, оскільки дозволяє розв'язок шукати у комплексній формі $\tilde{y}_2(\Omega, t) = y_m(\Omega) \exp(i\Omega t + \phi)$. Тоді підставивши його в диференціальне рівняння одержимо, що зміна амплітуди вимушеного коливання осцилятора від частоти дії зовнішнього джерела Ω має резонансний характер: $y_m(\Omega) = \frac{F_0 / m}{\omega_0^2 - \Omega^2}$. Дисипативні втрати енергії в коливному контурі відсутні, тому уявна частина $\text{Im} \tilde{y}_m = 0$, а дійсна $\text{Re} \tilde{y}_m = \frac{F_0 / m}{\omega_0^2 - \Omega^2}$ змінюватиме свій знак. Приймаючи до уваги властивість тангенса фази $\text{tg}[\phi] = \frac{\text{Im} \tilde{y}_m}{\text{Re} \tilde{y}_m}$ приходимо до висновку, що коли частота $\Omega < \omega_0$, то $\text{Re} \tilde{y}_m > 0$ і вимушені коливання відбуваються у фазі з дією зовнішньої сили.

В ділянці частот $\Omega > \omega_0$, дійсна частина $\text{Re} \tilde{y}_m < 0$ і вимушені коливання відстають за фазою на π від дії зовнішньої сили, як це зображено на рис.2.2.4а. Стрибкоподібна зміна фази під час переходу через резонансну частоту зумовлена лише обмеженістю розглядуваної моделі, як консервативної. Якщо ж присутні сили опору коливному процесу, то фазова характеристика набирає плавного вигляду (пунктирна крива). Провести аналіз загального розв'язку рівняння доцільно для конкретних початкових умов. Нехай такими є: $y_0 \neq 0$, $\dot{y}_0 \neq 0$, $F_0 \neq 0$.

Тоді ввівши позначення $\Theta = \frac{F_0 / m}{\omega_0^2 - \Omega^2}$ для косинусної сили $F(\Omega, t) = F_0 \cos \Omega t$

вимушені коливання описуватимуться рівняннями:

$$y(\Omega, t) = y_0 \cos \omega_0 t + \frac{dy/dt|_{t=0}}{\omega_0} \sin \omega_0 t - \Theta \cdot \left(\cos \phi \cos \omega_0 t + \frac{\Omega}{\omega_0} \sin \phi \sin \omega_0 t \right) + \Theta \cos(\Omega t - \phi).$$

Тут перші два доданки визначають вільні коливання осцилятора без врахування дії зовнішнього джерела, третій і четвертий – вимушені. Причому третій доданок описує відоме у фізиці явище інтерференції вільних і вимушених коливань, а четвертий – ті, які встановлюються в усталеному режимі. Отже, спочатку актуальні власні коливання, які з часом загасають і в системі встановлюються усталені коливання – це так званий режим стаціонарних вимушених коливань. На відміну від власних коливань, система під час вимушених не пам'ятає початкових умов. Ця властивість дозволяє суттєво спростити алгоритм аналізу складних задач акустики.

Сформулюємо основні закономірності вимушених коливань. **1.** Якщо дія зовнішнього джерела має статичний характер $\Omega = 0$, то величина зміщення осцилятора з положення рівноваги визначає статичну амплітуду:

$$y_m(\Omega = 0, \gamma) = \frac{F_0}{m\omega_0^2}.$$
 При статичній дії зовнішнього джерела, фаза вимушених

коливань збігається з фазою сили. **2.** При дуже великій частоті Ω дії коливальна система внаслідок її інерції не може зміщуватись із положення рівноваги, тому якщо загасання відсутнє, то амплітуда зміщення прямує до нуля $y_m(\Omega \gg \Omega_{рез}) \rightarrow 0$, а фаза коливань протилежна до фази зовнішньої сили. Однак, амплітуда резонансних коливань стає скінченою, якщо присутнє загасання в коливній системі. **3.** Під час резонансу сила випереджає за фазою коливання на $\frac{\pi}{2}$, якщо присутнє загасання. Тому вона весь час збільшує розмах коливань. Це відбуватиметься доти, доки вся робота сили не буде витрачена на переборення сил тертя і встановиться амплітуда вимушених коливань.

Резонанс коливань. Резонанс – явище зростання амплітуди вимушених коливань системи, коли частота дії зовнішньої сили наближається до частоти власних коливань. Такий стан – це один із проявів нестійкості коливної системи, що зумовлена ідеалізацією коливної моделі і зникає, якщо врахувати незворотні втрати енергії на пере додання опору. Це досягається тому, що із-за опору частота

резонансних коливань змінюється, тому зникає необмежене зростання амплітуди $y_m(\Omega, \gamma)$. Резонансне зростання амплітуди вимушених коливань описується

функцією $y_m(\Omega, \gamma) = \frac{F_0}{m} \left[1 + \left(\frac{2\gamma\Omega}{\Omega^2 - \omega_0^2} \right)^2 \right]^{-1/2}$. Однак, резонансний контур має скінчену

висоту, якщо коефіцієнт загасання $\gamma \neq 0$ (рис.2.10.4). Чим більше загасання ($\gamma = 0.25$ - контур з найвищою амплітудою; $\gamma = 0.5$ - середньою і $\gamma = 0.75$ - найменшою), тим менша резонансна частота (контур 1) і менша амплітуда резонансу.

Розв'язавши рівняння $\frac{dy_m(\Omega, \gamma)}{d\Omega} = 0$ знаходимо, що резонанс спостерігається на так званій резонансній частоті*)

$$\Omega_{рез} = \sqrt{\omega_0^2 - 2\gamma^2} = \omega_0 \sqrt{1 - \left(\frac{\sqrt{2}\gamma}{\omega_0} \right)^2}.$$

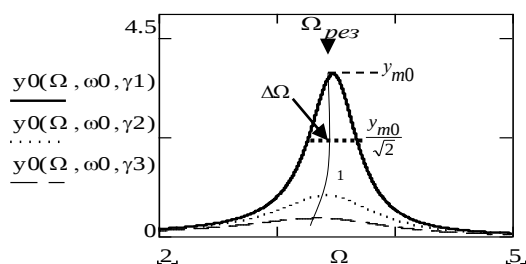


Рис.2.10.4

Ширина резонансного контуру і добротність. Важливою характеристикою резонансних коливань є відносна ширина контуру $\Delta\Omega$ на піввисоті від максимальної енергії коливної системи в стані резонансу, яка пропорційна квадрату амплітуди y_{m0}^2 . Тому, якщо вздовж вертикалі відкласти амплітуду вимушених коливань, то ширина $\Delta\Omega$ відповідає висоті резонансного контуру $\frac{y_{m0}}{\sqrt{2}}$. Відносна ширина його $\frac{\Delta\Omega}{\Omega_{рез}}$ дорівнює $\frac{2\gamma}{\Omega_{рез}}$. Добротність резонансного контуру дорівнює відношенню $\frac{\Omega_{рез}}{2\gamma} = \Xi$, або відношенню амплітуди коливань при резонансній частоті до величини статичного зміщення визначає добротність

*) Поняття про резонанс ввів Галілей, продемонструвавши, що амплітуду коливань вантажу на нитці можна поступово збільшувати, якщо дути на маятник в такт його руху

коливної системи: $Q_\omega = \frac{y_m|_{\Omega_{рез}}}{y_m|_{\Omega=0}}$. Тому чим більша добротність, тим вузкий і вищий резонансний контур.

Реактанси коливної системи. Механічний імпеданс та рухливість одиночного елемента. Поведінка коливної системи під дією зовнішньої періодичної сили залежить як від її амплітуди, так і частоти. Якщо прийняти до уваги співвідношення $(\Omega^2 - \omega_0^2) = \frac{\Omega}{m} \left(\Omega m - \frac{\mu}{\Omega} \right)$, то амплітуди вимушених коливань можна записати як $\tilde{y}_m(\Omega, \gamma) = \frac{F_0}{\Omega} \left[r^2 + \left(\Omega m - \frac{\mu}{\Omega} \right)^2 \right]^{-1/2}$, $r^2 = 4\gamma^2 m^2$. Величини Ωm і $\frac{\mu}{\Omega}$ називаються реактансами. Вони вносять не лише ефективний опір в коливну систему, але й зумовлюють зсув фази між зміщенням та силою. Тому, таку коливну систему характеризують, в загальному комплексним, механічним імпедансом: $\tilde{Z}_m = r + i \left(\Omega m - \frac{\mu}{\Omega} \right) = \text{Re } \tilde{Z}_m + i \text{Im } \tilde{Z}_m$. Обернена величина до механічного імпедансу називається рухливістю механічного елемента. Для систем з зосередженими параметрами механічні імпеданси і рухливість подані в таб 2.10.1.

Табл.2.10.1

Елемент	Сила	Імпеданс	Рухливість
Маса	$F = ma = i\omega m \mathcal{G}$	$i\omega m$	$i / \omega m$
Пружність	$F = -\mu x = -i \frac{\mu}{\omega} \mathcal{G}$	$-i \frac{\mu}{\omega}$	$-i \frac{\omega}{\mu}$
Тертя	$F = -r \mathcal{G}$	r	$1/r$

Вираз $\text{Re } \tilde{Z}_m$ відіграє роль так званого активного опору коливанию осцилятора, а другий – реактивного. Повний опір рухові буде дорівнювати

$Z_m = \sqrt{r^2 + \left(\Omega m - \frac{\mu}{\Omega} \right)^2}$ і прийнявши до уваги запис $\tilde{Z}_m = Z_m \exp(i\phi)$, одержимо, що

тангенс зсуву фази виражається через реактанси як $\tan \phi = \left(\Omega m - \frac{\mu}{\Omega} \right) / r$. Отже, дійсна

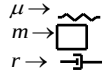
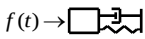
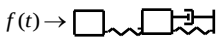
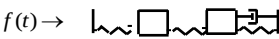
частина комплексного механічного імпедансу пов'язана з дисипацією енергії в коливній системі і з витратами енергії на збудження коливань, а уявна – з

реакцією сил пружності та інерції. В області низьких частот зовнішньої дії $\omega_0 \gg \Omega$ імпеданс має пружний характер, тоді як на високих – інерційний.

2.11. Мікроакселерометри, мікрогіроскопи і мікроприводи механічного руху

Нагадаємо, що електромеханічні системи поділяють на два типи: сенсори (акселерометри, гіроскопи) – вимірювальні пристрої, які перетворюють механічний рух в електричний сигнал, і актюатори (виконавчі пристрої) – системи, які вирішують обернену задачу, тобто перетворюють електричний сигнал в механічну дію. В основу їх дії покладені основні закони динаміки поступального і обертового рухів, а базовою математичною моделлю моделювання процесів є пружний осцилятор. В табл.2.11.1 приведені ряд

Таблиця 2.11.1

			
Додаткові моделі	$f(t) \rightarrow \begin{matrix} m + \mu \\ m + r \\ m + 0 \end{matrix}$	$f(t) \rightarrow \begin{matrix} m + \mu + m + \mu \\ m + \mu + m + r \\ m + \mu + m + 0 \end{matrix}$	$f(t) \rightarrow \begin{matrix} \mu + m + \mu + m + \mu \\ \mu + m + \mu + m + r \\ \mu + m + \mu + m \end{matrix}$
$x(t),$ $F(t),$ $g(t).$	$m \frac{d^2 x}{dt^2} + r g + \mu x = f$	$\begin{matrix} m_1 \frac{d^2 x_1}{dt^2} + \mu_2 x_1 - \mu_2 x_2 = f \\ m_2 \frac{d^2 x_2}{dt^2} + \kappa g_2 - \mu_2 x_1 + (\mu_2 + \mu_3) x_2 = 0 \end{matrix}$	$\begin{matrix} m_1 \frac{d^2 x_1}{dt^2} + (\mu_1 + \mu_2) x_1 - \mu_2 x_2 = f \\ m_2 \frac{d^2 x_2}{dt^2} + \kappa g_2 - \mu_2 x_1 + (\mu_2 + \mu_3) x_2 = 0 \end{matrix}$
$x(s),$ $F(s),$ $g(s).$	$\frac{x(s)}{F(s)} = \frac{1}{ms^2 + rs + \mu}$	$\frac{x_1(s)}{F(s)} = \frac{ms^2 + rs + \mu_2 + \mu_3}{D(s)}, \frac{x_2(s)}{F(s)} = \frac{\mu_2}{D(s)}$ $D(s) = (m_1 s^2 + \mu_2)(m_2 s^2 + rs + \mu_2 + \mu_3) - \mu_2^2$	$\frac{x_1(s)}{F(s)} = \frac{ms^2 + rs + \mu_2 + \mu_3}{D(s)}, \frac{x_2(s)}{F(s)} = \frac{\mu_2}{D(s)}$ $D(s) = (m_1 s^2 + \mu_1 + \mu_2)(m_2 s^2 + rs + \mu_2 + \mu_3) - \mu_2^2$

поширених в практиці конструювання МЕМСмоделей сенсорів руху, представляючи собою суперпозицію базових осциляторів. В другому рядку таблиці знаки суми використані в розумінні додавання дій фізичних чинників. В цій же таблиці подані відповідні часові диференційні рівняння руху робочого інерційного тіла в t -області ($x(t), g(t), f(t)$), та їх Лапласове ($X(s), V(s), F(s)$) [2.35;2.36].

Акселерометри.

Акселерометри – це прилади для реєстрації та перетворення інформації про прискорення з метою одержання кількісного результату у формі, яка зручна для

подальшого використання. Акселерометри призначені для вимірювання прискорення інерційного тіла (кутового чи лінійного), що рухається. За їх допомогою вимірюють кути нахилу в автомобільних сигналізаціях, вони вимірюють кут нахилу ліжок пацієнтів, захищають диски комп'ютерів типу Notebook, слугують системами захисту від вібрацій, використовуються в чутливих елементах систем безпеки, в системах контролю роботи ліфта, вібростендів, в космічних та авіаційних технологіях, тощо.

Мікромеханічні ємнісні акселерометри. Класичним представником мікромеханічного акселерометра є кремнієвий серії ADXL фірми Analog Devices. Вихідний електричний сигнал акселерометра через підсилювач і комутатор подається на АЦП, а потім в цифровому вигляді на мікроконтроллер, який керує всією схемою, в тому числі формуванням сигналу для тестування акселерометра.

Конструктивно він складається з тягарця масою m , закріпленого до пружини жорсткістю μ , розташованого в середовищі з опором, пропорційним швидкості руху. Коливне тіло закріплене до рухомої пластини конденсатора. Відносно нерухомої системи відліку (Землі), маятник здійснює коливання амплітудою $y(t)$, відносно верхньої пластини конденсатора - $u(t)$ так, що відстань між пластинами дорівнює $x(t) = y(t) - u(t)$. Якщо коливання маятника гармонічні, то

описуються ДР $m \frac{d^2 y(t)}{dt^2} + r \frac{du(t)}{dt} + \mu u(t) = 0$. Приймаючи до уваги перетворення

$$x(t) = y(t) - u(t) \quad \text{одержимо рівняння} \quad m \frac{d^2 u(t)}{dt^2} + r \frac{du(t)}{dt} + \mu u(t) = -m \frac{d^2 x(t)}{dt^2} \quad \text{або}$$

$$\frac{d^2 u(t)}{dt^2} + 2\gamma \frac{du(t)}{dt} + \omega_0^2 u(t) = -\frac{d^2 x(t)}{dt^2}, \quad \text{де} \quad \omega_0^2 = \frac{\mu}{m} \quad \text{і} \quad 2\gamma = \frac{r}{m}. \quad \text{Застосуємо до рівняння}$$

перетворення Лапласа, після чого одержимо $s^2 U(s) + 2\gamma s U(s) + \omega_0^2 U(s) = -s^2 X(s)$ або

$$(s^2 + 2\gamma s + \omega_0^2) U(s) = -s^2 X(s), \quad \text{якщо} \quad u(0) = 0 \quad \text{і} \quad \left. \frac{du}{dt} \right|_{t=0} = 0. \quad \text{Права частина} \quad s^2 X(s) \quad \text{за суттю}$$

прискорення є вхідним сигналом, $U(s)$ -вихідний (реакція акселерометра), тому перетворення Лапласа автори [2.38;2.39] акселерометра визначають як

відношення $G(s) = \frac{U(s)}{s^2 X(s)} = -\frac{1}{s^2 + 2\gamma s + \omega_0^2}$. Передаточна функція $G(\omega)$ дорівнює

$$|G_L(\omega)| = \frac{1}{\sqrt{(\omega_0^2 - \omega^2)^2 + (2\gamma\omega_0)^2}}.$$

Одержаний вираз описе амплітудно-частотну характеристику так званого низькочастотного (важкого) акселерометра, робочий діапазон якого зосереджений

області низьких частот $\omega \ll \omega_0$ (заштрихована ділянка), де передаточна функція

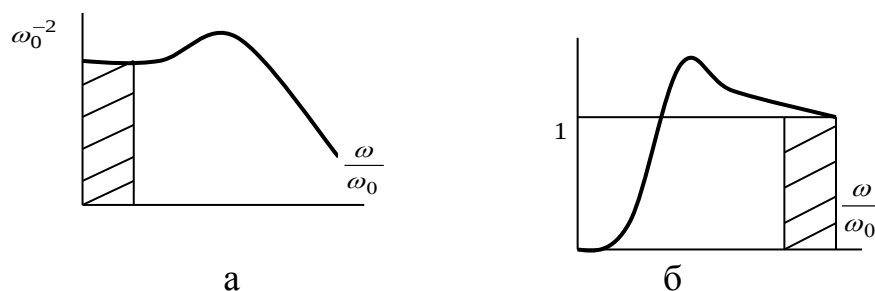


Рис.2.11.1 [58]

прямує до сталого значення $|G_L(\omega)| \rightarrow \omega_0^{-2} = \text{const}$ (рис.2.11.1а). В області високих

частот $\omega \gg \omega_0$ $|G_h(s)| = \frac{U(s)}{X(s)} = -\frac{s^2}{\omega_0^2} \left[1 + \frac{s^2}{\omega_0^2} + \frac{2\gamma\omega}{\omega_0^2} s \right]^{-1}$, звідки

$G_h(\omega) = \frac{\omega^2}{\omega_0^2} \left[\left(1 - \frac{\omega^2}{\omega_0^2} \right)^2 + \left(\frac{2\gamma\omega}{\omega_0^2} \right)^2 \right]^{-1/2} \rightarrow 1$ (рис.2.11.1,б). Фактично, такий пристрій є

віброметром, вхідний сигнал, якого є вихідним, оскільки маса робочого тіла порівняно мала і діапазон робочих параметрів зосереджений в області високих частот. В таких системах затухання незначне, а маса інертного тіла відносно велика.

Мікромеханічні акселерометри серії ADXL використовуються для вимірювання дійсної швидкості автомобілів. Він розташовується якомога ближче до його центру з віссю чутливості, орієнтованою у напрямку поздовжньої вісі автомобіля. Давач прискорення електрично пов'язується з інтегруючим обчислювальним блоком та блоком корекції.

Пристрій для вимірювання швидкості працює так. За допомогою давача вимірюється прискорення автомобіля, сигнал про яке подається на інтегратор,

який шляхом інтегрування за часом вираховує швидкість $\dot{x}(t) = C \int_{t-\Delta t}^t a_x(t) dt + \dot{x}_K$, де C - стала інтегрування, $\dot{x}_K$ - вихідний сигнал з давача швидкості в момент корекції протягом часу Δt .

Вирахувана швидкість автомобіля періодично кригується за інформацією давача швидкості за допомогою блоку корекції. Момент здійснення корекції обирається блоком керування корекцією, який видає команду на здійснення корекції у кожному випадку, коли за період часу Δt зміна вихідного сигналу з давача швидкості автомобіля не перевищувала $\Delta \dot{x}$, а сигнал з давача прискорень автомобіля не перевищувала величину Δa . Крім того, блок керування корекцією змінює сталу інтегрування в залежності від режиму руху, який визначається за сигналами, що надходять від давачів прискорення.

На рис.2.11.2,а, показаний мікромеханічний акселерометр маятникового типу. З різними температурними властивостями шарів. Чутливим елементом в ньому є шарувата пластина, що підкріплена до основи пружними підвісками.

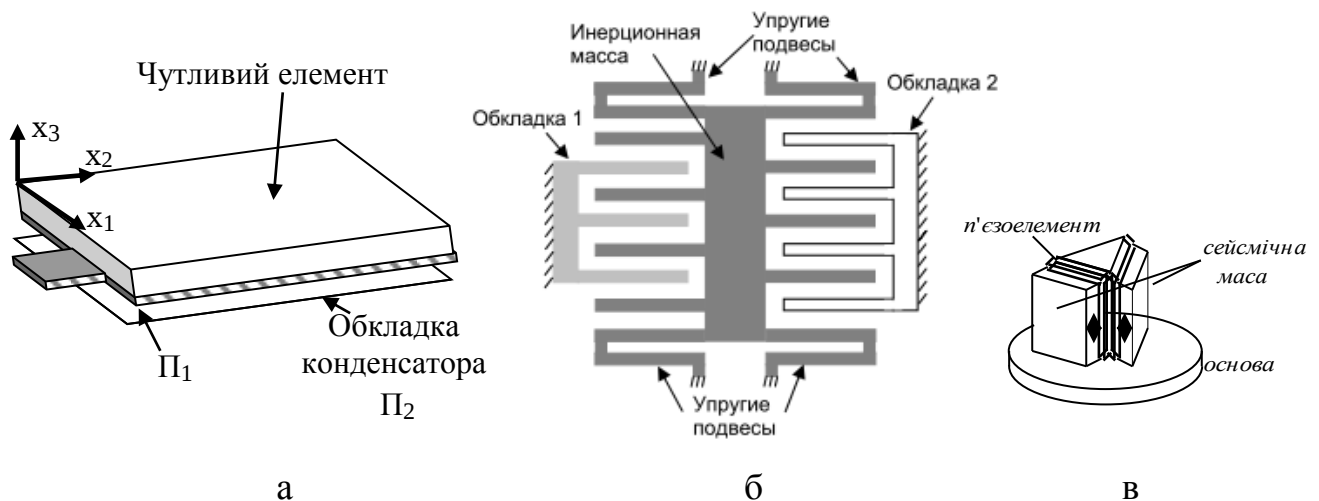


Рис.2.11.2 [2.38]

процесі вимірювання проекції вектора прискорення на вісь x_3 визначають положення інерційної маси (чутливого елемента) за величиною ємності конденсатора, утвореного струмопровідними поверхнями на чутливих елементах (поверхнями Π_1 і Π_2).

В МЕМС переважно використовують багатогребінчаті акселерометри (рис.2.11.2,б). Комутація рухомих пластин, відносно нерухомих утворюють конденсатори $C_{1,2} = \frac{\varepsilon_0 S}{x_{1,2}}$, де $x_1 \neq x_2$, якщо прискорення не дорівнює нулеві і в

протилежному разі $x_1 = x_2$ ($x_{1,2}$ - відстані рухомої пластини відносно нерухомих, відповідно верхньої та нижньої). Значення ємностей пов'язані з прискоренням руху через зміщення як $C_2 - C_1 = 2\Delta C = 2\varepsilon_0 \frac{x}{d^2 - x^2}$, звідки розв'язок рівняння $\Delta Cx^2 + \varepsilon_0 x - \Delta Cd^2 = 0$, дає можливість знайти зміщення x .

В першому наближенні $x \approx \frac{d^2}{\varepsilon_0} \Delta C = d \frac{\Delta C}{C_0}$, де $C_0 = \frac{\varepsilon_0 S}{d}$ і з рівняння електричного балансу $(U_x + U_0)C_1 + (U_x - U_0)C_2 = 0$ динамічна складова напруги вихідного електричного сигналу дорівнює $U_x = \frac{C_2 - C_1}{C_2 + C_1} = \frac{x}{d} U_0$, де координата x і прискорення $\frac{d^2 x}{dt^2}$ пов'язані між собою ньютонівським рівнянням руху $m \frac{d^2 x}{dt^2} = \mu x$. Тут U_0 - статична напруга між обкладинками 1 та інерційним тілом.

П'єзоелектричні акселерометри. П'єзоелектричні акселерометри є одними з найкращих датчиків для вимірювання параметрів механічних коливань та рухів [2.40]. Такі переваги зумовлені наступними властивостями акселерометрів: -широкий частотний діапазон; - лінійна характеристика в широкому динамічному діапазоні; - вихідний електричний сигнал пропорційний прискоренню механічних коливань; - забезпечення стійкості експлуатаційних характеристик в несприятливих умовах; - активний характер перетворювача, відсутня потреба в джерелі напруги; - відсутність механічно-рухомих деталей, що значно продовжує термін експлуатації; - можливість мініатюризації.

За класифікацією [2.41], розроблені датчики на прямому п'єзоелектричному ефекті в моно- і полікристалічних матеріалах, з моно, бі- і триморфними п'єзоелементами, резонансного типу, на об'ємних і поверхневих акустичних хвилях. Кращі характеристики мають монолітні датчики і вигляді циліндра з дисковими електродами на кінцях. Датчик має нульову поперечну чутливість. Однак, вимога кріпити його в області електрода, ускладнює конструкцію, що негативно впливає на діапазон динамічних характеристик. Щоб уникнути цього недоліку, внутрішня частина циліндра видаляється. Тоді зовнішня частина відіграє роль одночасно п'єзоелемента і інерційної маси, що зменшує вимоги до кріплення циліндра до основи.

Традиційна механічна модель п'єзоелектричного акселерометра зображена на рис.2.11.2,в. До станини у вигляді трикутної призми послідовно закріплюються п'єзоелемент та інерційна (сейсмічна) маса у вигляді пластин. Діючі на основу акселерометра механічні коливання зумовлюють на п'єзоелементи силову дію сейсмічної маси на прискорення.

Пієзоелементи генерують електричний заряд, пропорційний діючий на них силі. Оскільки амплітуда і фаза прискорення сейсмічних мас в широкому діапазоні частот ідентичні амплітуді і фазі прискорення основи акселерометра, електричний заряд пропорційний механічному прискоренню коливань поверхні зразка, на який закріплений акселерометр. Найвідоміші відомі виробники акселерометрів є фірми “Kistler”, “Instrumente AG”, “Analog Devices”, таганрозький завод “Виброприбор”, “Brüel & Kjaer”, “Genisco”, “SFIM”, “Endevco”, “BAE System” та ін.

Порівняльний аналіз акселерометрів. Оскільки акселерометри відносяться до найважливіших елементів систем інерційної навігації, то до їх розробки та виготовлення висуваються високі вимоги. Вони повинні мати високу точність та чутливість, необхідний діапазон вимірюваних прискорень, їх лінійність, стабільність показів, нечутливість до вібраційних і ударних дій, до зміни температури та тиску навколишнього середовища. Аналіз робочих характеристик різних типів свідчить про таке.

Оскільки в диференційних струнних акселерометрах інерційний елемент закріплений між двома розтягнутими струнами, які слугують одночасно пружними елементами, то шляхом зворотного зв'язку, досягається безперервність коливань. Електричний сигнал пропорційний прискоренню руху тіла, генерується і піддається подальшій обробці електронним перетворювачем. Лінійність робочої характеристики досягається тим, що рівень пружної деформації струни контролюється і втримується на сталому рівні, що дозволяє частотний сигнал перетворювати з точністю до 0,001–0,01 % та в подальшому піддавати точному вимірюванню і інтегруванню.

Струнні акселерометри мають високу точність вимірювання, високу вібраційну та ударну міцність, надійність, частотно-модульований вихідний

сигнал, високу потужність вихідного сигналу, малі габарити та вагу. Недоліками їх є висока вартість, обумовлена технологічними труднощами їх виготовлення. Крім цього, їх робочі параметри чутливі до коливань температури, вологості. Діапазон вимірюваних прискорень розпочинається із значень від 0.001g в діапазоні частот 3–20 000 Гц, а точність вимірювання складає 3–5 %.

П'єзоелектричні акселерометри мають меншу точність вимірювань, однак вони простіші та надійніші, не вимагають використання зовнішніх джерел живлення, тому дешевші. Однак, вони також позбавлені недоліків: треба підсилювати вихідний сигнал; чутливі до впливу температури; мають нелінійну характеристику, яка підлягає деградації, зміщується нуль, тощо. Вони вимірюють прискорення від діапазону від 0,01 до 1000 g в діапазоні частот до 1 500 Гц з точністю ± 1 %. У зв'язку з тим, що п'єзорезистивні акселерометри мають малий діапазон вимірювань, малу чутливість, однак витримують високі перевантаження, найбільшого поширення вони здобули у віброметрії та ударних вимірюваннях.

Ємнісні акселерометри набули широкого використання завдяки їх високій чутливості, стабільності при детектуванні статичного прискорення, малому дрейфу нуля, низькій температурній чутливості, малій споживаній потужності, високій надійності, хорошим шумовим характеристикам, задовільній точності, низькій ціні. Завдяки простоті конструкції та малим габаритам, ємнісні акселерометри набули широкого застосування в мікротехнологіях.

Гіроскопи

Це датчики кутової швидкості. В основу їх роботи можуть покладатися різні фізичні процеси. Так, гіроскопи механічного типу базуються на використанні сили Коріоліса, тоді як оптичні – ефекту Сеньяка. Запропоновані гіроскопи на квантових фізичних принципах.

Механічний гіроскоп. Гіроскопи – це давачі інерційної інформації про механічні рухи. З появою MEMS технологій їх почали застосовувати в автомобільній промисловості, електроніці, роботах і інтелектуальних системах. Довгий час під гіроскопом розумілося симетричне тверде тіло, що швидко обертається. Тепер до гіроскопів відносять широкий клас приладів, в тому числі

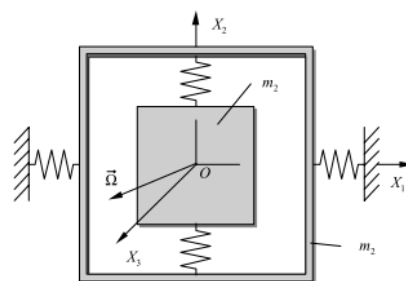
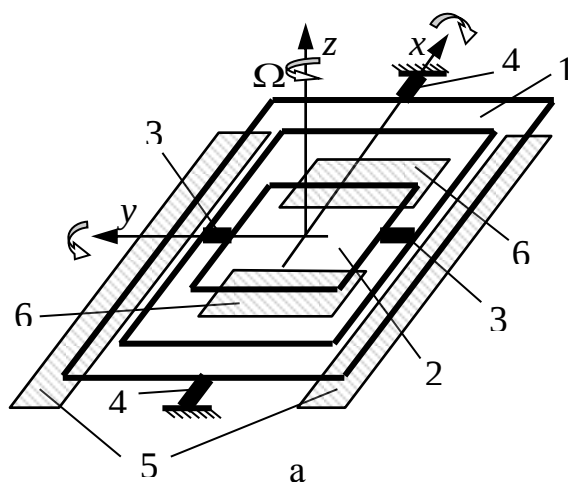
тих, які містять в собі матеріальний об'єкт, що здійснює швидкі періодичні рухи, в результаті яких стає чутливим до обертання в інерційному просторі. Гіроскопічні прилади можна поділити на вимірювальні і силові. Силові прилади служать для створення моментів сил, прикладених до основи на яку встановлено гіроприлад, а вимірювальні призначені для визначення параметрів руху основи, як кут повороту основи, проекції вектора кутової швидкості і інші. До силових гіроприладів відносять механічні гіроскопи, всі решта – вимірювальні.

Механічний гіроскоп використовує інерційні властивості ротора (рис.1.2.3а), що швидко обертається. Ним є симетричне тверде тіло, яке швидко обертається навколо своєї вісі симетрії, яка також може обертатися в просторі. Ротор підтримується двома круглими кільцями, які називаються “карданним підвісом”. Одне кільце змонтоване всередині іншого так, що пара кілець може обертатися в межах зовнішньої рамки. Вісі обертання ротора, внутрішнього і зовнішнього кілець є взаємно перпендикулярні. Такий гіроскоп спроможний повертатися навколо трьох взаємно перпендикулярних осей, що пересікаються в одній точці О. Тому він має три ступені вільності.

В основі теорії механічного гіроскопу покладено рівняння моментів: $\frac{d\vec{L}}{dt} = \vec{M}$, де $\vec{M}$ – сумарний момент сил, що діють на систему, момент імпульсу якої $\vec{L}$. Важливо, що $\vec{M}$ і $\vec{L} = \int \vec{M} dt$ повинні бути обчислені відносно однієї і тієї ж (довільної) точки простору. Для симетричного тіла, що обертається навколо власної осі, його власний момент імпульсу (момент кількості руху): $\vec{M} = J \vec{\omega}$, де J – момент інерції відносно осі обертання з кутовою швидкістю $\vec{\omega}$. Гіроскоп, що швидко обертається, має великий момент імпульсу, який спрямований вздовж вісі обертання. Якщо на гіроскоп не діють зовнішні сили або їх сумарний момент рівний нулю і центр тяжіння всієї системи спіпадає з нерухомою точкою гіроскопа, то такий гіроскоп називається вільним. Згідно з рівнянням моментів, при $\vec{M} = const$, $\vec{L} = const$ (закон збереження моменту імпульсу, або моменту кількості руху), тобто момент імпульсу такого гіроскопу залишається незмінним, як за величиною, так і за напрямком, тобто вільний гіроскоп зберігає в просторі

незмінним напрямком вісі обертання. Суть цього полягає в тому, що якщо зовнішня рамка може бути повернута на кут в будь-якому напрямку, однак положення власної вісі обертання (спіна гіроскопа) буде залишатися в просторі незмінним.

Майже всі механічні гіроскопи належать до класу вібраційних. Суть цього полягає в тому, що в кожному з них інерційне тіло здатне здійснювати зворотно-поступальний рух в одній площині. Якщо поставити такий робочий орган на платформу, що обертається, площина якої співпадає з площиною коливань, то на віброуючу масу почне діяти сила Коріоліса $\vec{F}_c = 2m[\vec{\Omega} \times \vec{g}]$, де $\vec{\Omega}$ – вектор кутової швидкості (перпендикулярний площині обертання), $\vec{g}$ – вектор лінійної швидкості робочого органа відносно платформи, що обертається, m – маса робочого органа. Сила Коріоліса спрямована до площини коливань і вісі обертання. При протилежно напрямлених рухах, сила Коріоліса також діє в протилежних напрямках. На цьому і побудований принцип дії вібраційного гіроскопа.



б [2.42]

Рис.2.11.3

На точність вимірювання механічних гіроскопів впливають реальні відхилення осі обертання гіроскопа внаслідок механічних похибок при виготовленні (тертя в підвісах, не збалансований ротор, тощо), а також видимі відхилення викликані обертанням Землі. Ці відхилення треба враховувати при використанні гіроскопів в інерційних платформах і системах навігації. Видимі (аперентні) відхилення мають постійну інтенсивність і залежать від розташування гіроскопа на поверхні Землі (широти місця). Похибка гіроскопа вимірюється швидкістю відхилення його осі від початкового положення (її називають

дрейфом). Основними недоліками цих гіроскопів є високе співвідношення ціна/точність, відносно великі розміри і маса, що обмежує широку область їх застосування, відносно велике енергоспоживання, порівняно високий час підготовки до роботи, малий ресурс роботи, чутливість до механічних ударів і вібрацій, втрата точності з часом.

Мікромеханічні гіроскопи (рис.2.11.3,а). Чутливий елемент мікромеханічного гіроскопа складається із зовнішньої рамки 1 і внутрішньої пластини 2, яка фіксується за допомогою двох торсіонів 3. Зовнішня рамка 1 з'єднується з основою двома торсіонами 4.

Система координат XYZ зв'язана з гіроскопом. Вісь Z завжди перпендикулярна до пластини 2. Між обкладками 5 і на рамці 1 створюється змінна в часі різниця потенціалів. Зовнішня рамка вимушено коливається, генеруючи змінний в часі електричний сигнал. Якщо коливна рамка обертається з кутовою швидкістю Ω , на чутливий елемент починають діяти сили Кориоліса, які його повертають, змінюючи при цьому ємність конденсатора, утвореного пластинами 6 і 2.

Вібраційні гіроскопи. Вібраційні гіроскопи (ВГ) поділяються за характером руху чутливого елемента і принципу дії, в яких виникаючий момент (сила) зумовлює відхилення інерційного тіла відносно основи, що здійснює переносний рух. Характерною ознакою ВГ є зв'язок чутливого елемента з основою, на яку він встановлений.

ВГ поділяють на осциляторні і роторні. В основу цього поділу покладений характер руху відносно чутливого елемента. В осциляторних інерційна маса здійснює коливальні рухи, тоді як в роторних – обертові. Осциляторні ВГ поділяються в свою чергу на зосередженими і розподіленими параметрам. В роторних гіроскопах такий поділ здійснюється за характером зв'язку чутливого елемента з основою: пружний, псевдо пружний і комбінований. Прояв того чи іншого типу зв'язку в роторних ВГ класифікується ступінню функціональної залежності між позиційного члена в рівняннях руху з кутовою швидкістю власного обертання гіроскопа. Якщо вона відсутня, то зв'язок пружний, якщо існує – псевдо пружна, а коли проявляються оба види – комбінований. В

осциляторних гіроскопах із зосередженими параметрами, динаміка руху інерційного елемента описується звичними диференціальними рівняннями, тоді як із розподіленими- відповідними в часткових похідних. Відзначимо також, що роторні ВГ ще називаються динамічно налаштовуваними гіроскопами, в яких виникаючий пружний момент компенсується шляхом відповідного підбору основних його параметрів [2.43].

Кінематична модель одновісного гіроскопа була запропонована авторами [111] і зображена на рис.2.11.3,б. Розглянемо спрощену математичну модель, яка характеризує роботу камер тонного ВГ. Нехай система координат $X_1X_2X_3$ жорстко зв'язана з гіроскопом і обертається відносно іншої інерційної системи відліку з абсолютною кутовою швидкістю Ω та рухається з абсолютним лінійним прискоренням a . Якщо початки координат систем відліку співпадають з центром маси чутливого елемента (точкою його рівноваги), то положення системи $X_1X_2X_3$ відносно інерційної визначається кутом $\varphi(t)$ і при цьому $\frac{d\varphi}{dt} = \Omega_3(t)$. Якщо жорсткості обох пружин однакові та відсутня анізотропія коефіцієнтів тертя, то за вибраних умов $\Omega_1 = \Omega_2 = a_1 = a_2 = 0$ і рівняння руху запишуться у вигляді:

$$\begin{cases} \frac{d^2x_1}{dt^2} + \gamma \frac{dx_1}{dt} + (\omega_0^2 - \Omega_3^2)x_1 - \Omega_3x_2 - 2\Omega_3 \frac{dx_2}{dt} = f_{m1}; \\ \frac{d^2x_2}{dt^2} + \gamma \frac{dx_2}{dt} + (\omega_0^2 - \Omega_3^2)x_2 + \Omega_3x_1 + 2\Omega_3 \frac{dx_1}{dt} = f_{m2}. \end{cases} \quad \text{Конструкція вібраційного гіроскопа}$$

дозволяють інформацію про значення параметрів $x_1, x_2, \frac{dx_1}{dt}, \frac{dx_2}{dt}$, на основі якої вібраційний гіроскоп вимірює кутову швидкість Ω_3 , або кут повороту

$$\varphi(t) = \int_{t_0}^t \Omega_3(\tau) d\tau.$$

1. Режим вимірювання сталої кутової швидкості. В цьому випадку $\Omega_2 = const$ і якщо періодична сила діє вздовж вісі X_2 : $f_{m1} = 0, f_{m2} = f_{m0} \sin(\omega t + \varphi_0)$. Тоді із-за лінійності диференціальних рівнянь, розв'язком є сума відповідного для однорідного і неоднорідного рівнянь. Часткові розв'язки шукаємо у вигляді $x_1(t) = x_{m1} \sin(\omega t + \varphi_1)$ і $x_2(t) = x_{m2} \sin(\omega t + \varphi_2)$. Якщо φ_0 таке, що $\varphi_2 = 0$, то

$$x_{m1} = \frac{2\Omega_3 x_{m2} \omega}{\sqrt{(\omega^2 + \Omega_3^2 - \omega_0^2)^2 + \gamma^2 \omega^2}}, \quad \operatorname{tg} \varphi_1 = \frac{-\omega^2 - \Omega_3^2 + \omega_0^2}{\gamma \omega} \quad [2.44].$$

На частоті $\omega = \omega_0$, якщо $\Omega_3^2 \ll \gamma \omega_0$, то відношення амплітуд коливань вздовж вісей x_1 і x_2 пропорційне кутової швидкості обертання основи вібраційного гіроскопа: $\frac{x_1}{x_{m2}} \approx \frac{2\Omega_3}{\gamma}, \varphi_2 \approx 0$.

2. Режим вимірювання кута повороту. В цьому випадку також вважається, що $\Omega_2 = \text{const}$. Якщо тертям знехтувати $\gamma \approx 0$ і зовнішня сила відсутня $f(t) = 0$, то гіроскоп працює в режимі власних коливань. Якщо $\Omega_3^2 \ll \gamma \omega_0$, то

$$\text{диференційні рівняння руху набудуть вигляду} \quad \begin{cases} \frac{d^2 x_1}{dt^2} + \omega_0^2 x_1 - 2\Omega_3 \frac{dx_2}{dt} = 0; \\ \frac{d^2 x_2}{dt^2} + \omega_0^2 x_2 + 2\Omega_3 \frac{dx_1}{dt} = 0. \end{cases} \quad \text{За цих}$$

умов, кут повороту основи гіроскопа відносно початкового положення дорівнює

$$\operatorname{tg} 2\varphi = \frac{2\left(\omega_0^2 x_1 x_2 + \frac{dx_1}{dt} \frac{dx_2}{dt}\right)}{\omega_0^2 (x_1^2 - x_2^2) + \left[\left(\frac{dx_1}{dt}\right)^2 + \left(\frac{dx_2}{dt}\right)^2\right]}.$$

Мікромеханічні гіроскопи камертонного типу. Можливість використання камертона (*tuning fork*) для одержання інформації про інерцію руху тіла відома давно. Гіроскопи з чутливим елементом у вигляді камертона вперше розроблені компаніями Charles Stark Draper Laboratory, Fujitsu, Honeywell, Samsung Electro Mechanics, тощо.

Розглянемо принцип дії камертонного гіроскопа. Нехай його штири коливаються у протифазі в площині рис.2.11.4,а (робочі коливання). Якщо коливний камертон обертати навколо вертикальної вісі, на штирі камертона діють сили Коріоліса, які призводить до виникнення додаткових коливань у перпендикулярному напрямку до площини рисунка. Отже, вільні кінці камертона приймають участь у двох взаємноперпендикулярних коливаннях і результуюча траєкторія руху ніжок в просторі визначатиметься зсувом фаз між взаємно перпендикулярними коливаннями.

В реальних мікромеханічних гіроскопах робочі коливання здійснюються змінним електростатичним полем, яке прикладається до гребінчатих електродів (вібраторів). Причому, використовується принцип електростатичного

відштовхування, оскільки пружна сила, що виникає під час вигину ніжок вібратора, повертає їх до положення рівноваги.

Напрямок орієнтації сили Коріоліса відносно електростатичної, показаний на рис.2.11.4,б на прикладі мікромеханічного гіроскопа фірми The Charles Stark Draper Laboratory. Якщо платформа гіроскопа обертається з кутовою швидкістю Ω

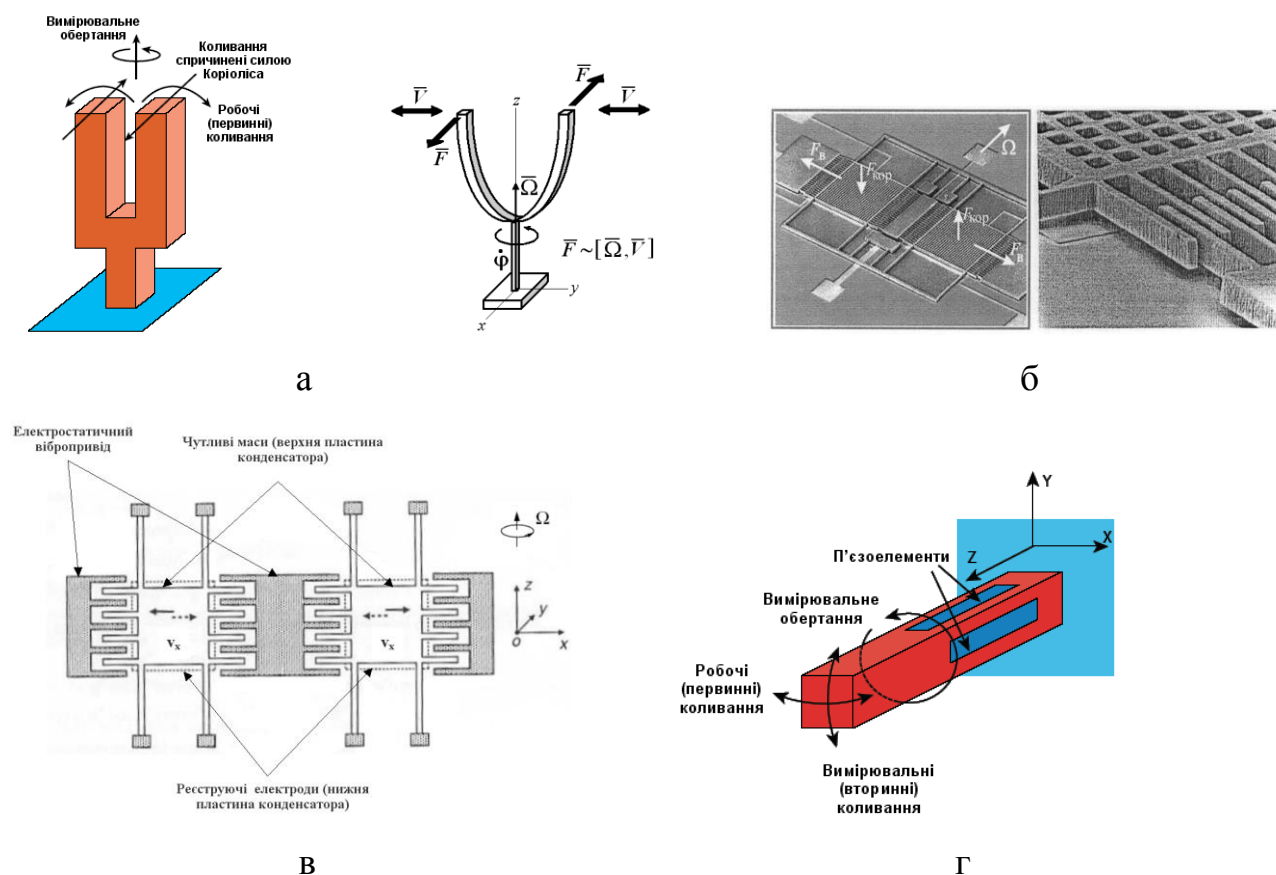


Рис.2.11.4

відносно вісі z , то на інерційне тіло діють протилежно напрямлені сили

Коріоліса, яка за величиною пропорційна Ω . Коріолісові сили інерції призводять до виникнення поступальних коливань чутливих мас в площині YOZ . Саме ці коливання є мірою кутової швидкості обертання основи Ω .

В даній системі координат, вирази для кінетичної, потенціальної і дисипативної

енергій мають вигляд: $W_k = \frac{1}{2}m \left[\left(\frac{dx}{dt} - y\Omega \right)^2 + \left(\frac{dy}{dt} + x\Omega \right)^2 \right]$; $W_{II} = \frac{1}{2}(\mu_x x^2 + \mu_y y^2)$,

$W_Q = \frac{1}{2} \left[r_x \left(\frac{dx}{dt} - y\Omega \right)^2 + r_y \left(\frac{dy}{dt} + x\Omega \right)^2 \right]$. Підставивши ці вирази в рівняння Лагранжа

другого роду, одержимо систему диференційних рівнянь другого порядку [73]:

$$\begin{cases} m \frac{d^2 x}{dt^2} + r_x \frac{dx}{dt} + (\mu_x - m\Omega^2)x - 2\Omega \frac{dy}{dt} - \left(r_x \Omega - \frac{d\Omega}{dt} \right) y = F \sin \omega t, \\ m \frac{d^2 y}{dt^2} + r_y \frac{dy}{dt} + (\mu_y - m\Omega^2)y + 2\Omega \frac{dx}{dt} + \left(r_y \Omega + \frac{d\Omega}{dt} \right) x = 0, \end{cases} \quad \text{де } m - \text{маса робочого тіла; } \mu_{x,y} - \text{коефіцієнти жорсткості та } r_{x,y} - \text{коефіцієнти тертя вздовж вісей } x \text{ і } y; \Omega - \text{кутова швидкість.}$$

Якщо кутова швидкість обертання гіроскопа навколо вісі Z незмінна і задовольняється нерівність $\sqrt{\frac{\mu_y}{m}} \gg \Omega, \sqrt{\frac{\mu_x}{m}} \gg \Omega$ та можна знехтувати впливом руху вздовж Y на рух вздовж вісі X , то диференційні рівняння

запишуться у наближеному вигляді:

$$\begin{cases} m \frac{d^2 x}{dt^2} + r_x x + \mu_x \frac{dx}{dt} = F \sin \omega t, \\ m \frac{d^2 y}{dt^2} + r_y y + \mu_y \frac{dy}{dt} = -2m \frac{dx}{dt} \Omega, \end{cases} \quad \text{що суттєво спрощує подальший аналіз.}$$

Дійсно, якщо, як перше наближення розглянути лише вимушені коливання, то розвязки цієї системи диференційних рівнянь матимуть вигляди:

$$x(t) = x_m \sin(\omega t - \varphi_x) = \frac{f_0}{\sqrt{(\omega_{0x}^2 - \omega^2)^2 + 4\gamma_x \omega^2}}; \quad y(t) = y_m \cos(\omega t - \varphi_x - \varphi_y) = x_m \frac{2\Omega \omega}{\sqrt{(\omega_{0y}^2 - \omega^2)^2 + 4\gamma_y \omega^2}}, \quad \text{де } \operatorname{tg} \varphi_{x,y} = \frac{2\gamma_{x,y} \omega}{\omega_{0x,y}^2 - \omega^2}.$$

Бачимо, що амплітуда вертикальних коливань камертона $y_m \cong \frac{x_m}{\omega_{0y} - \omega} \Omega$, якщо знехтувати дисипативними процесами в коливній системі.

Викладений аналіз у спрощеному вигляді не враховує роль перехідних процесів. Це можна врахувати шляхом застосування перетворень Лапласа. Для цього приймемо до уваги традиційну заміну параметрів та застосуємо перетворення Лапласа для нульових початкових умов $x(0)=0$ і $\left. \frac{dx}{dt} \right|_{t=0} = 0$, в результаті чого

$$\text{одержимо систему алгебраїчних рівнянь: } \begin{cases} s^2 X(s) + 2\gamma_x s X(s) + \omega_{0x}^2 X(s) = f_m \frac{\omega}{s^2 + \omega^2}, \\ s^2 Y(s) + 2\gamma_y s Y(s) + \omega_{0y}^2 Y(s) = -2s X(s) \Omega. \end{cases}$$

Для простоти знехтуємо дисипативними процесами. Тоді

$$\begin{cases} s^2 X(s) + \omega_{0x}^2 X(s) = f_m \frac{\omega}{s^2 + \omega^2}, \\ s^2 Y(s) + \omega_{0y}^2 Y(s) = -2sX(s)\Omega. \end{cases} \quad \text{звідки} \quad X(s) = \frac{\omega f_m}{(s^2 + \omega_{0x}^2)(s^2 + \omega^2)} \quad \text{і}$$

$$Y(s) = \frac{-2\omega s \Omega f_m}{(s^2 + \omega_{0x}^2)(s^2 + \omega^2)(s^2 + \omega_{0y}^2)}.$$

Застосувавши в подальшому зворотнє Лапласа,

одержимо кінцеві наближені вирази для амплітуд коливань підвіски:

$$x(t) \cong \frac{f_0(\omega_0^2 - \omega^2)}{m\Theta} \sin \omega t; \quad y(t) \cong \Re \frac{(\omega_0^2 - \omega^2)}{m\Theta} \cos \omega t, \quad \text{де знехтувано анізотропністю}$$

пружно-дисипативними характеристиками коливного осцилятора вздовж вісей

$$X \text{ і } Y, \quad \Re = -2\Omega f_0 \frac{\omega(\omega_0^2 - \omega^2)}{\Theta}, \quad \Theta_{x,y} = \omega^4 + 4\alpha_{x,y}^2 \omega^2 + \omega_{0x,y}^4 - 2\omega^2 \omega_{0x,y}^2.$$

Отже відношення

амплітуд коливань підвіски гіроскопа дорівнює

$$\frac{y_m(t)}{x_m(t)} = 2\Omega \frac{\omega(\omega_0^2 - \omega^2)}{\omega^4 + 4\alpha^2 \omega^2 + \omega_0^4 - 2\omega^2 \omega_0^2} \cong \frac{2\Omega \omega}{(\omega_0^2 - \omega^2)}, \quad \text{тобто лінійно пропорційне кутовій}$$

швидкості Ω . Як бачимо з рис.2.11.5,а, за наявності силових дій на підвіску у двох взаємно перпендикулярних напрямках, в площині її траєкторія руху має форму еліпса, який відтворює відповідний еліпс самих сил. На рис.2.11.5,б і в, схематично показані так звані кінематичні схеми камертонного гіроскопа (б) стосовно використання їх для вироблення траєкторії руху літака.

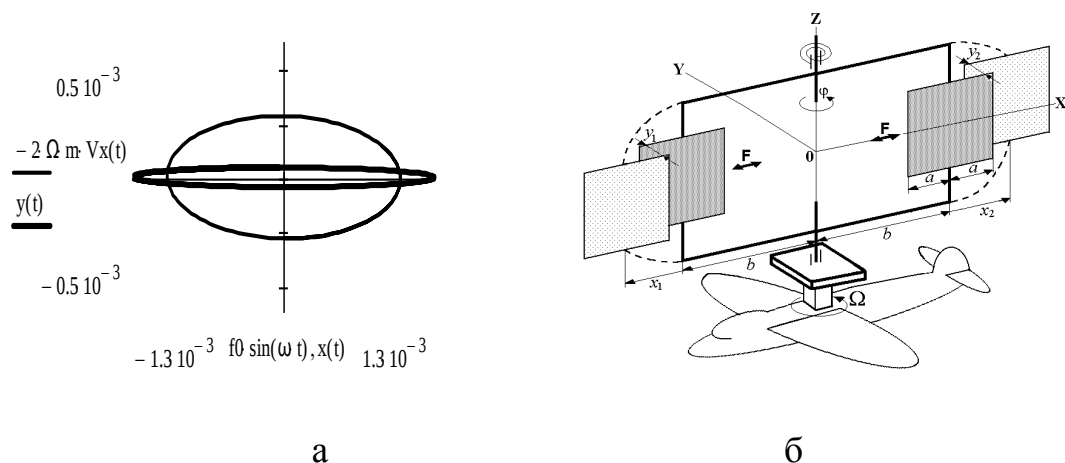


Рис.2.11.5

Мікромеханічні гіроскопи за технологією iMEMS. Це клас гіроскопів серії ADXRS, які виконані на кристалі кремнію і включають в себе всі необхідні електронні схеми формування сигналу. Вони розроблені фірмою Analog Devices [2.45].

Схеми чутливого елементу мікромеханічного гіроскопу показані на 2.11.6. Принцип роботи гіроскопів серії ADXRS базується на вимірюванні коріолісового прискорення, який діє на резонуючу масу. Чутливий елемент гіроскопа складається внутрішньої і зовнішньої рамок, чутливої маси, пружних елементів резонуючої маси, пружних елементів внутрішньої рамки, гребінок, які виконують роль ємнісного давача коріолісового прискорення.

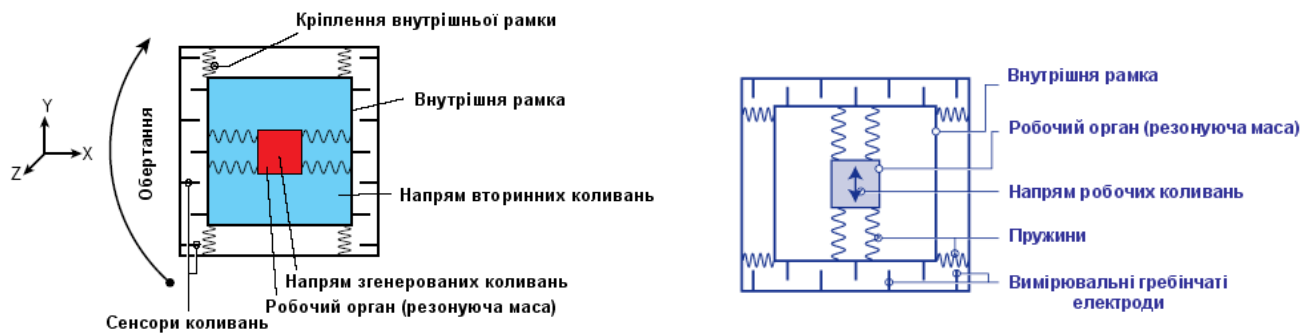


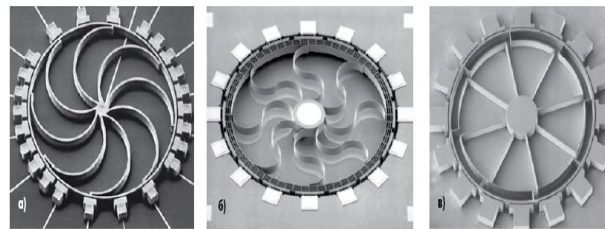
Рис.2.11.6

Коливання робочого тіла збуджуються електростатично за допомогою гребінчатого конденсатора, які розташовані на зовнішній поверхні рамки. Коли рамка коливається відносно основи, відстань між гребінками змінюється, відповідно змінюється і ємність конденсатора. Вимірювальні сигнали знімаються з ємностей обох структур та подаються на диференційну схему. При цьому сигнали, викликані коливаннями під дією коріолісовго прискорення, будуть додаватися, а синфазні складові, зумовлені лінійними прискореннями, - відніматися. Нечутливість мікрогіроскопа до вібрацій і лінійного прискорення досягається завдяки наявності всередині мікросхеми двох перпендикулярно розташованих однотипних мікроскопів чутливих механічних елементів.

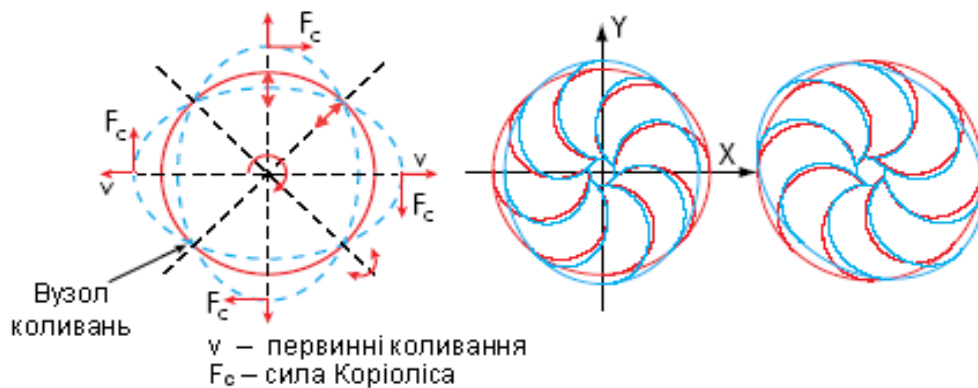
Гіроскопи з диском-вібратором

Ще одна різновидність вібраційних гіроскопів – пристрої з диском-резонатором. Одні з перших гіроскопів з дисковим резонатором (рис.2.11.7, а) нікелевий, б) полікремнієвий, в) монокристалічний). Нікелевий диск діаметром 1 мм – обруч (кільце) з вісьмома напівкруглими спицями, жорстко закріплених в центрі на полікремнієвій підкладці. Із зовнішньої відносно кільця сторони з невеликим зазором розташовані приводні, вимірювальні та коректуючі електроди.

Під дією прикладеної до приводних електродів напруги генерується основна мода коливань (наприклад, вздовж вісі Y) – диск розтягується вздовж вісі Y і стискається вздовж вісі X (рис.2.11.7,б). Якщо основа (а з нею і диск - кільце) починає обертатися навколо вісі Z , сила Коріоліса намагається викликати коливання в напрямку вісі X . Виникає друга (вимірювальна) мода коливань – головна вісь еліпса обертається на кут 45° . Ці коливання реєструються ємнісними сенсорами.



а



б

Рис.2.11.7

Вібраційні гіроскопи обертового типу

Основним їх робочим елементом є диск-резонатор, що закріплений на торсійній підвісці, крутильні коливання вертикальної навколо вісі Z якого збуджуються електричним методом. Один з таких гіроскопів, створений в Берклі, показаний на рис.2.11.8. Якщо гіроскоп обертається навколо вісі X , то під дією сили Коріоліса виникають коливання перпендикулярні площині диска вздовж вісі Y . Оскільки на поверхні диска нанесений електрод, а на основі під ним – чотири секторних електроди (відповідно вздовж вісей X і Y , то коливання супроводжуються зміною ємності конденсаторної системи, що пов'язується з амплітудою вертикальних коливань і кутовою швидкістю обертання гіроскопа.

Оптичні гіроскопи. На відміну від механічних, у них відсутні рухоми деталі, вони стійкі до прискорення, удару і механічних вібрацій, мають просту конструкцію, короткий час запуску, високу чутливість та лінійність робочих характеристик, мале енергоспоживання, високу надійність. Точність вимірювання ними знаходиться в межах $0.001^\circ/\text{год} \dots 0,01^\circ/\text{год}$, а

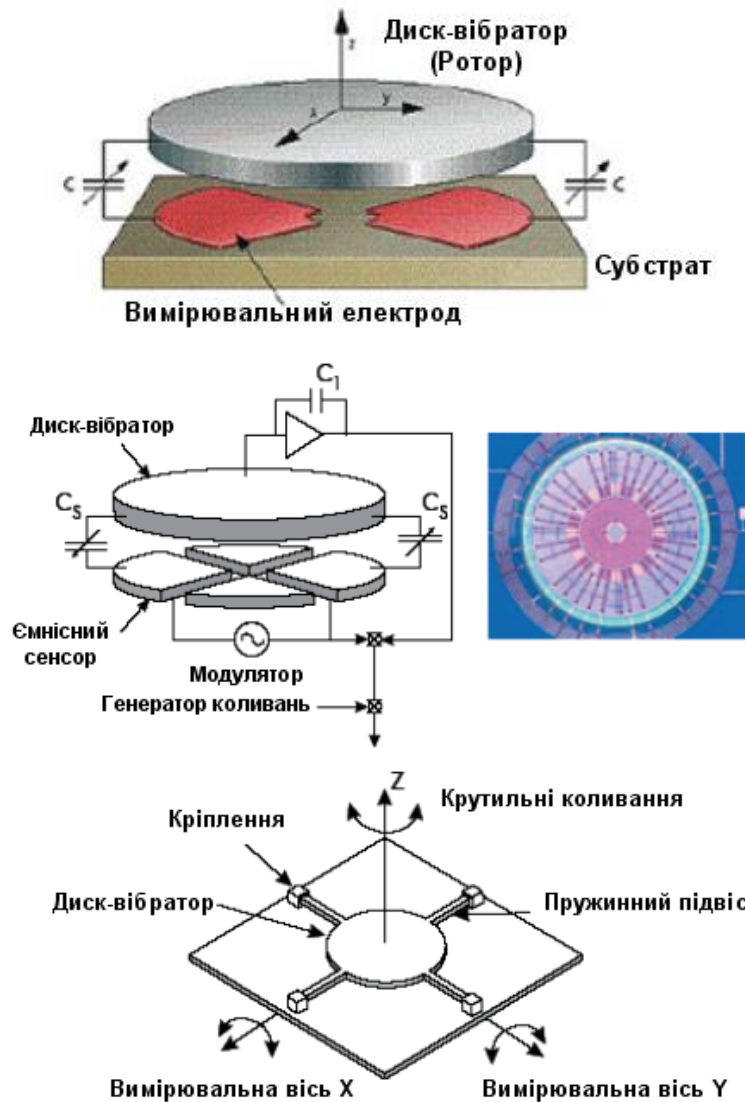


Рис.2.11.8

діапазон вимірювання кутових швидкостей від $1^\circ/\text{год} \dots 300^\circ/\text{сек}$. Завдяки відсутності механічних деталей, оптоволоконні гіроскопи можуть використовуватися в навігації, для контролю за рухом бура під час буріння нафтових свердловин, для прогнозу погоди, вимірювання флуктуацій власного обертання Землі, тощо.

В основу оптичних гіроскопів покладений ефект Саняка (або вихровий оптичний ефект, тобто здатність виявляти кутову швидкість обертання), який відкритий в 1913 році і його суть полягає в наступному. Розщеплений лазерний промінь світла поширюється в двох протилежних напрямках (рис.2.11.9,а). В стані спокою обидва промені проходять до точки суперпозиції однаковий оптичний шлях, тоді як під час обертання станини гіроскопа з кутовою швидкістю Ω між виникає різниця фаз, пропорційна кутовій швидкості обертання контура.

Широке застосування знайшли кільцеві гіроскопи (рис.2.11.9,б). Датчиком в ньому служить кільцевий лазер, який генерує дві біжучі назустріч одна одній світлові хвилі, що поширюються вздовж одного оптичного каналу. Резонатор кільцевого лазера складається як мінімум з трьох дзеркал, що змонтовані на жорсткій основі і утворюють замкнуту оптичну систему. Частина світла проходить крізь напівпрозоре дзеркало і попадає на фотодетектор. Якщо гіроскоп не обертається, то лазер генерує дві зустрічні хвилі однакової частоти

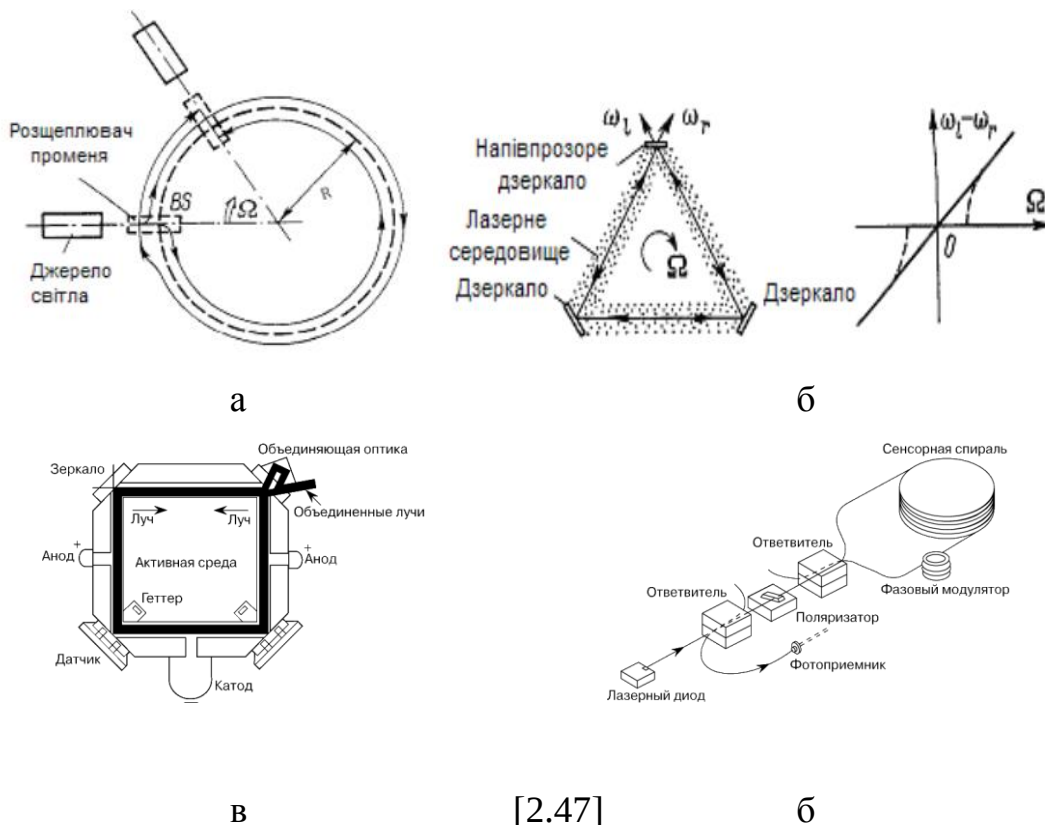
$$\nu_0 = \frac{c_0}{\lambda_0} = \frac{c_0}{L} m, \text{ де } m - \text{ціле число довжин хвиль, що формують стоячу хвилю на}$$

кільцевому шляху довжиною L . Якщо ж гіроскоп обертається з кутовою швидкістю Ω , то за час обходу променями шляху до зустрічі, гіроскоп встигає повернутися на деякий кут. Із-за ефекту Доплера частоти зустрічних пучків хвиль площею поперечного перерізу S відрізнятимуться між собою: $\nu_{\pm} = \nu_0 \pm 2 \frac{\nu_0}{c_0} S \frac{\Omega \cos \nu}{L}$,

ν - кут між напрямком спостереження. Фотодетектор фіксуватиме сигнал биття з частотою $\Delta \nu = \nu_+ - \nu_-$. Конструкція кільцевого гіроскопа приведена на рис.2.11.9,в. Однак, оптичні гіроскопи не позбавлені недоліків: вихідний сигнал нелінійний при малій кутовій швидкості; має місце дрейф сигналу із-за газових потоків; оптичний шлях може змінюватися внаслідок теплового розширення, зміни тиску газу і механічних деформацій гіроскопа [2.46].

За методом вимірювання кутової швидкості, гіроскопи Саняка поділяють на групи. До першої групи відносяться так звані кільцеві лазерні гіроскопи, в яких зустрічні промені, поширюючись в обертовому активному середовищі, зазнають різниці частот, за величиною якої роблять висновок про швидкість обертання. Нагадаємо, що під час обертання починає зміщуватися

інтерференційна картина із швидкістю, що пропорційна швидкості обертання. Друга група методів вимірювання кутової швидкості базується на використанні оптоволоконного котушки (аналогічної за своєю конструкцією до соленоїда). Третя група методів зв'язана з використанням в структурі гіроскопа кільцевого резонатора пасивного типу, в яких інформація про інерцію гіроскопа добувається із величини фазового зсуву між інтерферуючими променями.



[2.47]

Рис.2.11.9

Широкого застосування знайшли волоконно-оптичні гіроскопи (рис.2.11.9,в). Він такж діє по принципу інтерферометра Саньяка, однак в ньому використовується зовнішнє лазерне джерело. Для збільшення довжини оптичного шляху і підвищення чутливості гіроскопа оптичне волокно намотано у вигляді спіралі. Якщо гіроскоп обертається з кутовою швидкістю $\Omega = \frac{d\theta}{dt}$, то різниця шляхів для зустрічних хвиль дорівнює $\Delta\ell \cong \frac{4\pi R^2}{g}\Omega$, а різниця фаз між цими променями $\Delta\varphi = \frac{4\pi L}{c_0\lambda}R\Omega$, де L – довжина волокна, R – радіус оптоволоконного

кільця. Система інерційної навігації з застосуванням оптичного гіроскопа показана на рис.2.11.9,б.

Мікромеханічні приводи руху – актюатори

Мікроактюатор – це механічний пристрій, який перетворює електричну, теплову, хімічну та інші види енергії в механічну роботу, електромагнітне (переважно теплове) випромінювання, в тому числі світло. Це невід’ємна частина МЕМС і належить до виконавчих елементів. В залежності від фізичного принципу, що використовується в актюаторі, мікроактюатори поділяють на п’єзоелектричні, електростатичні, термомеханічні, електромагнітні та пневматичні [2.49].

Термоелектричні актюатори. Розроблені актюатори, робочим елементом в яких є металевий стрижень, що закріплений з одного або обох кінців. Якщо крізь нього протікає електричний струм, то внаслідок ефекту Джоуля-Ленца стрижень нагрівається і розширюється. Якщо кінці стрижня закріплені, то ефект термічного розширення проявляється у вигляді його вигину (рис.2.11.10,а) [2.48]. Отже, коли на актюатор подається змінна електрична напруга, то стрижень здійснює згинні коливання, аналогічно закріплений з обох кінців струни.

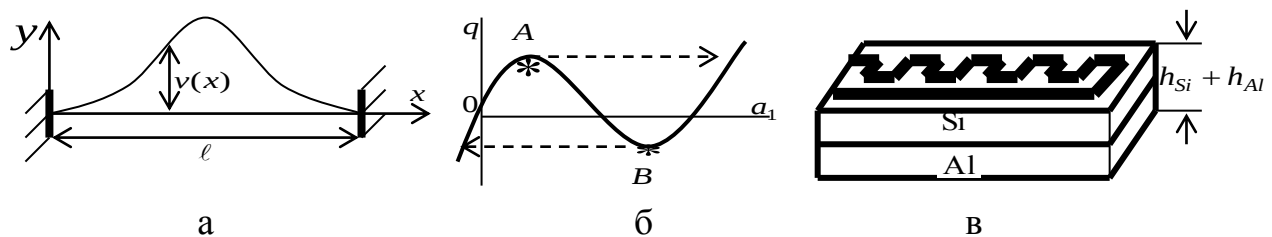


Рис.2.11.10

Нехай балка має форму прямокутного паралелепіпеда площею поперечного перерізу S , момент інерції J , модуль Юнга E_σ і закріплена двома анкерами, відстань між якими ℓ . Поперечні коливання балки в напрямку вісі y амплітудою

$v(x)$ описуються диференціальним рівнянням $E_\sigma J \frac{d^4 v}{dx^4} + P_\ell \frac{d^2 v}{dx^2} = 0$, де P_ℓ - вісьове навантаження.

В точках кріплення виконуються граничні умови

$v(0) = v(\ell) = 0$, $\left. \frac{dv}{dx} \right|_{x=0} = \left. \frac{dv}{dx} \right|_{x=\ell} = 0$, тому зміщення балки жорсткістю μ

виражатиметься через гармоніки з просторовими частотами $\frac{n}{2\ell}$ як

$$v(x) = v_0(x) - \sum_{n=1}^{\infty} a_n \sin^2\left(\frac{n\pi x}{\ell}\right).$$

Якщо $v_0(x) = h \sin^2\left(\frac{\pi x}{\ell}\right)$, а рівномірно розподілена навантаження q вздовж y , то вищі гармоніки «затиснуті», тобто не збуджуються і рівняння для зміщення можна записати як $v(x) = v_0(x) - a_1 \sin^2\left(\frac{\pi x}{\ell}\right) = v_0(x) - \frac{a_1}{2} \left(1 - \cos\left(\frac{2\pi x}{\ell}\right)\right)$. Енергія

пружних деформацій дорівнює $W_p = \int_0^{\ell} \left[\frac{\{P_{\ell}(x)\}^2}{2E_{\sigma}S} + \frac{\{M(x)\}^2}{2E_{\sigma}J} + q(v(x) - v_0(x)) \right] dx$, де

$$P_{\ell}(x) = E_{\sigma}S\varepsilon(x) = \frac{E_{\sigma}S\pi^2 a_1(1-2h)}{4\ell^2} \left[1 - \cos\frac{4\pi x}{\ell}\right], \quad M(x) = E_{\sigma}J \frac{2\pi^2 a_1}{\ell^2} \cos\frac{2\pi x}{\ell}. \quad \text{Розв'язуючи}$$

рівняння $\frac{\partial W_p}{\partial a_1} = 0$ одержимо, що $q = \frac{3E_{\sigma}S\pi^4(a_1^3 - 3ha_1 + 2h^2a_1)}{8\ell^4} + \frac{4E_{\sigma}J\pi^4 a_1}{\ell^4}$. Графік

цього рівняння поданий на рис.2.11.10,б. Це відома у фізиці залежність S- типу з двома бістабільними станами A і C .

Біморфний термоактюатор

Розроблені термоактюатори на основі бішарової пластини, математична модель якого наведена на рис.2.11.10,в. Тут актюатор складається з двошарової пластини завдовжки l , завширишки w та товщиною $h = S_{Si} + S_{Al}$. На верхній площині актюатора розташований нагрівний елемент.

Принцип дії цього пристрою такий. Під час нагрівання конструкції актюатора нагрівним елементом пристрій виконує переміщення в вертикальному напрямку за рахунок різних різних коефіцієнтів теплового розширення обох пластин. коефіцієнтів теплового розширення, оскільки один матеріал розширюється швидше ніж інший.

Сформулюємо основні переваги теплових мікроактюаторів: - проста конструкція, оскільки робочими елементами є резистори; - швидкодія при зменшенні розмірів зростає за квадратичним законом; - конструктивно виготовлення актюатора задовольняє вимогам планарної технології. До недоліків можна віднести: -нагрівний елемент споживає значну електроенергію, тому к.к.д.

низький; - нагрівний елемент вимагає охолодження, щоб досягнути вихідного стану. Це природно займає деяку кількість часу і обмежує швидкість теплових актюаторів.

Електростатичні актюатори. Принцип дії цього типу актюатора базується на електростатичній взаємодії між рухомим і нерухомим електродом. В електричному полі напруженістю E на елемент з зарядом dQ діє сила $dF = ECdU$, звідки на підставі базових співвідношень електричного конденсатора ємністю C інтегральна сила дорівнює $F = \int dF$.

Однак, щоб електростатична сила притягання була значною, треба збільшити ємність, чого можна досягнути лише шляхом збільшення площі взаємодіючих поверхонь, тобто використання гребінчастих електродів. При цьому зростання ємності пропорційне кількості стрижнів, а такого двигуна досягає 50%.

П'єзоелектричні актюатори. Більш простим варіантом термопневматичного актюатора є п'єзоелектричний актюатор. Цей тип актюатора є найбільш поширеним типів МЕМС та НЕМС-пристроїв і складається з п'єзоелектричного диска і пружної мембрани. В основу його роботи покладений обернений п'єзоелектричний ефект: під дією електричного поля в п'єзоелектрика (кварц, оксид цинку чи танталат літію) виникає механічна деформація (на рис.2.11.11 вона зображена у вигляді вигину).



Рис.2.11.11

Нагадаємо, що суть прямого п'єзоелектричного ефекту полягає в тому, що під час деформації п'єзоелектрика, на його протилежних поверхнях з'являються електричні заряди різного знаку і навпаки, якщо до п'єзоелектрика прикласти електричну напругу, то він деформується. Із-за малої зміни параметра ґратки, за допомогою п'єзоелектричного актюатора можна здійснювати, в тому числі нанорухи в зондових мікроскопах. За їх допомогою вдається реалізувати різні типи активацій за рахунок того, що зміна полярності прикладеної напруги.

Призводить до зміни типу деформації – від розтягу до стиску і навпаки. На їх основі сконструйовані блочні актюатора, в яких пластини п'єзоелектрика розташовуються послідовно; двигуни обертання стоячою хвилею, в яких ротор обертається шляхом вібрації; двигуни обертання біжучою хвилею, в яких ротор обертається шляхом його стискання в напрямку, що протилежний статору; крокові двигуни, біморфні актюатори, тощо.

Електромагнітні актюатори [2.7;2.7;2.48]

За аналогічним фізичним принципом діє електромагнітний актюатор. В ньому коливання робочого тіла збуджуються силою Ампера, що діє на провідний стрижень, що закріплений з обох кінців, під час проходження крізь нього електричного струму зі сторони зовнішнього магнітного поля (рис.2.11.12,а). Під дією магнітного поля напруженістю H на провідник із струмом діє сила Ампера $F = q_m H$, де q_m – заряд нерухомого кінця коливного тіла, - так званий магнітний заряд. Відносно нерухомої точки струмопровідної частини (або без струму, якщо матеріал магнітний, як сплав $NiFe$, сила Ампера створює обертовий механічний момент $M = q_m \ell H \sin \alpha$, де α - кут між напрямком $\vec{H}$ і віссю актюатора. Недоліком магнітних актюаторів є вимога для їх функціонування магнітного поля.

Функціональна схема іншого типу електромагнітного актюатора приведена на рис.2.11.12,б [2.50]. При обертанні залізного диска в зовнішньому електромагнітному полі, в ньому індукується ЕРС, величина якої пропорційна кутовій швидкості $E_i = K_j \omega$. З наявності струму в колі мотора I_a , останній генерує обертовий момент $M_\omega = K_{dc} I_{dc}$.

Для мікроелектронних технологій розроблений електромагнітний актюатора пластинчатого типу. В ньому тонка пермалоева пластина в магнітному полі генерує механічний обертовий момент, який намагається зменшити повну енергію в системі актюатора шляхом вирівнювання вектора намагніченості з силовими лініями зовнішнього магнітного поля. На рис.2.11.12,в показана конструкція лінійного мотора на базі магнітного актюатора. Шляхом перемикавання електричного сигналі між обвитками, повзунковий магніт займатиме те чи інше положення. Розроблені так звані магнітострикційні

актюатори [2.51]. В них використовується ефект магнітострикції, тобто зміни форми феромагнетика, яка спричинена індукцією магнітного поля.

Має місце загальна проблема для магнітних актюаторів: обвитки котушок двохвимірні (тривимірні обвитки дуже важко виготовити на мікрорівні). Додаткове обмеження пов'язане з вибором магнітного матеріалу. Вибираються тільки ті матеріали, які легко обробити на мікрорівні, тобто виходить, що не завжди матеріал магніту вибирається оптимально. Тому часто магнітні мікроактюатори споживають значну кількість енергії та розсіюють багато тепла. З таких міркувань, електростатичні мікропристрої можуть бути вигідніші ніж магнітні. Інший тип магнітогідродинамічного актюатора показаний на рис.2.11.12,г.

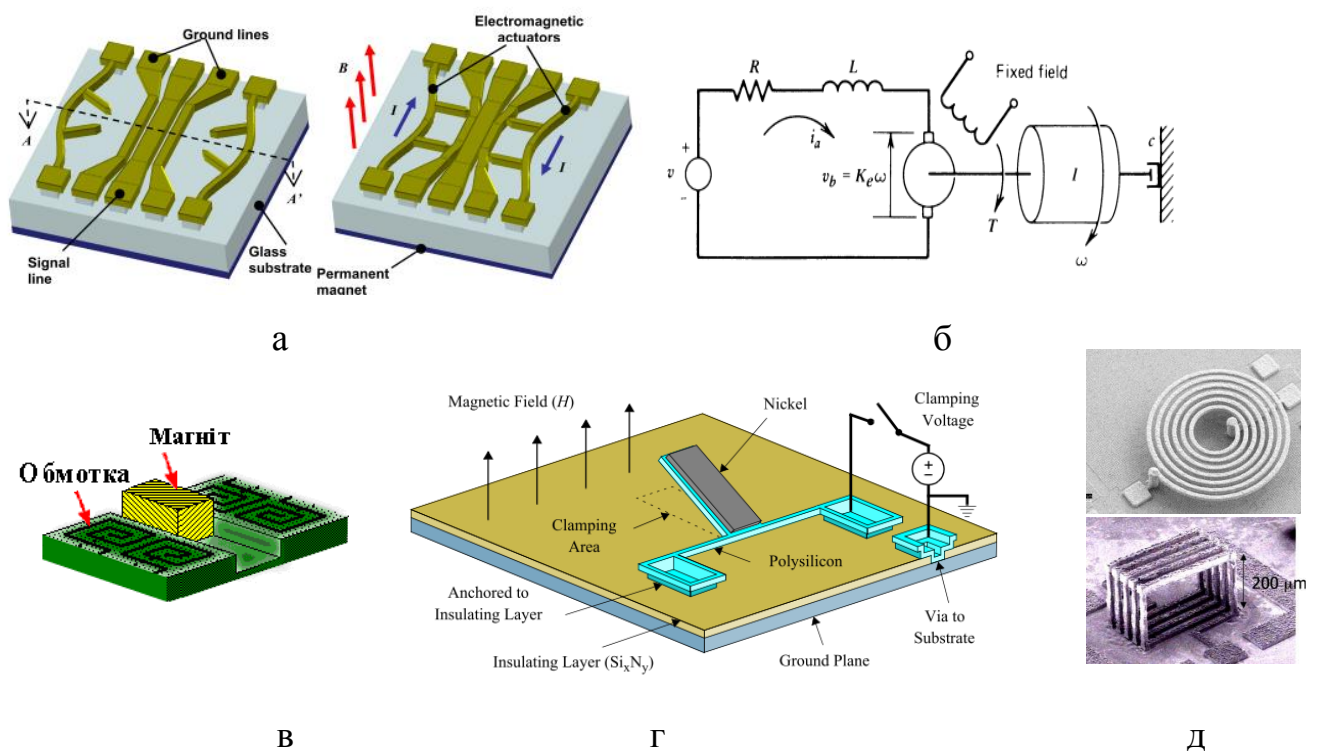


Рис.2.11.12

Для електромагнітних актюаторів чи сенсорів важливим є інтеграція в сучасні технології, що вимагає конструювання планарних та об'ємних мікрокотушок індуктивності. Так для створення електромагнітних мікроактюаторів виникає потреба інтегрувати в котушки магнітом'яккі структури, які забезпечують концентрацію магнітного потоку. Схеми безових конструкцій мікрокотушок, показані на рис.2.11.12,д. Зліва схематично показаний зразок планарної котушки, а справа- тривимірної гвинтового типу («меандрового

типу»), які створюють горизонтальне магнітне поле. Технологічно, виготовляються вони з застосуванням УФ-літографії, гальванічного нанесення провідної структури (міді), нікеля (магнітної структури) та полімерної ізоляції [2.52].

Термопневматичні актюатори. Термоактюаторні елементи МЕМС виготовляють по технології поверхневої обробки і MUMPs-технології. Так, на рис.2.11.14,а, показаний термопневматичний актюатор, який складається з



Рис.2.11.14

нагрівного елемента і герметичної порожнини з пружною мембраною, заповненого газом. Всередині об'єму інтегрований електричний нагрівник, який, змінюючи температуру газу, призводить до зміни його об'єму, тобто деформації мембрани.

В інтегральних технологіях нагрівником є напівпровідниковий резистор меандрового типу. Електричний струм, що тече крізь нього, за елементарний проміжок часу dt призводить до виділення $dQ = \frac{U^2}{R} dt$. Термопневматичні актюатори виготовляються за LIGA-технологією разом з об'ємною мікроборкою.

Якщо газ замінити рідиною, то одержимо терморідинний актюатор. Рідина також нагрівається за допомогою електричного резистивного нагрівача. Розширюючись, вона тисне на тонку мембрану деформуючи її при цьому. Якщо допустити, що рідина густиною ρ має властивості ідеального газу, то нове значення тиску можна визначити як з виразу: $p = (\gamma - 1)\rho C_v (T_2 - T_1)$, $\gamma = C_p / C_v$, де C_p – питома теплоємність рідини (газу) при сталому тиску; C_v – питома теплоємність при сталому об'ємі; T_1, T_2 – початкове та кінцеве значення

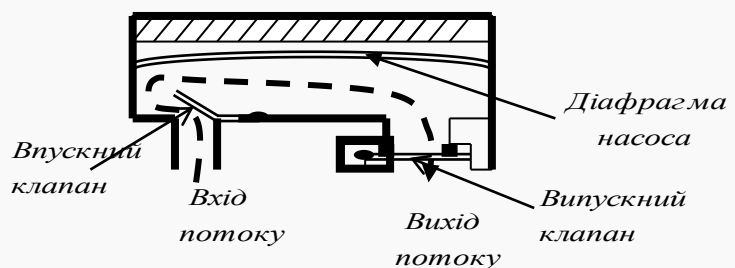
температури рідини (газу). Інший тип термоактюатора без пневмоефекта показаний на рис.2.11.14,б.

Оптичні актюатори. Електростатичний принцип реалізований в інтегральних мікродзеркалах, які використовуються в МЕМС чи робо технічних системах для створення оптичних ключів і системах обробки зображень для відхилення напрямку поширення лазерного променя (рис.2.11.15,а).

Принцип дії такого дзеркала базується на тому, що шляхом об'ємної мікрообробки виготовляється електрод у вигляді мікродзеркала на торсійних підвісках. Рухоме дзеркало заземляється, а на відхиляючий електрод підводиться змінна електрична напруга, яка відхиляє мікродзеркало. Пружна деформація кручення торсійних балок з обертовим моментом $M = \frac{1}{2} U^2 \frac{dC}{d\phi}$ повертає мікродзеркало у вихідне положення, якщо електрична напруга відсутня. де $\frac{dC}{d\phi}$ - зміна ємності конденсатора за рахунок кутового відхилення рухомого електрода (мікродзеркала). Мікродзеркала можуть здійснювати механічні рухи вздовж двох чи трьох напрямках.



а



б

Рис.2.11.15

Електрогідродинамічні мікронасоси. За принципом функціонування мікронасоси поділяються на два класи: немеханічні, які не мають рухомих механічних частин, дія яких базується на електродинамічному, електроосмотичному та ультразвуковому ефектах, та мікронасоси з рухомими механічними частинами, в тому числі з активними та пасивними мікроклапанами. Електрогідродинамічні насоси – це мікронасоси інжекторного типу. Вони складаються з двох електродів у вигляді сіток, простір між якими заповнений діелектриком. Якщо такий насос вставити в полярну рідину, що має іони і диполі,

і між електродами-сітками створити різницю потенціалів, то внаслідок взаємодії іонів та диполів із електричним полем, на частинки рідини діють сили, що зумовлюють їх прискорений рух від одного електрода до іншого. Отже, в одному напрямку виникає рух рідини крізь сопло сітки. Зміна полярності електричної напруги на електродах, призведе до зміни напрямку руху рідини. Зазначимо, що перекачувані рідини повинні мати високий питомий електричний опір (не менше $10^{10} \div 10^{14} \text{ Ом} \cdot \text{см}$) і процес перекачки не повинен супроводжуватися її електролізом.

Мембранні насоси

В даний час основним силовим елементом механічної частини актюатора є мембрана, вигін якої здійснюється електростатичними або п'єзоелектричними силами. На рис.2.11.15,б, показана схема дії мембранного актюатора-мікронасоса, який виготовляється методами планарної технології із кристалів кремнію, що склеюються між собою. Верхня пластина-електрод разом із мембранною діафрагмою утворюють електропривід. Якщо їх металізовані поверхні заряджені зарядами різних знаків, то вони притягуються один до одного і під мембраною виникає розрідження, внаслідок чого мембрана вхідного клапана прогинається вгору і камера під діафрагмою наповнюється рідиною чи газом. Коли електростатичне поле між електродами електроприводу відсутнє, то внаслідок пружності мембрана діафрагми повертається у вихідне положення і в камері утворюється надлишковий тиск. Коли електроди електроприводу заряджаються зарядом однакового знаку, то мембрана діафрагми вигинається вниз, виштовхуючи при цьому рідину чи газ.

3. Моделі та методи статистичного аналізу мікросистем

3.1. Вступні поняття

Випадкова подія. Події, перебіг яких передбачити неможливо, називають випадковими. Події, що не можна розкласти на простіші, називають елементарними. Сукупність усіх елементарних подій утворюють простір елементарних подій, в якому самі елементарні події є його точками. До випадкових подій відносяться імпульси лічильника Гейгера, попадання електронів на анод вакуумної лампи чи поява покупця в крамниці, тощо. Всі ці події можна відмітити точками на вісі часу. Випадковий характер мають значення енергій космічних частинок на вісі енергій розташування яких призводить до вивчення особливого класу змінних – стохастичних. Відношення числа випадків m , в яких подія реалізується, до числа всіх можливих випробувань $P = \frac{m}{n}$ ($0 \leq \frac{m}{n} \leq 1$) називається ймовірністю події. Це класичне означення ймовірності [3.1-3.3], згідно з яким випадкові події порівнюють за певною мірою можливості їх появи, яке дав Лаплас. Однак воно не відповідає на питання, що таке ймовірність випадкової події, а дає лише метод її обчислення. Таке означення непридатне для не дискретного простору елементарних подій і навіть для дискретного простору нерівноймовірних подій.

Випадкова величина, її частота . Випадковою називається величина, яка непередбачувано змінюється від досліду до досліду і її подальші значення на основі попередніх досліджень можна прогнозувати лише з певною ймовірністю. Випадкова величина (ВВ) відрізняється від випадкової події тим, для випадкової події є випадковим її поява, тоді як для випадкової величини поява її не викликає сумніву і випадковим є її значення. ВВ поділяють на неперервні, які можуть приймати безмежну множину значень, що неперервно змінюються, та дискретні ВВ, які можуть приймати скінченне число значень. Дискретні ВВ поділяються на безмежнозначні та скінченнозначні. Для дискретної ВВ задана ймовірність її здійснення, тоді як для неперервної ВВ існує густина або щільність ймовірностей. З ймовірнісної точки зору ВВ описана повністю, якщо встановлений закон її

розподілу. Відношення $\nu = \frac{m}{n}$ називається частотою або відносною частотою.

Намагання [3.2] поняття ймовірності випадкової події означити як границю частоти привели до деяких математичних труднощів, тому уникають точного його визначення і вважають основним поняттям.

Статистичне визначення ймовірності. Класичний спосіб визначення ймовірності має обмежений характер - множина елементарних подій повинна бути скінченною, а всі події – рівноможливими. Якщо хоч одна з цих вимог порушена, то треба застосовувати статистичний метод. Його суть полягає в тому, що за ймовірність події приймають її відносну частоту при достатньо великій кількості випробувань. Відношення $\nu = \frac{m}{n}$ називають статистичною ймовірністю ВВ. З статистичної точки зору, ймовірність - це її стале значення, навколо якої коливається значення частоти реалізації при досить значній кількості випробувань. Переважно ВВ позначається великою буквою, а її можливі значення –відповідно маленькою (може бути і з індексом) або числами. Так, ймовірність того, що випадкова величина X може приймати одне із значень x_i послідовності $x_1, x_2, x_3, \dots, x_n$, якщо верхня межа невідома, дорівнює $p_i = P(X = x_i) = \lim_{n \rightarrow \infty} \frac{m_i}{n}$.

Статистичне поняття ймовірності широко використовується в моделюванні методом Монте-Карло.

Геометрична ймовірність. Щоб позбутися недоліку класичного визначення ймовірності стосовно того, що його не можна використати до випробувань з безмежною кількістю реалізацій, вводять геометричну ймовірність – ймовірність «попадання» точки в область (відрізок, частину площини тощо). Так, коли відрізок ℓ складає частку від довжини L , і на відріжку L навмання «поставлена» точка, то вона може бути поставлена в довільне його місце. Ймовірність потрапляння точки на відрізок ℓ пропорційна його довжині і не залежить від його розташування відносно $L : P = \frac{\ell}{L}$. Це геометрична ймовірність.

Сформулюємо основні властивості ймовірностей та правило їх складання:**1.**

Ймовірність випадкової події є число невід’ємне. **2.** Ймовірність достовірної події дорівнює одиниці. **3.** Ймовірність випадкових подій підлягає правилу

аддитивності. Випадкові події незалежні, якщо ймовірність їх суміщення дорівнює добутку їх ймовірностей. Для незалежних ВВ X і Y , розподіл величини X не змінюється від того, що нам уже відоме значення, яке прийняла інша ВВ Y .

Псевдовипадкові числа. Той факт наскільки доцільно використовувати випадкові числа, визначається критерієм того, наскільки вони задовольняють спеціальним тестам. Числа, що обчислюються за якоюсь формулою і можуть бути використані замість випадкових, називаються псевдовипадковими числами. Псевдовипадкові числа задовольняють таким вимогам: **1.** Вони повинні бути рівномірно розподілені на деякому інтервалі. **2.** Між ними не повинні бути ніякої кореляції. **3.** Період послідовності псевдовипадкових чисел мусить бути великий. **4.** Послідовність псевдовипадкових чисел повинна відтворюватися. Більшість алгоритмів для отримання псевдовипадкових чисел будуються на рекурентних співвідношеннях, в яких початкове значення задане.

3.2. Числові характеристики та закони розподілу дискретних випадкових величин [3.1-3.3]

Випадкова величина X дискретна, якщо вона може приймати дискретну множину значень $(x_1, x_2, \dots, x_n)$. Допускається і безмежна кількість значень $(x_1, x_2, \dots, x_n, \dots)$. Дискретна ВВ є визначеною, якщо дані всі її можливі значення та ймовірності цих значень p . Якщо дискретну ВВ записують у вигляді таблиці

x_1	x_2	x_3	$\dots$	x_i	$\dots$	x_n
p_1	p_2	p_3	$\dots$	p_i	$\dots$	p_n

то в ній вказується, якої ймовірності p_i набуває подія $(X = x_i)$: $p_i = p(X = x_i)$. Це ряд розподілу ймовірностей для дискретної ВВ. Якщо ВВ X приймає лише одне із значень $x_1, x_2, x_3, \dots, x_n$, то сума ймовірностей (частот) усіх можливих значень ВВ дорівнює одиниці: $\sum_{i=1}^n p_i = 1$. Це умова нормування ймовірностей.

Числові характеристики дискретних ВВ

Статистичний ряд. Елементи $x_1, x_2, \dots, x_n$ називають варіантами ВВ X . Вони утворюють ряд варіант або простий статистичний ряд. Числові характеристики статистичного ряду виражають найбільш важливі властивості закону розподілу ВВ.

Так, величина $\langle x \rangle = \frac{n_1 x_1 + n_2 x_2 + \dots + n_n x_n}{n}$ називається середнім арифметичним

значенням ВВ X . Приймаючи до уваги статистичне означення ймовірності

одержуємо, що $\langle x \rangle = \sum_{i=1}^n x_i P(X = x_i) = \sum_{i=1}^n x_i p_i = \bar{x}$. За своїм змістом $\langle x \rangle = \frac{1}{n} \sum_{i=1}^k n_i x_i$ – це

середнє значення ВВ, яке обчислюється як сума добутків всіх можливих значень

ВВ на відповідні їм ймовірності. Оскільки $\sum_{i=1}^n p_i = 1$, то середнє значення можна

записати у вигляді $\langle x \rangle = \frac{\sum_{i=1}^{\infty} x_i p_i}{\sum_{i=1}^n p_i}$. Тому середнє $\bar{x}$ – це середнє статистичне

значення, що співпадає з середнім арифметичним.

Математичне сподівання випадкової величини. Число $M[X] = \sum_{i=1}^n x_i p_i$,

якщо кількість значень ВВ скінченне або $M[X] = \sum_{i=1}^{\infty} x_i p_i$, якщо кількість значень

ВВ зліченне, для відомого ряду розподілу ймовірностей, називають

математичним сподіванням або центром розсіювання ймовірностей дискретної

ВВ X . Тут і в подальшому розглядатимемо такі ВВ, для яких безмежний ряд

$M[X] = \sum_{i=1}^{\infty} x_i p_i$ збігається абсолютно. Математичне сподівання може і не належати

до множини можливих значень ВВ.

Сформулюємо основні властивості математичного сподівання. Якщо C – довільна стала, а X і Y – дві незалежні ВВ (для них незалежні випадкові події з ймовірностями p_k та q_l , а ймовірність суміщення визначається як $p_{kl} = p_k q_l$), то:

1. Математичне сподівання лінійної функції $M[C + X] = C + M[X]$. **2.** $M[CX] = CM[X]$.

3. $M[X + Y] = M[X] + M[Y]$. **4.** $M[X \cdot Y] = M[X] \cdot M[Y]$. Одному і тому ж середньому

значенню ВВ може відповідати безмежна кількість ВВ, що відрізняються між

собою не лише значенням, але й фізичною природою. Тому вводиться інша

характеристика – відхилення ВВ від її середнього значення $x - \langle x \rangle$. Оскільки

середнє значення $\langle x - \langle x \rangle \rangle = \sum_{i=1}^{\infty} (x - \langle x \rangle) p_i = \sum_{i=1}^{\infty} x_i p_i - \langle x \rangle \sum_{i=1}^{\infty} p_i = \langle x \rangle - \langle x \rangle = 0$, то вводять характеристику розподілу ймовірностей ВВ X як середнє значення квадрату відхилення $\langle (x - \langle x \rangle)^2 \rangle$.

Дисперсія випадкової величини. Дисперсією дискретної ВВ називається число $D[X] = M[(X - M[X])^2] = M[X^2] - (M[X])^2$. Дисперсія виражає міру розкиду окремих значень ВВ навколо математичного сподівання: $D[X] = \sum_{i=1}^k (x_i - M[X])^2 p_i$.

До основних властивостей дисперсії належать: **1.** $D[X + C] = D[X]$. **2.**

$D[X + Y] = D[X] + D[Y]$, якщо ВВ незалежні. **3.** $D[X \cdot C] = C^2 M[X]$. Властивість

$D[X + Y] = D[X] + D[Y]$ справджується для незалежних ВВ.

Середнє квадратичне відхилення

Більш природною характеристикою ВВ є її середнє квадратичне відхилення (СКВ), або стандарт, або флуктуація, що обчислюється як $\sigma_X = \sigma = \sqrt{D[X]}$ і використовується для оцінки розсіяння ймовірностей можливих значень ВВ відносно $\langle x \rangle$. Середнє відхилення ніколи не перевищує середнє квадратичне відхилення. Сформулюємо основні властивості СКВ:

1. $\sigma(C \cdot X) = C \sigma_X$; **2.** $\sigma(C + X) = \sigma_X$; **3.** $\sigma^2(X + Y) = \sigma_X^2 + \sigma_Y^2 \Rightarrow \sigma(X + Y) = \sqrt{\sigma_X^2 + \sigma_Y^2}$;

4. Якщо $\sigma_{X_1}^2 = \sigma_{X_2}^2 = \sigma_{X_3}^2 = \dots = \sigma_{X_n}^2$, то $\sigma^2\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{\sigma}{\sqrt{n}}$. Дисперсія

розподілу ВВ має вимірність квадрату вимірюваної величини, тому вона незручна для користування. Коли зручно, щоб оцінка розсіяння мала вимірність ВВ, в розрахунках використовується позитивне значення квадратного кореня з дисперсії $\sigma_X = +\sqrt{D[X]}$.

Функція розподілу ймовірностей. Функцією розподілу ймовірностей дискретної ВВ X називають функцію $F(x)$, яка визначає ймовірність того, що дискретна ВВ X в результаті випробовування прийме менше або рівне допустимому її значенню x : $F(x) = P(X \leq x)$. Функція розподілу дорівнює ймовірності того, що випадкова величина прийме значення, менше або рівне її аргументу x і обчислюється як інтеграл від густини ймовірностей в межах від $-\infty$ до x . Це інтегральна функція розподілу (*the comulative distribution function*). Вона

виражає ймовірність того, що ВВ X приймає значення, менші за конкретне число x . Густина розподілу дискретна, тому функція розподілу має ступінчатий вигляд (рис.3.2.1а). До основним властивостей функції розподілу відносяться:

1. $0 \leq F(x) \leq 1$. **2.** $F(x_2) \geq F(x_1)$ - неспадна, якщо $x_2 > x_1$. **3.** $P(x_1 \leq X < x_2) = F(x_2) - F(x_1)$ - ймовірність попадання ВВ X в інтервал $[x_1; x_2]$. **4.** Коли ВВ X приймає значення в інтервалі $[a; b]$, тобто $(a \leq X < b)$, то виконуються умови: $F(x) = 0$ для $x \leq a$ і $F(x) = 1$ для $x > b$.

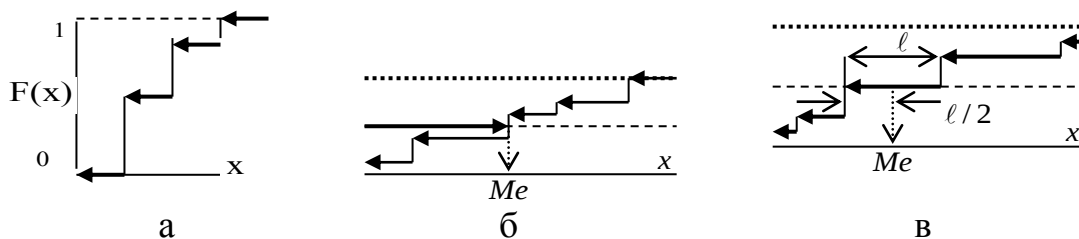


Рис.3.2.1

Медіана і мода дискретного ряду варіант. Математичне сподівання – важлива, але не єдина характеристика «усередненого» значення ВВ.

«Усереднене» значення може характеризуватися також медіаною і модою.

Медіаною Me ряду дискретних ВВ називається таке число, яке ділить цей ряд на дві рівні за кількістю варіант частин. Якщо число варіант непарне $2m+1$, то

$Me = x_{m+1}$. У випадку парної кількості варіант, $Me = \frac{1}{2}(x_m + x_{m+1})$. Таким чином, для

випадкової величини прийняття значення, меншого медіани, рівно ймовірно

прийняттю значення, більшого медіани: $P(x < Me) \approx P(x > Me)$. Модою Mo ряду

варіант називається варіанта, якій відповідає найбільша частота, тобто мода - це

найбільш ймовірне значення ВВ, для якого ймовірність приймає максимальне

значення. Для дискретної ВВ медіана може попадати в точку розриву графіка

функції $F(x)$ (3.2.1б). Якщо значення функції 0.5 попаде в горизонтальну ділянку

графіка функції $F(x)$, то за медіану беруть серединне значення x_M горизонтальної ділянки (рис.3.2.1в).

Центрована, нормована і зведена випадкові величини. За кожною ВВ визначають ще три величини – центровану Y , нормовану V і зведену U .

Центрована ВВ Y - це різниця між ВВ X та її математичним сподіванням

$Y = X - M[X]$. Математичне сподівання центрованої ВВ $M[Y] = 0$, а дисперсія

$D[X] = D[Y]$. Функція розподілу $F_Y(x) = F(x + M[X])$, а її густина

$f_Y(x) = f(x + M[X])$. Нормована ВВ U - це відношення даної ВВ X до її СКВ:

$V = \frac{X}{\sigma}$. Математичне сподівання і дисперсія нормованої ВВ дорівнюють:

$M[V] = \frac{M[X]}{\sigma}$ і $D[V] = 1$. Функція розподілу нормованої ВВ дорівнює $F_V(x) = F(\sigma x)$, а

її густина $f_V(x) = \sigma f(\sigma x)$. Зведена ВВ U - це центрована і нормована ВВ:

$U = \frac{X - M[X]}{\sigma}$. Для зведеної ВВ: $M[U] = 0$, $D[U] = 1$, $F_U(x) = F(\sigma x + M[X])$,

$f_U(x) = \sigma f(\sigma x + M[X])$.

Коефіцієнт варіації. Коливання окремих варіант характеризують розмахом R та коефіцієнтом ν варіації. Розмах варіації визначається як різниця між найбільшою та найменшою варіантами розподілу: $R = x_{\max} - x_{\min}$. Коефіцієнт варіації обчислюється як відношення $Var = \frac{\sigma}{M[X]}$, використовується при $M[X] > 0$, характеризує розкид у відносних одиницях.

Моменти випадкових величин. Розкид випадкових похибок вимірювань навколо істинного значення результатів вимірювання характеризує функція дисперсії. Чебишев показав, що математичне очікування і дисперсія розподілу ВВ є частковими випадками інших числових характеристик – моментів розподілу. В багатьох задачах замість визначення функції розподілу розглядають початковий і центральний моменти, означення яких нагадують означення механічних моментів у фізиці. Початковим моментом κ -го порядку ВВ X називається математичне сподівання κ -ого степеня ВВ X , яке обчислюється за формулою

$m_k = M[X^\kappa] = \sum_{i=1}^n x_i^\kappa p_i$. Центральним моментом κ -го порядку ВВ X називається

величина μ_k , яка обчислюється як $\mu_k = M[(X - M[X])^\kappa] = \sum_{i=1}^n (x_i - M[X])^\kappa p_i$. Перші три

з них дорівнюють: $\mu_1 = M[(X - M[X])^1] = 0$; $\mu_2 = M[(X - M[X])^2] = M[X^2] - (M[X])^2 = M[X^2] - m_1^2$; $\mu_3 = M[(X - M[X])^3] = M[X^3] - 3M[X] \cdot M[X^2] + 2m_1^2$.

Асиметрія і ексцес закону розподілу ВВ. Третій центральний момент характеризує асиметрію закону розподілу ВВ. Якщо $\mu_3 = 0$, то ВВ X розподілена

симетрично відносно її математичного сподівання $M[X]$. Оскільки вимірність μ_3 співпадає із відповідною для X^3 , то вводять безвимірну характеристику $As = \frac{\mu_3}{\sigma^3}$ - асиметрією закону розподілу ВВ. Якщо $As=0$, то крива розподілу симетрична; при $As<0$ - більш пологий спуск справа відносно максимум, а при $As>0$ - зліва. На практиці визначають знак асиметрії за розташуванням кривої розподілу відносно моди (точки максимуму функції густини розподілу ймовірностей): якщо «плавна» частина кривої розташована правіше моди, то асиметрія додатна, а якщо лівіше – від'ємна.

Центральний момент μ_4 характеризує вид плоскої вершини (гостроту) щільності ймовірності $f(x)$. Крутизну, тобто ділянку з швидшим чи повільнішим ходом кривої розподілу в порівнянні з нормальним, визначають за ексцесом. Характеристика, що обчислюється як: $Es = \frac{\mu_4}{\sigma^4} - 3$, називається ексцесом. Від правої частини віднімається 3 тому, що для нормального розподілу виконується умова $\frac{\mu_4}{\sigma^4} = 3$ і ексцес $Es = 0$. Для $Es>0$ максимум функції розподілу щільності перевищуватиме відповідний для $Es = 0$ і буде нижчий у випадку $Es<0$. Якщо для деякого розподілу ексцес не дорівнює нулеві, то цей розподіл відрізняється від нормального. Для додатного значення ексцесу, графік матиме більш «гостру» вершину, тоді як для від'ємного значення, вершина буде більш плоска.

Квантилі розподілів. Під час розв'язування задач та при встановленні довірчих меж для параметрів розподілу ВВ часто використовують поняття про «квантиль порядку p » (позначається як x_p), де $0<p<1$. Квантиль порядку p - значення ВВ, для якого функція розподілу приймає значення p або має стрибок від значення меншого p до значення більшого p . Якщо ця умова виконується для всіх значень x , то розглядають «квантиль порядку p ». Для неперервної функції, переважно існує один квантиль x_p порядку p , причому: $F(x_p) = p$.

Квантилі можуть визначатись неоднозначно. Для деяких p рівняння $F(x_p) = p$ може мати декілька розв'язків, для деяких – жодного. Це означає, що

для відповідної ВВ деякі квантилі визначені неоднозначно, а деякі квантилі не існують. Для дискретних розподілів, переважно не існує x_p , що задовольняє умову $F(x_p) = p$. Деякі квантилі мають свої назви: медіана – квантиль рівня 0.5; нижня квантиль – квантиль рівня 0.25; верхня квантиль – квантиль рівня 0.75; децили – квантилі рівнів 0.1, 0.2, ..., 0.9; проценти – квантилі рівнів 0.01, 0.02, ..., 0.99. Для тих розподілів, для яких в MathCAD є вбудовані функції розподілів і їх густини, визначені і вбудовані функції обчислень квантилів.

Статистичні характеристики вибірок.

Вибіркове середнє або її математичне очікування. Нехай деяка величина X в процесі вимірювань приймає ряд значень: $x_1, \dots, x_n$. Нехай вимірювані значення мають нормальний розподіл з математичним очікуванням $M[x]$, що співпадає з істинним значенням, та дисперсією σ^2 . Ймовірність того, що всі значення $x_1, \dots, x_n$ попадуть в безмежно малий інтервал $[x - \delta/2, x + \delta/2]$ за теоремою перемноження ймовірностей дорівнює добутку ймовірностей того, що кожне виміряне значення

попаде в інтервал:
$$P(X, \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{k=1}^n (x_k - X)^2\right).$$

Встановимо, при якому зв'язку X, σ^2 із значеннями $x_1, \dots, x_n$ правдоподібність максимальна, тобто функцію треба дослідити на екстремум. Для цього зручно досліджувати на екстремум не саму функцію, а логарифм:

$$L = \ln\left(\frac{P}{\delta^n}\right) = -\frac{n}{2} \ln(2\pi) - n \ln \sigma - \frac{\sum_{k=1}^n (x_k - X)^2}{2\sigma^2}.$$

При фіксованому σ максимум L

наступає при $\frac{\partial L}{\partial X} = 0$, тобто $\sum_{k=1}^n (x_k - X)^2 = 0$, звідки $X = \frac{1}{n} \sum_{k=1}^n x_k = \langle x \rangle$. Отже, вибіркове середнє значення є максимально правдоподібною оцінкою істинного значення вимірюваної величини.

Вибіркове середнє квадратичне. Якщо фіксованим є аргумент X , то

максим функції L настає при $\frac{\partial L}{\partial \sigma} = 0$, звідки $\sigma = \sqrt{\frac{\sum_{k=1}^n (x_k - X)^2}{n}}$. Отже,

максимальна правдоподібна оцінка стандартного квадратичного відхилення дорівнює вибірковому середньому квадратичному відхиленню вимірюваних

значень від істинного значення. Так як в процесі вимірювань істинне значення невідоме, то одержана формула не годиться для розрахунку похибки. Виразимо σ

через $\langle x \rangle$:
$$\frac{\sum_{k=1}^n (x_k - X)^2}{n} = \frac{\sum_{k=1}^n (x_k - \langle x \rangle + \langle x \rangle - X)^2}{n} = \frac{\sum_{k=1}^n (x_k - X)^2}{n} + 2(\langle x \rangle - X) \frac{\sum_{k=1}^n (x_k - X)}{n} + (\langle x \rangle - X)^2.$$

В цьому виразі другий доданок рівний нулю.

Розглянемо третій доданок:
$$(\langle x \rangle - X)^2 = \frac{\left(\sum_{k=1}^n (x_k - X) \right)^2}{n^2} = \frac{\sum_{k=1}^n (x_k - X)^2}{n^2} +$$

$\sum_{k=1}^n \sum_{m=1}^n (x_k - X)(x_m - X)$. Другий доданок при $n \rightarrow \infty$ прямує до нуля, тобто відхилення вимірюваних значень від істинного зустрічаються з різними знаками однаково

часто. Отже,
$$\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n (x_k - X)^2}{n} = \lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n (x_k - \langle x \rangle)^2}{n-1}.$$
 Величина
$$S_x = \sqrt{\frac{\sum_{k=1}^n (x_k - \langle x \rangle)^2}{n-1}}$$

називається вибірковою середнім квадратичним відхиленням одиночного спостереження, яке в границі дає максимально правдоподібну оцінку стандартного квадратичного відхилення: $\sigma = \lim_{n \rightarrow \infty} S_x$. Якщо $n \neq \infty$, то $\sigma \approx S_x$.

Вибіркове середнє є сумою N нормально розподілених випадкових величин, що мають однакову дисперсію. Воно представляє випадкову величину з дисперсією в n разів меншу, ніж дисперсія складуваних. Тому вибіркоче середнє квадратичне відхилення середнього $S_{\langle x \rangle}$ в $\sqrt{n}$ разів менше ніж S_x , тобто: $S_{\langle x \rangle} = \frac{S_x}{\sqrt{n}}$. екстремуму:

$$\frac{\partial L}{\partial X} = 0 \text{ і } \frac{1}{\sigma_1^2} (\langle x \rangle_1 - X) + \frac{1}{\sigma_2^2} (\langle x \rangle_2 - X) = 0. \text{ Тоді } \langle x \rangle = \left[\frac{\langle x \rangle_1}{\sigma_1^2} + \frac{\langle x \rangle_2}{\sigma_2^2} \right] \left\{ \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right\}^{-1}.$$
 Дисперсія

цієї оцінки дорівнює
$$\sigma^2 = \left(\frac{\partial \langle x \rangle}{\partial \langle x \rangle_1} \right)^2 \sigma_1^2 + \left(\frac{\partial \langle x \rangle}{\partial \langle x \rangle_2} \right)^2 \sigma_2^2, \text{ або } \sigma^2 = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^{-1}.$$

Довірчий інтервал для математичного очікування. В теорії ймовірності доводиться і це було зроблено вище, що найкращою оцінкою математичного очікування $M[x]$ є вибіркоче середнє значення $\langle x \rangle$, яке представляє собою випадкову нормально розподілену величину з густиною ймовірності $f\left(x | M, \frac{\sigma}{\sqrt{n}}\right)$.

Довірчим інтервалом для $M[x]$ називається інтервал $[M - \Delta x; M + \Delta x]$, в який з

ймовірністю $\alpha = \int_{M-\Delta x}^{M+\Delta x} f\left(x \mid M, \frac{\sigma}{\sqrt{n}}\right) dx$ попадає величина M . Величину α називають довірчою ймовірністю. За заданим значенням α можна завжди розрахувати ширину $2\Delta x$ довірчого інтервалу, якщо відомі значення M і σ (рис.3.2.2). При обробці результатів вимірювань M і σ невідомі, тому можуть бути обчислені лише їх оцінки $\langle x \rangle$ і $S_{\langle x \rangle}$. Тому задачу побудови довірчого інтервалу необхідно сформулювати по-іншому. Нехай задана довірча ймовірність α , з якою довірчий інтервал шириною $2\Delta x$ покриває невідоме математичне очікування $M[x]$. В цьому випадку межі довірчого інтервалу є випадковими. Для їх визначення треба із

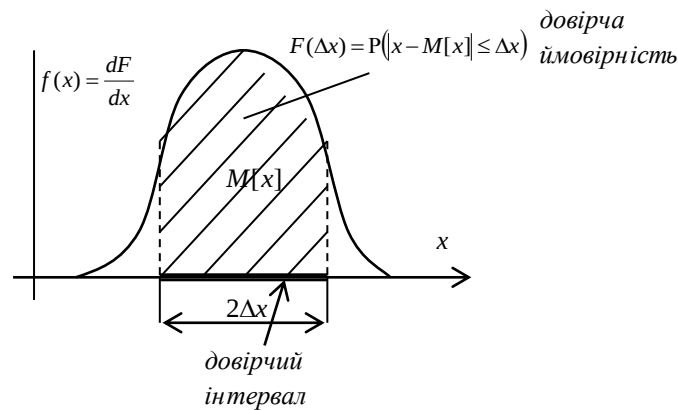


Рис.3.2.2

функції розподілу виключити невідомий параметр $M[x]$. Використаємо для цього

лінійне перетворення $\xi = \frac{x - \langle x \rangle}{\sigma_{\langle x \rangle}}$ (нормована випадкова величина) випадкової

величини $\langle x \rangle$, де $\sigma_{\langle x \rangle} = \frac{\sigma}{\sqrt{n}}$. Випадкова величина $\xi = \frac{x - \langle x \rangle}{\sigma_{\langle x \rangle}}$ розподілена по

нормальному закону з нульовим математичним очікуванням і одиничною

дисперсією. Тоді $\alpha = \int_{-\Delta x / \sigma_{\langle x \rangle}}^{\Delta x / \sigma_{\langle x \rangle}} f(\xi | 0,1) d\xi$. У багатьох випадках обробка результатів

вимірювань дисперсія σ , а отже і $\sigma_{\langle x \rangle}$ також невідомі. Тому при обробці

використовують їх оцінки S_x , $S_{\langle x \rangle}$ і $\frac{\langle x \rangle - M}{S_{\langle x \rangle}}$. Випадкова величина $\frac{\langle x \rangle - M}{S_{\langle x \rangle}}$ вже

розподілена не по нормальному закону, а по так званому закону Стюдента з

густиною ймовірності $f(x | 0,1, N)$, що залежить від числа вимірювань N величини

x , який в загальному розглядатиметься нижче.

Густина або щільність функції розподілу. В практиці замість функції розподілу часто використовують інший спосіб подання закону розподілу неперервної ВВ у вигляді густини ймовірностей $f(x) = \frac{dF}{dx}$. Тоді інтегральна функція розподілу дорівнює інтегралу $F(x) = \int_{-\infty}^x f(x)dx$.

Відома теорема: Ймовірність того, що неперервна ВВ X набуде значення з інтервалу $[a;b]$ дорівнює означеному інтегралу від щільності розподілу в межах від a до b : $P(a < X < b) = \int_a^b f(x)dx$.

Інтегральними числовими характеристиками розподілу неперервних ВВ є математичне сподівання $M[X] = \int_{-\infty}^{+\infty} x \cdot f(x) dx$ або $M[X] = \frac{\int_{-\infty}^{+\infty} x \cdot f(x) dx}{\int_{-\infty}^{+\infty} f(x) dx}$ та дисперсія

$D[X] = \langle x^2 \rangle - \langle x \rangle^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - \left(\int_{-\infty}^{\infty} x f(x) dx \right)^2 = \int_{-\infty}^{\infty} (x - M[X])^2 f(x) dx$. Прикладом неперервної ВВ може бути швидкість молекул ідеального газу g , яка приймає значення в інтервалі $(0; \infty)$. Густина ймовірностей розподілу цієї ВВ називається розподілом Максвелла. Завдання статистичної фізики – визначити відносну кількість частинок в цьому елементарному об'ємі $\frac{dn}{n}$.

В 1860 р. Максвелл вперше методами статистичної фізики обґрунтував закон розподілу частинок ідеального газу за швидкостями, що дозволило визначити відносну кількість частинок $\frac{dN}{N}$ в цьому елементарному об'ємі. Нагадаємо, що внаслідок хаотичного характеру молекулярного руху встановлюється рівноважний стан, в якому тиск газу пропорційний $\langle g^2 \rangle$. Величина $\sqrt{\langle g^2 \rangle}$ виражає макростан системи в стані термодинамічної рівноваги, хоч насправді ця величина є її мікропараметром.

З іншої сторони безперервний хаотичний рух ніяким чином не відображає «безладдя» ансамблю частинок. Внаслідок співударів змінюються як швидкості молекул, так і напрям їх рухів тому при цьому встановлюється стійкий розподіл

молекул за швидкостями та їх напрямом. Причому розподіл молекул за швидкостями встановлюється не довільний, а цілком природний, в якому в стані рівноваги всі напрямки швидкостей рівноправні між собою і характеристикою цього розподілу є функція розподілу за швидкостями, яку вперше встановив Максвелл.

Максвелл першим увів поняття ймовірності в фізиці. Розробляючи теорію ідеального газу, він задає не положення і швидкість частинки, а ймовірність знаходження її з визначеною швидкістю у визначеному об'ємі. За Максвеллом, нас цікавитиме не відносна кількість $\frac{dN}{N}$ молекул у газі з визначеними швидкостями, а та відносна їх кількість, значення швидкостей яких знаходяться в певному проміжку їх значень $d\vartheta$. Такий розподіл молекул за абсолютними швидкостями відображає функція $f(\vartheta) = \frac{1}{n} \frac{dn}{d\vartheta}$, де $f(\vartheta) = A^3 \exp\left(-\frac{1}{2}\alpha\vartheta^2\right)\vartheta^2$ при $0 < \vartheta < \infty$ і задовольняє умову нормування $\int_0^{\infty} f(\vartheta) d\vartheta = 1$. За своїм змістом умова нормування означає закон збереження кількості частинок в системі $\int dn = n$, а також, що ймовірність виявити молекулу з довільним значенням швидкості, в тому числі з значенням абсолютної швидкості $\vartheta = 0$, є достовірною подією і повна ймовірність дорівнює одиниці.

3.3. Основні моделі та закони розподілу дискретних випадкових величин

Рівномірно розподілена дискретна ВВ. Дискретна ВВ рівномірно розподілена, якщо події $X = x_i$ відбуваються з однаковими ймовірностями:

$$P(X = x_i) = p_i = \begin{cases} \frac{1}{n}, & i \in \{1, 2, 3, \dots, n\}; \\ 0, & i \notin \{1, 2, 3, \dots, n\}. \end{cases} \text{ Тут } [1, \dots, n] - \text{множина цілих чисел від 1 до } n.$$

Послідовність значень ВВ із рівномірним розподілом в інтервалі $[a; b]$

називається рівномірною випадковою послідовністю, міра якої визначається як

$$[a; b]: p(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b; \\ 0, & x \notin [a; b], \end{cases} \text{ а інтегральна функція } F(x) = \begin{cases} 0, & x < a \\ \frac{[x] - a}{b - a}, & a \leq x < b; \\ 1, & x \geq b, \end{cases} \text{ де } [x] -$$

найбільше ціле число $\leq x$, a, b - цілі числа, для яких $a \leq b$. Рівномірно розподілена послідовність ВВ символічно позначається як $Ud(n)$. Для рівномірно розподіленої ВВ в інтервалі $[a; b]$, $M[X] = \frac{a+b}{2}$, $D[X] = \frac{(b-a)^2}{12}$.

Випробування Бернуллі або бінарний розподіл. Часто проводиться не одне випробування, а багато, які називаються повторними, а їх сукупність – схемою повторних випробувань, або схемою Бернуллі. Якщо m - кількість «успіхів» у n дослідах Бернуллі з імовірністю успіху p , то m - випадкова величина, множина значень якої збігається з множиною $\{0, 1, 2, \dots, n\}$, а розподіл має вигляд $P(X_n = m) = C_n^m p^m (1-p)^{n-m} = C_n^m p^m q^{n-m}$, ($m = 1, 2, 3, \dots, n$). Кажуть, що ВВ X_n підлягає розподілу Бернуллі з параметром p ($0 < p < 1$), якщо вона може набувати двох значень 0 та 1, і $P(m=1) = p = 1 - P(m=0)$. Щільність та інтегральна функція

мають вигляд: $p(x) = \begin{cases} 1-p, & x=0; \\ p, & x=1; \\ 0, & \text{otherwise.} \end{cases}$ і $F(x) = \begin{cases} 0, & x < 0; \\ 1-p, & 0 \leq x < 1; \\ 1, & x \geq 1. \end{cases}$ Символічно розподіл

Бернуллі записується так: $L(X) = Bi[1; p]$. Його використовують для генерування інших дискретних ВВ. Математичне сподівання розподілу Бернуллі і дисперсія дорівнюють: p і $p(1-p)$.

Біноміальний розподіл ймовірностей дискретних ВВ. Якщо дискретна ВВ X при n випробуваннях може набувати значення $0, 1, 2, 3, \dots, n$ разів з частотами $0, \frac{1}{n}, \frac{2}{n}, \dots, 1$ і кожному випробуванню протиставляється ВВ m із множини

$m = 0, 1, 2, 3, \dots, n$, то ймовірність m того, що в даному випробуванні настане успіх при n випробувань також ВВ і обчислюється як $P_n(m) = C_n^m p^m (1-p)^{n-m}$.

Ймовірнісна міра та функція біноміального розподілу мають вигляд:

$$p(x) = \begin{cases} \frac{n! p^m (1-p)^{n-m}}{m!(n-m)!}, & x \in [0; n], \\ 0, & x \notin [0; n], \end{cases} \text{ і } F(x) = \begin{cases} 0, & x < 0, \\ \sum_0^m \frac{n! p^m (1-p)^{n-m}}{m!(n-m)!}, & x \in [0; n], \\ 1, & x > n, \end{cases}$$

$0 \leq m \leq n$, 0 і n - цілі числа,

для яких $0 \leq n$. Множина значень ймовірностей $\{P_n(m)\}$ називається біноміальним розподілом. Для нього $M[m] = np$, $D[m] = np(1-p)$ і справджуються два

співвідношення: $\sum_{m=0}^n m \cdot P_n(m) = np$ і $\sum_{m=0}^n m^2 \cdot P_n(m) = n^2 p^2 + np(1-p)$. Символьно

біноміальний розподіл позначається як $Bi[n; \theta]$, $\Theta = (\theta: 0 < \theta < 1)$. Його граничні форми мають вигляд: **1.** Якщо p фіксоване, то розподіл прямує до відомого в статистиці нормального розподілу при $n \rightarrow \infty$. **2.** У випадку фіксованого np , розподіл прямує до розподілу Пуассона (коли $n \rightarrow \infty$). Цьому закону підпорядковані кількість зіпсованих деталей в партії розміром n , кількість успішних реалізацій в незалежних випробуваннях, коли результатом кожного з них успіх або неуспіх.

Від’ємний біноміальний розподіл ймовірностей дискретної ВВ .

Від’ємна біноміальна ВВ X визначається через схему Бернуллі як кількість всіх «неуспіхів» до k -го «успіху». Для неї закон розподілу має вигляд:

$$P(X = j) = C_{k+j-1}^j p^k (1-p)^j \text{ з параметрами } M[X] = \frac{(1-p)k}{p}, D[X] = \frac{(1-p)k}{p^2}.$$

Геометричний розподіл дискретної ВВ. Ймовірнісна міра та функція

геометричного розподілу мають вигляд: $p(x) = \begin{cases} p(1-p)^x, & x \geq 0, \\ 0, & x < 0, \end{cases}$ і

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - (1-p)^{[x]+1}, & x \geq 0, \end{cases} \text{ де } [x] - \text{найбільше ціле число } \leq x, \text{ а параметри розподілу}$$

$$M[X] = \frac{1-p}{p}, D[X] = \frac{1-p}{p^2}. \text{ Геометричний розподіл ще називають розподілом}$$

Фаррі, його використовують у випадку, коли треба встановити кількість деталей, що піддаються контроль до того, як виявлений перший пошкоджений виріб.

Розподіл Паскаля. ВВ з цим розподілом визначається як число всіх випробувань в схемі Бернуллі до появи k -го «успіху». Для неї закон розподілу має вигляд: $P(Y = j) = C_{j-1}^{k-1} p^k (1-p)^{j-k}$, $j = k, k+1, \dots$, а його параметри $M[X] = \frac{Y-X}{p}$ і

$$D[X] = \frac{(Y-X)(1-p)}{p^2}, \text{ де ВВ } X \text{ розподілена від’ємно біноміально.}$$

Розподіл Пуассона дискретної ВВ . Якщо число випробувань Бернуллі велике (декілька сотень і більше), а ймовірність p успішної реалізації події при одиничному випробуванні мала (близька до нуля), то добрим наближенням є розподіл Пуассона, який називають законом рідкісних явищ. Пуассон задався

питанням, що буде коли збільшувати не n , залишаючи при цьому незмінними не m і p , а m і np , тобто яким буде розподіл $P_n(m)$ для великих n , однак малих значень ймовірності p і довів граничу теорему про те, що розподіл за

сформульованої умови прямує до вигляду $P_n(m) = \frac{(np)^m}{m!} \exp(-np)$ з параметрами:

$$M[m] = \langle m \rangle = np, \quad D[m] = \lim_{n \rightarrow \infty} np(1-p) = np, \quad Var = \frac{1}{\sqrt{np}}, \quad \text{асиметрії} - As = \frac{1}{\sqrt{np}}.$$

Ймовірнісна міра і інтегральна функція розподілу Пуассона з параметром

$$\lambda \text{ мають вигляд: } f(m) = \begin{cases} \frac{\lambda^m \exp(-\lambda)}{m!}, & 0 \leq m; \\ 0, & m < 0, \end{cases} \quad \text{і } F(m) = \begin{cases} \exp(-\lambda) \sum_{k=0}^{[m]} \frac{\lambda^k}{k!}, & 0 \leq m, \\ 0, & m < 0. \end{cases} \text{ звідки}$$

математичне сподівання і дисперсія дорівнюють: $\lambda = np$, $D = np$.

Розподіл Пуассона визначений на множині натуральних чисел, описує розподіл нечастих випадкових подій або граничний випадок (при $n \rightarrow \infty$, $np = \text{const}$) біноміального закону. Закону Пуассона підлягають розподіл кількості розривів нитки в прядильному цеху, кількість появи бракованих деталей на конвеєрній лінії або кількість розмов, що фіксуються на телефонній станції протягом деякого часу. Розподіл Пуассона використовують для моделювання шумів оптичних зображень, які можуть бути спричинені, наприклад, зернистістю структури фотоплівки чи фотоелемента, яку моделюють за допомогою кліткового графу у вигляді шахматної дошки із n квадратних елементів. Якщо число прозорих клітинок m , то ймовірність влучення в нього розсіяного фотона

дорівнює $p = \frac{m}{n}$. Якщо площа екрану L^2 , а клітинки - l^2 , то $n = \frac{L^2}{l^2}$ і ймовірність

$p = \frac{l^2}{L^2}$. Ймовірність того, що серед n прозорих і непрозорих клітинок буде m

прозорих описується функцією $P_n(m) \approx \frac{(\bar{m})^m \exp(-\bar{m})}{m!}$, де $\bar{m} = pn$ [3.3].

Дискретний потік Пуассона. Послідовність результатів n випробувань, в кожному з яких ймовірність успіху дорівнює $p_m = \frac{m}{n}$, утворює дискретний набір ВВ X . Вона розподілена за законом Пуассона, якщо приймає значення

0,1,2,...,m,... із ймовірністю $P_n(X=m) = C_n^m p^m (1-p)^{n-m} = \lim_{n \rightarrow \infty} \frac{n!}{k!(n-m)!} p^m (1-p)^{n-m}$, де $m = 0,1,2,3,...,n$. Такий потік називається пуассоновим.

3.4. Основні моделі, числові характеристики та закони розподілу неперервних випадкових величин

Випадкова величина X називається неперервною, якщо вона не може приймати довільне значення із деякого інтервалу $[a;b]$. Допускається випадок $a = -\infty$ і $b = +\infty$. Неперервна ВВ X визначається заданням інтервалу $[a;b]$, що містить можливі значення цієї величини і функція $p(x)$, яка називається густиною ймовірностей ВВ X (або густиною розподілу X) [3.2-3.3].

На відміну від дискретної ВВ, неперервна має безмежну множину можливих значень із певного проміжку і ймовірність появи довільного конкретного із них дорівнює нулеві. Неперервна або континуальна (лат. *continuum* - неперервний) ВВ набуває неперервних значень у заданому інтервалі, тому для неї використовується інший спосіб встановлення відповідності між значеннями ВВ та ймовірностями їх появи у вигляді функції розподілу ймовірностей $F(x)$, яка має такі властивості: $F(-\infty) = 0$ і $F(+\infty) = 1$. Отже, якщо ВВ визначена лише в області додатних значень, її функція розподілу дорівнює нулеві на всьому проміжку від мінус безмежності до нуля. Ймовірність того, що ВВ прийме значення із деякого інтервалу $[a;b]$, визначається через функцію розподілу $P(a < X < b) = F(b) - F(a)$.

Густина або щільність функції розподілу. В практиці замість функції розподілу часто використовують інший спосіб подання закону розподілу неперервної ВВ у вигляді густини ймовірностей $f(x) = \frac{dF}{dx}$. Тоді інтегральна функція розподілу дорівнює інтегралу $F(x) = \int_{-\infty}^x f(x) dx$.

Теорема. *Ймовірність того, що неперервна ВВ X набуде значення з інтервалу $[a;b]$ дорівнює означеному інтегралу від щільності розподілу в межах від a до b : $P(a < X < b) = \int_a^b f(x) dx$.* Прикладом неперервної ВВ може бути

швидкість молекул газу $\mathcal{U}$, яка приймає значення в інтервалі $(0; \infty)$. Густина ймовірностей розподілу цієї ВВ називається розподілом Максвелла.

Інтегральні числові характеристики розподілу неперервних ВВ.

Математичним сподіванням неперервної ВВ X називається число

$$M[X] = \int_{-\infty}^{+\infty} x \cdot f(x) dx \quad \text{або} \quad M[X] = \frac{\int_{-\infty}^{+\infty} x \cdot f(x) dx}{\int_{-\infty}^{+\infty} f(x) dx}, \quad \text{а дисперсією число}$$

$$D[X] = \langle x^2 \rangle - \langle x \rangle^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - \left(\int_{-\infty}^{\infty} x f(x) dx \right)^2 = \int_{-\infty}^{\infty} (x - M[X])^2 f(x) dx.$$

Числові характеристики функції розподілу. Модою неперервної ВВ X ($Mo(X)$) є таке її можливе значення x_m , при якому існує максимум густини розподілу або таке значення ВВ, для якого ймовірність $P(X = x_m)$ має найбільше значення. Залежно від кількості вершин (максимумів), розподіли поділяються на одномодальні (унімодальні) та багатомодальні (полімодальні). Для ВВ, які підпорядковані рівномірному розподілу, моди відсутні.

Медіаною неперервної ВВ X ($Me(X)$) називається таке її можливе значення x_p , для якого виконується рівність $P(X < x_p) = P(X > x_p) = \frac{1}{2}$, тобто ймовірності того, що величина X буде меншою x_p або більшою x_p , однакові. Медіані відповідає найвища точка на кривій густини розподілу.

Моменти, як інтегральні характеристики функцій розподілу неперервних ВВ. Початковий момент k -го порядку ВВ X - це інтеграл

$$m_k = \int_{-\infty}^{+\infty} x^k f(x) dx = M[X^k] = \langle X^k \rangle, \quad \text{що дорівнює математичному сподіванню ступеня}$$

X^k . Для прикладної статистики інтерес представляють початкові моменти першого і другого порядку. Центральний момент k -го порядку неперервної ВВ

X - це величина μ_k , що дорівнює інтегралу $\mu_k = \int_{-\infty}^{+\infty} (x - M[X])^k f(x) dx$. Перший

центральний момент дорівнює нулеві $\mu_1 = \int_{-\infty}^{+\infty} (x - m_1) f(x) dx = \int_{-\infty}^{+\infty} x f(x) dx - m_1 \int_{-\infty}^{+\infty} f(x) dx$

$= m_1 - m_1 = 0$, а другий центральний момент співпадає з дисперсією розподілу

$$D[X] \text{ ВВ: } \mu_2 = \int_{-\infty}^{+\infty} (x - M[X])^2 f(x) dx = \langle (X - \langle X \rangle)^2 \rangle = D[X].$$

Стандартний розподіл. Якщо ввести заміну $Y = aX + b$ (a, b - деякі числа), то

$$M[Y] = aM[X] + b, D[Y] = \sigma^2 D[X], F_Y(x) = F\left(\frac{x-b}{a}\right), f_Y(x) = \frac{1}{a} f\left(\frac{x-b}{a}\right). \text{ Отже, з кожною ВВ}$$

X можна пов'язати безліч ВВ Y , заданих формулою $Y = aX + b$ при різних значеннях a, b . Ця множина випадкових чисел називається масштабно-зсунутою множиною, що породжена величиною X . Функцією розподілу масштабно-зсунутої множини ВВ є масштабно-зсунутий набір розподілів, що породжений функцією розподілу $F(x)$. Замість $Y = aX + b$ часто використовують запис

$$Y = \frac{X - c}{d}, \text{ де } d = \frac{1}{a} > 0, c = -\frac{b}{a}. \text{ Число } c \text{ називають параметром зсуву, а число } d -$$

параметром масштабу. Для масштабно-зсунутого набору величин $Y = \frac{X - c}{d}$

розподіл X називають стандартним. Відомі стандартні розподіли Вейбулла-Гніденко, стандартний гама-розподіл та інші [3.4]. Використовують й інші перетворення ВВ, наприклад, $Y = \lg X$. Тоді $F_Y(x) = P(\lg X \leq x) = P(X \leq 10^x) = F(10^x)$ пов'язує розподіли X і Y .

Рівномірний розподіл ВВ на інтервалі $[a; b]$ ($R_{[a;b]}$). Неперервна ВВ X

називається рівномірно розподіленою на інтервалі $[a; b]$, якщо міра або густина

ймовірностей має вигляд: $p(x) = f(x) = \begin{cases} 1, & a \leq x \leq b, \\ 0, & x < a, x > b \end{cases}$, а інтегральна функція

розподілу $F(x) = \begin{cases} 1, & x \geq b, \\ \frac{x-a}{b-a}, & a \leq x < b, \\ 0, & x < a. \end{cases}$ Математичне сподівання і дисперсія дорівнюють:

$$M[X] = \frac{b+a}{2}; D[X] = \frac{(b-a)^2}{12}, \text{ а інші параметри - } Me = \frac{a+b}{2}, Var = \frac{b-a}{\sqrt{3}(b+a)}, As = 0. \text{ В}$$

інтервалі $[0; 1]$ $M[X] = \frac{1}{2}; D[X] = \frac{1}{12}$. Рівномірному розподілу підлягають похибки

грубих вимірювань, які допускають при заокругленні показів вимірювальних приладів до цілих поділок або до цілого числа часток поділки приладу.

Допущену при цьому похибку можна розглядати як ВВ X , яка з сталою густиною розподілу ймовірностей може набувати будь-яких значень між сусідніми поділками шкали.

Експоненційний розподіл. Неперервна ВВ X розподілена за показниковим законом з параметром $\lambda > 0$, якщо її густина ймовірностей має вигляд: $f(x) = \begin{cases} 0, & x < 0; \\ \lambda \exp(-\lambda x), & x \geq 0, \end{cases}$ де λ - стала розподілу. Інтегральна функція показникового розподілу має вигляд:

$F(x) = \begin{cases} 0, & x < 0; \\ 1 - \exp(-\lambda x), & x \geq 0. \end{cases}$ Ймовірність влучення ВВ X в заданий інтервал $[\alpha; \beta]$ дорівнює $P(\alpha < X < \beta) = \exp(-\lambda \alpha) - \exp(-\lambda \beta)$, а параметри $M[X] = \lambda^{-1}$ і $D[X] = \lambda^{-2}$. Експоненційному закону підлягають тривалість телефонних розмов, термін справної роботи електронних компонентів, час надходження замовлень на підприємства обслуговування, час прибуття літаків до аеропорту тощо.

Розподіл Гаусса. Встановлено, що коли в моделі Бернуллі кількість вимірювань велика $n \rightarrow \infty$, то розподіл описується наближеною формула Лапласа:

$P_n(m) \approx \frac{1}{\sqrt{npq}} f(x)$, $x = \frac{m - np}{\sqrt{npq}}$, $q = 1 - p$, функція $f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$ - це щільність розподілу, яка називається функцією Гаусса. Вона непарна, приймає лише додатні значення і має єдиний максимум з координатою $x = 0$: $f(x) > 0$,

$\max f(0) = \frac{1}{\sqrt{2\pi}}$. З умови нормування маємо $\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx = \sqrt{2\pi}$.

Нормальний розподіл. Це найважливіший в статистиці закон розподілу ймовірностей. Нормальною ВВ називається ВВ $X \in N_{m,\sigma}$, яка визначена на всій числовій вісі і має густину ймовірностей $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$, $-\infty \leq x < +\infty$ з математичним сподіванням m і дисперсією σ^2 . Принявши до уваги зведену випадкову змінну $U = \frac{X-m}{\sigma}$, функція розподілу набирає вигляду

$\varphi(u) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$. Заміна параметра не впливає на форму кривої густини ймовірностей $\varphi(u)$. Параметри нормально розподіленої ВВ

дорівнюють: $M[U] = 0$, $D[U] = 1$. Інтегральна функція розподілу має вигляд

$$\Phi(u) = F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{u^2}{2}\right) du - \text{це інтегральна функція Лапласа, яка табульована.}$$

При $m = 0$ і $\sigma = 1$ нормальний розподіл перетворюється в стандартний $X \in N[0;1]$.

За суттю інтегральна функція Лапласа виражає ймовірність того, що ВВ приймає значення від $-\infty$ до x (рис.3.4.1а). Оскільки відповідний інтеграл не береться в загальному вигляді, то значення інтегральної функції Лапласа задаються таблично для параметрів розподілу $M[X] = 0$ і $D[X] = 1$. Вона симетрична відносно точки з координатами ($x = 0$): $\Phi(-x) = 1 - \Phi(x)$, для неї має

$$\text{місце перетворення } \Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{u^2}{2}\right) du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \exp\left(-\frac{u^2}{2}\right) du + \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{u^2}{2}\right) du =$$

$\frac{1}{\sqrt{2\pi}} \sqrt{\frac{\pi}{2}} + \Phi_0(u) = \frac{1}{2} + \Phi_0(u)$, де функція $\Phi_0(u)$ табулюється. В літературі можна зустріти позначення $\text{erf}(x) = \Phi(x)$, де $\text{erf}(x)$ - функція похибок. Обернена до неї позначається як $\text{erfc}(x)$. За функцію похибок ще приймають інтеграл

$$\text{erf}(x) = \frac{2}{\sqrt{2\pi}} \int_0^x \exp(-u^2) du.$$

Правило трьох сигм. Яки б не були m і σ , значення інтегралу наближено

дорівнює $\int_{m-3\sigma}^{m+3\sigma} p(x) dx = P(m - 3\sigma < X < m + 3\sigma) \cong 0.997$. Однак величина 0.997 настільки мало

відрізняється від одиниці, що часто записаний інтеграл інтерпретують так: *при одному випробуванні практично неможливо отримати значення X , що відрізняється від $M[X]$ більше, ніж на 3σ .*

Доведено, що при нормальному розподілі ймовірностей, який широко використовується в статистиці, теорії похибок і тощо, практично весь розподіл ймовірностей припадає на проміжок $(m - 3\sigma; m + 3\sigma)$, що відоме як правило трьох сигм. Суть цього правила полягає в тому, що коли в інтервалі $(m - \sigma; m + \sigma)$

зосереджена ймовірність $-\int_{-\sigma}^{+\sigma} f(x) dx = 0.6827$, то в $(m - 2\sigma; m + 2\sigma)$ - $\int_{-2\sigma}^{+2\sigma} f(x) dx = 0.9545$,

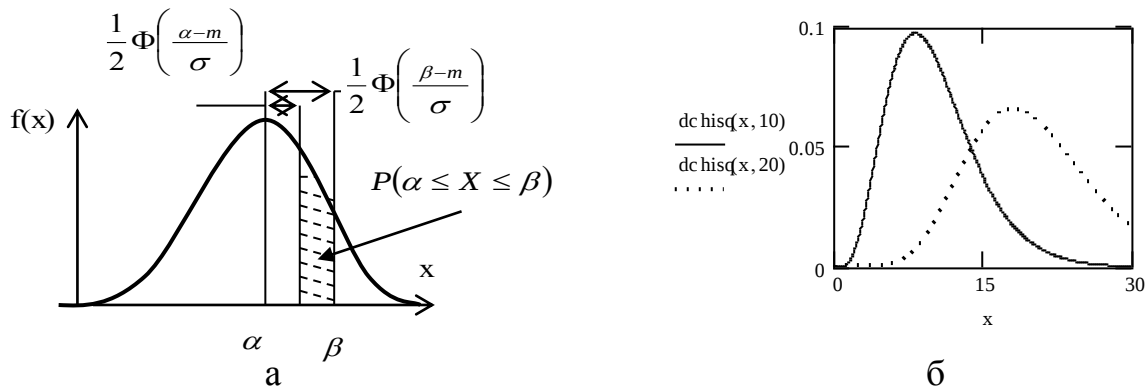


Рис.3.4.1

а в інтервалі $(m - 3\sigma; m + 3\sigma)$ - $\int_{-3\sigma}^{+3\sigma} f(x)dx = 0.9973$.

Розподіл Пірсона або χ^2 - (хі-квадрат) розподіл. У багатьох випадках виникає потреба обчислити не середнє значення вимірюваної величини, а розсіяння окремих вимірювань відносно цієї середньої величини, яке характеризується дисперсією розподілу σ^2 . З іншої сторони при зростанні кількості вимірювань, зменшується відмінність між СКВ і σ . Тому треба встановити інтервал із певною надійністю для значення σ при конкретному числі вимірювань n .

Нехай X_k ($k = 1, 2, \dots, n$) - незалежні ВВ, де $X_i \in N(x, \sigma, m)$ для всіх i . Тоді сума квадратів цих ВВ $\chi_n^2 = X_1^2 + X_2^2 + \dots + X_n^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - m)^2$ розподілена за законом *хі-квадрат* з $k = n$ ступенями вільності^{*)}. Густина розподілу для *хі-квадрат* (χ_n^2) для

$$\chi_n^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - m)^2 \text{ має вигляд: } f(x) = \begin{cases} \frac{\exp\left(-\frac{x}{2}\right) x^{\frac{n-2}{2}}}{2^{n/2} \Gamma(n/2)}, & x \geq 0, \text{ де } x = \chi_n^2 > 0, \\ 0, & x < 0. \end{cases}$$

$\Gamma(x) = \int_0^\infty \exp(-t) \cdot t^{x-1} dt$ для всіх $x > 0$ - означення функції розподілу за Ейлером або

$$\Gamma(x) = \lim_{n \rightarrow \infty} \frac{n! n^{x-1}}{x(x+1)(x+2)\dots(x+n-1)} - \text{за Гауссом, яка відноситься до класу спеціальних}$$

функцій. Графік χ^2 - розподілу показаний на рис.3.4.1,б, він асиметричний, причому її величина тим більша, чим менше число n . Для великих n *хі-квадрат*

^{*)} Число ступеней вільності визначають як $\nu - 1 - r$, де ν - число груп (інтервалів) вибірки, r - число параметрів розподілу. Так для нормального розподілу $r = 2$ (мат сподівання і СКВ)

розподіл характеризується одним параметром n і має числові параметри:

$$M[\chi_n^2] = n; \quad D[\chi_n^2] = 2n, \quad Me \cong n - \frac{2}{3}, \quad Mo = n - 2 \quad \text{для } n > 2, \quad Var = \sqrt{\frac{2}{n}}, \quad As = \frac{2^{1.5}}{\sqrt{n}}. \quad \text{Інтегральна}$$

функція χ^2 - розподілу з n степенями вільності дорівнює інтегралу від щільності в межах від $-\infty$ до χ_n^2 . $100\alpha_{\chi^2}$ - відсоткові точки χ^2 розподілу позначаються через

$$\chi_{n,\alpha_{\chi^2}}^2 \text{ і обчислюються як } \int_{\chi_n^2}^{\infty} f(\chi^2) d\chi^2 = P\left\{\chi_n^2 > \chi_{n,\alpha_{\chi^2}}^2\right\} = \alpha_{\chi_n^2}.$$

В практиці важливими є такі особливості χ^2 -розподілу. **1.** χ^2 - розподіл є частковим випадком більш загального гамма-розподілу. **2.** Розподіл ВВ, що дорівнює кореню квадратному з випадкової величини, що має χ^2 -розподіл з двома степенями вільності відомий як розподіл Релея. **3.** Із χ^2 -розподілом пов'язаний також відомий у класичній фізиці ідеального газу розподіл Максвелла, як χ^2 -розподіл з трьома степенями вільності. **4.** При зростанні числа степенів вільності χ^2 -розподіл наближається до нормального.

Розподіл Стюдента або t -розподіл. Нехай X і Y - незалежні ВВ такі, що X має χ_n^2 - розподіл, а Y - нормальний розподіл з нульовим середнім та одиничною дисперсією. ВВ вигляду $t_n = \frac{X}{\sqrt{Y/n}} = \frac{X}{\sqrt{\chi^2}} \sqrt{n}$ підлягає статистичному

t -розподілу Стюдента (t - статистиці) з n степенями вільності, має щільність ймовірностей $f(t) = \frac{\Gamma[(n+1)/2]}{\sqrt{\pi n} \Gamma(n/2)} \left[1 + \frac{t^2}{n}\right]^{-(n+1)/2}$, інтегральна функція якої

обчислюється інтегруванням щільності від $-\infty$ до χ_n^2 . Для t -розподілу, $100\alpha_t$ -

відсоткові точки позначаються як t_{n,α_t} : $\int_{t_{n,\alpha_t}}^{\infty} f(t) dt = P\{t_n > t_{n,\alpha_t}\} = \alpha_t$. Параметри

розподілу дорівнюють: $M[t_n] = 0$; $D[t_n] = \frac{n}{n-2}$, $n > 2$. Для $n = 1$ і $n = 2$ дисперсії не

існує. При зростанні n розподіл Стюдента наближається до нормального розподілу. Студентове відношення часто використовують для аналізу похибок

фізичних вимірювань. Розподіл Стюдента^{*)} застосовують для малих вибірок, коли нормальний розподіл непридатний для опрацювання результатів [3.5].

Розподіл Фішера-Снедекора або F – розподіл. Цей розподіл описує відношення двох незалежних ВВ Y_1 і Y_2 , що підлягають розподілу Пірсона χ_n^2 із ступенями вільності n_1 і n_2 : $F_{n_1, n_2} = \frac{Y_1}{Y_2} \cdot \frac{n_2}{n_1}$. Тоді ВВ F_{n_1, n_2} – це F – розподілена ВВ із n_1 і n_2 степенями вільності. Причому, за першу вибирають той, для якого оцінка дисперсії більша, тобто статистика відношень їх незміщених оцінок завжди більша одиниці. Щільність F – розподілу описується функцією

$$f(F) = \frac{\Gamma\left(\frac{n_1 + n_2}{2}\right) \left(\frac{n_1}{n_2}\right)^{n_1/2} F^{n_1/2-1}}{\Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right) [1 + (n_1 F / n_2)]^{(n_1 + n_2)/2}}. \text{ Інтегральна функція розподілу}$$

обчислюється інтегруванням $f(F)$ в межах $[-\infty, F_{n_1, n_2}]$. $100\alpha_F$ – відсоткові точки

F – розподілу позначаються $F_{n_1, n_2; \alpha_F}$: $\int_{F_{n_1, n_2; \alpha_F}}^{\infty} f(F) dF = P\{F_{n_1, n_2} > F_{n_1, n_2; \alpha_F}\} = \alpha_F$. Параметри

F – розподілу дорівнюють: $M[F_{n_1, n_2}] = \frac{n_2}{n_2 - 2}; n_2 > 2$ та $M[(F_{n_1, n_2} - M[F_{n_1, n_2}])^2]$

$= \sigma_F^2 = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)}; n_2 > 4$. Для усіх трьох розподілів побудовані ймовірнісні

таблиці.

Розподіл Вейбулла. Цьому закону підлягає розподіл час роботи до відмови систем з багатокомпонентними елементами. Для $x > 0$ густина ймовірностей розподілу Вейбулла $f(x) = nkx^{n-1} \exp(-kx^n)$, а інтегральна функція -

$F(x) = 1 - \exp(-kx^n)$, де k, n – довільні додатні числа. Математичне сподівання

$M[X] = \Gamma\left(1 + \frac{1}{n}\right) k^{-1/n}$, а дисперсія $D[X] = k^{2/n} \left[\Gamma\left(1 + \frac{2}{n}\right) - \Gamma^2\left(1 + \frac{1}{n}\right) \right]$, де $n = \left(\frac{\Gamma(1 + n^{-1})}{M[X]} \right)^n$.

Функція Вейбулла використовується для моделювання часу роботи досліджуваної системи до її відмови, який можна моделювати методом оберненої функції за формулою $t_i = \sqrt[n]{-\lambda^{-1} \ln(1-r)}$, де $r \in [0; 1]$. Характер кривої густини

^{*)} У.С.Госсетом будучи студентом підписував свої праці під псевдонімом «Студент».

розподілу залежить від значення n . Якщо $n > 1$, то крива має вершину, при $n \leq 1$ - асиметрична і наближена до вісі X . При $n = 1$ - розподіл Вейбулла перетворюється в експоненційний, а при $n = 2$ - розподіл Релея.

Розподіл Ерланга. Розподіл Ерланга описує час відновлення системи у випадку пошуку і усунення причин її відмови оператором. Функція розподілу і густина ймовірностей часу встановлення визначаються виразами:

$$F(t) = 1 - \left(1 + \frac{2t}{\tau}\right) \exp\left(-\frac{2t}{\tau}\right), \quad f(t) = \frac{4t}{\tau^2} \exp\left(-\frac{2t}{\tau}\right), \quad \text{де } \tau - \text{середній час відновлення. Для}$$

функції розподілу Ерланга не існує оберненої функції, тому для моделювання часу t_i використовувати метод оберненої функції не можна.

Логістичний розподіл. Логістичний розподіл задається функцією

$$F(x) = \frac{1}{1 + \exp(-x)}, \quad -\infty \leq x \leq \infty. \text{ Логістична функція була введена Верхюлстом,}$$

симетрична $F(-x) = F(x)$, має густину $f(x) = \frac{\exp(-x)}{(1 + \exp(-x))^2}$ (функція парна, що є

наслідком симетричності функції $F(x)$). Для логістичної ВВ X існують моменти всіх порядків, причому початкові моменти непарного порядку дорівнюють нулеві, а $M[X] = 0, D[X] = \frac{\pi^3}{3}$.

Розподіл Максвелла. Цей розподіл пов'язаний з χ^2 -квадрат розподілом.

Його густина дорівнює $f(x) = \frac{2x^2}{a^3 \sqrt{2\pi}} e^{-\frac{x^2}{2a^2}}$, де $x > 0, a > 0$, а інтегральна функція

розподілу $F(x) = \operatorname{erf}\left(\frac{x}{\sqrt{2}a}\right) - \frac{x\sqrt{2}}{a\sqrt{\pi}} \exp\left(-\frac{x^2}{2a^2}\right)$. Якщо ВВ X, Y, Z незалежні і мають

один і той же нормальний розподіл $N(0, \sigma^2)$, то модуль випадкового числа

$(X, Y, Z) \quad \Lambda = \sqrt{X^2 + Y^2 + Z^2}$ має розподіл Максвелла з параметром

$\sigma^2: f_{\Lambda}(\vartheta) = \frac{\vartheta^2}{\sigma^3} \sqrt{\frac{2}{\pi}} e^{-\frac{\vartheta^2}{2\sigma^2}}$, при $\vartheta > 0$ і $f(\vartheta) = 0$ при $\vartheta \leq 0$. Отже, розподіл Максвелла

здається одним параметром $\sigma^2 > 0$, а числові характеристики ВВ V

дорівнюють: $M[V] = 2\sqrt{\frac{\pi}{2}}\sigma$ і $D[V] = \left(3 - \frac{8}{\pi}\right)\sigma^2$. Символьно розподіл Максвелла з

параметром $\sigma^2 > 0$ позначають як $M(\sigma^2)$. В просторі швидкостей він аналогічний

розподілу Гаусса за похибками декартових координат. Тому коли система складається з N різних частинок, кожна з яких має m дозволених станів (комірок) незалежно від того, де при цьому перебувають решта частинок, то кількість станів для N частинок в m комітках дорівнює m^N . Оскільки всі ці стани (розташування по комітках) рівноймовірні, то ймовірність кожного стану дорівнює $p = m^{-N}$.

Розподіл Коші або Лоренца. Розподіл Коші відомий у фізиці ще як розподіл Лоренца [3.6]. Його переважно використовують для опису подій, що пов'язані з резонансними явищами. Густина ймовірностей розподілу Лоренца описується функцією $f(x; Me(X), \Delta) = \frac{1}{\pi} \frac{\Delta/2}{(x - x_0)^2 + (\Delta/2)^2}$, а інтегральна функція розподілу $F(x) = \frac{1}{2} + \frac{1}{\pi} \arctg \frac{2(x - x_0)}{\Delta}$. Для розподілу Лоренца математичне сподівання і дисперсію визначити не можливо, оскільки інтеграли розбіжні. Тому розподіл Лоренца характеризують медіаною $Me(X)$ та півшириною Δ , які визначаються з рівності $x - Me(X) = \pm 0.5\Delta$, коли густина ймовірностей досягає половині максимального значення.

Розподіл Бозе-Ейнштейна. Розподіл Бозе-Ейнштейна описує розподіл частинок у стані теплової рівноваги із середовищем, в якому вони поширюються і характеризує середнє число заповнень ними квантових станів

$\langle x \rangle : P(x; \langle x \rangle) = \frac{1}{(1 + \langle x \rangle)} \left(\frac{1}{1 + \langle x \rangle} \right)^x, x = 0, 1, 2, 3, \dots$ Для великих значень середнього числа заповнень $P(x; \langle x \rangle) = \frac{1}{\langle x \rangle} \exp\left(-\frac{x}{\langle x \rangle}\right)$. Функція розподілу $F(x)$, математичне сподівання і дисперсія розподілу рівні: $F(x) = \frac{1}{\langle x \rangle} \sum_{j=0}^x \left(1 + \frac{1}{\langle x \rangle}\right)^{-(1+j)}, M[X] = \langle x \rangle, \sigma^2 = \langle x \rangle^2 + \langle x \rangle$.

Частинки Бозе-Ейнштейна – квантово нерозрізніювані. Тому якщо система складається з N різних частинок, кожна з яких має m дозволених станів (комірок), то кожний з цих станів можна задати «числами заповнення» $N_1, N_2, \dots, N_n$, де N_j - кількість частинок в j -ій комірці. Кількість різних станів дорівнює C_{m+N-1}^{m-1} , тому ймовірність заповнення кожного з цих станів дорівнює

$\frac{1}{C_{m+N-1}^{m-1}}$. Якщо допустити, що жодна з комірок не є пустою, то кількість можливих

станів буде дорівнювати C_{N-1}^{m-1} , причому $N \geq m$.

Розподіл Фермі-Дірака. Система частинок Фермі-Дірака розглядається як система Бозе-Ейнштейна, в якій додатково діє принцип заборони Паулі (в одній комірці може перебувати не більше одної частинки). Оскільки і в цьому випадку частинки квантово нерозрізніювані, то стани для них характеризуються числами $N_j = 0, j = 1, \dots, m$, однак $N \leq m$. Розподіл таких частинок описується функцією Фермі, що належить до класу логістичних функцій (експоненційних сигмоїдів – гладких монотонних нелінійних S - подібних функцій $f(x) = \frac{1}{1 + \exp(-x)}$), які часто використовують для згладжування значень флуктуаційних параметрів та в нейронних мережах.

Розподіл Релея. Важливе значення має розподіл ВВ, який чисельно дорівнює квадратному кореню з випадкової величини, що має хі-квадрат розподіл з двома степенями вільності. Це розподіл Релея і застосовується переважно в електро - та радіотехніці .

Розподіл Релея використовується для розв'язування задач, пов'язаних із влучанням у двохвимірну мішень [3.7]. Так, якщо X і Y дві нормально розподілені ВВ $N(0, \sigma^2)$, то ВВ $R = \sqrt{X^2 + Y^2}$ підлягає розподілу Релея з параметром σ^2 , який символічно позначається як $Rl(\sigma^2)$. Для X , нормально розподіленої з параметрами $D[X]$ і $\langle x \rangle = 0$, що періодично змінюється з сталою частотою, однак випадковими амплітудою R та фазою, зміна амплітуди підлягає розподілу Релея $f(r) = \frac{r}{D[X]} \exp\left(-\frac{r^2}{2D[X]}\right)$, де $R > 0$. Інтегральна функція розподілу

має вигляд $F(r) = \int_0^r f(r) dr = 1 - \exp\left(-\frac{r^2}{2D[X]}\right)$, де $r \geq 0$. Параметри розподілу

дорівнюють: $M[R] = \langle r \rangle = \int_0^r r f(r) dr = \sqrt{\frac{\pi}{2} D[X]}$, середній квадрат $\langle R^2 \rangle = 2\sigma^2$ і

дисперсія $D[R] = \int_0^\infty (r - \langle r \rangle)^2 f(r) dr = \frac{4 + \pi}{2} D[X] = 0.429\sigma^2$.

Для моделювання законів розподілу безвідмовної роботи оптоелектронних елементів, часову функцію Релея використовують у вигляді $f(t) = 1 - \exp\left(-\frac{t^2}{2C^2}\right)$,

де C - параметр розподілу, що визначає час безвідмовної роботи $\tau_c = C\sqrt{\frac{\pi}{2}}$.

Моделювання часу t_i часу роботи елементів до відмови, методом оберненої функції здійснюється за формулою $t_i = \sqrt{-2C^2 \ln(1-r)}$, де $r \in [0;1]$. Розподіл Релея є граничним випадком розподілу Гаусса екстремальних значень у вигляді вузької смуги випадкового гауссового сигналу, коли його ширина прямує до нуля.

3.5. Моделювання розподілу випадкових величин та критерії перевірки згоди між дослідними і теоретичними розподілами

Найважливішою характеристикою ВВ є її закон розподілу. Знання його дозволяє прогнозувати наступне значення ВВ, оцінити ймовірність попадання значення ВВ в заданий інтервал та моделювати ВВ за допомогою стандартизованих програм ЕОМ.

Нехай дослідним шляхом одержана вибірка значень ВВ, причому кількість n вимірювань досить велика. Побудуємо гістограму, відклавши вздовж вісі

ординат $\frac{\Delta n}{n} \frac{1}{\Delta x}$, де $\Delta x = x_{\max} - x_{\min}$, а по всі абсцис - x , n – повна кількість

вимірювань, Δn - кількість результатів, що попали в інтервал $[x, x + \Delta x]$.

Частка $\frac{\Delta n}{n}$ виражає долю результатів, що попадають в інтервал $[x, x + \Delta x]$ і

характеризує ймовірність попадання в нього результату окремого вимірювання.

Відношення цієї величини до ширини інтервалу $\frac{\Delta n}{n} \frac{1}{\Delta x}$, характеризує густину

ймовірності. При $n \rightarrow \infty$ можна зробити заміну $\Delta t \rightarrow dt$ і перейти до іншої змінної $\Delta n \rightarrow dn$. Тоді замість ступінчатої функції розподілу матимемо плавну (рис.3.5.1):

$\rho(x) = \lim_{n \rightarrow \infty} \frac{\Delta n}{n} \frac{1}{\Delta x} = \frac{1}{n} \frac{dn}{dx}$ - закон розподілу досліджуваної випадкової величини.

Нормальний розподіл описується функцією Гаусса $\rho(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \langle x \rangle)^2}{2\sigma^2}\right)$.

Параметр $MC[X] = \langle x \rangle$ обчислюється за формулою $\langle x \rangle = \frac{1}{n} \sum_{i=1}^n x_i$, а дисперсія σ , як

параметр ширини нормального розподілу як $\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \langle x \rangle)^2}$. Бачимо, що

густина нормального розподілу має максимум при $x = \langle x \rangle$ і симетрична відносно цього значення. Найбільшу висоту гістограми $\rho_{\max}(x)$ можна пов'язати з

параметрам дисперсії σ як: $\rho_{\max}(t) = \frac{1}{\sqrt{2\pi}\sigma}$.

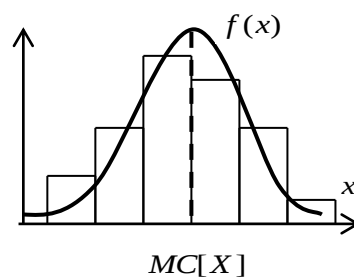


Рис.3.5.1

Для кількісної перевірки того, наскільки добре отримані результати описуються нормальним розподілом, використаємо співвідношення

$P_{x_1, x_2} = \int_{t_1}^{t_2} \rho(x) dx \approx \frac{n_{x_1, x_2}}{n}$. Тут записано, що імовірність P_{x_1, x_2} попадання результату

кожного вимірювання в інтервал $[x_1; x_2]$, з одної сторони, може бути обчислений як інтеграл функції розподілу в цих межах, а з іншої – визначена, як відносна кількість вимірювань, результати яких попали в цей інтервал. Доцільно використати значення імовірностей для найбільш вживаних в техніці вимірювань

$$x \in [\langle x \rangle - \sigma, \langle x \rangle + \sigma], \quad P_\sigma = 0.68;$$

$$\text{меж: } x \in [\langle x \rangle - 2\sigma, \langle x \rangle + 2\sigma], \quad P_\sigma = 0.95;$$

$$x \in [\langle x \rangle - 3\sigma, \langle x \rangle + 3\sigma], \quad P_\sigma = 0.97.$$

Отже, щоб ідентифікувати закон розподілу ВВ, треба перш за все побудувати гістограму частот і на підставі неї сформулювати гіпотезу стосовно закону розподілу, потім оцінити значення параметрів розподілу та перевірити відповідність за критерієм згоди.

Параметричний критерій χ^2 . Цей критерій ґрунтується на групуваннях. Статистична його характеристика – це сума стандартизованих квадратів відхилень емпіричних частот від теоретичних. Висновки за критерієм згоди χ^2 значною мірою залежать від кількості груп, частота яких має бути не менша 5. Окрім того, Критерій згоди χ^2 не враховує послідовність знаків відхилень частот (часток), за наявності серій знаків надійність висновку зменшуються.

В практиці критерій застосовується для встановлення емпіричного закону розподілу ВВ і виконується за алгоритмом: - будують гістограму – оцінку функції щільності розподілу ймовірностей. Спершу визначають мінімальне $x_{\min}$ і максимальне $x_{\max}$ значення елементів вибірки, після чого розбивають інтервал $x_{\max} - x_{\min}$ на m рівних підінтервалів довжиною кожний $h = \frac{x_{\max} - x_{\min}}{m}$; - висувають гіпотезу про згоду емпіричного закону із вибраним теоретичним розподілом; - обчислюють критерій $\chi^2 = \sum_{i=1}^m \frac{(n_i - Np_i)^2}{Np_i}$, де p_i - ймовірність попадання ВВ в i -інтервал (кількість підінтервалів m). Якщо $N \rightarrow \infty$, то закон розподілу хі-квадрат, як міра розбіжності, залежить тільки від кількості підінтервалів і наближається до закону розподілу χ^2 із $m-r-1$ ступенями вільності, де r - число параметрів теоретичного закону розподілу, де m - кількість інтервалів. Так для двох параметричного нормального розподілу кількість оцінюваних параметрів $r = 2$ (μ - математичне сподівання і σ - СКВ), тоді як для експоненційного розподілу лише один оцінюваний параметр розподілу λ - число, обернене математичному сподіванню, тому для нього $r = 1$. Два оцінювані параметри a і b ($r = 2$) характерні для рівномірного розподілу: $f(x) = \frac{1}{b-a}$.

Як впливає із цієї формули, основною ідеєю критерію χ^2 є вимірювання розбіжності між спостережуваною і очікуваною за теоретичним законом розподілу кількість попадань в i -й інтервал. Практична реалізація перевірки гіпотези здійснюється так. 1. Вибірка результатів $x_1, x_2, \dots, x_N$ групують в інтервальний варіаційний ряд: - визначається довжина інтервалу $\Delta x = \frac{x_{\max} - x_{\min}}{1 + 3.21 \cdot \lg N}$;

- обчислюється кількість інтервалів $m = \frac{x_{\max} - x_{\min}}{\Delta x}$; - розраховується кількість значень n_i ВВ X в кожному інтервалі i визначається її попадання в даний інтервал $p_i = \frac{n_i}{N}$; - будується варіаційний (статистичний) ряд у вигляді таблиці, полігону частот для дискретних значень ВВ чи гістограми для неперервних ВВ. Полігон розподілу бується за таким алгоритмом. У вибраних масштабах на вісі абсцис наноситься шкала для фактичних значень ВВ X , на вісі ординат – для частот $p^* = \frac{n}{N}$. Наносять точки A_i з координатами $\frac{n_i}{N}$ і точки $A_0(x_0, 0), A_1(x_1, p_1), \dots, A_i(x_i, p_i), \dots, A_m(x_m, p_m), A_{m+1}(x_{m+1}, p_{m+1})$ сполучають між собою ламаною лінією. Одержаний багатокутник – це полігон розподілу. Об'єм комп'ютерних реалізацій повинен бути більший 20. Якщо виявиться, що частота спостережень $n_i < 5$, то інтервал об'єднують з сусіднім так, щоб частота попадання значень ВВ в підінтервал $n_i > 5$. 2. За виглядом обвідної гістограми, висовують гіпотезу про вид розподілу. 3. Задають рівень довіри α . 4. Обчислюють теоретичну ймовірність $P_i = F(x_{i+1}) - F(x_i)$ попадання ВВ в кожний підінтервал. 5. Обчислюють міру розбіжності законів χ^2 . 6. Визначають кількість ступеней вільності $m - r - 1$. 7. За обчисленими значеннями χ^2 і $m - r - 1$, з таблиць визначають ймовірність $P\{\chi^2_{\text{теор}} \geq \xi\}$. Якщо вона перевищує рівень значимості $\xi > \alpha$, то вважається, що гіпотеза про вид розподілу відхиляється.

Непараметричні критерії

Критерій Колмогорова. Нехай дослідним шляхом отримана вибірка значень $x_1 \leq x_2 \leq \dots \leq x_n$. Спершу обчислюють значення статистики S^*

критерію за формулами: $S^* = \frac{6nD_n + 1}{6\sqrt{n}}$, $D_n = \max\{D_n^+, D_n^-\}$, $D_n^+ = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F(x_i, \theta) \right\}$,

$D_n^- = \max_{1 \leq i \leq n} \left\{ F(x_i, \theta) - \frac{i-1}{n} \right\}$. Потім знаходять значення $P\{S > S^*\} = \int_{S^*}^{\infty} g(s | H_0) ds = 1 - K(S^*)$,

$K(S) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 S^2}$. Якщо $P\{S > S^*\} > \alpha$, де α - рівень довіри, то гіпотеза H_0

приймається. В протилежному випадку гіпотеза H_0 відхиляється.

Можна обчислене значення статистики S^* із рівняння $\int_{S_\alpha}^{\infty} g(s | H_0) ds = \alpha$

порівняти з критичним S_α , де $g(s|H_0)$ – щільність ймовірностей розподілу статистики критерію при справедливості гіпотези H_0 . Критичне значення критерію S_α при заданому α можна взяти із таблиці Б.2 [3.18]. Гіпотеза про згоду відхиляється, якщо значення статистики попадає в критичну область, тобто при $S^* > S_\alpha$.

Критерій Смірнова. Формулюють гіпотезу H_0 . Випадковим методом відбирають вибірку об'ємом n , в якій значення сортують в напрямку $x_1 \leq x_2 \leq \dots \leq x_n$, після чого обчислюють значення статистики S^* критерію за формулами: $S^* = \frac{(6nD_n + 1)^2}{9n}$, $D_n = \max\{D_n^+, D_n^-\}$, $D_n^+ = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F(x_i, \theta) \right\}$, $D_n^- = \max_{1 \leq i \leq n} \left\{ F(x_i, \theta) - \frac{i-1}{n} \right\}$.

Потім обчислюють значення $P\{S > S^*\} = \int_{S^*}^{\infty} g(s | H_0) ds = e^{-\frac{S^*}{2}}$. Якщо $P\{S > S^*\} > \alpha$,

де α – рівень довіри, то нема підстав відхиляти гіпотензу H_0 . В противному випадку гіпотеза відхиляється.

Критерій ω^2 -Крамера-Мізеса-Смірнова. Формулюють гіпотезу H_0 . Потім випадковим методом відбирають вибірку об'ємом n , в якій значення сортують в напрямку $x_1 \leq x_2 \leq \dots \leq x_n$, після чого обчислюють значення статистики S^* :

$S^* = \frac{1}{12 \cdot n} + \sum_{i=1}^n \left(F(x_i, \theta) - \frac{2 \cdot i - 1}{2 \cdot n} \right)^2$. Тоді обчислюють значення $P\{S > S^*\} = \int_{S^*}^{\infty} g(s | H_0) ds = 1 - a_1(S^*)$,

$$a_1(S) = \frac{1}{\sqrt{2 \cdot S}} \sum_{j=0}^{\infty} \frac{\Gamma\left(j + \frac{1}{2}\right) \cdot \sqrt{4j+1}}{\Gamma\left(\frac{1}{2}\right) \cdot \Gamma(j+1)} \cdot e^{-\frac{(4j+1)^2}{16 \cdot S}} \cdot \left\{ I_{-\frac{1}{4}} \left[\frac{(4j+1)^2}{16 \cdot S} \right] - I_{\frac{1}{4}} \left[\frac{(4j+1)^2}{16 \cdot S} \right] \right\},$$

$$I_\nu(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{\nu+2 \cdot k}}{\Gamma(k+1) \cdot \Gamma(k+\nu+1)}, \quad |z| < \infty, \quad |\arg z| < \pi. \text{ Якщо } P\{S > S^*\} > \alpha, \text{ де } \alpha -$$

рівень довіри, то нема підстав відхиляти гіпотензу H_0 . В противному випадку гіпотеза відхиляється.

Можна обчислене значення статистики S^* порівняти з критичним значенням S_α , яке визначається із умови $\int_{S_\alpha}^{\infty} g(s | H_0) ds = \alpha$, де $g(s|H_0)$ – щільність

ймовірностей розподілу статистики критерію при справедливості гіпотези H_0 . Критичне значення критерію S_α при заданому α можна взяти із таблиці Б4 [3.18]. Гіпотеза про згоду відхиляється, якщо значення статистики попадає в критичну область, тобто при $S^* > S_\alpha$.

Емпіричні тести.

Тест 1. Перевірка вибірки на випадковість. З цим тестом перевіряється кількість перестановок (Q), коли $x_i > x_{i+1}$ при $i = \overline{1, n-1}$. Якщо їх число Q при

заданому рівні довіри α попадає в довірчий інтервал $\left[\frac{P}{2} - \frac{U_{1-\alpha/2} \cdot \sqrt{P}}{2}; \frac{P}{2} + \frac{U_{1-\alpha/2} \cdot \sqrt{P}}{2} \right]$,

де P – довжина вибірки, $U_{1-\alpha/2}$ – квантиль рівня $1-\alpha/2$ нормального розподілу $N(0,1)$ то вибірка випадкова.

Тест 2. Частотний тест на рівномірність. Тест ґрунтується на законі великих чисел і виконується за таким алгоритмом. **2.1.** Розбивається $[0; m)$ на K рівних інтервалів. **2.2.** Генерується n чисел $x_1, x_2, \dots, x_n$. **2.3.** Підраховується кількість елементів, які попали в кожний інтервал $(n_1, n_2, \dots, n_K)$. **2.4.**

Розраховуються відносні частоти попадань в інтервали: $v_1 = \frac{n_1}{n}, v_2 = \frac{n_2}{n}, \dots, v_K = \frac{n_K}{n}$.

2.5. Будується

гістограма частот $v_1, v_2, \dots, v_K$ на K відрізках інтервалу $[0; m)$. **2.6.** Обчислюється

МС ВВ ξ : $M^{(n)} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$. **2.6.** Обчислюється оцінка дисперсії ВВ ξ :

$D^{(n)} = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - M^{(n)})^2$. **2.7.** Повторюються пункти 2–6 для великих значень n .

2.8. Оцінюється збіжність кожної частоти v_i до теоретичної ймовірності попадання в i -й інтервал ($P = 1/K$). Згідно із законом великих чисел повинно бути

$v_i \xrightarrow{P} \frac{1}{K}$ при $n \rightarrow \infty$. Це означає, що з ростом n висоти столбців гістограми, в

цілому повинні наближатись до рівня $1/K$. Вважається, що всі частоти v_i збігаються до теоретичної ймовірності попадання в i -й інтервал, якщо при заданому рівні довіри α кожна частота v_i попадає в довірчий інтервал

$\left[\frac{1}{K} - \frac{U_{1-\alpha/2}}{K} \cdot \sqrt{\frac{K-1}{n}}; \frac{1}{K} + \frac{U_{1-\alpha/2}}{K} \cdot \sqrt{\frac{K-1}{n}} \right]$, де $U_{1-\alpha/2}$ – квантиль рівня $1-\alpha/2$ нормального

розподілу $N(0,1)$. **2.9.** Оцінюється збіжність МС і дисперсії до теоретичних

значень: $M\xi = \frac{m}{2}$, $D\xi = \frac{m^2}{12}$. Якщо при заданому рівні довіри α при різних n оцінка

МС $M^{(n)}$ попадає в довірчий інтервал $\left[\frac{m}{2} - \frac{U_{1-\alpha/2} \cdot \sigma}{\sqrt{n}}; \frac{m}{2} + \frac{U_{1-\alpha/2} \cdot \sigma}{\sqrt{n}} \right]$, де $\sigma^2 = m^2 / 12$,

то вважається, оцінка МС збігається до $M\xi$. Якщо при заданому α при різних n

оцінка дисперсії $D^{(n)}$ попадає в довірчий інтервал $\left[\frac{(n-1) \cdot \sigma^2}{\chi_{1-\alpha/2}^2(n-1)}; \frac{(n-1) \cdot \sigma^2}{\chi_{\alpha/2}^2(n-1)} \right]$,

де $\sigma^2 = m^2 / 12$, $\chi_{\zeta}^2(r)$ – квантиль χ^2 -розподілу з рівнем ζ і кількістю ступеней

вільності r , то вважаємо, що оцінка дисперсії збігається до $D\xi$. Тест вважається пройденим успішно, якщо всі оцінки попали у відповідні довірчі інтервали.

Тест 3 на незалежність. Тест статистичної незалежності чисел $x_1, x_2, \dots, x_n$ можна проводити, оцінюючи кореляцію між числами x_i і x_{i+s} , між яким відстань у вибірці дорівнює $s > 0$. Для цього розглядаються дві ВВ ξ та η і визначається коефіцієнт кореляції між ними за формулою: $R(\xi, \eta) = \frac{M(\xi\eta) - M(\xi) \cdot M(\eta)}{\sqrt{D(\xi) \cdot D(\eta)}}$. Якщо

числа ξ і η розподілені рівномірно на інтервалі $[0; m)$, то $M\xi = M\eta = \frac{m}{2}$ і

$$D\xi = D\eta = \frac{m^2}{12} \text{ і } R(\xi, \eta) = \frac{12}{m^2} \cdot M(\xi\eta) - 3.$$

Розглянемо пару чисел (x_i, x_{i+s}) як реалізацію пари ВВ ξ і η . Тоді в послідовності $x_1, x_2, \dots, x_n$ маємо всього $n - s$ реалізацій цієї пари:

$(x_1, x_{1+s}), (x_2, x_{2+s}), \dots, (x_{n-s}, x_n)$. За ними можна розрахувати оцінку $R^{(n)}$ коефіцієнта

кореляції $R(\xi, \eta)$: $R^{(n)} = \frac{12}{m^2} \cdot \frac{1}{n-s-1} \sum_{i=1}^{n-s} (x_i \cdot x_{i+s}) - 3$. З ростом n оцінка $R^{(n)}$ повинна

прямувати до нуля, інакше згенерована послідовність не відповідає вимогам незалежності. Отже, якщо при заданому рівні довіри α для різних n оцінка коефіцієнта кореляції $R^{(n)}$ попадає в довірчий інтервал

$\left[th\left(-\frac{U_{1-\alpha/2}}{\sqrt{n-3}}\right); th\left(\frac{U_{1-\alpha/2}}{\sqrt{n-3}}\right) \right]$, де $U_{1-\alpha/2}$ – квантиль рівня $1-\alpha/2$ нормального

распределения $N(0,1)$, то вважаємо, що послідовність ВВ відповідає вимозі незалежності.

Тест 4 на випадковість, рівномірність і незалежність. Для цього згенерована послідовність довжиною P розбивається на r підпослідовностей:

$$x_1, x_2, \dots, x_P \Rightarrow \begin{cases} x_{0-r+1}, x_{1-r+1}, \dots, x_{t-r+1}, \\ x_{0-r+2}, x_{1-r+2}, \dots, x_{t-r+2}, \\ \dots \\ x_{0-r+r}, x_{1-r+r}, \dots, x_{t-r+r} \end{cases}, \text{ де } t = \left\lfloor \frac{P-1}{r} \right\rfloor.$$

Потім для кожної підпослідовності виконуються тести 1-3. Якщо всі вони виконуються, то і тест 4 в цілому вважається успішно завершеним.

Методи моделювання розподілів ВВ. За встановленим законом розподілу $f(x)$, як обвідної гістограми, можна встановити закон зміни ВВ X . Це можна отримати трьома методами: із зарань складених таблиць випадкових чисел; фізичним генератором та за допомогою ЕОМ шляхом генерування так званих псевдовипадкових чисел.

Майже всі стандартні бібліотечні програми обчислення рівномірних ВВ ґрунтуються на конгруентних методах, розроблених Лемером. Якщо закон розподілу відомий, то переважно використовується інверсний метод. Так, якщо ВВ розподілені рівномірно за законом $f(x) = \frac{1}{b-a}$, то рівняння іверсії має вигляд

інтегралу $\int_0^x \frac{dx}{b-a} = r$, звідки $r = \frac{x-a}{b-a}$ і ВВ числа змінюються за законом

$x = a + r(b-a)$, де $r \in [0;1]$ і розподілені рівномірно. У випадку експоненційного

закону матимемо $\int_0^x \exp(-\lambda x) dx = r$, звідки аналогічно $x = -\frac{1}{\lambda} \ln(1-r)$. Оскільки в

інтервалі $[0;1]$ ВВ $(1-r)$ розподілені так само, як ВВ r , то $x = -\frac{1}{\lambda} \ln r$.

Алгоритм Метрополіса [3.13]. Схема Метрополіса ґрунтується на принципі детальної рівноваги і використовується для моделювання еволюції системи до рівноважного стану [3.14]. Якщо в певному стані S ймовірність перебування системи $P(S)$, то середнє значення фізичного параметру A

обчислюється як $\langle A \rangle = \sum_S A(S)P(S)$ і алгоритм релаксації системи будується як набір кроків випадкових зсувів кожної її підсистеми (частинки) із одного стану в інший до досягнення потрібного ансамблю конфігурацій. Під час випадкових зсувів контролюється вибраний параметр. Так, коли моделюється канонічний ансамбль при температурі T , то контролюється зміна енергії стану за алгоритмом: якщо $\Delta W \leq 0$, то зсув приймається; якщо $\Delta W \geq 0$, то зсув приймається з ймовірністю $\exp(-\Delta W / k_B T)$; якщо $\Delta W = 0$, то зсув не приймається. Значення $\exp(-\Delta W / k_B T)$, якщо $\Delta W \geq 0$ лежить в інтервалі $[0;1]$. Тому вимагається, щоб два можливі вихідні стани після зміщення приймалися або з ймовірністю $p = \exp(-\Delta W / k_B T)$, або відхилялись із ймовірністю $p = 1 - \exp(-\Delta W / k_B T)$.

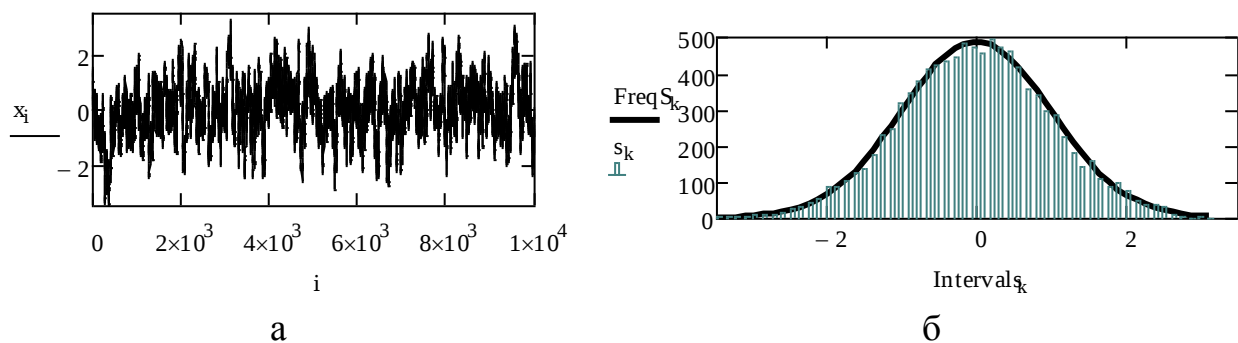


Рис.3.5.2

Щоб цього досягнути, використовується чергове випадкове число ξ із інтервалу $[0;1]$. Якщо виявиться, що $\xi < \exp(-\Delta W / k_B T)$, то зсув приймається, інакше відхиляється. Метрополіс встановив, що послідовне застосування зсувів $\exp(-\Delta W / k_B T)$ поступово наближає модель до стану рівноваги за цих фізичних умов. Критерієм цього є мінімізація флуктуаційних коливань параметра навколо його середнього значення. Для канонічного ансамблю, додатковим фактором наближення системи до стану рівноваги є те, що енергія описується розподілом, пропорційним множнику Больцмана. Наприклад, якщо x змінна стану, то потенціальну енергію можна записати у вигляді $W_p(x) = -\log p(x) + C$. Тоді застосовуватиметься принцип мінімальної потенціальної енергії, за яким стан системи тим ймовірніший, чим менше значення потенціальної енергії в стані.

Наведемо приклад практичного застосування алгоритму Метрополіса

$x_0 := 0$ $N := 1000$ $K := 55$ $\sigma := \text{stdev}(x)$

$$\begin{aligned}
 i &:= 0..N - 1 \\
 p(x) &:= \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) \quad \mu := \text{mean}(x) \\
 x_{i+1} &:= \text{Metropolis}(x_i, 0.5, p) \\
 k &:= 0..K - 1 \\
 \text{Metropolis}(x, \delta\text{max}, p) &:= \begin{array}{l} x_t \leftarrow x + \delta\text{max} \cdot (2 \cdot \text{md}(1) - 1) \\ w \leftarrow \frac{p(x_t)}{p(x)} \\ x_{\text{new}} \leftarrow x_t \text{ if } w \geq 1 \\ \text{if } w < 1 \\ \quad r \leftarrow \text{md}(1) \\ \quad x_{\text{new}} \leftarrow x_t \text{ if } r \leq w \\ \quad x_{\text{new}} \leftarrow x \text{ otherwise} \end{array} \\
 \text{Intervals}_k &:= \text{mir}(x) + \frac{\text{max}(x) - \text{mir}(x)}{K} \cdot k \\
 dx &:= \frac{\text{max}(x) - \text{mir}(x)}{K} \quad s_{xx} := \text{hist}(\text{Intervals}, x) \\
 \text{FreqS}_k &:= p(\text{Intervals}_k) \cdot N \cdot dx \\
 x_{\text{new}} &
 \end{aligned}$$

На рис.3.5.2,а, наведений графік побудованої ВВ за допомогою алгоритму Метрополіса, а на рис.3.5.2,б – графік асимптотичного розподілу ймовірності випадкових значень вибірки.

Практична реалізація методики статистичного аналізу результатів прямих багаторазових вимірювань. Нехай за допомогою лічильника Гейгера вимірюється активність радіоактивних джерел в повітрі, результати 100 вимірів яких, представлені в таблиці 1.

Вважається, що результати вимірювань задовольняють таким вимогам: **1.** Для окремих вимірювань вони повинні відображати статистичну природу процесу, тобто випадково змінюватися від одного вимірювання до іншого. **2.** Незважаючи на статистичну природу флуктуацій окремих вимірювань, результати з довільною заданою точністю повинні давати змогу одержувати інформацію про середні значення вимірюваних величин. **3.** Результати вимірювань повинні підлягати статистичним закономірностям. **4.** Характерний параметр часу $\tau = 1/\lambda$ значно перевищує час окремих вимірювань t_i і повного часу спостереження T .

Таблиця 3.5.1.

i	n_i	$n_i - \langle n \rangle$	i	n_i	$n_i - \langle n \rangle$	i	n_i	$n_i - \langle n \rangle$
1	84	6.53	35	74	-3.47	69	74	-3.47
2	77	-0.47	36	83	5.53	70	83	5.53
3	85	7.53	37	79	1.53	71	72	-5.47
4	79	1.53	38	85	7.53	72	73	-4.47
5	68	-9.47	39	67	-10.47	73	69	-8.47

6	88	10.53	40	82	4.53	74	87	9.53
7	92	14.53	41	73	-4.47	75	69	-8.47
8	79	1.53	42	89	11.53	76	70	-7.47
9	66	-11.47	43	89	11.53	77	75	-2.47
10	70	-7.47	44	70	-7.47	78	77	-0.47
11	67	-10.47	45	80	2.53	79	71	-6.47
12	80	2.53	46	79	1.53	80	64	-13.47
13	75	-2.47	47	90	12.53	81	74	-3.47
14	82	4.53	48	74	-3.47	82	61	-16.47
15	79	1.53	49	74	-3.47	83	84	6.53
16	94	16.53	50	98	20.53	84	79	1.53
17	60	-17.47	51	72	-5.47	85	75	-2.47
18	67	-10.47	52	73	-4.47	86	79	1.53
19	70	-7.47	53	70	-7.47	87	78	0.53
20	94	16.53	54	89	11.53	88	86	8.53
21	67	-10.47	55	70	-7.47	89	73	-4.47
22	68	-9.47	56	77	-0.47	90	76	-1.47
23	86	8.53	57	84	6.53	91	83	5.53
24	60	-17.47	58	85	7.53	92	81	3.53
25	81	3.53	59	77	-0.47	93	94	15.53
26	77	-0.47	60	93	14.53	94	77	-0.47
27	73	-4.47	61	74	-3.47	95	74	-3.47
28	90	12.53	62	78	0.53	96	74	-3.47
29	90	12.53	63	72	-5.47	97	79	1.53
30	78	0.53	64	83	5.53	98	77	-0.47
31	88	10.53	65	58	-19.47	99	86	8.53
32	89	11.53	66	70	-7.47	100	69	-8.47
33	65	-12.47	67	76	-1.47			
34	69	-8.47	68	84	6.53			

Всі 100 незалежних вимірів проведені в ідентичних умовах. Як бачимо із рис.3.5.3,а, результати вимірювань, як функція номера вимірювання i , зосереджуються поблизу середнього значення $\langle n \rangle = \frac{1}{n} \sum_{i=1}^n n_i \cong 77.8$, зображеного прямою лінією.

Нагадаємо, що генеральна сукупність ВВ – це множина всіх можливих їх значень. Однак, в результаті експериментальних вимірювань ми ніколи не маємо можливості одержати повний набір усіх. В будь-якому дослідженні кількість вимірювань завжди обмежена, тому замість генеральної сукупності одержуємо вибірку, середнє по якій $\langle n \rangle = \frac{1}{n} \sum_{i=1}^n n_i$ лише наближено співпадає з математичним очікуванням $M[N]$, оскільки усереднення проводиться не по всій генеральній сукупності, а по вибірці. Незважаючи на те це, найкращою оцінкою істинного

значення вимірюваної величини із вибірки є математичне очікування $M[N]$ і нас цікавитиме в подальшому, за яким законом розподіляються результати вимірювань навколо значення $M[N]$.

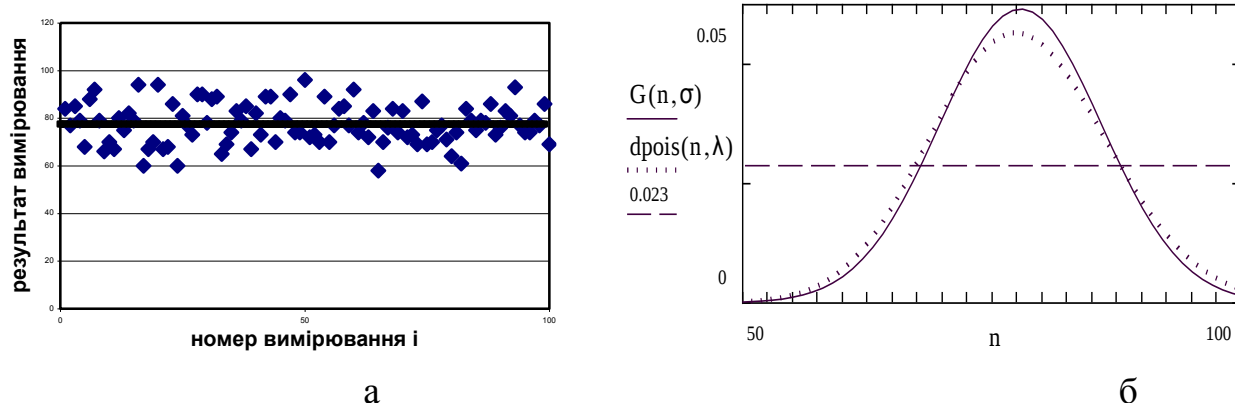


Рис. 3.5.3

Як бачимо із рис.3.5.3,а, результати вимірювань розсіюються з певною ймовірністю по обидві сторони відносно $\langle n \rangle$, причому частота спостережень похибок $n_i - \langle n \rangle$ при віддаленні від $\langle n \rangle$ зменшується. Це означає, що функція густини розподілу має екстремум і нам треба встановити, якому закону підлягає розподіл ймовірності одержання результату відносно $\langle n \rangle$.

Як впливає з таблиці 1, середнє значення похибки $\langle \Delta n \rangle = \frac{1}{K} \sum_{i=1}^K \Delta n_i = \frac{8}{100} \cong 0$, що свідчить про правильність значень відхилень, тому треба обчислювати значення середньо квадратичного відхилення $\delta_n = \sqrt{\frac{1}{K-1} \sum_{i=1}^K (\Delta n_i)^2}$. Значення $(n_i - \langle n \rangle)^2$ приведені поряд з виміряними результатами в таблиці 3.5.2.

Таблица 3.5.2.

i	n_i	$(n_i - \langle n \rangle)^2$	i	n_i	$(n_i - \langle n \rangle)^2$	i	n_i	$(n_i - \langle n \rangle)^2$
1	84	42.64	35	74	12.04	69	74	12
2	77	0.22	36	83	5.53	70	83	30.6
3	85	56.7	37	79	30.6	71	72	30
4	79	2.34	38	85	56.7	72	73	20
5	68	89.7	39	67	109.6	73	69	71.7
6	88	110.9	40	82	20.5	74	87	91
7	92	211.12	41	73	20	75	69	71.7
8	79	2.34	42	89	133	76	70	55.8
9	66	131.56	43	89	133	77	75	6.1
10	70	55.8	44	70	55.8	78	77	0.22

11	67	109.6	45	80	6.40	79	71	-6.47
12	80	40.96	46	79	2.34	80	64	41.8
13	75	6.1	47	90	157	81	74	12
14	82	20.5	48	74	12	82	61	271
15	79	2.3	49	74	12	83	84	42.6
16	94	273.24	50	98	421	84	79	2.34
17	60	305	51	72	30	85	75	6.1
18	67	109.6	52	73	20	86	79	2.34
19	70	55.8	53	70	55.8	87	78	0.28
20	94	273.24	54	89	133	88	86	72.7
21	67	109.6	55	70	55.8	89	73	20
22	68	89.7	56	77	0.22	90	76	2.16
23	86	8.53	57	84	42.6	91	83	30.6
24	60	72.76	58	85	56.7	92	81	12.5
25	81	12.46	59	77	0.22	93	94	241
26	77	0.22	60	93	211	94	77	0.04
27	73	20	61	74	12	95	74	12
28	90	157	62	78	0.28	96	74	12
29	90	157	63	72	-5.47	97	79	1.53
30	78	0.28	64	83	30	98	77	2.34
31	88	110.8	65	58	379	99	86	72.7
32	89	133	66	70	55.8	100	69	71.7
33	65	155.5	67	76	2.16			
34	69	71.74	68	84	42.6			

Як впливає з таблиці 2, середньо квадратичне відхилення результатів

спостережень дорівнює $\delta_n = \sqrt{\frac{1}{K-1} \sum_{i=1}^n (n_i - \langle n \rangle)^2} \cong 8.037$, $K = 100$. Відзначимо також,

що стандартне відхилення σ характеризує середню похибку окремих вимірювань

$n_1, n_2, \dots, n_i, \dots, n_k$, тобто відображає ймовірність виявлення спостережуваних значень

вимірюваних величин в інтервалі $[\langle n \rangle - \sigma, \langle n \rangle + \sigma]$. Крім цього, в практичних

розрахунках вираз $D = \frac{1}{N} \sum_{i=1}^N (n_i - \langle n \rangle)^2$ «модифікують» так, щоб в знаменнику

замість K була величина $K - 1$: $D = \frac{1}{K-1} \sum_{i=1}^K (n_i - \langle n \rangle)^2$.

Це практична оцінка дисперсії вимірювання, яка є точною при $k \gg 1$. Дисперсія вимірювань представляє собою усереднення середньоквадратичного відхилення результатів окремих вимірювань $n_1, n_2, \dots$ від середнього значення $\langle n \rangle$. Виявляється, що при $k \gg 1$ в інтервалі $[\langle n \rangle - \sigma, \langle n \rangle + \sigma]$ зосереджено біля 70 % випадкових результатів незалежних вимірювань в однакових умовах.

Перевірка експериментального розподілу. Кажуть, що випадкова величина підлягає розподілу Пуассона, якщо ймовірність одержання значення n

при середньому $\langle n \rangle$ дорівнює: $p(n, \langle n \rangle) = \frac{\langle n \rangle^n}{n!} e^{-\langle n \rangle}$. Нагадаємо, що розподіл Пуассона - це розподіл дискретних ВВ, максимум якого формується в околі значення $\langle n \rangle$.

На рис. 3.5.4 показані розподіли Пуассона для різних середніх значень $\langle n \rangle$: 0.2, 2, 6. Переконаємось, що з ростом $\langle n \rangle$ формується екстремум розподілу так, що крива розподілу набирає більш симетричної форми, зменшується амплітуда максимуму, зменшується відносна дискретність випадкових результатів вимірювань.

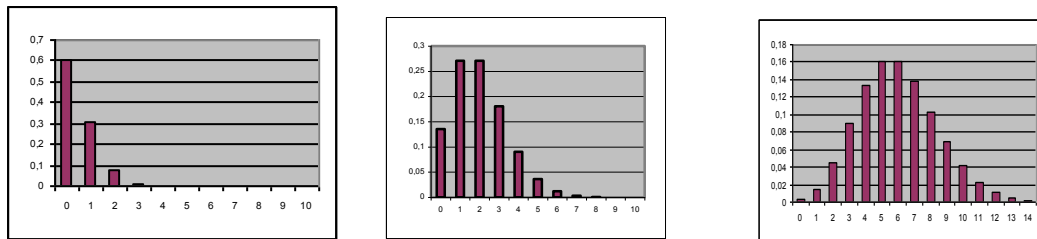


Рис. 3.5.4

На рис.3.5.3,б, приведена розрахована в MathCAD швидкість зміни функції розподілу (густина або щільності) для $\langle n \rangle = \text{const} = 77.8$. Контур дещо несиметричний і має ширину на половині висоти порядку $\Delta n \approx 20$. Отже, з точки зору розподілу Пуассона, для $\langle n \rangle = \text{const} = 77.8$, корінь квадратний з ширини контуру дорівнює $\sqrt{\Delta n} \approx 4.47$. Це вдвічі менше значення $\sqrt{\langle n \rangle} \approx 8.8$, тобто виконується рівність $\sqrt{\Delta n} \cong \frac{1}{2} \sqrt{\langle n \rangle} \cong \frac{1}{2} \delta_n$.

Однак відомо, що для великих значень $\langle n \rangle$ розподіл Пуассона переходить в розподіл Гаусса $G(n, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(n - \langle n \rangle)^2}{2\sigma^2}\right)$. Тому накладемо на криву Пуассона (рис.3) криву Гаусса для параметрів: $\langle n \rangle \cong 77.8$ і $\sigma_n \cong 8.1$. Щоб дати відповідь остаточно на поставлене питання стосовно ідентифікації експериментального розподілу теоретичним, треба на рис.3.5.3,б, нанести експериментальний розподіл частоти експериментального значення і всі криви сумістити максимумами.

Побудуємо гістограму експериментального розподілу. Для цього перепишемо таблицю 3.5.1 в порядку зростання значень вимірюваних величин:

Таблиця 3.5.3

58	67	70	73	74	77	79	83	86	90
60	68	70	73	75	77	79	83	86	90
60	68	70	73	75	77	79	84	87	92
61	69	70	73	75	78	80	84	88	92
64	69	70	74	76	78	80	84	88	93
65	69	71	74	76	78	81	84	89	93
66	69	72	74	77	79	82	85	89	94
67	70	72	74	77	79	82	85	89	94
67	70	72	74	77	79	83	85	89	94
67	70	73	74	77	79	83	86	90	98

Покладемо $C_1 = n_{\min}$, а $C_k = n_{\max}$ та розіб'ємо інтервал $[C_1; C_k]$ на k рівних проміжків $\Delta_{\text{int}} = \frac{C_k - C_1}{k}$. Початок довільного інтервалу C_i , а кінець - C_{i+1}

($i=1, \dots, k$). Нехай $k = 10$. Тоді $\Delta_{\text{int}} = \frac{C_k - C_1}{k} = \frac{95 - 58}{10} = \frac{37}{10} = 3.7 \cong 4$ приймається за ширину інтервалу.

Визначимо частоти і відносні частоти попадань результатів окремих вимірювань в кожний інтервал (в таблиці 3.5.2.) Побудована на представлених в рис.3.5.5. За формою обвідної можна наближено стверджувати, що деяка асиметрія експериментального контуру $Ef(i)$ схиляє до прийняття гіпотези про те, що експериментальний розподіл ближчий до розподілу Пуассона.

Таблиця 3.5.4

Інтервал	Межі інтервалу	Число попадань результату, V_i	Відносна частота попадань, p_i
$[C_1; C_2)$	$[58, 62)$	4	0,04
$[C_2; C_3)$	$[62, 66)$	2	0,02
$[C_3, C_4)$	$[66, 70)$	11	0,11
$[C_4, C_5)$	$[70, 74)$	17	0,17
$[C_5, C_6)$	$[74, 78)$	19	0,19
$[C_6, C_7)$	$[78, 82)$	13	0,13
$[C_7, C_8)$	$[82, 86)$	13	0,13
$[C_8, C_9)$	$[86, 90)$	10	0,10
$[C_9, C_{10})$	$[90, 94)$	7	0,07
$[C_{10}, C_{11}]$	$[94, 98)$	4	0,04

таблиці 3 значеннях параметрів гістограма частот та їх обвідна $Ef(i)$, приведені на

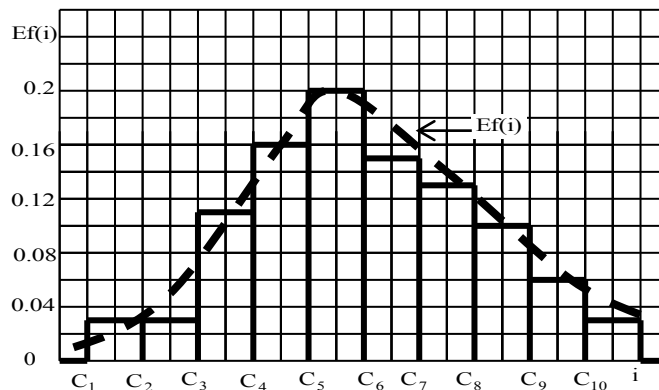


Рис.3.5.5

Кількість інтервалів розбиття експериментальної вибірки менша 20, тому правомірно провести оцінку можливості промахів за Романовським. Щоб його застосувати, визначимо серединні значення кожного інтервалу:

i	1	2	3	4	5	6	7	8	9	10
$\Delta n_{i, midl}$	59.5	63.6	67.5	71.5	75.5	79.5	84.5	88.5	92.5	96.5

Тоді оцінимо виконання критерію Романовського для максимального вимірюваного експериментального значення $n_{\max}=96.5$ і знайдемо для нього

$$\text{значення } \beta \text{ за формулою } \beta_i = \frac{|n_i - \langle n \rangle|}{S_n}: \beta_{10} = \frac{|96.5 - 77.5|}{8.06} = 2.36.$$

За таблицею

k	4	6	8	10	12	15	20
β	1.71	2.10	2.27	2.41	2.52	2.64	2.78

критерій Романовського при об'ємі вибірки (кількості інтервалів) 10 і довірчій ймовірності $P=0.95$ складає 2.41. Оскільки для даного об'єму вибірки $k=10$ задовольняється умова $\beta_{10} < \beta$, то результат вимірювання $n=96.56$ не вважається промахом.

Перевірка висунутої гіпотези за критерієм згоди Пірсона. Основна мета збільшення кількості вимірювань полягає в зменшенні випадковості результату вимірювання i , отже, в найкращому наближенні результату до істинного значення величини. Але збільшувати число вимірювань лише з метою знайти істинне значення величини безпідставно.

Однак, якщо кількість проведених вимірювань $n > 50$, то найкращим способом оцінки близькості експериментального розподілу даних до нормального, використовується критерій згоди Пірсона (хі-квадрат). За ним перевіримо нормальність розподілу виміряних значень з рівнем значущості $\alpha = 0.05$

Розіб'ємо вибірку випадкових значень вимірюваної величини із K вимірювань в порядку зростання на $\sqrt{K} = \sqrt{100} = 10$ (або $3 \ln K + 1$) інтервалів та визначимо експериментальні і теоретичні частоти (таблиця 3.5.5). Як критерій

Таблиця 3.5.5

Інтервал вибірки	Межі інтервалу	Відносні частоти, ν_i	Теоретичні частоти, $\nu_{теор,i} = G(n_i, \sigma) \cdot 100$
$[C_1; C_2)$	$[58,62)$	2	0.45
$[C_2; C_3)$	$[62,66)$	2	1.2
$[C_3, C_4)$	$[66,70)$	11	2.4
$[C_4, C_5)$	$[70,74)$	17	3.8
$[C_5, C_6)$	$[74,78)$	19	4.8
$[C_6, C_7)$	$[78,82)$	13	4.7
$[C_7, C_8)$	$[82,86)$	13	3.7
$[C_8, C_9)$	$[86,90)$	10	2.2
$[C_9, C_{10})$	$[90,94)$	7	1.1
$[C_{10}, C_{11}]$	$[94,98)$	4	0.4

розбіжності між теоретичними ймовірностями і експериментально спостережуваними статистичними відносними частотами p_i , розглянемо

величину $\chi^2 = \sum_{i=1}^K \frac{(\nu_i - \nu_{теор,i})^2}{\nu_{теор,i}}$, де ν_i , $\nu_{теор,i}$ - експериментальне і теоретичне

значення відносних частот в i -му інтервалі розбиття. Оскільки довжина інтервалів розбивки однакова, то різними є відносні частоти попадання в них значень випадкових величини, якщо звичайно розподіл їх не є рівномірний.

Для симетричного закону розподілу, як нормальний, число незалежних умов, яким повинні задовольняти статистичні ймовірності p_i , приймається рівним 3. Тоді за допомогою таблиці про теоретичний χ^2 -розподіл з параметром розподілу $k = s - r - 1 = 10 - 2 - 1 = 7$, визначаємо теоретичне значення $\chi_{теор}^2 \cong 14.6$.

Номер інтервалу, i	1	2	3	4	5	6	7	8	9	10
χ^2	5.34	0.53	30.8	45.8	46.7	14.6	24.4	27.6	31.6	32.4

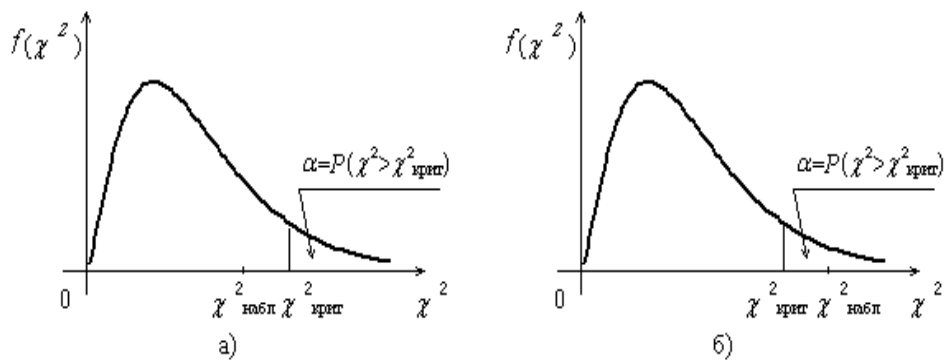


Рис.3.5.6

Висновок про прийняття чи відхилення гіпотези про те, що експериментальний розподіл нормальний, робимо на підставі наступного: 1. Якщо експериментальне значення $\chi^2_{\text{експ}}$ попало в область прийняття гіпотези ($\chi^2_{\text{експ}} < \chi^2_{\text{крит}}(\alpha, k)$, рис.3.5.6,а), то нема підстави відхиляти нульову гіпотезу і розбіжність між експериментальними і теоретичними частотами випадкова. 2. Якщо експериментальне значення $\chi^2_{\text{експ}}$ критерію попало в критичну область ($\chi^2_{\text{експ}} > \chi^2_{\text{крит}}(\alpha, k)$, рис.3.5.6,б), то нульова гіпотеза відхиляється, а справджується конкурентна гіпотеза про те, що розбіжність між експериментальними і теоретичними частотами значима.

Висновок. Розподіл частоти розкиду результатів вимірювань навколо середнього $\langle n \rangle = 77.8$, найімовірніше підлягає розподілу Пуассона.

3.6. Статистичний аналіз флуктуацій в мікросистемах [3.15;3.16]

Відомо, що основою шумів та спричинених ними похибок вимірювань є ймовірнісна природа флуктуацій фізичних величин. Бернуллі обґрунтував закон розподілу (біноміальний) дискретних ВВ, якщо зв'язок між подіями відсутній. Якщо ймовірність окремої події значно менша одиниці, то розподіл Бернуллі спрощується і перетворюється в розподіл Пуассона. Для розподілу Пуассона характерний максимум функції. В цьому параграфі приділена увага найпростішій фізичній системі –системі з багатьма частинками, як ідеальний газ із N окремих частинок, що безперервно хаотично рухаються і газ в цілому перебуває в стані

теплової рівноваги. З точки зору класичної механіки, щоб описати його стан, треба знати значення $3N$ їх координат і $3N$ їх імпульсів частинок у визначений момент часу $t = t_0$, щоб потім, використовуючи рівняння Ньютона розрахувати траєкторії руху кожної з них в довільний інший момент часу t : $x_i = x_i(t)$; $p_i = p_i(t)$ Так стверджує так званий лапласовий детермінізм. Однак, як показав Крилов, зіткнення частинок (атомів чи молекул) одна з одною призводить до того, що рух багаточастинкової системи стає нестійким за Ляпуновим, тобто не детермінованим, а хаотичним. Це означає, що для опису руху інший, що ґрунтується на використанні методів теорії ймовірностей. Висновок Крилова узгоджується з результатами спостережень в дослідах Броуна і Перрена. Шляхом усереднення одержаних ними результатів було встановлено, що *рух частинок, що належать макроскопічним системам, має випадковий, хаотичний характер і їх опис вимагає використання статистичних методів*, в основу яких покладені уявленні про те, що:

1. Система з багатьма частинками проявляє колективні властивості. Такими є температура газу, тиск, в'язкість тощо, роль яких зменшується із зменшенням кількості частинок в системі.
2. Існує зв'язок між колективними властивостями системи та середніми значеннями параметрів, які характеризують поведінку кожної частинки зокрема, як тиск та температура, що виражаються через кінетичну енергію руху молекул.
3. У встановленні цього зв'язку нема потреби точно знати положення та швидкість кожної частинки, досить знати ймовірність того, що значення швидкості попадає на певний інтервал шириною $d\vartheta$. З іншої сторони система багатьма частинками, як макроскопічна, може обмінюватися екстенсивними (об'єм і маса) та інтенсивними (тиск, температура) параметрами з зовнішнім середовищем.

Повний опис мікростану означає, що знаючи координати та імпульси всіх частинок у певний момент часу, можна розрахувати будь-яку механічну величину, що характеризує систему, наприклад, енергію, а за допомогою рівнянь руху – визначити її стан у довільний момент часу. Макроскопічна природа параметрів V , p і T різна. Так під час ізотермічно повільного процесу змінюється об'єм газу, тоді як його тиск і температура залишаються сталими. Якщо тиск і температура одержуються

завдяки усередненню мікроскопічних величин, то об'ємом володіє і порожня посудина. Тому об'єм - макроскопічний параметр, але в геометричному розумінні.

Статистична модель релаксації макроскопічної системи до рівноважного стану. Згідно з одним із основних положень статистичної фізики, система з фіксованими об'ємом і кількістю частинок, будучи в рівновазі з термостатом при температурі T , може перебувати в різних станах, реалізація кожного з яких пропорційна больцманівському множнику $p(n) \approx \exp\left(-\frac{W_n}{k_B T}\right)$. Це канонічний розподіл, тому ансамбль системи, що ним описується, називається канонічним ансамблем. Задача статистичного моделювання полягає в тому, щоб імітувати контакт ансамблю з термостатом, що має температуру T . Допускається, що ансамбль на термостат не впливає, оскільки останній значно перевищує розмірами. Навпаки, стан термостату впливає на стан моделі (ансамблю). Під дією цього впливу структура ансамблю релаксує до рівноважного стану з термостатом при температурі T . Перехід макроскопічної системи з нерівноважного стану в рівноважний називається релаксацією, а відповідний час, протягом якого відбувається такий перехід, називається часом релаксації. Час релаксації – це макроскопічна величина. В статистичній фізиці відоме так зване нульове начало, згідно з яким: *В довільній макроскопічній ізольованій системі встановлюється термодинамічна рівновага (стан теплової рівноваги) за скінчений проміжок часу.* Суть цього постулату в тому, що в ізольованій макроскопічній системі з плином часу відбувається перехід із довільного початкового стану в скінчений стаціонарний стан. Якщо протягом переходу система продовжує залишатися ізольованою. Серед усіх стаціонарних станів існує один виділений стаціонарний стан, який характеризується максимально можливим при заданих умовах хаосом в макроскопічній системі. В статистичній фізиці ця фізична величина називається ентропією. Отже, перехід відбувається в той стан, для якого ентропія максимальна.

Для імітації цієї релаксації використовують алгоритм Метрополіса, за яким вона проводиться шляхом послідовних випадкових зсувів усіх атомів по одному

з метою досягнення ансамблю конфігурацій при даній температурі. Атом для зсуву вибирається довільним чином. Так, якщо він опинився в стані з індексом «1», то в результаті зсуву координати атома дорівнюють: $x'_i = x_i + \ell_{\max} r_1$, де $i = 1, 2, 3$, r_i - три випадкові числа із інтервалу $[-0.5; +0.5]$, $\ell_{\max}$ - максимальна зміна координати. Спершу зсув розглядається як «пробний». Потім визначається зміна повної енергії ΔW моделі, зумовлена даним зсувом, як різниця енергій до зсуву і після зсуву. Якщо $\Delta W \leq 0$ - зсув приймається; якщо $\Delta W \geq 0$ - зсув приймається з ймовірністю таких систем детермінований підхід не підходить, а треба використовувати $\exp\left(-\frac{\Delta W}{k_B T}\right)$, значення якої знаходиться в інтервалі $[0; 1]$. Отже, що прийняти або відхилити зсув атома, треба згенерувати наступне випадкове число із інтервалу

$[0; 1]$. Тоді якщо $r \leq \exp\left(-\frac{\Delta W}{k_B T}\right)$, то зсув приймається, в противному разі –

відхиляється. Шляхом здійснення багатьох аналогічних послідовних зсувів модель поступово наближається до рівноважного стану, що відповідає даній температурі. Існує два тести, що дозволяють перевірити чи досягла досліджувана модель рівноважного стану: **1.** Енергія моделі не зменшується монотонно, а флуктуаційно коливається навколо середнього значення. **2.** Значення середньої енергії, одержаної внаслідок серії зсувів після переходу в рівноважний стан, підлягає канонічному розподілу.

Ентропійний зміст статистичної вибірки . Як впливає з класичної терії ймовірностей, при $n \rightarrow \infty$ відношення $\frac{\sqrt{n}}{n} \rightarrow 0$. Це означає, що при значній кількості випробувань, результати випробувань групуються навколо найбільш ймовірного середнього значення. Це означає, що великий масив випробувань над одиничним об'єктом дослідження можна замінити одним випробуванням над статистичним масивом, тобто системи із великої кількості однакових об'єктів. Наприклад, багаторазове кидання одного грального кубика можна змінити одноразовим киданням багатьох кубиків і при цьому результат прямуватиме до найбільш ймовірного значення.

Розглянемо такий приклад. Нехай маємо колоду з дуже великої кількості карт порядку $n \cong 10^{23}$, серед яких αn білих, а βn -чорних ($\alpha + \beta = 1$). Якщо за стан такої системи вважати порядок розташування в стопі чорних і білих карт, то кількість таких станів із n елементів по αn дорівнює $\Omega = \frac{n!}{(\alpha n)!(\beta n)!}$, звідки за формулою Стирлінга $\Omega = \frac{1}{\sqrt{2\pi n \alpha \beta}} \frac{1}{\alpha^{\alpha n} \beta^{\beta n}}$. Виявляється, що стан стопи повинен буди таким, що довільна її частина, що складається із значної кількості карт, повина мати в собі наближено αn білих карт і βn чорних карт з похибкою $\sqrt{n}$. Величина S , що описує інформаційний стан великої системи як логарифма із числа можливих станів системи $S = \ln \Omega$ виражає ентропію (безладність або хаос) системи. Ентропія системи із декількох не взаємодіючих складових дорівнює сумі ентропій цих складових. Отже, якщо ймовірність реалізації кожного із станів позначити як p_i , то ентропія статистичної системи обчислиться за формулою $\sum_i p_i \ln \frac{1}{p_i}$. У фізиці для статистичної системи, ентропія обґрунтовується як міра її хаосу. Встановлено, що ентропія статистичної системи може лише зростати або залишатись сталою, тому процеси, при яких ентропія зростає є необоротними.

Ентропія і інформація. У 1929 р. *Сцїлард* встановив зв'язок між ентропією та інформацією. Суть зв'язку полягає в тому, що інформація також зв'язана з ймовірністю. Дійсно, за *Больцманом* ймовірність зв'язана з ентропією, тому ентропія зв'язана з інформацією. Дійсно, нехай «алфавіт» складається із N рівноймовірних букв. Відома формула Хартлі $H = \log_2 N$, яка встановлює зв'язок між кількістю інформації H і N , та формула Шенона $H = \sum_{i=1}^N p_i \log_2 \frac{1}{p_i}$, що виражає ентропію «букв» повідомлення, де $p_1, p_2, \dots, p_N$ - ймовірності букв алфавіту $x_1, x_2, \dots, x_N$.

Розглянемо процес пересилання інформації каналом зв'язку. Нехай на вході каналу поступає дискретна ВВ X вибіркою $x_1, x_2, \dots, x_N$ та розподілом імовірностей $p_1, p_2, \dots, p_N$. Тоді на виході каналу матимемо вибірку $y_1, y_2, \dots, y_N$ з розподілом імовірностей $q_1, q_2, \dots, q_N$. Тому, пославши сигнал x_i , на виході каналу зв'язку, в

принципі можемо мати довільний сигнал із вибірки $y_1, y_2, \dots, y_N$. Більше того, прийнявши сигнал y_j , споживач не впевнений, який саме сигнал із вибірки $x_1, x_2, \dots, x_N$ вислав відправник. Ця невизначеність зростає, якщо в каналі наявні шуми. Це означає, що шуми збільшують різницю між ентропіями вхідної і вихідної інформацій. Щоб зменшити вплив шумів, сигнал кодують. Якщо шуми відсутні, то ентропії вхідної і вихідної інформації співпадають між собою.

Відома теорема Шенона, яка стверджує про те, що існує можливість побудови «завадостійкого» коду. Першим таким кодом був код Хеммінга, в якому кодувались не окремі букви, а цілі блоки з багатьох букв. На завершення зазначимо, що оскільки з зростанням ентропії міра хаосу в статистичній системі збільшується, то зменшується інформація про мікростан системи, через які може реалізуватися її макростан і навпаки, якщо система впорядковується, її безладдя зменшується і збільшується інформація про цей макростан. Закон про незменшення ентропії має ймовірністний характер і зумовлений тим, що кожен стан в природі зустрічається тим частіше, чим він більш імовірний.

На завершення, наведемо аналогію між деякими поняттями, що описують інформаційні та теплові потоки, що, в принципі, може дозволити застосувати для моделювання інформаційних систем термодинамічних процесів, які досить добре вивчені та перевірені практикою. Нагадаємо, що коли деякий тепловий баланс підлягає диференційному рівнянню $C \frac{\partial T}{\partial t} = \kappa \Delta T + Q(x, y, z, t)$, то аналогічно можна поступити для системи з спрямованим обміном потоками інформацій I :

$$\frac{\partial J}{\partial t} = \zeta \Delta J + I(k, t), \quad k - \text{номер підсистеми. Тоді приймаючи до уваги зв'язок між}$$

ентропією та інформацію, можна навести аналогію між основними поняттями процесів обміну між системами термодинамічним та інформаційним потоками:

Кількість інформації $I \Leftrightarrow$ кількість теплової енергії Q .

Потік інформації $\Leftrightarrow$ потік теплоти.

Міра кількості інформації $J \Leftrightarrow$ температура T .

Коефіцієнт інформаційної провідності $\zeta \Leftrightarrow$ коефіцієнт теплопровідності κ

Ймовірність мікростану та статистичне усереднення параметрів за часом і ансамблем. В принципі, не обов'язково для опису системи з багатьма

частинками використовувати класичний метод Ньютона. Виявляється, що макроскопічні властивості колективу просто не залежить від початкових умов. Тому нема потреби спостерігати за рухом індивідуальних складових системи, тобто спиратись на мікроскопічний підхід. Систему з багатьма частинками зручніше описувати, використовуючи набір макроскопічних параметрів. Адже в стані рівноваги, поведінка системи найпростіша, а її макроскопічні параметри не залежать від часу, тобто виражають суть усереднених макроскопічних (статистичних) закономірностей системи. Саме в цьому полягає перевага статистичного підходу, який дозволяє моделювати поведінку системи з багатьма частинками, використовуючи принципово нового поняття у фізиці – ймовірності перебування системи у певному мікростані. Макроскопічний стан системи можна описати деяким мінімальним набором фізичних (термодинамічних) параметрів, таких як тиск, температура, об'єм, внутрішня енергія. Макроскопічний стан системи з багатьма частинками складається з мікростанів, якими є стани окремо взятих частинок з певними швидкостями і координатами. В рівноважному стані мікростани весь час можуть змінюватись, приймаючи хаотичні значення, тоді як макростан залишається незмінним. Макростан тим стійкіший, чим більшим її числом він реалізується. В рівноважний макростан система прямує через проміжні, кожен з яких реалізується з певною ймовірністю. Причому поняття ймовірності первинне, оскільки ймовірність не встановлюється, а постулюється і виражає найбільш правдоподібну (усереднену) фізичну картину багаторазових повторень одного і того ж випадкового мікростану системи. Усереднення проводять за часом (середнє за часом) і за ансамблем (середнє за ансамблем).

Усереднення за часом. Нехай маємо дискретний ряд таких мікростанів. Якщо в s – стані система може перебувати час t_s , то величина

$P_s = \lim_{\tau \rightarrow \infty} \frac{t_s}{\tau}$ визначатиме ймовірність цього стану. Сукупність ймовірностей в

різних станах утворюватиме розподіл ймовірності для різних станів системи.

Рівноважне значення певної фізичної величини обчислюється як

$\langle a \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} a(t) dt$. Суть цієї формули означає, що рівноважне значення дорівнює

середньому значенню фізичної величини за достатньо великий проміжок часу вимірювання. Набір таких середніх значень різних фізичних величин і характеризує стан термодинамічної системи. Як впливає з виразу

$$\langle a \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} a(t) dt, \text{ щоб знайти середні значення фізичних величин, треба знати}$$

закон їх зміни з часом. Це надто складний шлях обчислень. Статистична фізика пропонує значно простіший – в заміні середніх за часом значень макроскопічних величини їх середніми статистичними значеннями, які одержують за допомогою відповідної функції розподілу.

Усереднення за ансамблем. Нехай n разів вимірюється координата атома x і допустимо, що із них $\Delta n(i)$ разів її значення попадало в проміжок $[x_i; x_i + \Delta x_i]$.

Відношення $\Delta \Omega(i) = \frac{\Delta n(i)}{n}$ відоме як частота випадань координати атома. Через

частоту випадань обчислюється середнє значення $\langle x \rangle = \frac{1}{n} \sum_i \Delta n(i) \cdot x_i$. Якщо таких

вимірювань багато $n \rightarrow \infty$, то частота випадань в граничному випадку частота

випадань перетворюється в їх ймовірність $\Delta P(i) = \frac{\Delta n(i)}{n} \Rightarrow \Delta n(i) = \Delta P(i) \cdot n$ і середнє

значення буде дорівнювати $\langle x \rangle = \frac{1}{n} \sum_i \Delta n(i) \cdot x_i = \frac{1}{n} \sum_i n \cdot \Delta P(i) \cdot x_i = \sum_i \Delta P(i) \cdot x_i$. У випадку

тривимірного представлення, ймовірність того, що в момент часу t координати атома в трьохвимірному просторі перебувають в сферичній оболонці інтервалу

значень $[\vec{r}; \vec{r} + d\vec{r}]$, дорівнює $dP(\vec{r}, t)$. За величиною $dP(\vec{r}, t)$ пропорційна ширині

сферичного інтервалу $d\vec{r} : dP(\vec{r}, t) = f(\vec{r}, t) d\vec{r}$, тому середнє значення просторової

координати буде дорівнювати $\langle \vec{r} \rangle = \int \vec{r} \cdot P(\vec{r}, t) d\vec{r}$, де функція пропорційності $P(\vec{r}, t)$

називається функцією розподілу ймовірностей густиною $f(\vec{r}, t)$, а сам метод

визначення середніх ще називають обчисленням за допомогою функції розподілу.

В стаціонарному стані функція розподілу не залежить від часу $P(\vec{r})$.

Як свідчить аналіз результатів спостережень Броуна і Перрена, в багаточастинкових системах із хаотичним рухом частинок, середнє значення

вектора зміщення частинки в ансамблі $\langle \vec{\Delta r} \rangle = 0$. Середнє квадратичне значення $\sqrt{\langle \vec{\Delta r}^2 \rangle} \neq 0$, а його квадрат лінійно зростає з часом спостереження за процесом $\langle \vec{\Delta r}^2 \rangle = D \cdot t$, де D - коефіцієнт дифузії. Цей висновок справджується лише під час хаотичного руху атома, оскільки при рівномірному $r^2 \approx t^2$. Це означає, що детермінований підхід для опису хаотичного руху не справджується, в тому числі до опису явищ молекулярної фізики, а в цьому випадку треба використовувати ймовірністний підхід.

Модель дисперсії як флуктуація флуктуацій. Нехай додуються N коливань. Ймовірність того, з цієї кількості коливань m має амплітуду $+a$, визначається біноміальним законом розподілу $P_n(m) = C_n^m p^m q^{n-m}$. Ймовірність результуючої інтенсивності дорівнює $I = I(n, m) = (2m - n)^2 a^2$ [3.20], тому середня інтенсивність результуючого коливання буде дорівнювати

$$\langle I \rangle = \langle (2m - n)^2 \rangle a^2 = [4\langle m^2 \rangle - 4\langle mn \rangle - \langle n^2 \rangle] a^2 = n^2 a^2 \left[1 + 4pq \left(\frac{1}{n} - 1 \right) \right].$$

Якщо $p = q = \frac{1}{2}$, то $\langle I \rangle = na^2$, тобто результуюче коливання повністю некогерентне.

При $p = q = 1$ маємо $\langle I \rangle = n^2 a^2$ повну когерентність результуючого коливання. Незалежно від принципу усереднення фізичних параметрів статистичної системи, завжди актуальним постає питання про відхилення значень середніх фізичних величин в обидві сторони від їх середніх значень. Такі відхилення називаються флуктуаціями.

Статистична флуктуація адитивних фізичних величин. Незалежно від принципу усереднення фізичних параметрів статистичної системи, завжди актуальним постає питання про відхилення значень середніх фізичних величин в обидві сторони від їх середніх значень. Такі відхилення називаються флуктуаціями.

Нагадаємо, що за міру відхилення довільно фізичної величини a від її середнього значення $\langle a \rangle$ можна було б прийняти середнє значення. Однак воно може дорівнювати нулеві $\Delta \langle a \rangle = \langle a - \langle a \rangle \rangle = 0$, оскільки відхилення значення

розглядуваної величини в бік більших і менших значень від середнього відбуваються із однаковою частотою. Тому за міру флуктуацій зручніше розглядати середній квадрат різниці $\langle(\Delta a)^2\rangle = \langle(a - \langle a \rangle)^2\rangle = \langle a^2 \rangle - \langle a \rangle^2$, яка називається квадратичною флуктуацією. Тоді відношення $\varepsilon_a = \frac{\sqrt{\langle(\Delta a)^2\rangle}}{\langle a \rangle}$ називається відносною флуктуацією.

Різниця $D[\xi] = \langle \xi^2 \rangle - \langle \xi \rangle^2$ називається флуктуацією флуктуації випадкової величини або її дисперсією, про що йшла мова вище. Так у випадку суперпозиції коливань дисперсія – це флуктуація флуктуації інтенсивності. Тут a^2 виражає інтенсивність процесу, як, наприклад, це має місце для хвилі.

Ця числова характеристика проявляє важливу для фізики властивість. Так, для повністю некогерентних коливань дисперсія дорівнює $D[\xi] = 2N(N-1)a^4$.

Це означає, що відносна флуктуація інтенсивності буде рівною $\sqrt{\frac{D[\xi]}{\langle \xi \rangle^2}} = \sqrt{2\left(1 - \frac{1}{N}\right)}$ зовсім не зменшується, а навпаки, зростає,

наближаючись до значення $\sqrt{2}$. Таким чином, ніякого згладжування відносних флуктуацій тут нема: *при складанні двох коливань відносна флуктуація інтенсивності менша, ніж при складанні багатьох*. Це означає, що віднесення явища до флуктуаційного чи не флуктуаційного суттєво залежить від того, якою фізичною величиною характеризувати це явище. Так для електромагнітної хвилі, $\langle E \rangle = \langle H \rangle = 0$, тому $D[E] = \langle E^2 \rangle$, $D[H] = \langle H^2 \rangle$ і середня енергія $\langle W \rangle = k \int_V (\langle E^2 \rangle + \langle H^2 \rangle) dV = k \int_V (D[E] + D[H]) dV$, тобто сама середня енергія є міра інтенсивності флуктуації напруженостей E, H . Таким чином, маємо розгодження в термінологіях: енергія електромагнітних хвиль W має середнє значення $\langle W \rangle \neq 0$ і міру флуктуації $\langle \Delta W^2 \rangle$, тоді як інші характеристики електромагнітних хвиль E, H із середніми значеннями $\langle E \rangle = \langle H \rangle = 0$ мають флуктуацію поля $\langle W \rangle \neq 0$ та флуктуацію флуктуацій $\langle \Delta W^2 \rangle$. Наведений приклад засвідчує наскільки важливим є коректний опис систем з багатьма частинками, навіть з точки зору методом опису їх статистичних характеристик.

Для макроскопічної системи адитивними фізичними величинами є енергія, імпульс, маса і кількість частинок. Так, середнє значення енергії системи з N частинок дорівнює: $\langle W_N \rangle = N \langle W_1 \rangle$. Тоді похибка $\Delta W = W - \langle W \rangle$, а її квадрат $(\Delta W)^2 = (W - \langle W \rangle)^2$, звідки середнє $\langle (\Delta W)^2 \rangle = \langle (W - \langle W \rangle)^2 \rangle = \langle W^2 \rangle - \langle W \rangle^2$. Тоді відносне відхилення дорівнює $\varepsilon_w = \frac{\sqrt{\langle (\Delta W)^2 \rangle}}{\langle W \rangle} = \sqrt{\frac{\langle W^2 \rangle}{\langle W \rangle^2} - 1}$. Оскільки $\langle W_N \rangle = N \langle W_1 \rangle$, то $\varepsilon_w = \frac{1}{\sqrt{N}} \sqrt{\frac{\langle W^2 \rangle}{\langle W_1 \rangle^2} - 1}$. Отже, при $N \rightarrow \infty$ відносні флуктуації адитивних величин макросистеми прямують до нуля. Фізично, цей висновок підтверджує те, що рівноважний стан найбільш стабільний.

Моделювання методом випадкової координати і випадкової сили. При програмуванні випадкових процесів для генерування ВВ використовується давач випадкових змінних. У відношенні до моделювання броунівського руху в якості зміщення задається величина $\Delta x = d\ell \cdot (2.0 \cdot \text{rnd}(1) - 1.0)$, де rnd - квазівипадкове число, що генерується комп'ютером, $d\ell$ - параметр випадкового «блукання» - параметр кроку. Тоді випадкова координата задаватиметься як $x := x + d\ell \cdot (2 \cdot \text{rnd}(1) - 1)$. Тут ВВ розподілені рівномірно в інтервалі $[0;1]$, тому розподіл величин Δx також рівномірний в інтервалі

$$[-d\ell; +d\ell], \quad \frac{dP}{d\Delta x} = \begin{cases} 0 & \text{при } |\Delta x| > d\ell, \\ \frac{1}{2d\ell} & \text{при } |\Delta x| \leq d\ell. \end{cases}$$

Середній квадрат зміщення для

одного кроку дорівнює: $\langle (\Delta x)^2 \rangle = \int_{-\infty}^{\infty} (\Delta x)^2 \frac{dP}{d\Delta x} d\Delta x = \frac{1}{3} d\ell^2$. Розглянемо метод випадкової сили. Якщо броунівська частинка зазнає опору $F = r \cdot dx/dt$, то характерним може бути час гальмування τ , оцінений як: $\frac{1}{\tau} m \frac{dx}{dt} \approx r \frac{dx}{dt}$, звідки $\tau = m/r$ [1.6]. Отже, випадкове зміщення броунівської частинки розглядається в інтервалі руху $\Delta t \ll \tau$.

В той же час, величина Δt не дуже мала, щоб навіть в розрідженому газ і кількість співударів між броунівськими частинками була порівняно велика $L \gg 1$. Тоді зміна імпульсу частинки $\Delta p = \langle \Delta p \rangle + \Delta p_{\text{вин}}$. Перший доданок описує зміну

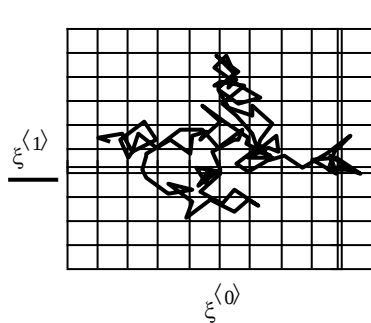
імпульсу $\Delta p = m\Delta g$ за час взаємодії броунівської частинки з іншою протягом часу Δt . Оскільки сила $F = \Delta p / \Delta t$, то середнє значення $\langle \Delta p \rangle$ визначає силу тертя: $r g = \langle \Delta p \rangle / \Delta t$. Другий доданок зумовлений тим, що за рахунок хаотичного теплового руху $m\langle g^2 \rangle / 2 = 3k_B T / 2$ в системі броунівських частинок встановлюється середня рівновага між випадковими поштовхами, які між ударами прискорюють частинку, і гальмуючою дією сил тертя, внаслідок чого встановлюється швидкість $\sqrt{\langle g^2 \rangle}$. Тоді $\langle \Delta p \rangle_{\text{вун}} = m\sqrt{\langle g^2 \rangle} / \Delta t$, добо ця добавка пропорційно зростає з температурою і з коефіцієнтом тертя.

Отже, щоб оцінити випадкову силу, треба оцінити зміну частинки броунівської частинки за час Δt : $g \rightarrow g - \frac{r}{m} g \Delta t + \frac{1}{m} \langle \Delta p_{\text{вун}} \rangle$. Зміна швидкості не може привести до зміни середнього від її квадрату: $\langle g^2 \rangle \rightarrow \langle g^2 \rangle (1 - r \cdot \Delta t / m) + \langle (\Delta p_{\text{вун}})^2 \rangle / m^2$, звідки $\sqrt{\langle (\Delta p)^2 \rangle} = \sqrt{2mr\Delta t \langle g^2 \rangle}$, де знехтувано доданком з малим множником $(\Delta t / \tau)^2 \ll 1$. Приймавши до уваги вираз $\frac{1}{2} m \langle g^2 \rangle = \frac{3}{2} k_B T$ одержуємо випадкову силу $F_{\text{вун}} = \sqrt{6rk_B T / \Delta t}$.

Броунівський рух. 1827 р. Броун спостерігав в мікроскоп суспензію квіткового пилу у воді. Він встановив, що пилинки перебувають в безперервному русі і ці рухи не зв'язані з біологічним походженням частинок або з якимись рухами рідини. Виявилось, що їх здійснюють будь які дрібненькі частинки, завислі в рідині або в газі. Цей безладний рух називається **броунівським** і має свої закономірності. Розглянемо їх.

Як будь яка молекула в газі, броунівська частинка зазнає ударів зі сторони інших рухомих частинок. При цілковитій безладності такого руху можна, здавалося б, сподіватись, що кількість ударів в частинку з однієї сторони дорівнює кількості ударів з іншої і завдяки хаотичності в середньому вона з усіх сторін однакова.

Однак в системі з багатьма частинками можливі флуктуації фізичних величин від середніх значень. З цих позицій броунівські рухи^{*)} – це не молекулярні рухи. Завдяки тому, що кількість ударів молекул рідини об частинку з різних сторін має випадковий характер, у певному напрямку виникає деяка рівнодійна сила. Оскільки флуктуації короткочасні, то через короткі проміжки часу напрям рівнодійної зміниться, а разом із нею зміниться напрям руху частинки. Саме в цьому суть хаотичності броунівських рухів, фрагмент якого зображений схематично на рис.3.6.1а, змодельований в MathCAD за алгоритмом, наведеного на рис.3.6.1б.



а

```

n := 100
Brown(n, step) :=
    step := md(1)
    x0,0 ← 0
    x0,1 ← 0
    for i ∈ 1..n - 1
        xi,0 ← xi-1,0 + step · cos(2 · π · md(1))
        xi,1 ← xi-1,1 + step · sin(2 · π · md(1))
    x
    ξ := Brown(n, step)

```

б

Рис.3.6.1

Нехай в початковий момент всі N частинок розташовані на вісі Y і зміщуються послідовно кроками Δx вздовж вісі X . Кожний крок такої частинки вибирається випадковим чином і незалежно від інших кроків. Однак розподіл ймовірностей при виборі довільного кроку один і той же.

Прийmemo, що зміщення в протилежні напрямки рівно ймовірні. Це означає, що середнє значення зміщення частинки $\langle \Delta x \rangle$. Суть цього твердження полягає в тому, що із збільшенням кількості частинок ця величина залишаючись малою насправді не дорівнює нулеві. Інколи так і середні називаються **апріорними**. Крім апріорних середніх розглядають спостережувані середні – середнє арифметичне

^{*)} Молекулярно-кінетичне пояснення броунівського руху дав Ейнштейн. Незалежно від нього, математичну теорію броунівського руху розробив Смолуховський, того часу професор Львівського університету

для необхідно розглядати середнє значення задної кількості частинок $\langle \Delta x \rangle_e$, яке мале, але не дорівнює нулеві.

Нехай після k -го кроку координата деякої частинки $x(k)$. Тоді $x(k+1) = x(k) + \Delta x$, яке після чергового усереднення набере вигляду $\langle x(k+1) \rangle = \langle x(k) \rangle$, тобто середнє значення $\langle x(k) \rangle$ не змінюється від кроку до кроку і, отже, рівне нулю $\langle x(0) \rangle = 0$.

Ширину інтервалу, в якому розподіляються частинки після k -го кроку, зручно характеризувати середнім від квадрату координати зміщення $\langle x^2(k) \rangle$. Щоб його знайти, визначимо середнє від квадрату: $\langle x^2(k+1) \rangle = \langle x^2(k) \rangle + 2\langle x(k) \cdot \Delta x \rangle + \langle (\Delta x)^2 \rangle$.

Оскільки $x(k)$ і Δx незалежні, то $\langle x(k) \cdot \Delta x \rangle = \langle x(k) \rangle \cdot \langle \Delta x \rangle = 0$ і, якщо позначити $\langle (\Delta x)^2 \rangle = \zeta^2$, то маємо: $\langle x^2(k+1) \rangle = \langle x^2(k) \rangle + \zeta^2$, тобто середній квадрат координати зростає з кожним кроком на величину ζ^2 . Тоді можна записати, що $\langle x^2(k) \rangle \cong k\zeta^2$,

тоді як середнє від квадрату спостережуваного значення $\langle x^2 \rangle_e = \frac{1}{N} \sum_{j=1}^N x_j(k)$ змінюється пропорційно кількості кроків.

Усереднена координата броунівської частинки в середовищі з температурою T дорівнює $\langle x^2 \rangle = \frac{k_B T}{r} t$. Ця формула справджується для усереднення за всіма частинками і всіма в просторі напрямками. Коефіцієнт пропорційності між $\langle x^2 \rangle$ і t у вигляді $D = \frac{k_B T}{r}$ за суттю виражає коефіцієнт дифузії. Аналогічну $\langle x^2 \rangle = \frac{k_B T}{r} t$ формулу одержали Смолуховський і *Ейнштейн*, допустивши повну хаотичність броунівського руху молекул. Ідею перевірки цього постулату запропонував Ланжевен. Суть її полягає в тому, що положення кінців зміщень молекул будуть відтворювати картинку, яка нагадуватимемо відповідну під час обстрілу мішені списками.

Статистка похибок Аллана. Нагадаємо, що похибки вимірювальних приладів поділяються на детерміновані і випадкові. Детерміновані похибки описуються математичною моделлю, яка пов'язує між собою вихідний і вхідний

сигнали датчика. Випадкові похибки, як зазначалось вище, зумовлені різними шумами, природа яких оцінюється шляхом аналізу варіацій Аллана $\sigma^2(\tau)$.

В гіроскопі вихідний сигнал від датчика повинен залежити лише від кутової швидкості Ω , причому ця залежність повинна бути строго визначена і описуватись рівнянням вигляду $U_{вих} = K_1\Omega + K_0$, де $K_{0,1}$ - нульовий і масштабний коефіцієнти, відповідно. Однак, насправді в реальних пристроях вона нелінійна, яку можна записати як функцію змінної Ω , випадкової змінної ξ та прискорення a : $U_{вих} = F(\Omega, \xi, a)$ за допомогою полінома канонічного вигляду:

$U_{вих} = K_0 + K_1\Omega^1 + K_2\Omega^2 + \dots + K_n\Omega^n$. Переважно вважається, що максимальна ступінь нелінійності не повинна перевищувати 0.2%, тобто $U_{вих} \cong K_0 + K_1\Omega^1 + K_{Nelin}\Omega^2$, де K_{Nelin} - приймає значення від $-K_{Nelin,max}$ до $+K_{Nelin,max}$, а $K_{Nelin,max} \leq 0.2\%$.

Вихідний сигнал сенсора інерційного датчика залежить не лише від кутової швидкості, але від прискорення. Причому прискорення діє вздовж вісі чутливості і перпендикулярно до неї, тому її компоненти мають різні впливи і вихідний сигнал виражається як $U_{вих}(a) = K_{ax}a_x + K_{ay}a_y$. Реальний електричний сигнал насправді має декілька складових, як білий шум, тренд (повільна нерегулярна складова, яка переважно моделюється алгебраїчним поліномом невисоких порядків) нульового сигналу, нестабільність нульового сигналу типу фліккер або $1/\nu$, які можуть бути зумовлені механікою чутливого елемента, електронікою схеми та схемами обробки сигналів. Суперпозиція цих випадкових складових фактично виражає статистику шуму в вихідному сигналі:

$U_{вих}(\xi) = \xi_{тренд} + \xi_{нестаб} + \xi_{шум}$, тому в цілому вихідний сигнал дорівнює

$U_{вих} = K_0 + K_1\Omega + K_{Nelin}\Omega^2 + K_{ax}a_x + K_{ay}a_y + \xi_{тренд} + \xi_{нестаб} + \xi_{шум}$. Встановлено [3.8;3.9],

що на вихідний сигнал гіроскопа в основному впливають чотири типів шумів: випадкове блукання кута (angle random walk- ARW), нестабільність нульового сигналу (bias instability – BI), випадкове блукання кутової швидкості (rate random walk -RRW) і дрейф кутової швидкості (rate ramp -RR). Якщо прийняти до уваги шум квантування (quantization noise), який переважно виникає при перетворенні аналогового сигналу в дискретний, і вважати, що ці п'ять джерел шумів статистично незалежні, то їх параметри можна визначити варіацією Аллана, як

суму дисперсій кожного типу похибок:

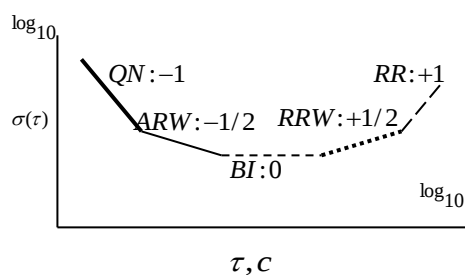
$$\sigma^2(\tau) = \sigma_{QN}^2(\tau) + \sigma_{ARW}^2(\tau) + \sigma_{BI}^2(\tau) + \sigma_{RRW}^2(\tau) + \sigma_{RR}^2(\tau). \text{ Із-за різних асимптотичних}$$

властивостей кожної із складових, в логарифмічному ($\log-\log$) масштабі вздовж вісі τ нахили графіків $\sigma^2(\tau)$ будуть різні (рис.3.6.2,а).

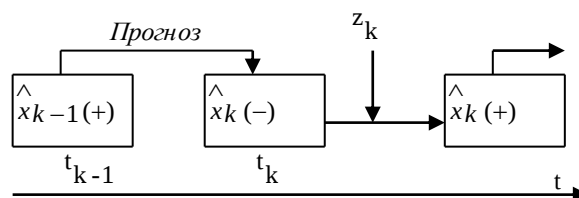
В загальному випадку, для аналізу шумових складових вихідного сигналу із вимірювального пристрою будується в логарифмічному масштабі $\sigma(\tau)$ -графік

зміни квадратного кореня із варіації Аллана $\sigma^2(\tau) = \frac{1}{2(n-1)} \sum_{j=1}^{n-1} (x_{j+1} - x_j)^2$, який

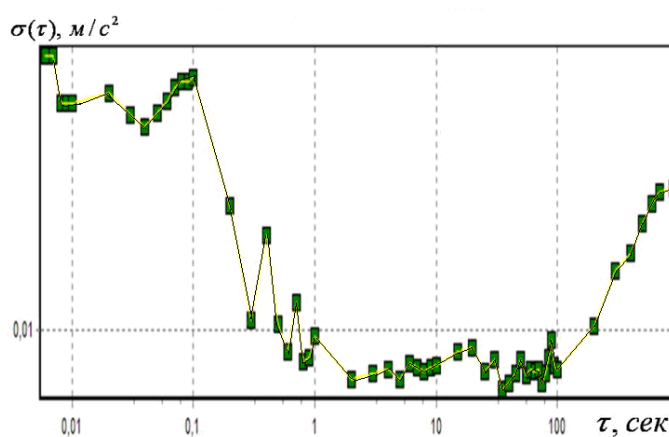
моделюється поліномом $\sigma^2(\tau) \cong p^2(\tau) = R^2 \frac{\tau^2}{2} + K^2 \frac{\tau}{3} + B^2 \frac{2}{\pi} \ln 2 + N^2 \frac{1}{\tau} + Q^2 \frac{3}{\tau^2}$.



а



б



в [3.8]

Рис.3.6.2

Значення коефіцієнтів R, K, B, N, Q полінома характеризують інтенсивність шумових складових вихідного сигналу (ВС), які визначаються вони методом найменших квадратів. В таблиці 3.6.1, подані співвідношення між параметрами розкладу варіації Аллана і спектральної щільності шуму $S(\nu)$, яка з дисперсією

$$\sigma^2(\tau) \text{ зв'язана співвідношенням } \sigma^2(\tau) = 4 \int_0^{\infty} S(\nu) \frac{\sin^4(\pi \tau \nu)}{(\pi \tau \nu)^2} d\nu.$$

Таблиця 3.6.1

Тип похибки в термінах $\frac{S(\nu)}{\sigma^2(\tau)}$	Спектральна густина шуму $S(\nu)$	Варіація Аллана $\sigma^2(\tau)$	Коефіцієнт полінома $p^2(\tau)$	Нахил $\sigma(\tau)$ - графіка
1. Білий шум/ шум квантування ВС	$S(\nu) = (2\pi\nu)^2 Q^2 \tau_0$ при $\nu < 1/2\tau_0$	$\sigma_{QN}^2(\tau) = 3Q^2 / \tau^2$	Q -похибка квантування	-1
2. Білий шум/випадкове блукання ВС	$S(\nu) = N^2$	$\sigma_{ARW}^2(\tau) = N^2 / \tau$	N -коефіцієнт випадкового блукання ВС	-1/2
3. Флікер-шум/ нестабільність зміщення нуля	$S(\nu) = \frac{B^2}{2\pi\nu}$ при $\nu < \nu_0$; 0 при $\nu > \nu_0$	$\sigma_{BI}^2(\tau) = \frac{2B^2}{\pi}$ $[\ln 2 - \frac{\sin^3 x}{2x^2} \cdot$ $(\sin x + 4x \cos x) +$ $+ Ci(2x) - Ci(4x)]$	B -коефіцієнт нестабільності зміщення нуля ВС	0
4. Білий шум вимірювання швидкості зміни ВС/ випадкове блукання зміни ВС	$S(\nu) = \frac{K^2}{2\pi\nu^2}$	$\sigma_{RRW}^2(\tau) = \frac{K^2 \tau}{3}$	K -коефіцієнт випадкового блукання швидкості зміни ВС	+1/2
5. Шум уходу ВС/тренд ВС	$S(\nu) = \frac{R^2}{(2\pi\nu)^3}$	$\sigma_{RR}^2(\tau) = \frac{R^2 \tau^2}{2}$	R -коефіцієнт випадкового уходу зміни ВС	+1
6. Синусоїдний шум		$\sigma_S^2(\tau) = \omega_0^2 \left(\frac{\sin^4(\pi\tau\nu)}{(\pi\tau\nu)^2} \right)^2$		-1

Наведемо алгоритм ви значення статистики Аллана:

$$\sigma^2(\tau) = \frac{1}{2\tau^2(N-2n)} \sum_{k=1}^{N-2n} (x_{k+2n} - 2x_{k+n} + x_k)^2, \text{ де } N - \text{кількість вимірювань, } n - \text{кратність}$$

інтервалу усереднення даних ($n=1,2,3,\dots$; при $n \leq (N-1)/2-1$), $\tau = n \cdot \tau_0$ - інтервал

усереднення даних, τ_0 - параметр приладу, $t = k \cdot \tau_0$ - дискретні моменти

вимірювань ($k=1,2,3,\dots,N$), θ - значення вихідного сигналу з усередненням на

інтервалі τ [3.10]. Залежність квадратного кореня із варіації Аллана від інтервала

усереднення для ММА для нульового значення вимірюваного прискорення

наведена на рис.3.6.2, в, який відображає динаміку блукання нульового сигналу.

В інтернеті відкритий доступ до програми *AlaVar* 5.2 обробки масиву даних методом варіацій Аллана [3.9].

Статистична мінімізація шумів інформаційних сигналів. Всі вимірювальні систем видають не точне значення інформаційного сигналу, а з деякою похибкою. Тому важливим завданням є мінімізація шумової складової сигналу. Одні з таких методів статистичні і ґрунтуються на мінімізації

функціоналу $\sum_i (x_i - a)^2$ МНК. Однак такі вимагають проведення значного об'єму обчислень, а відповідна модель нестійка до дії яких-небудь факторів.

Ефективнішим є фільтрація Калмана. Відповідний алгоритм відноситься до послідовних рекурсивних алгоритмів оцінки і спроможний функціонувати в режимі реального часу та підлаштовуватись у випадку неточно вибраних параметрів математичної моделі. Фільтр Калмана працює за принципом прогноз-корекція, як це показано на рис.3.5.1,б. Так, нехай в момент часу t_{k-1} одержана оцінка вектора стану $\hat{x}_{k-1}$ і треба отримати оцінку в момент t_k . Для цього будується прогноз оцінки $\hat{x}_{k-1}(-)$, ґрунтуючись на $\hat{x}_{k-1}$ одержуємо вимірювання z_k , а потім коректуємо оцінку в момент t_k , спираючись на прогноз і вимірювання, і одержуємо кінцеву оцінку вектора стану $\hat{x}_{k-1}(+)$ [3.11].

3.7. Метод статистичного моделювання Монте-Карло [3.5]

Використання МДЕ для моделювання наносистем обмежене межами просторово-часових масштабів. В ММК відсутня потреба враховувати повний ансамбль частинок із врахуванням їх взаємодії. Замість цього, перехід до нового стану здійснюється шляхом розігрування випадкових величин, які відображають реальну природу фізичного процесу. Розглянемо основну суть ММК.

Ще Фермі в 1930 р., а пізніше в 1940 р. фон Нейман і С.Улам в Лос-Аламосі успішно використали зв'язок між стохастичними (випадковими, ймовірними) процесами та ДР «в зворотному напрямку» і запропонували використати стохастичний підхід для апроксимації багатовимірних інтегралів у рівняннях, що описують дифузійний перенос нейтронів в середовищі.

Нагадаємо, що у фізиці до стохастичних відносять процеси, що описуються функцією від двох змінних $f(x,t)$, одна з яких є часова змінна t , а іншою – стохастична змінна - X . Саму функцію $f(x,t)$ називають випадковою функцією. Якщо замість X підставити будь-яке з можливих її значень x , то отримаємо

функцію від t : $Y_x(t) = f(x, t)$, яку ще називають вибірковою функцією або функцією реалізації процесу. Стохастичний процес уявляють як «ансамбль» вибірових функцій.

За ідеєю Улама, оцінити значення означеного інтегралу $\int_a^b f(x)dx$ в інтервалі $[a; b]$ означає обчислити величину $h(b-a)n_k/n$, де $h(b-a)$ - площа прямокутника, всередині якого лежить графік функції $f(x)$, n_k/n - відношення долі випадкових точок під кривою $f(x)$, до загальної кількості n випадково згенерованих пар із координатами x_i ($a \leq x_i \leq b$) і y_i ($0 \leq y_i \leq h$) з рівномірним розподілом. Для кількості розбиттів $n \rightarrow \infty$ інтеграл прямує до точного значення середнього. Для квантової системи ФРЩ $f(\vec{r}) = |\psi(\vec{r})|^2$.

В класичному підході, принцип інтегрування реалізується методом прямокутників, в якому точки $x_1, x_2, \dots, x_n$ задаються як $x_i = x_1 + (i-1)\Delta x$, а площа обчислюється за формулою $F(a, b) = \int_a^b f(x)dx \approx \sum_{i=1}^n f(x_i)\Delta x_i$. З точки зору ММК, точки $x_1, x_2, \dots, x_n$ - вибираються випадково на інтервалі $[a; b]$, а площа обчислюється як $F(a, b) = (b-a) \frac{1}{n} \sum_{i=1}^n f(x_i)$.

Важливо застосовуючи ММК знати алгоритм моделювання ВВ. Поширеним є метод обернених функцій. Дійсно, нехай треба одержати значення ВВ ξ , що розподілена в інтервалі $[a; b]$ із щільністю ймовірностей $f(x)$.

Стандартний метод моделювання ґрунтується на тому, що інтегральна функція розподілу довільної неперервної ВВ рівномірно розподілена в інтервалі $[0; 1]$,

тобто для довільної ВВ x з щільністю розподілу $f(x)$ ВВ $F(y) = \int_a^y f(x)dx$

рівномірно розподілена в інтервалі $[0; 1]$. Тоді ВВ ξ з довільною щільністю розподілу $f(x)$ можна знайти так: **1.** Генеруємо рівномірно розподілену ВВ

$r \in [0; 1]$. **2.** Покладаючи $F(\xi) = \int_a^\xi f(x)dx = r$ і розв'язуючи рівняння $\xi = F^{-1}(r)$,

знаходимо значення ξ . Графічна суть методу обернених функцій наведена на

рис.3.7.1,а. Для генерування ВВ x з відомим розподілом щільності $p(x)$ використовується метод режекції (відкидання). Суть якого відображена на рис.3.7.1,б.

Важливо в ММК контролювати оцінку похибок обчислень. Так, нехай N ВВ нерівномірно розподілені в об'ємі V і мають значення $x_1, x_2, \dots, x_n$. Тоді за основною суттю ММК $\int f(x)dV \approx \langle f \rangle V \pm V \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{N}}$. За визначенням математичне сподівання або середнє значення $\langle f \rangle = \frac{1}{N} \sum_{i=1}^N f(x_i)$ і $\langle f^2 \rangle = \frac{1}{N} \sum_{i=1}^N f^2(x_i)$ і за основною статистичною теоремою $\sigma_x = \frac{S}{\sqrt{N}}$, де $S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2}$. Оскільки, тому $\sigma_x \approx N^{-1/2}$. В межах $[\langle x \rangle \pm \sigma_x]$ з достовірністю 68.27% перебуває шукана величина; в межах $[\langle x \rangle \pm 2\sigma_x]$ - 95.45%, $[\langle x \rangle \pm 2\sigma_x]$ - 99.73%.

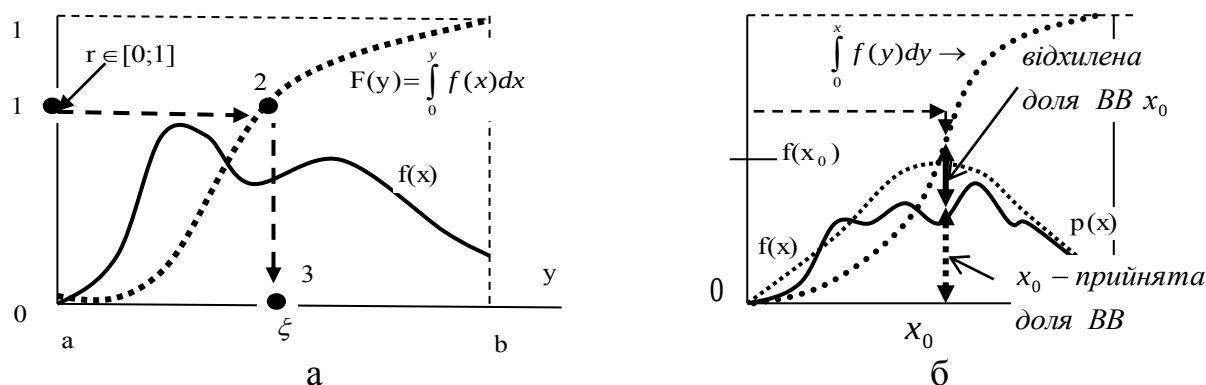


Рис.3.7.1

Пряме моделювання ММК. Пряме моделювання ММК будь-якого фізичного процесу передбачає моделювання поведінки окремих елементарних частин фізичної системи. За суттю пряме моделювання близько до розв'язування задач з перших принципів, як в методі МД, де система описується через поведінку її елементарних складових частин і полі емпіричного потенціалу міжчастинкової взаємодії. Приклади прямого моделювання ММК є : - моделювання опромінення твердих тіл іонами в наближенні бінарних зіткнень; - розріджених газів; - моделювання процесів молекулярно-пучкової епітаксії кінетичним ММК.

Квантовий метод Монте-Карло . Квантовий ММК широко застосовується для дослідження складних молекул і твердих тіл. Квантовий ММК ґрунтується на варіаційному принципі, який за суттю є чисельною інтеграцією багатовимірних інтегралів. Їх треба розв'язувати у зв'язку із розв'язуванням рівняння Шредінгера. Іншим різновидом ММК є статистичне моделювання дифузного процесу.

Кінетичний ММК. Одночастинковий ММК використовується для моделювання електронних процесів в н/п. [3.12]. Ідея підходу полягає в тому, що рух електронів між процесами розсіювання моделюється в імпульсному $\vec{k}$ - просторі. Моменти часу, в які відбуваються акти розсіювань, визначаються випадково і вони залежать від сумарної частоти розсіювань: $\Gamma(W) = \sum_{i=1}^n \nu_i(W)$, де $\nu_i(W)$ - частота для різних механізмів розсіювання, $W = (\hbar k)^2 / 2m$.

Повна ймовірність розсіювань електронів в інтервалі dt дорівнює $p \cdot dt = \Gamma \cdot \exp(-\Gamma \cdot t) dt$. Після того, як обчислений час вільного пробігу t , треба встановити процес розсіювання, що перериває вільний пробіг. Імовірність попадання i -го процесу залежить від енергії електрона і пропорційна $\nu_i(W)$, причому $\frac{1}{\Gamma} \sum_{i=1}^n \nu_i(W) = 1$. Кінетичний ММК ґрунтується на розв'язуванні класичних кінетичних рівнянь. Точний аналіз перенесення електронів через канал транзистора нанорозмірів здійснюється з рівняння Шредінгера.

Суть квантового підходу ММК полягає в тому, що для розрахунку багаточастинкових квантових систем із сильною взаємодією, електрони випадковим чином розкладають за станами і знаходять мінімум енергії відкритої квантової системи, до якої належить канал транзистора [3.12]. Статистична похибка ММК від кількості випробувань оцінюється як $1/\sqrt{n}$. На величину похибки не впливає вимірність задачі, тому коли в кінцево-різницевому методі число обчислень при переході від одновимірної моделі до тривимірної зростає на два порядки, то в ММК вона не змінюється.

4. Методи статистичної оптимізації мікросистем

4.1. Елементи математичних основ оптимізації мікросистем

Вступні поняття. Використання мікро- і наносистем, наприклад в космічній техніці, висуває підвищені вимоги до оптимізації розташування елементів, трасування електричних провідників, тощо, що дає можливість мінімізувати взаємний вплив різних фізичних ефектів, особливо в 3D-компановці сенсорних блоків. В електричних сенсорах похибка ширини зазору між електродами складає $10 \div 100 \text{ мкм}$ і його зміна навіть на незначну величину як 0.1 мкм може спричиняти суттєву похибку вимірювань. Тому під час їх проектування актуальним є питання оптимізації конструкції сенсора приймаючи до уваги вплив різних факторів, перш за все температурні зміни лінійних розмірів електродів. В ємнісних сенсорах приєднувані ємності значно переважають власні ємності сенсорів. Тому щоб уникнути порушення градування із-за дрейфу нуля, треба забезпечити якісну електричну ізоляцію електродів один від одного та від землі.

З іншої сторони, на етапі проектування електричних сенсорів часто виникає потреба оптимізувати протирічиві вимоги – малих габаритів і великої початкової ємності, як це має місце в сенсорах тиску. Стала інтегрування зміни заряду на вимірювальних електродах пропорційна добутку RC , тому збільшення ємності C зменшує вхідний опір давача R , що послаблює вимоги до якості електричної ізоляції. В сенсорах руху треба врахувати малі значення пружності реактивних сил електростатичного притягання між пружно-динамічними електродами.

Аналіз фізичних принципів роботи сенсорів руху свідчить, що компенсувати сили інерції і моменти можна найкраще досягнути в конструкціях чутливого елемента маятникового типу «коромисло», оскільки в них у різні сторони відносно вісі підвісу рознесені інерційні елементи. Крім цього, для конструкцій типу «коромисло» найбільш ефективно застосування електростатичного перетворювача. Однак в електродинамічних акселерометрах з електростатичним зворотним зв'язком треба враховувати суттєвий недолік – невелику за величиною силу зворотного зв'язку, що обмежує діапазон вимірювань. Тому сформульовані особливості електрично-ємнісних сенсорів

поряд з їх високою чутливістю $\sim 500 \frac{B}{мм}$ за умови малих зміщень рухомих електродів $0.2 \div 0.4 мм$ вимагають проведення попередньої перед виготовлення оптимізації їх конструкції, щоб одночасно забезпечити малу інерцію, розміри, тощо.

В сенсорах механічного руху амплітуда коливань системи визначається демпфуючими властивостями, підвіски, тоді як демпфуючі властивості визначаються модулем Юнга і геометричними розмірами підвіски. Модуль Юнга для конкретного завдання переважно визначається проектувальником мікропристрою, тому для забезпечення оптимальної амплітуди коливань інерційної маси оптимізують геометричні розміри. Але в силу об'єктивних причин, не можна як завгодно збільшувати наприклад довжину підвіски, тому застосовують підвіски складної форми, наприклад конструкції із ламаних стрижнів. На відміну від прямих стрижнів, де напруження в основному зосереджуються в місцях розташування анкерів, в торсіонах деформації розподіляються більш-менш вздовж всієї довжини. Актуальність задач оптимізації геометричних і масових параметрів пов'язана не лише з особливостями технології проектування і виробництва пристроїв нано- і мікросистемної техніки, вимог до структурної досконалості функціональних матеріалів, високовартісної оцінки кінцевих виробів для організації рентабельного виробництва, але й забезпечення цільової підготовки висококваліфікованих фахівців для цієї галузі, про що свідчать численні закордонні джерела інформації, як [4.1-4.12].

Задачі оптимізації конструкції поділяються на: - визначення оптимального розподілу розмірів елементів при незмінній геометрії і топології конструкції; - оптимізація розмірів елементів і геометрії (форми) конструкції; - оптимізація топології конструкції з подальшим визначенням форми і розмірів елементів. На відміну від оптимізації форми і розмірів силових елементів топологічна оптимізація дозволяє визначити оптимальний розподіл матеріалу при певних силових навантаженнях і граничних умовах.

Задачі оптимізації стрижневих систем формально можна поділити на дві групи. До першої можна віднести задачі оптимізації зовнішніх дій на систему (задачі

оптимального проектування стрижневих конструкцій, як систем із розподіленими параметрами), а другої – задачі пошуку оптимального статичного або динамічного навантаження системи, включаючи задачі оптимізації параметрів систем, як розподіл маси, пружності, площ поперечних перерізів, тощо. Таким чином, в задачах першої групи задаються дії, а ставиться задача визначити такі параметри проектування, за яких мінімізуються перш за все матеріальні витрати (за об'ємом, масою та вартості) без порушення міцністних характеристик конструкцій, тоді як в задачах другої групи задається певний матеріальний ресурс, а визначається об'єм параметрів проектування, за яких мінімізуються навантаження на конструкції за умови не погіршення їх міцністних характеристик.

В більшості задач проектування стрижневих конструкцій параметри оптимального проектування^{*)}, що відповідають області допустимих розв'язків задачі, можуть підлягати обмеженням за міцністю, поперечного перерізу, стійкістю згинних форм, частот и власних коливань, критичної сили, тощо. Ефективний розв'язок відповідних задач став можливим в зв'язку введенням критеріїв оптимальності, теоретичні основи якого були закладені ще в роботах [4.13-4.15], в яких були розглянут прості задачі на прикладі суцільних тіл, що привели до умов оптимальності у вигляді ДР. Задача оптимізації формулюється в варіаційній формі у вигляді рівнянь Ейлера, а розв'язки задачі оптимізації у вигляді ДР представляють оптимальну форму конструкцій (топологічна оптимізація). В подальшому був розроблений більш практичний підхід, в якому умовою оптимальності у вигляді розв'язків має вигляд алгебраїчних рівнянь. Це дозволило спершу формулювати критерії оптимальності, а потім розробляти алгоритм на основі цих критеріїв.

Мета алгоритму оптимізації – проектування об'єкту за системою критеріїв оптимальності для одержання оптимального розв'язку (локального мінімуму). Процедура оптимізації із-за нелінійності задачі ітераційна. На першому етапі здійснюється розрахунок конструкції для визначення характеристик її поведінки під дією прикладених навантажень, тоді як на другому – перетворюються параметри

^{*)} В оптимальному проектуванні вектор параметрів конструкції визнається не довільно, а в межах певних обмежень рівнистого і нерівнистого типів, що не гарантує спів падіння між собою локальних і глобальних екстремумів.

проектування за допомогою рекурентних співвідношень, одержаних на основі критеріїв. В частковому випадку одного критерію або одного домінуючого критерію, рекурентне співвідношення перетворюється в формулу, яка дозволяє визначити оптимальні розміри конструкції. Якщо при цьому внутрішні сили не залежать від розмірів поперечного перерізу стрижня, як це має місце при статичному навантаженні, то одна або декілька ітерацій по вибору розмірів конструкції призводять до оптимального проекту.

Опрацювання та інтерпретація фізичних даних МСТ взагалі відносяться до трудомістких технологій. Часто для їх вирішення вимагається швидке їх опрацювання і знайти оптимальні значення, тому використовуються методи оптимізації. При розв'язуванні інженерних задач оптимальні значення параметрів прийнято називати проектними параметрами.

Методи оптимізації поділяються на детерміновані, стохастичні і комбіновані. Детерміновані – це градієнтні методи послідовних наближень. Їх ідея полягає в русі в напрямку найшвидшого спуску в області розв'язків задачі за алгоритмом: $x_i^{[j+1]} = x_i^{[j]} + \lambda^{[j]} v_i^{[j]}$, де $v_i^{[j]}$ - коефіцієнт що вказує напрям спуску, $\lambda^{[j]}$ - відповідає за довжину кроку. Довжина кроку може бути стала, ділитись на деяке число та обчислюватись формулою найшвидшого спуску. До стохастичних методів відноситься метод Монте-Карло, імітаційний відпал та генетичні алгоритми, які ґрунтуються на випадкових числах.

Задача оптимізації добре відома в математичному аналізі як задача про пошук екстремумів функції (основна задача варіаційного числення [4.16]), яка вперше була сформульована ще Й. Бернуллі про брахістрону- пошуку між двома точками такої траєкторії, вздовж якої час падіння тіла без тертя під дією сили тяжіння мінімальний шляхом мінімізації інтегралу $\int dl$ за визначених граничних умовах. У фізиці такий підхід відомий як перша варіація функціоналу

$$\int_{t_1}^{t_2} (W_{\text{кін}} - W_{\text{ном}}) dt = \int_{t_1}^{t_2} L dt \text{ шляхом прирівняння до нуля, який застосовується для}$$

встановлення оптимального переходу системи із одного стану в інший протягом

деякого проміжку часу, де L - функція Лагранжа. Це відомий принцип Гамільтона.

Задачі оптимізації поділяються на безумовні та умовні з обмеженнями на критерії (цільової функції) пошуку екстремуму. У першому випадку незалежні змінні x і y в функції $f(x, y)$ не зв'язані між собою ніякими співвідношеннями, тому екстремум називається безумовним (unconstrained), тоді як удругому екстремум умовний (constrained).

Цільова функція $f(X)$ є однозначна числова характеристика системи, аргументом якої є її параметри, що підлягають оптимізації. Вектор параметрів $X = \{x_1, \dots, x_k\}^T$, що задовольняє заданим обмеженням і формує екстремум цільової функції, називається оптимальною точкою, а пара X і $f(X)$ представляють собою оптимальний розв'язок.

Якщо графік цільової функції має один екстремум, то функція називається одноекстремальною. Розрізняють прямі і функціональні обмеження. Прямі накладають обмеження на параметри оптимізації, а функціональні на співвідношення між ними, тому екстремум цільової функції умовний.

Для розв'язування одно екстремальних задач оптимізації без обмежень часто використовують симплекс-метод деформованого багатогранника Недлера-Міда [4.13] *). Багатоекстремальна функція має глобальний екстремум і локальні. В одноmodalній функції глобальним екстремумом є локальний.

4.2. Умовна мінімізація за наявності обмежень

В мікросистемній техніці досліджуються задачі оптимізації з обмеженнями, які можуть бути у вигляді рівностей $g_i(x) = 0, i = 1, 2, 3, \dots, m$ та нерівностей

$h_i(x) \leq 0, i = 1, 2, 3, \dots, m$. Обмеження типу $h_i(x) = 0$ називаються активними, а $h_i(x) < 0$ –

Метод Недлера-Міда - це ефективний алгоритм для розв'язування задач параметричної оптимізації автоматичних систем. Він використовує не єдину, а $m + 1$ точок в процесі розв'язування задачі, що знижує його залежність від вибору початкового наближення. Однак, його залежність від рель'єфу цільової функції, а також можливі за циклювання в сідлових точках, і точках локального екстремума обмежує його застосування лише для одноmodalних функцій.

неактивними. В цьому, математича модель умовної мінімізації формулюється у

$$\begin{aligned} & \text{Minimize : } f(X) \\ & \text{Subject to : } g_i(X) \leq 0, \quad i = 1, m \\ & \text{вигляді: } h_j(X) \leq 0, \quad j = 1, p \\ & x_k^{\text{lower}} \leq x_k \leq x_k^{\text{upper}}, \quad k = 1, n \end{aligned} \quad , \text{ де } f(x) - \text{цільова функція (} F(X) -$$

вектор цільових функцій), $X = [x_1, \dots, x_n]^T$ – вектор змінних, значення яких змінюються в інтервалі $[x^{\text{lower}}, x^{\text{upper}}]$ із області визначення

$S^{\text{feasibleregion}} = \{X \in R^n \mid g_i(X) \leq 0, \quad i = 1, m; \quad h_j(X) = 0, \quad j = 1, p; \quad x_k^{\text{lower}} \leq x_k \leq x_k^{\text{upper}}, \quad k = 1, n \}$ із вектором обмежень $g(X), h(X)$. Тоді в скороченому вигляді математична модель оптимізації запишеться як $\text{Minimize}_{X \in S^{\text{feasible region}}} f(X)$. Так, математичну модель оптимізації

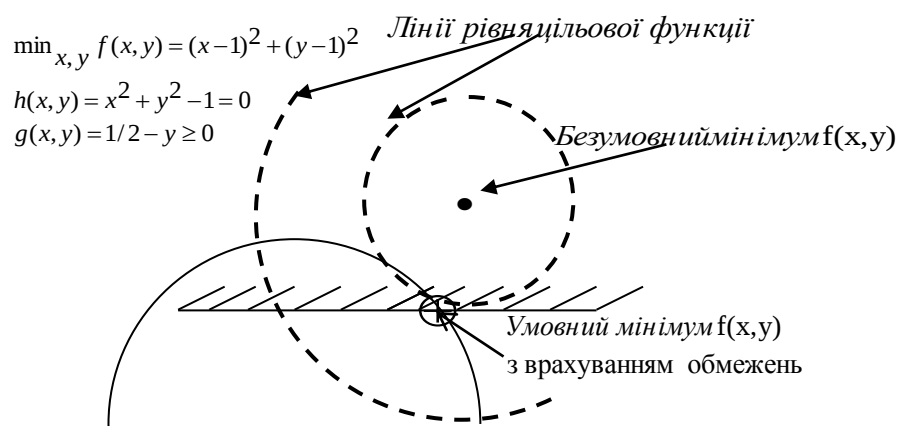


Рис.4.2.1

гладкої поверхні півсфери радіусом R можна сформулювати у вигляді:

$$\begin{aligned} & \text{Minimize : } f(X) \quad 0.5 \int \sigma \epsilon dV \\ & \text{Subject to : } g(X) \leq 0, \quad \text{напруга } \sigma \leq \sigma_{\text{max}} \\ & h(X) = 0, \quad \text{маса } m = m_{\text{def}} \\ & x_k, \quad k = 1, n \quad \text{змінні } l, R, x_k, y_k, z_k \end{aligned} \quad . \text{ Оптимізована поверхня матиме вигляд}$$

композиції із квадратних комірок розміром l . Суть впливу обмежень на цільову функцію зображена на рис.4.2.1[4.13].

Для задач оптимізації МСТ розроблені різні методи, як градієнтний метод [4.14]; метод множників Лангранжа [4.15]; метод ГА [4.18\$4.19]; метод лінеаризації функції суперпозиції дисперсії похибок [4.14], тощо. Метод множників Лагранжа дозволяє звести задачу умовної оптимізації до безумовної оптимізації. Це прямий метод пошуку мінімуму цільові функції, в якому явно і

неявно не використовується інформація про похідні), розв'язування якої знаходиться із необхідних умов існування екстремуму функції.

Розроблені методи чисельної оптимізації, в яких на підставі властивості дисперсії $D[X + Y + Z + \dots] = D[X] + D[Y] + D[Z] + \dots$, мінімізується середнє значення функції похибки, як середньо квадратичної або сумарної квадратичної. Оптимальність функціонування (якість) системи оцінюють за інтегральним СКВ $f(x) = a_1^2 \sigma_{x_1}^2 + \dots + a_n^2 \sigma_{x_n}^2$. Шляхом лінеаризації, нелінійне представлення $f(x)$ замінюють лінійним відображенням $L(x) = a_1 x_1 + a_2 x_2 + \dots + a_n x_n = Ax + b$, де A - матриця виміром $n \times n$. Тоді за певних граничних умов розв'язок x_L системи $L(x) = 0$ виявляється добрим наближенням до x_n - розв'язку рівняння $f(x) = 0$.

4.3. Оптимізація методами покоординатного і градієнтного спуску

Метод по координатного спуску відноситься до групи прямих методів і ґрунтується на багаторазовому застосуванні алгоритмів одновимірної оптимізації. Стратегія методу – поступове наближення до точки мінімуму функції шляхом послідовної варіації одної з координат при фіксованих значеннях інших. Так, якщо в n – вимірному просторі задана точка $X^{(0)}$ з координатами $x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, \dots, x_n^{(0)}$, яка є точкою початкового наближення до мінімуму функції $f(x_1, \dots, x_n)$. Зафіксуємо всі координати, крім першої - x_1 . Отримаємо $f_1(x_1) = f(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, \dots, x_n^{(0)})$ - функцію одної змінної. Для функції $f_1(x_1)$ розв'яжемо задачу одновимірної оптимізації і знайдемо значення $x_1^{(1)}$ - першу координату точки першого наближення до мінімуму. Зафіксуємо тепер усі координати, окрім x_2 , та розв'яжемо задачу одновимірної оптимізації для функції $f_2(x_2) = f(x_1^{(1)}, x_2^{(0)}, x_3^{(0)}, \dots, x_n^{(0)})$. Результатом розв'язування буде $x_2^{(1)}$ - друга координата точки першого наближення до мінімуму. Продовжуючи описаний процес перебору змінних, отримаємо усі координати $x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, \dots, x_n^{(1)}$ точки першого наближення $X^{(1)}$. На цьому перша ітерація алгоритму завершується.

Друга і подальші ітерації алгоритму конструюються аналогічним чином. Ітераційний процес завершується при виконанні умови близькості точок,

знайдених на двох послідовних ітераціях з номерами i та $i+1$: $\sqrt{\sum_{k=1}^n (x_k^{(i+1)} - x_k^{(i)})^2} \leq \varepsilon$,

де ε - мала величина. Задача одновимірної оптимізації в рамках методу координатного спуску може бути розв'язана будь-яким методом, наприклад методом золотого перерізу. Графічно метод по координатного спуску проілюстрований на рис.4.3.1 для мінімізації функції двох змінних $f(x_1, x_2)$, яка описує певну поверхню в тривимірному просторі, що дає можливість навести в

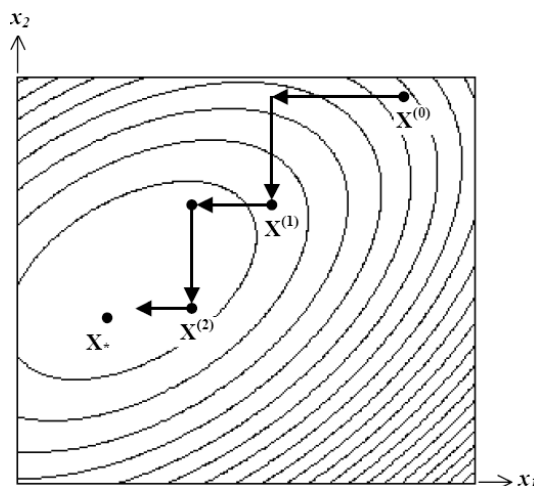


Рис.4.3.1

площині лінії сталих рівнів. Тоді процес мінімізації виглядає так. Точка $X^{(0)}$ задає початкове наближення. Рухаючись спершу паралельно вісі x_1 (спуск за координатою x_1), попадаємо в точку першого наближення $X^{(1)}$. Проводячи на другій ітерації ще два по координатні спуски, опинимось в точці другого наближення $X^{(2)}$, тощо.

Метод градієнтного спуску. Наведемо алгоритм одного із варіантів методу градієнтного спуску (із дробленням кроку): **Крок 0.** Задати параметр точності $\varepsilon > 0$, початковий крок $\alpha > 0$, підібрати $x \in E_n$. Обчислити $f(x)$. **Крок 1.** Знайти $f'(x)$ і перевірити умову досягнення точності: $\|f'(x)\| < \varepsilon$. Якщо вона виконана, то розрахунки завершені, покладаючи $x^* = x$, $f^* = f(x)$. Інакше – перейти до кроку 2.

Крок 2. Зайти $y = x - \alpha f'(x)$ і $f(y)$. Якщо $f(y) < f(x)$, то покласти $x = y$, $f(x) = f(y)$ і перейти до кроку 1, інакше – перейти до кроку 3. **Крок 3.** Положити $\alpha = \alpha/2$ и перейти к шагу 2.

4.4. Метод генетичних алгоритмів

Розроблені методи оптимізації генетичними алгоритмами (англ. genetic algorithm) та відпалу. Основна відмінність цих методів полягає в тому, що в процесі знаходження оптимального розв'язку відпадає потреба обчислювати похідні цільової функції, що найбільше впливає на швидкодію алгоритму. ГА - це евристичний алгоритм пошуку глобального екстремуму цільової функції шляхом випадкового підбору, комбінування та варіації шуканих параметрів з використанням механізмів, що нагадують біологічну еволюцію. Задача формалізується так, щоб її розв'язок можна було б закодувати у вигляді вектора («генотипу») генів, де кожен ген це біт інформації, число або довільний інший об'єкт. У класичних реалізаціях ГА передбачається, що генотип має фіксовану довжину. Однак існують варіації ГА, вільні від цього обмеження.

Певним випадковим чином створюється безліч генотипів початкової популяції. Вони оцінюються з використанням «функції пристосованості», в результаті чого з кожним генотипом асоціюється певне значення («пристосованість»), що визначає наскільки добре фенотип їм описуваний вирішує поставлене завдання. З одержаної множини розв'язків («покоління») із урахуванням значення «пристосованості» обираються розв'язки (зазвичай кращі особини мають велику ймовірність бути вибраними), до яких застосовуються «генетичні оператори» (у більшості випадків «схрещування» - crossover і «мутація» - mutation), що дає можливість отримати нові розв'язки. Для них також обчислюється значення пристосованості, а потім проводиться відбір («селекція») кращих розв'язків у наступне покоління.

Алгоритм роботи ГА зображений на рис.4.4.1: 1. Генерація проміжної популяції (intermediate generation) шляхом відбору (selection) поточного покоління. 2. Схрещування (recombination) осіб проміжної популяції шляхом кросовера (crossover) для формування нового покоління. Критерієм завершення процесу може бути задана кількість поколінь або збіжність (convergence) популяції, коли всі рядки популяції перебувають в області деякого екстремуму і майже однакові. В цьому стані кросовер слабо змінює популяцію, а особи, що призводять до мутації, схильні вимирати, як менш адаптовані. Така збіжність популяції означає,

що досягнутий розв’язок найближчий до оптимального. На завершальному етапі після кросовера і мутації, нове покоління формується двома стратегіями - діти замінюють батьків, та нове покоління формується із дітей та їх батьків на основі принципу елітизму, завдяки якому добрий розв’язок буде залишатись в популяції доти, доки не буде знайдений кращий розв’язок [9].

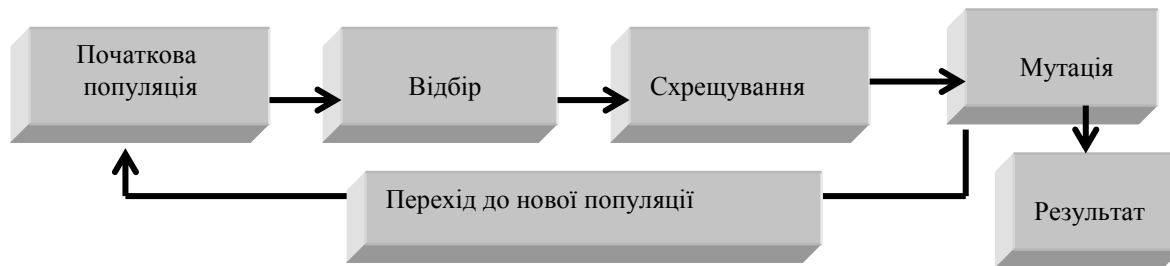


Рис.4.4.1

Для реалізації*) ГА треба певним чином закодувати змінні x_i , $i = \overline{1, n}$, які є компонентом вектора $\bar{x}$, наприклад за допомогою двійкового кодування у вигляді бітов рядка (хромосома), де кожен із елементів є геном, розташованим у різних позиціях. В бінарному рядку ген – це біт цього рядка, локус – позиція у рядку, а аллель – значення гена 0 або 1. Сукупність закодованих x_i , $i = \overline{1, n}$ - це популяція, яка складається з хромосом (особей), які формують в просторі розв’язків точки координатами x_i , $i = \overline{1, n}$.

В більшості, для тестування програмних алгоритмів оптимізації використовують пробну цільову функцію Розенброка: $f(x, y) = 300(y - x^2)^2 + (1 - x)^2$, з областю $|x, y| \leq 10$. Мова йде про умовний мінімум, оскільки при цьому приймаються до уваги обмежуючими співвідношення $g_1 = y + \frac{13}{7}x \leq 0$ і $g_2 = \frac{23}{31}(x - 1.25)^2 - (y - 1.65) \leq 0$, що накладаються на незалежні змінні x і y .

ГА для пошуку глобального мінімуму використовує ймовірністний підхід. Їх особливістю є те, що жоден із генетичних операторів (кросове, мутація,

Реалізація ГА у середовищі Matlab викладена, наприклад, в посібнику [4.20] Детальний опис ГА можна, наприклад, знайти в [4.21]

інверсія) не спирається на інформацію стосовно рель'єфу поверхні функції пристосування. Формування потомків відбувається випадковим чином, тому нема гарантії, що знайдені розв'язки завжди будуть кращі родинних і в процесі еволюції зустрічаються «невдалі» потомки, які збільшують кількість звертань до цільової функції, збільшуючи при цьому час пошуку глобального екстремума.

Відомий метод оптимізації методом відпалу. Він ґрунтується на фізичному ефекті заморожування структури кристалічної гратки матеріалу під час його швидкого охолодження, коли атоми намагаються зайняти найбільш оптимальне положення з мінімальною енергією. При високій температурі атоми рухаються хаотично і сили взаємодії між ними на цей рух впливають слабо. Коли температура знижується, то зменшується кінетична енергія хаотичного руху і актуальними стають сили притягання між атомами. Глобальний мінімум потенціальної енергії взаємодії формується при зниженні температури $T \rightarrow 0$ і відповідає ідеальному кристалу. Отже, якщо температуру знижувати повільно, то частинки будуть рухатись саме в області глобального мінімуму, незалежно від того, де знаходилась початкова точка траєкторії.

З точки зору оптимізації обчислювальних процедур в моделюванні теплового руху, найбільш ефективним є застосування випадкових блукань за алгоритмом Метрополіса, в якому частинка випадково зміщується і весь час порівнюються її енергія до і після зміщення. Якщо нове значення менше попереднього, то система вважається такою, що перейшла в новий стан, у протилежному випадку переміщення відбувається, якщо $r < \beta \cdot \exp(-\Delta W / k_B T)$. Потім процедура знову повторюється. Бачимо, що ймовірність переборення частинкою енергетичного бар'єру, залежить від значення параметра β , значення якого часто покладають рівним $\beta = 1/(k_B T)$. При високих температурах, атом здатен переборювати високі енергетичні бар'єри, тоді як із зниженням температури, зростає ймовірність затискання системи в області конфігураційного простору, що відповідає одному із мінімумів. Під параметрами моделювання «енергія» і «температура» розуміють середнє значення похибки моделі, а температура визначає амплітуду флуктуацій змінних параметрів в кожній ітерації.

На завершення треба відзначити таке. Пошук глобального, найбільш глибокого мінімуму на гіперповерхні відноситься до проблемних задач. В цілому його можна знайти лише шляхом перерахунку всіх її мінімумів. Однак, відповідна процедура може бути досить громіздкою. Так, в процесі пошуку геометричної конфігурації атомів, що відповідає мінімуму потенціальної енергії з функцією Леннарда-Джонса, кількість мінімумів зростає як $\exp(n^2)$: для $n = 13$, їх кількість порядку декілька сотень, тоді як для $n = 20$ - вже декілька мільярдів. Градієнтний метод в цьому випадку не ефективний, оскільки він багатократно знаходить одні і ті ж мінімуми. Методи, що дозволяють усунути лишні точки і підвищити ефективність пошуку, називаються методами кластеризації, одним з яким є саме «метод відпалу».

Векторний характер критеріїв оптимальності створює проблему формування цільової функції. Складність виражається як в кількості критеріїв, так їх взаємодії. Часто покращення одних супроводжується погіршенням інших, тому виникає потреба приймати компромісний (trade-off) розв'язок. В цьому випадку ведуть мову про Парето-оптимальний розв'язок. Розглянемо його.

4.5. Оптимізація багатокритеріальної задачі та крива компромісів Парето

Парето обґрунтував емпіричне правило про те, що *«20% зусиль дають 80% результату, а решта 80% зусиль — лише 20% результатів»*, на підставі якого будується діаграма Парето, яка дозволяє скоротити множину вихідних варіантів, тобто виключити із неформального аналізу ті варіанти розв'язків, які завідома будуть поганими. Принцип Парето дозволяє як розв'язок вибрати лише вектор $\bar{x}$ із множини Парето, але не виділяє єдиний розв'язок, лише скорочує множину альтернативів, що дозволяє полегшити процедуру вибору розв'язку. Якщо є домінуючий над іншими розв'язок, то крива Парето (фронт Парето) стискатиметься до одного. Випуклі границі^{*)} викривлені в сторону кращих розв'язків, а вгнуті — в протилежну.

^{*)} Існує клас задач оптимізації методом випуклого програмування і використанням випуклих функцій. Для них точки локального і глобального екстремумів співпадають, тобто розв'язування таких задач можна звести лише до знаходження локального екстремуму [Батищев Д.И. Методы оптимального проектирования. - М.: Радио и связь, 1984. - 248 с.]

Двохвимірною задачею оптимізації з двома критеріями f_1 і f_2 наведена на рис.4.5.1. Тут вектор характеристик $f(X)$ відображає параметричний простір в просторі цільових функцій. Оскільки $f(X)$ вектор, то будь-які його компоненти є конкуруючі і відсутній деякий єдиний розв'язок задачі. На множині оптимізації, крива Парето (жирна дуга $P(F)$) на множинах значень критеріїв $F = f(X)$ (область

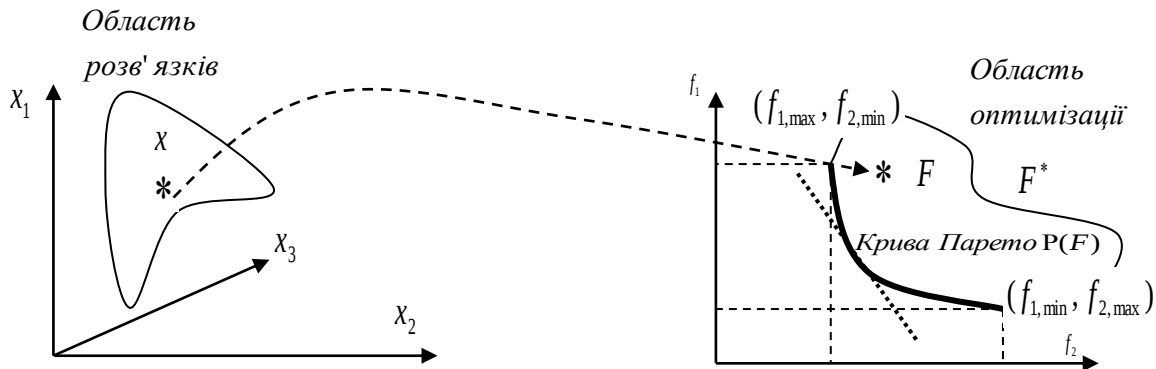


Рис.4.5.1

значень двох критеріїв оптимізації f_1 і f_2 , що відповідають змінним X у допустимій області) є сукупністю оптимальних точок Парето, як множини найкращих розв'язків в розумінні того, що подальшому покращенню вони не піддаються.

З практичної точки зору, інтерес представляє не множина Парето розв'язків $f(x) \in P(F)$, а єдина точка $x \in P(X)$. В цьому випадку застосовують різні методи оптимізації, в тому числі метод зважених сум, за допомогою якого оптимальним із Парето-множини розв'язків вибирають той, що відповідає точці дотику дотичної до кривої $P(F)$. Однак, цей підхід не можна використовувати у випадку, коли область Парето не випукла, оскільки може існувати декілька точок дотику прямої. Тут f - скалярні функції в моделі оптимізації $\min_{x \in S} f(x)$ системою обмежень (в неявному вигляді) S . У випадку однокритеріальної оптимізації область оптимізації має одновимірний вигляд. Наведемо загальний алгоритм програмної реалізації Парето оптимізації в кодах Матлаб^{*)} із двома критеріями :

Технологія оптимізації в кодах Matlab викладена в Messac A. Optimization in Practice with MATLAB. Cambridge University Press. 2015. Режим доступу: <https://books.google.com.ua/books?id=F62wBgAAQBAJ&pg=PA126&lpg=PA126&dq=Matlab+Code+for+Pareto+optimal+Optimization&source=bl&ots=Kg96q8LbSk&sig=MfDERM9FBpLix5dRs2CYhdN3RUUs&hl=ru&sa=X&ved=0CBsQ6AEwADgKahUKEwiOhoaX2aPHAhWDPRQKHWHJC-s#v=onepage&q=Matlab%20Code%20for%20Pareto%20optimal%20Optimization&f=false>

```

x0=[aa1 aa2 aa3 aa4 aa5] lowerb=[bb1 bb2 bb3 bb4 bb5] upperb=[ bb11 bb22 bb33 bb44 bb55]
Parameter1=[ a1 b1 c1 ] Parameter2=[ a2 b2 c2 ] fmin=[min1 min2]
for k=1:...
    fmax=[function(k)];
    [xoptim, foptim]=fminimax(@optim_1,x0,b,[],[],lowerb,upperb,@optim_constr,[], Parameter1 , Parameter2,fmin,fmax);
    for m=1:...
        ff(m)=foptim(m)*(fmax(m)-fmin(m))+fmin(m);
    end
    f1(k)=ff(1);    f2(k)=ff(2);
end
[f2sort,ind2]=sort(f2)
f1sort=f1(ind2)
plot(-f1sort,f2sort)    xlabel    ylabel
Function optim_1:
function f= optim_1 (x, Parameter1, Parameter2,fmin,fmax)
[Parameter1,ksi]=opim_2(x, Parameter1, Parameter2,);
f1=.....minimization function ;    f(1)=(f1-fmin(1))/(fmax(1)-fmin(1));
f2=..... minimization function;    f(2)=(f2-fmin(2))/(fmax(2)-fmin(2));
end
Function optim_2:
function [Parameter1,ksi]= optim_2(x, Parameter1, Parameter2,)
EsigmaI= functional Eq_1;
ksi=functional Eq_2;
end
Function optim_constr:
function [c,ceq]= optim_constr(x, Parameter1, Parameter2,c,fmin,fmax)
[Parameter1,ksi]= optim_2(x, Parameter1, Parameter2,);    c=f(ksi)*x(max(k))+const;
ceq=[];
end

```

4.6. Оптимізація пружних мікросистем кантилеверного типу

Аналіз фізичних умов коливань систем із просторово рознесеними динамічними параметрами, як жорсткості μ кріплення до нерухомої вісі тіла масою інерції m свідчить, що вони здатні здійснювати періодичні коливання з резонансною частотою $\omega_0^2 = \frac{\mu}{m}$. Ця фізична характеристика динамічної системи виражає собою так звану цільову функцію і в переважній більшості на її основі розробляється алгоритм оптимізації конструктивних закономірностей чутливого елемента. Обидва параметри пружність і інерція цільової функції є функціями геометричних параметрів конструкції, змінюючи які можна з одної сторони збільшити термін експлуатації сенсора, не зменшуючи його чутливості та без потреби не підвищуючи при цьому його вартісні характеристики.

На рис.4.6.1, наведені деякі конструкції інерційних елементів, які найчастіше використовуються в мікросенсорах руху. Їх та інші, об'єднує наявність коливної лінійної системи з розподіленими параметрами, для якої пружна сила пропорційна величині самої пружної деформації. Навіть для

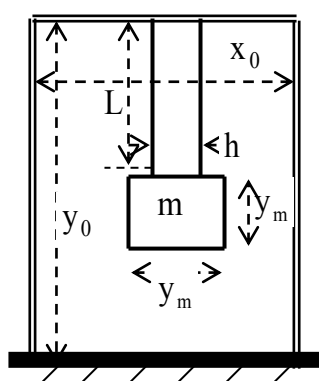
наведених конструкцій чутливих елементів можна зробити висновок про те, що на процес їх проектування накладаються два важливі обмеження, як конкретні значення фізичних параметрів матеріалів, які використовуються, та обмежені розміри замкнутого простору, в якому інерційне тіло рухається. Тому в мікро- і наносистемній техніці оптимізація цільові функції завжди супроводжується з врахуванням рівністних (типу h) і нерівністних (типу g) обмежень геометричних і масових параметрів, що накладаються на конструкцію чутливих елементів без зменшення їх чутливості, втрати механічної міцності, тощо. Розглянемо конкретні приклади.

Обґрунтуємо фізичний принцип оптимізації цільової функції акселерометра маятникового плоского типу (рис.4.6.1,а). Цільову функцію виберемо у вигляді

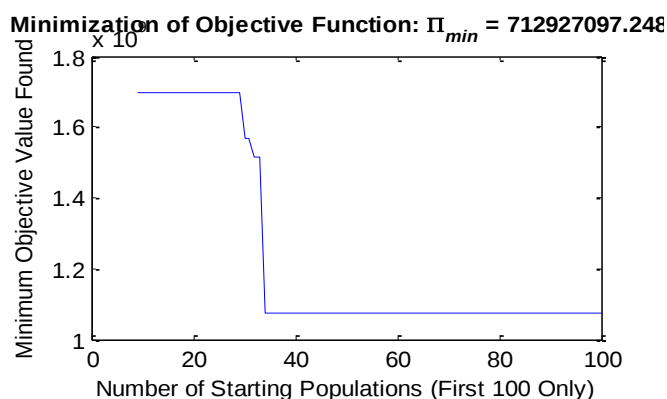
$$obj = \omega_0^2 = \frac{\mu}{m} = \frac{E_\sigma h t^2}{4 y_m^2 L^3}. \text{ Суть обмежень, що накладаються на функцію } obj, \text{ така: } \mathbf{1.}$$

Механічна міцність матеріалу конструкції обмежує мінімальне значення ширини консолі маятника $h_{\min}$, тому для цільової функції нерівністне обмеження формулюється як $g_1 : h_{\min} - h \leq 0$. **2.** Сумарна довжина маятника $L + 2r_m$ не може перевищувати внутрішні розміри платформи y_0 , тому друге нерівністне обмеження матиме вигляд: $g_2 : y_m + L - y_0 \leq 0$. **3.** Аналогічно обґрунтовується обмеження на поперечні розміри маятника: $g_3 : y_m - x_0 \leq 0$. Дві останні нерівності $g_{2,3}$ обмежують масу маятника.

За технічним завданням, акселерометр використовується для вимірювання прискорень, в тому числі ударних. Тому об'єктивними є обмеження на механічну міцність конструкції: $g_4 : \sigma_{\min} - \sigma \leq 0$ та амплітуду коливань маятника: $g_5 : \delta_{\min} - \delta \leq 0$ - обмеження на мінімальне відхилення маятника або чутливість і $g_6 : \delta - \frac{1}{2}(x_0 - y_m) \leq 0$ - обмеження на максимальне відхилення маятника.



а



б

Рис.4.6.1

Коефіцієнт жорсткості консолі дорівнює $\mu = \frac{3E_{\sigma}I}{L^3}$ - це перше рівнісне обмеження типу h_1 на цільову функцію. В цю формулу входить другий момент плоского перерізу $I = \frac{h \cdot t^3}{12}$, тому друге рівнісне обмеження матиме вигляд h_2 :

$I = \frac{ht^3}{12}$ - другий момент поперечного перерізу мікроконсолі. Маса плоского маятника товщиною t та густиною ρ_m дорівнює $m = \rho_m y_m^2 t$, тому одержимо третє

рівнісне обмеження на функцію obj матиме вигляд $h_3: m = \rho_m y_m^2 t$. Підвіска маятника жорстка і в процесі його коливання виникають згинного типу механічні напруження, тому четверте рівнісне обмеження записується у вигляді h_4 :

$\sigma_{bend} = \sigma(E_{\sigma}, h, L, \delta) = \frac{3E_{\sigma}h}{2L^2}$. Отже, на цільову функцію акселерометра маятникового

типу накладаються 6 нерівнісних обмежень та 4 – рівнісних.

Отже, цільова функція Pi представляється функцією від двох геометричних змінних $\{L, r_m\}$. В методі ГА значення кожного з них генеруються випадковим

чином в вибраному дослідником діапазоні $80 \text{ мкм} \leq L_1 \leq 200 \text{ мкм}$
 $50 \text{ мкм} \leq y_m \leq 150 \text{ мкм}$. Нижче наведена

програма моделювання цільової функції [4.7;4.8]:

```
clear
t = 2e-6;
rho = 2.33*100^3/1000;
Esigma = 170e9;
```

```

delta_min = 1e-7;
sigma_max = 1.6e9;
x_d = 1e-3;
y_d = 1e-3;
x_m = 1.8e-6;
x_m_min = 10e-6;
x_m_max = 0.5*x_d;
y_m = 1.8e-6;
y_m_min = 10e-6;
y_m_max = 0.5*y_d;
L = 200e-6;
L_min = 20e-6;
L_max = y_d - y_m_max;
h = 2e-6;
h_min = 10e-6;
h_max = x_d;
I = t*h^3/12;
start_pops = 1000;
iters = 50;
best1 = zeros(iters,start_pops);
best2 = zeros(1,iters);
for idx1 = 1:start_pops
    d = 100;
    clear Length
    random = rand(d,4);
    Length(:,1) = L_min + random(:,1) * (L_max - L_min);
    Length(:,2) = h_min + random(:,2) * (h_max - h_min);
    Length(:,3) = x_m_min + random(:,3) * (x_m_max - x_m_min);
    Length(:,4) = y_m_min + random(:,4) * (y_m_max - y_m_min);
    for idx2 = 1:iters
        L = Length(:,1);
        h = Length(:,2);
        x_m = Length(:,3);
        y_m = Length(:,4);
        m = x_m.*y_m.*b.*rho;
        mu = Esigma*t*h.^3./(4*L.^3);
        sigma = (x_d-x_m)*0.5*Esigma.*h.*(L+y_m/2)./(2*L.^3/3+3*y_m.*L.^2/2+ y_m.^2.*L);%3*(x_d-
x_m)*0.5*E.*h./(2*L.^2);
        obj = mu ./ m;
        Pi = obj;
        for idx3 = 1:d
            if(y_m(idx3)+L(idx3)-y_d > 0)
                Pi(idx3) = inf;
                disp('1')
            end
            if(h(idx3) - h_min <= 0)
                Pi(idx3) = inf;
                disp('2')
            end
            if(x_m(idx3) - x_d > 0)
                Pi(idx3) = inf;
                disp('3')
            end
            if(sigma(idx3) - sigma_max > 0)
                Pi(idx3) = inf;
                %disp('4')
            end
        end
        end
        [sorted_Pi, order_Pi] = sort(Pi);
        if (sorted_Pi(1) == inf)
            %disp('INFINITY')
        end
        best1(idx2,idx1) = sorted_Pi(1);
        nps = 10;
        for i = 1:nps

```

```

    parent(i,:) = Length(order_Pi(i,:),:);
end
random2 = rand(nps,1);
for j = 1:2:nps-1
    child(j,:) = random2(j,1) * parent(j,:) + (1 - random2(j,1))*parent(j+1,:);
    child(j+1,:)=random2(j+1,1)*parent(j,:)+(1- random2(j+1,1))*parent(j+1,:);
end
clear Length_new
random3 = rand(d-2*nps, 4);
Length_newrand(:,1) = L_min + random3(:,1) * (L_max - L_min);
% new L_1
Length_newrand(:,2) = h_min + random3(:,2) * (h_max - h_min);
% new L_2
Length_newrand(:,3) = x_m_min + random3(:,3) * (x_m_max - x_m_min);
% new L_3
Length_newrand(:,4) = y_m_min + random3(:,4) * (y_m_max - y_m_min);
% new y_m
Length_new = vertcat(parent, child, Length_newrand);
% Repeat Genetic Algorithm
Length = Length_new;
end
Length_final(idx1,:) = Length(1,:);
end
best2 = best1(iters,:);
[sorted_best2, order_best2] = sort(best2);
% Output the design variable values for top five
for kk = 1:5
    best_Length(kk,:) = Length_final(order_best2(kk),:);
    fprintf('Rank: %2.0f L = %6.3e h = %6.3e x_m = %6.3e y_m = %6.3e Pi = %6.5f\n',kk,best_Length(kk,1),best_Length(kk,2),best_Length(kk,3),best_Length(kk,4),sorted_best2(kk));
end
% Plot the minimum objective value as a function of generations
for idx4 = 1:100
    if (idx4 == 1)
        Pi_min(idx4) = best2(idx4);

```

Графік поведінки цільової функції для перших 100 популяцій наведений на рис.4.6.1,б. Аналіз залежності цільової функції від розміру початкової популяції свідчить, що із збільшенням її чисельності, середнє значення функції якості зменшується. Це означає, що збільшення розміру популяції альтернативних розв'язків за умови незмінної кількості ітерацій супроводжується збільшенням ймовірності знаходження оптимального розв'язку. Бачимо також, що аналогічна закономірність спостерігається із зростанням кількості ітерацій. Бачимо, що для розглядуваної задачі, оптимальний розв'язок досягається після 30 ітерацій.

Оптимізація за критерієм мінімуму середнього квадрату динамічної похибки, використовується для аналізу реакції чутливого елемента акселерометра на детерміновану ступінчасту дію. Розв'язується така задача з метою формулювання задачі на інформаційну надійність, тобто здатність приладу зберігати робочі характеристики в межах допусків. Саме СКВ вихідного сигналу слугує показником якості акселерометра. Якщо вектор параметрів

сенсора F , $X(t)$ - вектор стану процесів в момент часу t , $U(t)$ - вектор вхідних дій в момент часу t , то цільова функція формулюється у вигляді

$J = Q\{F, X(t), U(t)\}$, то в розглядуваному підході вектор параметрів F змінюють в допустимих межах випадковим чином і обчислюють в кожному статистичному випробуванні

функціонал якості $Q(F)$. Потім за умови досягнення достатнього обсягу статистичного матеріалу визначають екстремум функціоналу і виявляють при цьому оптимальні значення F за методикою [4.22; 4.23].

4.7. Оптимізація пружних мембран

Прямокутна мембран тонка за умови, коли відношення товщини до довжини $\frac{h}{a} < \frac{1}{20}$ [4.24-4.27]. Відношення $\frac{a}{b}$ визначає максимальне напруження $\sigma_{\max} = \beta \frac{pb^2}{h^2}$, максимальний прогин $w_{\max} = \alpha \frac{pb^4}{E_\sigma h^3}$, згинну $\sigma_i = E_\sigma \varepsilon_i = \frac{3E_\sigma h}{2a^2} w_{\max}$ та розтягуючи $\sigma_r = E_\sigma \varepsilon_r = \frac{E_\sigma h}{2a^2} w_{\max}^2$ механічні напруги, де сумарна механічна напруга не повинна перевищувати критичне значення σ_{kp} , а значення α, β подані нижче :

a/b	1	1.2	1.4	1.6	1.8	2.0	∞
α	0.0138	0.0188	0.0226	0.0251	0.0267	0.0277	0.0284
β	0.3078	0.3834	0.4356	0.4680	0.4872	0.4974	0.5000

В лінійних деформацій коефіцієнт жорсткості $\mu \cong 4 \frac{E_\sigma b h^3}{a^3}$, а прогин $w = \frac{p b a^3}{4 E_\sigma h^3}$.

Оптимізуємо пружні властивості в центрі прогину деформованої мембрани, для чого встановимо явний вигляд нелінійної функції $f(x) = c_1^2 \sigma_{x_1}^2 + \dots + c_n^2 \sigma_{x_n}^2$.

Оскільки $\mu \cdot w_{\max} = p a b$, то $w_{\max} = \alpha \frac{p b^4}{E_\sigma h^3}$ і коефіцієнт жорсткості $\mu = \frac{(a/b) E_\sigma h^3}{\alpha b^2}$.

Продиференціювавши формулу для μ одержимо, що явні вирази для СКВ σ

змінних параметрів мембрани дорівнюють: $\partial \mu_h = 3 \frac{(a/b)}{\alpha} E_\sigma \left[\frac{h}{b} \right]^2 \partial h = 3 \frac{(a/b)}{\alpha} E_\sigma \left[\frac{h}{b} \right]^2 \sigma_h$,

$$\partial \mu_b = -3 \frac{(a/b)}{\alpha} E_\sigma \left[\frac{h}{b} \right]^3 \sigma_h \partial b = -3 \frac{(a/b)}{\alpha} E_\sigma \left[\frac{h}{b} \right]^3 \sigma_b, \quad \partial \mu_{E_\sigma} = \frac{(a/b)}{\alpha} \left[\frac{h}{b} \right]^2 h \cdot \partial E_\sigma = \frac{(a/b)}{\alpha} \left[\frac{h}{b} \right]^2 h \cdot \sigma_{E_\sigma} \text{ і}$$

функціонал оптимізації як суперпозиція квадратів СКВ матиме вигляд

$$f = \left(3 \frac{(a/b)}{\alpha} E_\sigma \left[\frac{h}{b} \right]^2 \right)^2 \sigma_h^2 + \left(-3 \frac{(a/b)}{\alpha} E_\sigma \left[\frac{h}{b} \right]^3 \right)^2 \sigma_b^2 + \left(\frac{(a/b)}{\alpha} \left[\frac{h}{b} \right]^2 h \right)^2 \sigma_{E_\sigma}^2.$$

Для моделювання взяті гіпотетичні значення діапазонів змін товщини $h: 5-50 \text{ mkm}$ і ширини $b: 60-1000 \text{ mkm}$ мембрани. Для мінімальних значень товщини і ширини мембрани, коефіцієнт жорсткості $\mu \cong 3400 \text{ kg} \cdot \text{c}^{-2}$, взято до уваги значення модуля Юнга $E_\sigma = 17 \cdot 10^4 \text{ kg} \cdot \text{c}^{-2} \cdot \text{mkm}^{-1}$. Насправді мінімальне значення товщини визначається межею міцності матеріалу, а максимальне необхідною чутливістю величини деформації зі сторони зовнішньої дії. Розрахунки (без врахування розсіювання значень E_σ для кремнію, обчислені за лістингом програми:

```
h=2;b=60;
Esigma = 170*10^3; a_alpha=4;

% Standard deviation
Esigma_rms = 170*10^1;
h_rms = .1; % mkm
b_rms = .5; % mkm
h = 2:2:50; % mkm
b = 60:10:1000; % mkm
Esigma = 170*10^3; % kg*c^-2*mkm^-1
[X,Y] = meshgrid(h,b);
for n=1:length(h)
for m=1:length(b)
h1 = h(n);
b1 = b(m);
mu_v_h = (3*Esigma*a_alpha *h1^2/(b1^2))^2*h_rms^2;
mu_v_b = (3*Esigma*a_alpha *h1^3/(b1^3))^2*b_rms^2;
mu_v_Esigma = (a_alpha *h1^3/(b1^2))^2*Esigma_rms^2;
mu_v_res = mu_v_h + mu_v_b + mu_v_Esigma ;
mu1 = (a_alpha*Esigma*h1^3)/(b1^2);
mu(m,n) = mu1;
% Calculates score for the spring
val(m,n) = mu_v_res;
end % inner for loop
end % outer for loop
figure;
mesh(X,Y,log(val));
xlabel('height (microns)')
ylabel('length (microns)')
hold on
mesh(X,Y,log(mu));
zlabel('log(val),log(mu)')
```

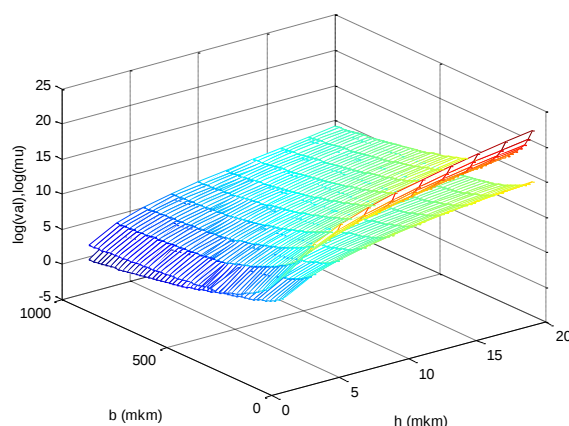


Рис.4.7.1

Значення параметрів функціонування вибираються із відповідних на контурі перетину поверхонь $\log(val)$ і $\log(mu)$ (рис.4.7.1).

4.8. Структурна оптимізація мікросистем методом змінної густини

В моделі тіла із змінною густиною допускається, що модуль пружності і міцнісні характеристики цієї моделі пропорційні її густині: $E = \rho \cdot \bar{E}$, $[\sigma] = \rho \cdot [\bar{\sigma}]$, де $\bar{E}$ і $[\bar{\sigma}]$ - модуль пружності і допустима напруга при одиничній густині – питомі характеристики матеріалу густиною ρ [4.28]. Розрахунок розподілу силового навантаження в суцільному тілі з врахуванням його деформації здійснюється за таким алгоритмом: **1.** Тіло, що піддається деформації, розбивається на скінченні елементи. **2.** Вибирається початкова, відмінна від нуля густина кожного елемента ρ_{0i} (0- номер ітерації, i – номер елемента). **3.** Чисельним методом розраховується розподіл тиску в абсолютно жорсткій конструкції тіла. **4.** Для навантаження по п.3, знаходиться оптимальний розподіл матеріалу в тілі змінної густини у вигляді густин окремих скінченних елементів з врахуванням питомої допустимої напруги реального матеріалу. **5.** За результатами п.4 назначаются по $E = \rho \cdot \bar{E}$ пружні характеристики кожного скінченного елемента. **6.** Розраховуються абсолютні деформації (переміщення) тіла з одержаним новим розподілом матеріалу. **7.** Розраховується розподіл тиску для заданої по п.6 деформованої конструкції. **8.** Одержані в п.7 навантаження передаються в п.4 замість навантажень по п.3, розрахунок повторюється до досягнення стабілізації.

Оптимізація розподілу матеріалу виконується за таким алгоритмом: **4.1.** Для заданого початкового розподілу густин елементів і відповідних модулів пружностей елементів E_{0i} , визначених за $E = \rho \cdot \bar{E}$, розраховується напружений стан тіла. Визначається еквівалентна напруга в кожному елементі $\sigma_{0i}^{екв}$ за прийнятою теорією міцності, за умовою текучості Генкі-Губера-Мізеса, $\sigma_{0i}^{екв} = (\sigma_1^2 + \sigma_2^2 + \sigma_3^2 - \sigma_1\sigma_2 - \sigma_2\sigma_3 - \sigma_3\sigma_1)^{1/2}$, де $\sigma_{1,2,3}$ - головні напруги. **4.2.** Обчислюються нові значення густин елементів ρ_{1i} за формулою: $\rho_{1i} = \frac{\sigma_{0i}^{екв}}{[\bar{\sigma}]}$. **4.3.** Нові значення густин елементів передаються в 4.1. замість вихідних, і розрахунок повторюється до повної стабілізації, тобто до отримання рівноміцнісної (повно напруженої) конструкції.

В результаті роботи алгоритму 4.1-4.3 в якості оптимального з'являється конструкція з змінною густиною. Розподіл напруг в балці із матеріалу з сталою і змінною густиною показаний на рис.1. Оптимальна конструкція із змінною густиною володіє властивістю граничної інтенсивності деформації в кожному елементі $\varepsilon_i = \frac{\rho_i \cdot [\bar{\sigma}]}{\rho_i \cdot E} = \frac{[\bar{\sigma}]}{E} = \frac{[\sigma]}{E} = const$, яке можна трактувати як повно напружену конструкцію.

Вище наведений алгоритм складається із двох циклів: внутрішнього (п.4), пов'язаного з оптимізацією розподілу матеріалу, і зовнішнього (пп.1-8), пов'язаного з поррахунком розподілу навантажень. Розрахунки свідчать, що внутрішній цикл збігається за 10-15 ітерацій, зовнішній 3-4. Для контролю за ваговою ефективністю отримуваних розв'язків на кожній ітерації зовнішнього і внутрішнього циклів обчислюється величина силового фактора:

$$G = \int_V \sigma^{екв} dV \approx \sum_{i=1}^n \sigma_i^{екв} V_i, \text{ де } V - \text{об'єм тіла, } V_i - \text{об'єм } i\text{-го елемента, } n - \text{кількість}$$

елементів а також величина без вимірного коефіцієнта силового фактора

$$C_k = \frac{G}{Y\sqrt{S}}, \text{ де } Y - \text{характерне навантаження на площу } S. \text{ З точністю до множника,}$$

цей коефіцієнт визначає масу силових елементів конструкції і дозволяє обчислювати абсолютну і відносну масу конструкції за ваговими формулами:

$$m_{kp} = \frac{\varphi}{\sigma} C_k n_p m_0 g \sqrt{S}, \quad \bar{m}_{kp} = \frac{m_{kp}}{m_0} = \frac{\varphi}{\sigma} C_k n_p g \sqrt{\frac{m_0}{p_0}}, \quad \varphi - \text{коефіцієнт повної маси, } p_0 - \text{питоме}$$

навантаження. Для обчислення розподілу навантаження в середовищі NASTRAN використовується програма APAME із вільним доступом в Інтернеті за адресою [4.29].

Рекомендована література

- 1.1. Gupta R.K., Zorian Y. Introducing in Core-Based System Design // Design & Test of Computer, Oct/Dec, 1997
- 1.2. Петренко А.И. Основы автоматизации проектирования. - К.: Техніка, 1982. 295 с.
- 1.3. Лобур М. В. Принятие оптимальных решений в процессе технологического моделирования ИС // Вопросы радиоэлектроники. Серия ТПП. 1986. №2. с. 47-52
- 1.4. Коваль В.О., Лобур М.В. Автоматизация технологического моделирования полупроводниковых ИС: Учебное пособие. - Львов: ЛПИ, 1987. - 84 с.
- 1.5. Арайс Е.А., Дмитриев В.М. Автоматизация моделирования многосвязных механических систем. – М.:
- 1.6. Senturia, S.D. *Microsystem Design*; Kluwer Academic Publishers: Boston, MA, USA, 2001.
- 1.7. Норенков И.П. Основы автоматизированного проектирования: Учеб. для вузов. 2-е изд. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2002. - 336 с. - ISBN: 5-7038-2090-1
- 1.8. Лысенко И.Е. Проектирование сенсорных и актюаторных элементов микросистемной техники. Уч. пособие. Таганрог, 2005. - 10 с.
- 1.9. Воронин А.В. Моделирование мехатронных систем. Учебное пособие. Издательство Томского политехнического университета 2008
- 1.10. Теслюк В.М., Денисюк П.Ю. Автоматизація проектування мікроелектромеханічних систем на компонентному рівні. Львів: НУЛП, 2011
- 1.11. Власов А.И., Зинченко Л.А., Макаrchук В.В. и др. Автоматизированное проектирование наносистем. Учебно-методический комплекс по тематическому направлению деятельности «Наноинженерия», М.: МГТУ, 2011
- 1.12. Зинченко Л.А. САПР НАНОСИСТЕМ. Учебно-методический комплекс по тематическому направлению деятельности «Наноинженерия», М.: МГТУ, 2011
- 1.13. Баррет С.Ф., Пак Д.Дж. «Встраиваемые системы. Проектирование приложений на микроконтроллерах семейства 68HC12/HCS12 с применением языка С (электронная книга) Режим доступа: <http://www.rulit.net/books/vstraivaemye-sistemy - proektirovanie - prilozhenij- na-mikrokontrollerah – semejstva – 68hc12-hcs12-s-primen-read-237830 – 1.html>»
- 1.14. Лобур М. В. Система наскрізного проектування вбудованих пристроїв. Радіоелектроніка та телекомунікації. Вісник НУ "Львівська політехніка". 2003. № 477. с. 192 – 200.
- 1.15. [http:// www.eurointech.ru/coventor](http://www.eurointech.ru/coventor)
- 1.16. Гладштейн М. Встраиваемые микроэлектронные системы: состояние, проблемы и перспективы // Электронные компоненты. 2008. №5. С. 20-32
- 1.17. Мочалов В.А., Турута Е.Н. Интеллектуальные САПР сенсорных сетей. Интеллектуальные САПР, 2009
- 1.18. Федосов П.В. Использование САПР при создании «умных домов» НТВ: Научные школы в СПбГУ ИТМО. Вып.51, 2008, С.39-43
- 1.19. Платунов А.Е., Постников Н.П. Высокоуровневое проектирование встраиваемых систем. (Часть 1). ИТМО Санкт-Петербург, Часть 1, 2011; Часть 2, 2013.

- 1.28. Жуков К.Г. Модельное проектирование встраиваемых систем в LabVIEW. Москва: ДМК Пресс, 2011.- 688 с.
- 1.29. Raghavan P., Lad A., Neelakandsn S. Embedded Linux System. Design and Development. Auerbach, 2005, 432 p.
- 1.30. Marwedel P. Embedded System. Design. 2005, 264 p.
- 2.1. Режим доступа: www.micromachine.narod.ru. Введение в МЭМС. Технология производства микроактюаторов
- 2.2. Трофимов Н.А. Наука за рубежом. №15, июль, 2012. Ежемесячное обозрение. Электронное издание: www.issras.ru/global_science_rewiew.
- 2.3. Borisenko V.E., Ossicini S. What is What in the Nano-world (Wiley-VCH, Weinheim, 2004), 335 p.
- 2.4. Обухов И.А. О возможности применения СТМ-АСМ литографии для создания новых типов квантовых приборов // Микросистемная техника. 2003. №6.- С.34-37.
- 2.5. Мир материалов и технологий. Нанотехнологии. Наноматериалы. Наносистемная техника. Мировые достижения-2008. Сб. под ред. Д.т.н., проф. Мальцева П.П., Техносфера, 2008.
- 2.6. Адамов Ю.Ф. Проектирование систем на кристалле. М.: 2005
- 2.7. Королев М.А., А.Ю. Красюков, С.А. Поломошнов. Учебное пособие по дисциплине «Современные проблемы технологии наноэлектроники». М.: МИЭТ, 2011. - 100 с.
- 2.8. Wakaki I., Keiei K., Takeshisa S. Physical Properties and Data of Optical Materials . 2007, Taylor & Francis Group
- 2.9. www.optic-glas.narod.ru
- 2.10. Волков С.В., Ковальчук Є.П., Огенко В.М. та інш. Нанохімія. Наносистеми. Наноматеріали. Київ, Наукова думка, 2008.
- 2.11. Павлиго Т.М., Сердюк Г.Г. Класифікація наноматеріалів у системі міжнародної стандартизації. Наноструктурное материаловедение. №4, С.92-99, 2010.
- 2.12. ISO/TS 80004-1:2010 Nanotechnologies- P.1: Core terms. Ed. 2010-10-15.-ISO, 2010.-4p.
- 2.13. Yee K.S. // IEEE Trans. Ant. Prop. — 1966. — V. 14. — P. 302–307.
- 2.14. Левченко В.Д. Информационные технологии и вычислительные системы. 2005. - № 1.-68 с.
- 2.15. Вольпян О.Д., Кузьмичев А.И. Наноразмерные электронно-фотонные устройства поверхностных плазмонных поляритонов. Наноструктуры и нанотехнологии в электронике. Электроника и связь. № 1, 2011. Тематический выпуск «Электроника и нанотехнологии». С.5-11
- 2.16. Springer Handbook of Electronic and Photonic Materials. Safa Kasap, Peter Capper (Eds.), Springer Science+Business Media, Inc., 2006.
- 2.17. Гуров В.С., Вихров С.П. Наноматериалы и методы их исследования. Конспект лекций Всероссийской школы-семинара для студентов, аспирантов и молодых ученых по направлению «Наноматериалы». Рязань 2010
- 2.18. Козлов В.Л. Оптоэлектронные датчики. Минск, БГУ, 2005.
- 2.19. Лепіх Я.І., Гордієнко Ю.О., Дзядевич С.В. та інш. Створення мікроелектронних дотчиків нового покоління для інтелектуальних систем. Одеса: Астропринт, 2010;

- 2.20. Козлов А.Г. Тепловые микросенсоры: классификация, основные типы. Нано- и микросистемная техника, № 4, 2006, С. 2-13
- 2.21. Van Herwaarden S. Physical principles of thermal sensors. Sensors and Materials. 1996. V. 8. P. 373—387.
- 2.22. Скришевский В.А. Фізичні основи напівпровідникових хімічних сенсорів. КДУ, ВПЦ, 2006.
- 2.23. Dzafarov T., Yuksel S. Nan-Porous Silicion for gas Sensor and Fuel Cell Applications. J. of Qafgas University, Number 25, P.20-35, 2009.
- 2.24. Котюк А.Ф. Датчики в современных измерениях. М.: Радио и связь, 2005.
- 2.25. Quandt E; Seemann K. Fabrication and simulation of magnetostrictive thinfilm actuators. Sens. Act. 1995; A 50: 105-109.
- 2.26. Amor A.V. et.al. A magnetoelastic microtransformer – based microstrain gauge. Sens. Actuat. A. vol.129.2006.P.41-44.
- 2.27. Herrea-May A.L., Aguilera-Cortes L.A., Carcia-Ramirez P.J. et.al. Resonant Magnetic Field Sensors Based on MEMS Technology; Sensors, 2009,9, PP.7785; Development of Resonant Magnetic Field Microsensors Challenges and Future Applications. P.66.
- 2.28. Ваганов В.И. Интегральные тензопреобразователи.-М.: Энергоатомиздат, 1983; Ильинская Л.С., Подмарьков А.Н. Полупроводниковые тензодатчики.-М.: Энергия, 1966.
- 2.29. Гридчин В.А., Драгунов В.П. Физика микросистем: Учеб. пособие. В 2 ч. Часть 1. Новосибирск: Изд-во НГТУ, 2004 г. 416 стр
- 2.30. Дружинін А.О., Островський І.П., Когут Ю.Р. Ниткоподібні кристали кремнію, германію та їхніх твердих розчинів у сенсорній електроніці. Львів: НУЛП, 2010; Дружинин А.А., Марьямова И.И., Лавитская Е.Н. і інш. Полупроводниковые сенсоры механических величин на основе монокристаллов кремния для экстремальных условий. Микросистемная техника, №9, С.3-8, 2001.
- 2.31. Режим доступу: <http://nano-edu.ulsu.ru/w/index.php>.
- 2.32. Binning G., Quate C.F., Gerber C. Atomic Force Microscope. PRL. 1986. Vol.56, P.930-933.
- 2.33. Кособуцький П.С. та інш. Моделювання коливань простих систем.– Львів: НУ"ЛП" – 2003.
- 2.34. Canham L. Properties of porous silicon. - London: INSPEC, 1997. - 416 p.
- 2.35. Friden J. Handbook of modern sensors: physics, design, and applications. AIP Press, N.Y. (1996).
- 2.36. Amir G., Asad D. Teaching Vibration and Control using Animation, Simulation, and Experimentation. Proceeding of the 2005 American Society for Engineering Education Annual Conference & Exposition, 2005.
- 2.37. Park S. A study on buckled-beam actuators for RF MEMS application. A thesis submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College, 2007.
- 2.38. Shirazi F., Velni J., Grogoriadis K. An LPV Design Approach for Voltage Control of an Electrostatic MEMS Actuator. J. of Microelectromechanical Systems, vol.20, N01, P.302, 2011

- 2.39. Layer E., Tomczyk K. Measurement, Modeling and Simulation of Dynamic Systems. Springer-Verlag Berlin Heidelberg, 2010
- 2.40. Шарапов В., Мусиенко М., Шарапова Е. Пьезоэлектрические датчики. Мир электроники. М.: Техносфера, 2006
- 2.41. Mochida Y., Tamura M., Ohwada K. A micromachined vibrating rate gyroscope with independent beams the drive and detection modes. Sensor and Actuators 80 (2000) 170-178
- 2.42. MEMS/NEMS Handbook. Techniques and Applications. Design Methods in MEMS/NEMS. Vol.1, Leondes C.L. Ed. Springer, 2005.
- 2.43. Лысов А.Н., Винниченко Н.Т., Лысова А.А. Прикладная теория гироскопов. Часть 3. Челябинск, Изд. Центр ЮУрГУ, 2009.
- 2.44. Painter C., Shkel A. Active structural error suppression in MEMS vibratory rate integrating gyroscopes. IEEE Sensors J.-2003.-Vol.3.-№5.-P.595-606.
- 2.45. www.analog.com
- 2.46. Шереметьев А.Г. Волоконно-оптический гироскоп. М.: Ради и связь, 1987.
- 2.47. Осадчук В.С., Осадчук О.В., Яремчук В.Ф. та інші. Волоконно-оптичні перетворювачі газу. Оптичні та оптико-електронні сенсори і перетворювачі в системах керування та екологічного моніторингу. С.173-181, 2009.
- 2.48. Park S. A study on buckled-beam actuators for RF MEMS application. A thesis submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College, 2007.
- 2.49. Лысенко И.Е. Проектирование сенсорных и актюаторных элементов микросистемной техники. Таганрог, 2005.
- 2.50. Kutz M. Mechanical Engineers Handbook. Third Edition. Instrumentation, Systems, Controls, and MEMS. WILEY JOHN WILEY & SONS, Inc, 2006
- 2.51. Friden J. Handbook of modern sensors: physics, design, and applications. AIP Press, N. Y. (1996).
- 2.52. Бутгенбах С., Сейдман В. Катушки индуктивности для МЭМС. Институт микротехнологий, Технический университет Брауншвейга, Лангер Камп 8, Брауншвейг, Германия.
- 3.1. Meiler P., Zund R. Statistical Methods in Analytical Chemistry. Chemical Analysis. A series of monographs on Analytical Chemistry and its Application. Ed. Winefordner J. Vol.153. A Wiley-Interscience Publication JOHN WILEY & SONS, INC. 2000
- 3.2. П.С. Кособуцький, М.В. Лобур. Статистичне моделювання. - Львів: НУЛП. - 2013.
- 3.3. Колмогоров А.Н., Журбенко И.Г., Прохоров А.В. Введение в теорию вероятностей. М.: Наука, 1978.
- 3.4. Хальд А. Математическая статистика с техническими приложениями. М.: ИЛ, 1956.
- 3.5. Кособуцький П.С. Статистичні та Монте-Карло алгоритми моделювання випадкових процесів у макро- і мікросистемах в MathCAD НУ ЛП: Львів, 2014.
- 3.6. Тейлор Дж. Введение в теорию ошибок. М.: Мир, 1985.
- 3.7. Феллер В. Введение в теорию вероятностей и её приложения. Т.1 и 2. - М.: Мир, 1984.

- 3.8.** Allan D. Statistics of Atomic Frequency Standards. Proceedings of the IEEE.1966.vol.52.№2. PP.221-230; IEEE Std 952-1997.IEEE standart specification format guide and test procedure for single-axis interferometric fiber optic gyros. IEEE Std 952-1997 (R2008).IEEE standart specification format guide and test procedure for single-axis interferometric fiber optic gyros.
- 3.9.** Федоров М. В. Динамика и погрешности микромеханических акселерометров при вибрационном воздействии / М. В. Федоров. Научная сессия ГУАП часть 1 технические науки: СПбГУАП,2009,С.208 – 212
- 3.10.** Кутовой Д.А., Ситников П.В., Федотов А.А. и др.. Оценка основных характеристик бесплатформенного инерциального блока с использованием вариации Алана. Вестник Самарского государственного аэрокосмического университета.2014,№1(43),С.201-209.
- 3.11.** S.Ivanov, M.Yu.Ovchinnikov.Determination of the Orientation of a Suspended Mock-up Using One-Axis Gyroscope, Preprint, Inst. Appl. Math., the Russian Academy of Science. 2008. Режим доступа: http://www.keldysh.ru/papers/2008/rep11/rep2008_11.html
- 3.12.** Jacoboni C. and Lugli P., The Monte Carlo Method for Semiconductor Device Simulation (Springer-Verlag, Vienna, 1989).
- 3.13.** Metropolis N., Rosenbluth A. W., Rosenbluth M. N. et.al. *J. Chem. Phys.* 1953.Vol. 21.P. 1087.
- 3.14.** Толстик А. М. Применение метода Метрополиса в учебном компьютерном эксперименте по молекулярной физике. Открытое образование • 5/2006. с.30-33
- 3.15.** Klaus R. Mecke,Dietrich Stoyan. Statistical Physics and Spatial Statistics: The Art of Analyzing and Modeling Spatial Structures and Pattern Formation. Springer-Verlag Berlin Heidelberg 2000]. Режим доступа: <http://ru.bookzz.org>
- 3.16.** L. Martinez and Angel R. Martinez. Computational Statistics Handbook with Matlab.pdf. Режим доступа: <http://www.ee.hacettepe.edu.tr/~solen/Matlab/MatLab/Matlab%20-%20Computational%20Statistics%20Handbook%20with%20Matlab.pdf>
- 3.17.**Varde P.V.Role of Stsistical VIS-A-VIS Physics-of-Failure Methods in Reliability Engineering. Journal of Reliability and Statistical Studies (ISSN: 0974-8024) Vol. 2, Issue 1(2009): 41 – 51. Режим доступа: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.301.8356&rep=rep1&type=pdf>
- 3.18.** Лемешко, Б.Ю. Некоторые приложения, рекомендуемые к использованию в задачах прикладной математической статистики [Электронный ресурс] / Б.Ю. Лемешко. – Режим доступа: <http://ami.nstu.ru/~headrd/applied/index.html>.
- 4.1.** Murthy K.N., Christiana P. Optimal Design for Prescribed Buckling Loads. - Journal of the Structural Division, ASCE, vol. 100, N 11, 1974.-p. 2175-2189.
- 4.2.** Segenreich S.A., McIntosh S.C. Weight minimization of structures for fixed flutter sped via an optimality criterion// Proc. AISS/ASME/SAE 16th 17th Structures, Structural Dynamics and Materials Conference, 1975, Denver, Colorado.
- 4.3.** Haftka T., Gurdal Z. Elements of Structural Optimization. In Gladwell G. Solid Mechanics and its applications. Vol.11. Kluwer Academic Publishers. 1992
- 4.4.** Vanderplaats G.N. Numerical Optimization Techniques for Engineering Design. Vanderplaats Research and Development Inc., 3rd Ed.2001

- 4.5. Bao M. Analysis and Design Principles of MEMS Devices. Elsevier Amsterdam, 2005
- 4.6. Ananthasuresh G. Optimal Synthesis Methods for MEMS. The Kluwer International Series in Microsystems. Ed.Senturia, S. 2003
- 4.7. Scott Moura. Parametric Design of a MEMS Accelerometer. 2006. <http://ebookbrowse.net/me128-2-doc-d85876869>
- 4.8. Pisano A. Course Materials. UC Berkeley. Spring 2004; Optimization of MEMS . Spring 2007;
Parametric and Optimal Design of MEMS. Course Materials: Optimization of MEMS, UC Berkeley. 2007. [Online]. Available: http://web.mit.edu/ndhillon/www/Publications/PDF/reports/4_pre_optimization.pdf; Optimization of MEMS (EE C246/ME C2110. Spring 2007
- 4.9. Liwei L. "Parametric design of MEMS accelerometer", ME-128, Apr. 2006. [Online]. Available: http://edge.rit.edu/content/P10010/public/PDF/ME128_2.pdf
- 4.10. Messac A. Optimization in Practice with Matlab for Engineering Students and Professionals. Cambridge University Press, NY, USA .2015
- 4.11. Wiak S., Napieralska-Juszczak E. Computational Methods for the Innovate Design of Electrical Devices. Studis in Computational Intelligence, Vol.327 Ed.Kacprzyk J. Polish Academy of Science Warschaw Poland 2010
- 4.12. Di Barba Paolo (2009). Evolutionary multiobjective optimization methods for the shape design of industrial electromagnetic devices. IEEE TRANSACTIONS ON MAGNETICS, vol. 45, p. 1436-1441, ISSN: 0018-9464.
- 4.13. Nelder J., Mead R. Simplex Method for Function Minimization. Comp. J., Vol. 7, 1965, 308-313.
- 4.14. Han J.S. and B.M. Kwak. Robust optimization using a gradient index: MEMS applications. Struct Multidisc Optim 27, 469–478 (2004)
- 4.15. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. М.: Мир, 1985-509 с.
- 4.16. Бронштейн И.Н., Семендяев К.А. Справочник по математике. М.: 1986.
- 4.17. Parametric Design of a MEMS Accelerometer. ME128 – Project 2. 7 April 2005. Режим доступа: <http://ebookbrowse.net/me128-2-pdf-d80359760>
- 4.18. Slawomir Koziel and Xin-She Yang (Eds.) Computational Optimization, Methods and Algorithms. 2011 Springer-Verlag Berlin Heidelberg
- 4.19. Hadjout L., Belgauerras Chapter 2. Multi-objective Design Optimization of Slotless PM Motors Using Genetic Algorithms Based on Analytical Field Calculation pp.19-37. In: Wiak S., Napieralska-Juszczak E. Computational Methods for the Innovate Design of Electrical Devices. Studis in Computational Intelligence, Vol.327 Ed.Kacprzyk J. Polish Academy of Science Warschaw Poland 2010
- 4.20. Горбійчук М.І. Алгоритми і методи обчислень. Івано-Франківськ, 2014.
- 4.21. Корначевський Я.І., Ладогубець В.В. Генетичні алгоритми в САПР. Вісник НТУУ «КПІ» Інформатика, управління та обчислювальна техніка, 2011.- №54.- С.48 -53
- 4.22. Волков, В.Л. Измерительные информационные системы. Новгород: НГТУ, 2009. – 242 с.
- 4.23. Волкова Н.В. Исследование датчиков первичной информации методом статистических испытаний Монте-Карло // Материалы Всероссийской научно-технической конференции "Приборостроение в аэрокосмической технике". – Арзамас. 1999. С. 24-28.

- 4.24.** Учебно-методические материалы для лабораторных работ по дисциплине «Микро- и наносистемная техника». «ЛЭТИ» » (СПбГЭТУ). Санкт – Петербург . 2009
- 4.25.** Ильинская Л.С., Подмарьков А.Н. Полупроводниковые тензодатчики.-М.: Энергия, 1966.
- 4.26.** Chia C.Y. Nonlinear analysis of plates. 1980. McGraw-Hill New York.
- 4.27.** Kosobutsky P., M. Komarnutsky, C. Maj, P. Zajac, M. Szermer, W. Zabierowski. Monte Carlo Modeling of Stiffness of MEMS Membrane. Mixdes,2014, Gdansk, Poland,2014
- 4.28.** Болдырев А.В., Комаров В.А. Применение моделей переменной плотности в задачах проектирования пространственных конструкций. Вестник Нижегородского университета им.Лобачевского.- №4(5).-с.2034-2036.-2011
- 4.29.** Режим доступа: <http://www.3dpanelmethod.com> .