



ВІСНИК

**КИЇВСЬКОГО НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

Серія: ФІЗИКО-МАТЕМАТИЧНІ НАУКИ

Випуск № 1

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

**ВІСНИК
КИЇВСЬКОГО УНІВЕРСИТЕТУ**

Серія: ФІЗИКО-МАТЕМАТИЧНІ НАУКИ

Наукова бібліотека
ім. М. Максимовича
КНУ

ім. ТАРАСА ШЕВЧЕНКА



13366JB

45 - чит. зал періодики та дисерт. Ц. 25.00

ВИПУСК №1 2008

З М І С Т

АЛГЕБРА, ГЕОМЕТРІЯ ТА ТЕОРІЯ ІМОВІРНОСТЕЙ

Бабич В.М. Групи Пуанкаре деякого класу бімодульних задач	7
Безущак О.О. Відношення Гріна інверсної напівгрупи частково визначених ко-скінченних ізометрій дискретної лінійки	12
Гончарова С.Я. Рівняння Беллмана для імпульсних процесів переносу	17
Дармосюк В. Властивості квазіфробеніусових кілець	21
Дяченко С.М. Ручні сагайдаки відносно поточно-проективних зображень	23
Жучок А.В. Оболонка зсувів Кліффордової напівгрупи	29
Заторський Р.А. Операції з формальними степеневими рядами з ненульовим вільним членом	36
Морозов Д. І. Спряженість автоморфізмів, що задаються лінійними функціями в групі скінченостанових автоморфізмів кореневого сферично однорідного дерева T_2	40
Пехтерєв В.О. Автоморфізми напівгрупи OT_n	44
Тарашанський М.Т. Екстремальні продовження міри та гомоморфізми	46

ДИФЕНЦІАЛЬНІ РІВНЯННЯ, МАТЕМАТИЧНА ФІЗИКА ТА МЕХАНІКА

Горошко О.О., Філімоніхіна І.І. Достатні умови усунення автобалансирами кута нутації незрівноваженого обертового тіла в ізолюваній системі	53
Король І.І. Дослідження періодичних розв'язків автономних диференціальних систем з імпульсною дією у критичному випадку	59
Таран Є.Ю., Кондрат Р.Я. Реологія розведеного розчину деформівних макромолекул у течії простого зсуву	67

КОМП'ЮТЕРНІ НАУКИ ТА ІНФОРМАТИКА

Grigoryeva L.V. Maple-exploring of a Free Flywheel Suspended by Superconductive Bearing	75
Антонюк М.А. Програмні рішення Microsoft для навчальних закладів	80
Власюк А.П., Мартинюк П.М. Математичне моделювання контактного розмиву неоднорідного масиву ґрунту фільтраційним потоком	86
Волинець В.І. Аналіз точності рекурентних методів обчислення багатовимірних дискретних перетворень Фур'є та Хартлі в арифметиці з плаваючою комою	92
Жук С.М. Гарантоване оцінювання у лінійних дескрипторних різницевих рівняннях	102
Заворотний А.Л., Касьянюк В.С. Про один підхід до задачі відновлення кривої за нечітко збуреними координатами її точок	106
Іванов І.Ю. Про оптимізацію алгоритмів узгодження ключів в динамічних рівноправних групах	112
Лялецький О.О. Про типи збіжності послідовностей та топологічні властивості частково упорядкованих множин	121
Музичук А.А. Оцінка вартості та моменту впровадження інвестиційних проектів в умовах ризику	127
Поляков С.А. Огляд стандартів мови SQL	132
Романюк С.І. Архітектура мережних засобів дистанційного навчання	138

Рубльов Б.В., Чурюмова Т.К., Шаманська В.Б. Площа взаємного розташування трьох фігур на площині	143
Сільвейструк Л.М. Розширена модель "сутність-зв'язок": уточнення типів сутностей суперклас і підклас та типу зв'язку суперклас/підклас	147
Філатов М.С. Автоматизована система адаптивного триангулювання із згущенням багатозв'язних та складених областей	155
Філатова І.А. Математичне моделювання напружено-деформованого стану багат шарового ґрунтового масиву при наявності рівня ґрунтових вод	161

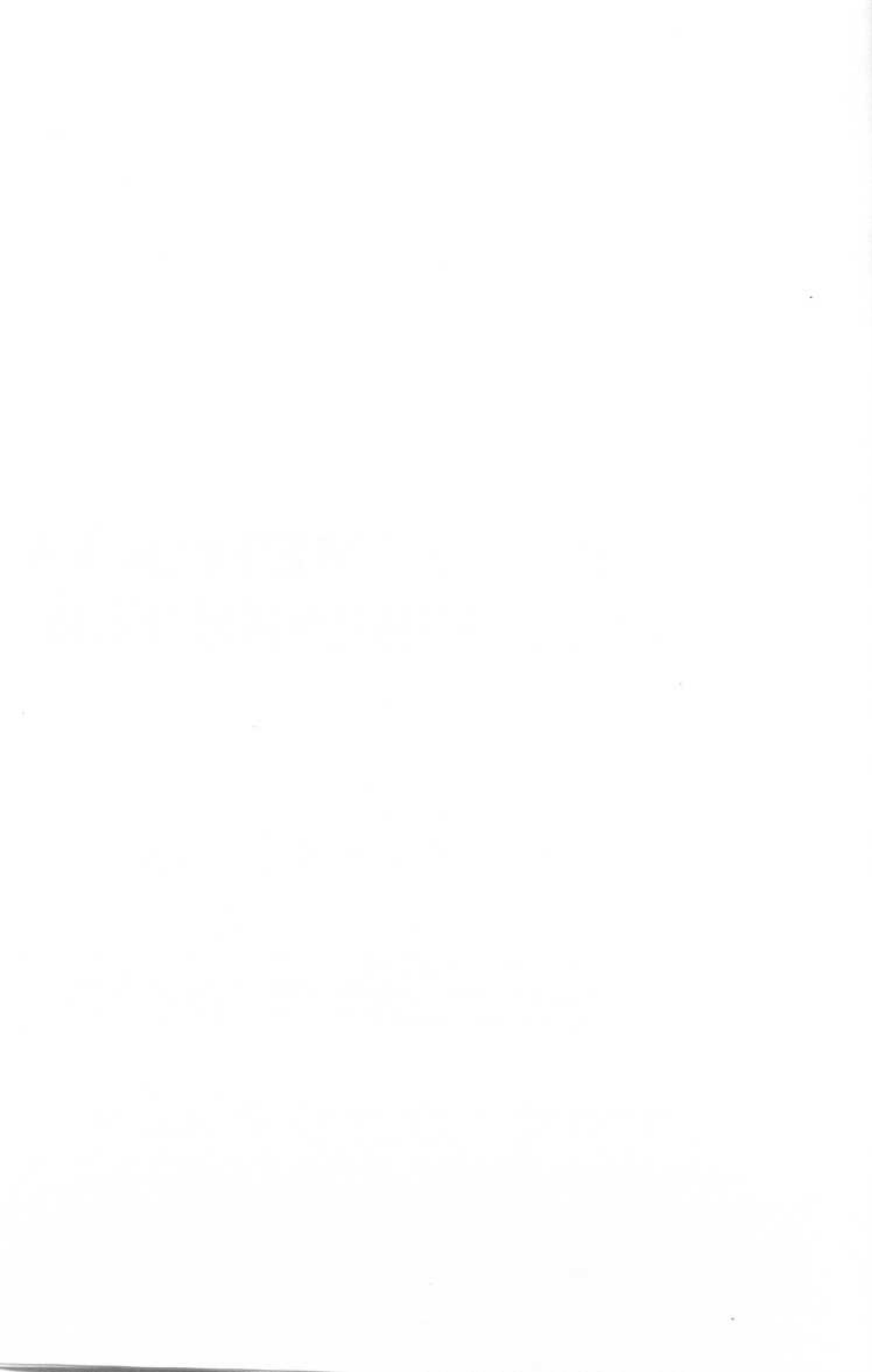
РАДІОФІЗИКА

Алексєєв О.С., Мусатенко Ю.С. Порівняння мір подібності кривих для задачі розпізнавання підписів	169
Голдаєвич Є.Л., Мартиш Є.В. Перехід у нерівноважний стан розряду з ковзаючою дугою	176
Іванюта О.М., Прокопенко О.В. Дослідження водних сольових розчинів за допомогою НВЧ аналізатора на основі резонатора поверхневої хвилі	180
Ільченко В.В. Моделювання рекомбінаційного струму в діодах Шотткі через шари квантових точок двох різних типів	183
Косаревич Б.П., Анісімов І.О. Числовий розрахунок характеристик темного розряду	188

СУЧАСНА ФІЗИКА

Альохін О.Д., Абдікарімов Б.Ж., Рудніков Є.Г. Взаємозв'язок величин критичних показників кривої співіснування, теплоємності та кореляційної функції	195
Булавін Л.А., Гаврюшенко Д.А., Сисоєв В.М., Якунов П.А. Розрахунок просторового розподілу компонентів бінарної суміші в плоскій порі з несиметричним пристінковим потенціалом	199
Заболотний М.А., Куліш М.П., Дмитренко О.П., Харкянен В.М., Барабаш Ю.М., Кобус О.С., Стасюк К.В. Спектральна залежність фотогенерації в фулеренмістких аморфних молекулярних напівпровідниках	202
Беженар А.А., Копань В.С. Вплив внутрішніх границь на електроопір порошкового фторопласту	208
Ніколенко А.С., Кондратенко С.В., Вакуленко О.В., Юхимчук В.О. Вплив термічних відпалів на фотопровідність Si/Ge гетероструктур з квантовими точками	212
Онанко А.П. Вказівна поверхня непружно-пружного тіла сплаву Ti_3Al	217
Погорелов В.Є., Голуб П.О., Школьнікова Н.І. Спектр комбінаційного розсіяння 4-(н-гексилосі)феніл-4'-(н-октилокси)бензоат в кристалічній фазі	220
Рогуль Т.Г. Термічна стабільність осажденного магнетронним розпиленням наддрібнозернистого хрому	223
Розуван К.П., Розуван С.Г., Шайкевич І.А. Лінійна і нелінійна спектроскопія π -зв'язків у політїофенів	227
Селищев П.О., Діденко Т.П. Перехідна повзучість матеріалу під опроміненням при умові генерації вакансій ковзаючими дислокаціями	231

АЛГЕБРА, ГЕОМЕТРІЯ ТА ТЕОРІЯ ІМОВІРНОСТЕЙ



УДК 512.643, 512.548

В'ячеслав М. Бабич, к. ф.-м. н., асистент

Групи Пуанкаре деякого класу бімодульних задач

За допомогою техніки групоїдів досліджуються групи Пуанкаре деякого класу бімодульних задач над кільцем степеневих рядів від однієї змінної над полем.

Ключові слова: бімодульна задача, групоїд, група Пуанкаре.

E-mail: vyacheslav.babych@gmail.com

Vyacheslav M. Babych, PhD, Assistant Professor

Poincare groups of some class of bimodule problems

Using groupoid technique we investigate Poincare groups of some class of bimodule problems over the ring of power series in one variable over a field.

Key words: bimodule problem, groupoid, Poincare group.

Статтю представив д.ф.-м.н., професор А.П.Петравчук

Вступ

Важливим і ефективним інструментом для вирішення задачі скінченності, ручності та дикості зображувального типу бімодульних задач є так званий «метод накриттів». Застосування цього методу до матричних задач скінченного зображувального типу наведені в [1, 2, 4, 9, 10], а до матричних задач і алгебр нескінченного зображувального типу — у роботі [8].

Ефективна побудова накриття бімодульної задачі залежить від структури фундаментальної групи. Особливо простою така побудова є у випадку, коли група Пуанкаре бімодульної задачі є вільною. Класичним прикладом, коли це виконується, є алгебри скінченного зображувального типу [3].

Стаття присвячена подальшому розвитку методів, розвинутих у [1]. У роботі ми наводимо основні факти з теорії фундаментальних груп (груп Пуанкаре) ([6]), адаптовані до потреб теорії накриттів для матричних задач.

1 Бімодульна задача

Нехай k — алгебраїчно замкнене поле. Бімодульною задачею називається пара $\mathcal{A} = (\mathcal{C}, \mathcal{M})$, де $\mathcal{C}$ — локально скінченновимірна категорія над k , а $\mathcal{M}$ — локально скінченновимірний бімодуль на категорію $\mathcal{C}$ (див. [1], а також [5, 7]).

Означення 1. Базисом бімодульної задачі $\mathcal{A} = (\mathcal{C}, \mathcal{M})$ називається біграф $\Gamma = \Gamma(\mathcal{A})$, в якому $\Gamma_0 = \text{Ob } \mathcal{C}$, $\Gamma_1^0(i, j)$ є базисом в $\mathcal{M}(i, j)$, $\Gamma_1^1(i, j)$

є базисом в $\mathcal{C}(i, j)$, якщо $i \neq j$, і $\Gamma_1^1(i, i)$ є базисом в $\text{Rad } \mathcal{C}(i, i)$, $s(x) = i$, $e(x) = j$ для кожної стрілки $x \in \Gamma_1^0(i, j) \cup \Gamma_1^1(i, j)$, $i, j \in \Gamma_0$.

Означення 2. Базис Γ бімодульної задачі $\mathcal{A}$ називається мультиплікативним, якщо добуток двох стрілок, які можна множити в $\mathcal{A}$, є стрілкою в Γ або нулем.

Означення 3. Бімодульна задача $\mathcal{A}$ називається зв'язною, якщо її базис зв'язний.

2 Групоїд Пуанкаре

Означення 4. Категорія $\mathcal{G}$ називається групоїдом, якщо всі її морфізми є ізоморфізмами.

Означення 5. Підкатегорія групоїда $\mathcal{G}$, яка є групоїдом, називається підгрупоїдом групоїда $\mathcal{G}$.

Кожній непорожній множині E можна поставити у відповідність групоїд $\text{Sc } E$, в якому $\text{Ob } \text{Sc } E = E$, $\text{Mor } \text{Sc } E = E \times E$, $\text{Sc } E(x, y) = \{(x, y)\}$, $x, y \in E$, а композиція $\text{Sc } E(y, z) \times \text{Sc } E(x, y) \rightarrow \text{Sc } E(x, z)$ визначена рівністю $(y, z) \circ (x, y) = (x, z)$.

Групоїд $\mathcal{G}$ називається однозв'язним, якщо він ізоморфний групоїду $\text{Sc } \text{Ob } \mathcal{G}$. Групоїд $\mathcal{G}$ називається зв'язним, якщо всі його об'єкти ізоморфні. Очевидно, що однозв'язний групоїд є зв'язним групоїдом, в якому для довільних об'єктів $i, j \in \text{Ob } \mathcal{G}$ множина морфізмів $\mathcal{G}(i, j)$ одноелементна.

Пов'яжемо з мультиплікативним базисом Γ бімодульної задачі $\mathcal{A}$ групоїд $\mathcal{G}_\Gamma$. Розглянемо

біграф Γ' , отриманий з Γ додаванням стрілки $x^{-1} : j \rightarrow i$ до кожної стрілки $x \in \Gamma_1(i, j)$, $i, j \in \Gamma_0$. Покладемо $\text{Ob } \mathcal{G}_\Gamma = \Gamma_0$, а $\text{Mor } \mathcal{G}_\Gamma$ вільно породимо множиною $\Gamma_1 \sqcup (\Gamma_1)^{-1}$, де $(\Gamma_1)^{-1}$ — множина доданих стрілок $x^{-1} : j \rightarrow i$ (а $\sqcup$ означає диз'юнктивне об'єднання), за модулем співвідношень $xx^{-1} = \mathbb{1}_j$, $x^{-1}x = \mathbb{1}_i$ для $x \in \mathcal{G}_\Gamma(i, j)$, $x^{-1} \in \mathcal{G}_\Gamma(j, i)$, $i, j \in \text{Ob } \mathcal{G}_\Gamma$, і $xy = z$ для всіх $x, y, z \in \text{Mor } \mathcal{G}_\Gamma$ за умови, що остання рівність виконується в $\mathcal{A}$. Групоїд $\mathcal{G}_\Gamma$ називається *групоїдом Пуанкаре* (фундаментальним групоїдом) бімодульної задачі Γ .

Означення 6 ([6]). Групою Пуанкаре (фундаментальною групою) групоїда $\mathcal{G}$ в об'єкті i називається група $\mathcal{G}(i, i)$ всіх автоморфізмів об'єкта i і позначається $\pi_1(\mathcal{G}, i)$.

У зв'язному групоїді група Пуанкаре не залежить від об'єкта.

Якщо $F : \mathcal{G} \rightarrow \mathcal{H}$ — функтор з групоїда $\mathcal{G}$ у групоїд $\mathcal{H}$, то для довільного об'єкта $i \in \text{Ob } \mathcal{G}$ позначимо символом $\pi_1(F, i)$ гомоморфізм груп

$$\pi_1(\mathcal{G}, i) = \mathcal{G}(i, i) \ni f \mapsto F(f) \in \mathcal{H}(F(i), F(i)) = \pi_1(\mathcal{H}, F(i)),$$

індукований функтором F .

Лема 1. Нехай H, G_1, G_2 — групи, φ_i гомоморфізми з H у G_i , $i = 1, 2$, і $G = G_1 \amalg_H G_2$ — амальгама над H груп G_1 і G_2 .

1. Якщо φ_1 — ізоморфізм, то $G \simeq G_2$.
2. Якщо φ_1, φ_2 — тривіальні гомоморфізми, то $G \simeq G_1 * G_2$.

Доведення. Друге твердження очевидне.

Для доведення першого твердження розглянемо комутативну діаграму

$$\begin{array}{ccc} H & \xrightarrow{\varphi_1} & G_1 \\ \varphi_2 \downarrow & & \downarrow \varphi_2 \varphi_1^{-1} \\ G_2 & \xrightarrow{\mathbb{1}_{G_2}} & G_2 \end{array}$$

Нехай G' — деяка група, $\psi_i : G_i \rightarrow G'$, $i = 1, 2$, — гомоморфізми груп такі, що діаграма

$$\begin{array}{ccc} H & \xrightarrow{\varphi_1} & G_1 \\ \varphi_2 \downarrow & & \downarrow \psi_1 \\ G_2 & \xrightarrow{\psi_2} & G' \end{array}$$

комутативна. Тоді існує єдиний гомоморфізм $\eta = \psi_2 : G_2 \rightarrow G'$ такий, що діаграма

$$\begin{array}{ccc} H & \xrightarrow{\varphi_1} & G_1 \\ \varphi_2 \downarrow & \varphi_2 \varphi_1^{-1} \downarrow & \downarrow \psi_1 \\ G_2 & \xrightarrow{\mathbb{1}_{G_2}} & G_2 \\ & \searrow \psi_2 & \downarrow \eta \\ & & G' \end{array}$$

комутативна. Отже, $G_1 \amalg_H G_2 \simeq G_2$. $\square$

Означення 7. Групою Пуанкаре (фундаментальною групою) зв'язної бімодульної задачі $\mathcal{A}$ називається група Пуанкаре фундаментального групоїда $\mathcal{G}_{\Gamma(\mathcal{A})}$ бімодульної задачі $\mathcal{A}$ і позначається $\pi_1(\mathcal{A})$.

3 Розглядуваний клас бімодульних задач

Розглянемо для кожного $m \in \mathbb{N}$ категорію $\mathcal{C}_m$, в якій

$$\text{Ob } \mathcal{C}_m = \{1, \dots, m\},$$

$$\mathcal{C}_m(i, j) = 0, \quad i \neq j,$$

$$\mathcal{C}_m(i, i) = k[[t]] \cdot \mathbb{1}_i, \quad i, j \in \text{Ob } \mathcal{C}_m.$$

Розглянемо цілі числа $n_1 > 0$, $n_2 > 0$, $n \geq 0$ і бімодуль $\mathcal{M}_{n_1, n, n_2}$ над категорією $\mathcal{C}_2$ такий, що $\mathcal{M}_{n_1, n, n_2}(1, 2)$ — векторний простір над полем k з базисом $v_1, v_2, \dots, v_n$, а $\mathcal{M}_{n_1, n, n_2}(1, 1) = \mathcal{M}_{n_1, n, n_2}(2, 1) = \mathcal{M}_{n_1, n, n_2}(2, 2) = 0$, причому

$$v_i(t\mathbb{1}_1) = \begin{cases} v_{i+n_1}, & \text{якщо } i + n_1 \leq n, \\ 0 & \text{в іншому разі,} \end{cases}$$

$$(t\mathbb{1}_2)v_i = \begin{cases} v_{i+n_2}, & \text{якщо } i + n_2 \leq n, \\ 0 & \text{в іншому разі.} \end{cases}$$

Бімодульна задача $\mathcal{A}_{n_1, n, n_2} = (\mathcal{C}, \mathcal{M}_{n_1, n, n_2})$, де $\mathcal{C} = \mathcal{C}_2 / \text{Ann}_{\mathcal{C}_2} \mathcal{M}_{n_1, n, n_2}$, описуватиметься наступним графом

$$\textcircled{1} \xrightarrow[n_1]{n} \xrightarrow[n_2]{} \textcircled{2},$$

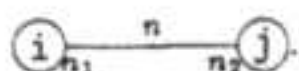
де кружечки означають об'єкти 1 і 2 відповідно, стрілка з 1 в 2 означає, що $\mathcal{M}_{n_1, n, n_2}(1, 2) \neq 0$, а мітки n, n_1, n_2 на стрілці визначені вище. Далі у тексті ми опускаємо напрямки стрілок, оскільки фундаментальна група не залежить від цих напрямків. Окрім того, якщо одне

з чисел n, n_1, n_2 дорівнює 1, то ми опускаємо відповідну мітку.

Базис бімодуля M_{n_1, n, n_2} разом з базисом радикала $\text{Rad } C$ категорії C , утвореним морфізмами $t^i \mathbb{1}_1, t^j \mathbb{1}_2, i \in \{1, \dots, [\frac{n-1}{n_1}]\}, j \in \{1, \dots, [\frac{n-1}{n_2}]\}$, задає мультиплікативний базис бімодульної задачі A_{n_1, n, n_2} (квадратні дужки означають цілу частину раціонального числа).

Розглянемо клас Q_r таких зв'язних бімодульних задач $A = (C, M)$ без циклів, де M — локально скінченновимірний бімодуль над категорією C_m , а $C = C_m / \text{Ann}_{C_m} M, m \in \mathbb{N}$, що для довільних $i, j \in \text{Ob } C, i \neq j$, підзадача $A_{i,j} = A_{\{i,j\}}$ еквівалентна A_{n_1, n, n_2} для деяких n, n_1, n_2 , причому $n_1 \leq n, n_2 \leq n$ і принаймні одне з чисел n_1, n_2 дорівнює одиниці.

Поставимо у відповідність кожній задачі $A \in Q_r$ розмічений орієнтований граф $\Delta = \Delta(A)$ з вершинами $1, \dots, m$, який в обмеженні на кожну пару точок $\{i, j\}, i \neq j$, збігається з визначеним вище розміченим графом



Граф Δ ми називатимемо розміщеною діаграмою бімодульної задачі. Кожному ребру цього графа відповідають три цілих числа n, n_1, n_2 , які, очевидно, визначають функції $w, s_1, s_2 : \Delta_1 \rightarrow \mathbb{Z}$ відповідно на ребрах діаграми бімодульної задачі. Функцію w назвемо вагою ребра, а s_1 і s_2 — початковим і кінцевим кроком відповідно. Набір функцій w, s_1, s_2 назвемо розміткою діаграми Δ . Ясно, що існує взаємно однозначна відповідність між бімодульними задачами з класу Q_r і відповідними розміченими діаграмами.

Розглянемо підклас Q_r^* класу Q_r , який складається з бімодульних задач, які задовольняють наступним умовам:

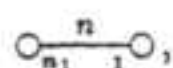
- кожна точка діаграми $\Delta(A)$ бімодульної задачі A має порядок, який не перевищує 3;
- кількість точок порядку 3 в $\Delta(A)$ не перевищує 1;
- якщо в діаграмі $\Delta(A)$ існує точка порядку 3, то усі ребра одного з ланцюгів $\Delta(A)$, який сполучає точку порядку 3 з точкою порядку 1, мають вагу 1;
- кількість неодиначних початкових і кінцевих кроків у діаграмі $\Delta(A)$ не перевищує 1.

Як показано у роботі [1], клас бімодульних задач скінченного зображувального типу з класу Q_r міститься в Q_r^* .

4 Група Пуанкаре

Лема 2. 1. Якщо $A = (k[[t]]/(t^n), 0)$, то фундаментальна група $\pi_1(A)$ бімодульної задачі A тривіальна при $n = 0$ і ізоморфна $\mathbb{Z}$ при $n > 0$.

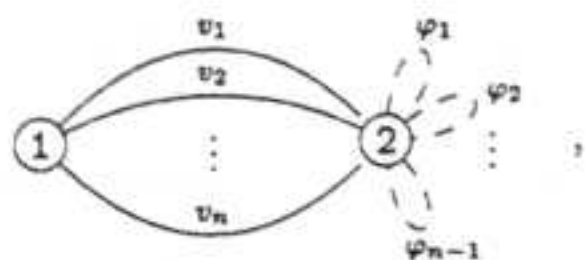
2. Якщо A — бімодульна задача з діаграмою



причому $n \geq n_1$, то фундаментальна група $\pi_1(A)$ бімодульної задачі A тривіальна при $n = 1$ і ізоморфна $\mathbb{Z}$ при $n > 1$.

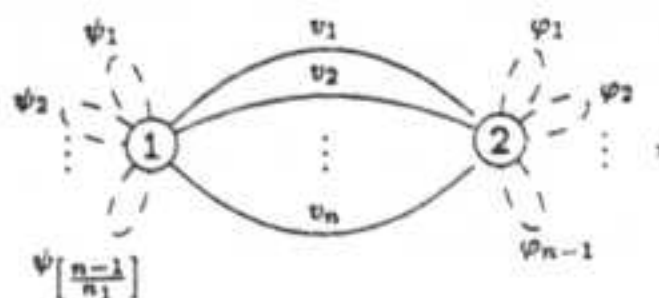
Доведення. У випадку 1 напівгрупа шляхів мультиплікативного базису Γ бімодульної задачі A породжена петлями $t \cdot 1, \dots, t^{n-1} \cdot 1$ зі співвідношеннями $t^i \cdot 1 \cdot t^j \cdot 1 = t^{i+j} \cdot 1$, якщо $i+j < n$, і, отже, є тривіальною при $n = 1$ і нескінченною циклічною з твірним t при $n > 1$. Тому в останньому випадку групоїд Пуанкаре буде нескінченною циклічною групою з твірним t .

Якщо у випадку 2 $n = n_1 = 1$, то група Пуанкаре бімодульної задачі A , очевидно, тривіальна. Якщо $n = n_1 > 1$, то мультиплікативний базис бімодульної задачі A має вигляд



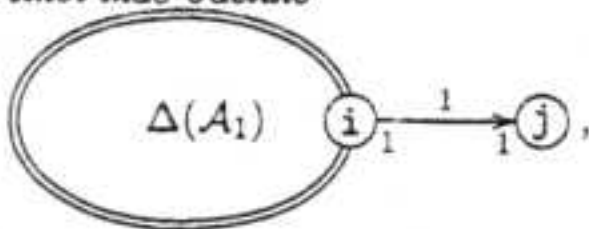
де $\varphi_i = t^i \cdot \mathbb{1}_2$. Тоді в групоїді Пуанкаре мають місце співвідношення $\varphi_i \varphi_j = \varphi_{i+j}$, якщо $i+j \leq n-1$, $\varphi_j v_i = v_{i+j}$, якщо $i+j \leq n$. Очевидно, що групоїд Пуанкаре породжується морфізмами φ_1, v_1 , оскільки $v_i = \varphi_1^{i-1} v_1, i = 2, \dots, n$. Отже, $\pi_1(A) \simeq \mathbb{Z}$.

Нарешті, якщо у випадку 2 $n > n_1 \geq 1$, то мультиплікативний базис бімодульної задачі A має вигляд



де $\varphi_i = t^i \cdot \mathbb{1}_2$, $i = 1, \dots, n-1$, $\psi_i = t^i \cdot \mathbb{1}_1$, $i = 1, \dots, [\frac{n-1}{n_1}]$. Тоді в групі Пуанкаре мають місце співвідношення $\varphi_i \varphi_j = \varphi_{i+j}$, якщо $i+j \leq n-1$, $\psi_i \psi_j = \psi_{i+j}$, якщо $i+j \leq [\frac{n-1}{n_1}]$, $\varphi_j v_i = v_{i+j}$, якщо $i+j \leq n$, $v_j \psi_i = v_{n_1 i+j}$, якщо $n_1 i+j \leq n$. Очевидно, що цей групі породжується морфізмами φ_1, ψ_1, v_1 . В свою чергу цю систему твірних і співвідношень можна спростити, оскільки $\varphi_{n_1} v_1 = v_1 \psi_1$. Очевидно, що група Пуанкаре в точці 2 є нескінченною циклічною, породженою φ_1 , оскільки $v_1 \psi_1^k v_1^{-1} = (v_1 \psi_1^k) v_1^{-1} = (\varphi_{n_1}^k v_1) v_1^{-1} = \varphi_1^{n_1 k}$. Отже, $\pi_1(\mathcal{A}) \simeq \mathbb{Z}$. $\square$

Лема 3. Нехай $\mathcal{A} \in Q_r$ — бімодульна задача, діаграма якої має вигляд



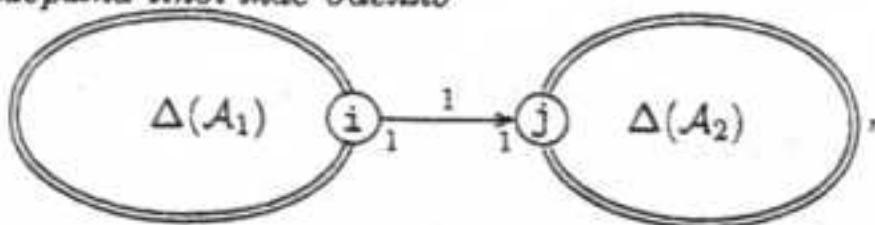
де $\mathcal{A}_1$ — підзадача $\mathcal{A}$. Тоді $\pi_1(\mathcal{A}) \simeq \pi_1(\mathcal{A}_1)$.

Доведення. Розглянемо бімодульні задачі $\mathcal{B}_1$ і $\mathcal{B}_2$, розміченими діаграмами $\Delta(\mathcal{B}_1)$ і $\Delta(\mathcal{B}_2)$ яких є точка $\textcircled{1}$ і ребро $\textcircled{k} \xrightarrow{1} \textcircled{1}$ відповідно. Розглянемо функтори $F : \mathcal{G}_{\Gamma(\mathcal{B}_1)} \rightarrow \mathcal{G}_{\Gamma(\mathcal{A}_1)}$, $G : \mathcal{G}_{\Gamma(\mathcal{B}_2)} \rightarrow \mathcal{G}_{\Gamma(\mathcal{A}_2)}$ такі, що $F(1) = i$, $G(1) = k$. Очевидно, $\mathcal{G}_{\Gamma(\mathcal{A})} \simeq \mathcal{G}_{\Gamma(\mathcal{A}_1)} \amalg_{\mathcal{G}_{\Gamma(\mathcal{B}_1)}} \mathcal{G}_{\Gamma(\mathcal{B}_2)}$. Тоді за теоремою ван Кампена ([6]) і лемами 2 і 1

$$\begin{aligned} \pi_1(\mathcal{A}) &= \pi_1(\mathcal{G}_{\Gamma(\mathcal{A})}, i) \simeq \\ &\simeq \pi_1(\mathcal{G}_{\Gamma(\mathcal{A}_1)} \amalg_{\mathcal{G}_{\Gamma(\mathcal{B}_1)}} \mathcal{G}_{\Gamma(\mathcal{B}_2)}, i) \simeq \\ &\simeq \pi_1(\mathcal{G}_{\Gamma(\mathcal{A}_1)}, i) \amalg_{\pi_1(\mathcal{G}_{\Gamma(\mathcal{B}_1)}, 1)} \pi_1(\mathcal{G}_{\Gamma(\mathcal{B}_2)}, k) = \\ &= \pi_1(\mathcal{A}_1) \amalg_{\pi_1(\mathcal{B}_1)} \pi_1(\mathcal{B}_2) \simeq \\ &\simeq \pi_1(\mathcal{A}_1) \amalg_{\{e\}} \{e\} \simeq \pi_1(\mathcal{A}_1) \end{aligned}$$

($\{e\}$ — тривіальна група). $\square$

Лема 4. Нехай $\mathcal{A} \in Q_r$ — бімодульна задача, діаграма якої має вигляд

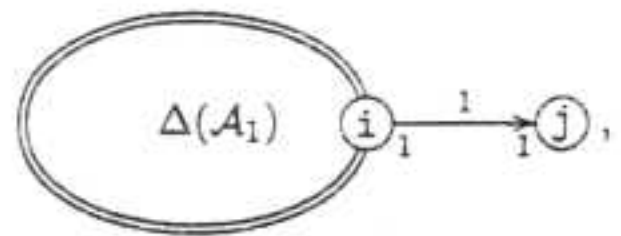


де $\mathcal{A}_1, \mathcal{A}_2$ — підзадачі $\mathcal{A}$. Тоді

$$\pi_1(\mathcal{A}) \simeq \pi_1(\mathcal{A}_1) * \pi_1(\mathcal{A}_2),$$

де $*$ означає вільний добуток.

Доведення. Розглянемо бімодульні задачі $\mathcal{B}_1$ і $\mathcal{B}_2$, розміченими діаграмами $\Delta(\mathcal{B}_1)$ і $\Delta(\mathcal{B}_2)$ яких є точка $\textcircled{1}$ і діаграма



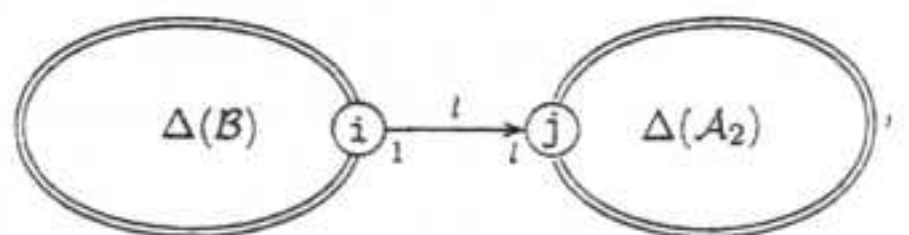
відповідно. Розглянемо функтори

$$F : \mathcal{G}_{\Gamma(\mathcal{B}_1)} \rightarrow \mathcal{G}_{\Gamma(\mathcal{B}_2)}, \quad G : \mathcal{G}_{\Gamma(\mathcal{B}_1)} \rightarrow \mathcal{G}_{\Gamma(\mathcal{A}_2)}$$

такі, що $F(1) = j$, $G(1) = j$. Очевидно, $\mathcal{G}_{\Gamma(\mathcal{A})} \simeq \mathcal{G}_{\Gamma(\mathcal{B}_2)} \amalg_{\mathcal{G}_{\Gamma(\mathcal{B}_1)}} \mathcal{G}_{\Gamma(\mathcal{A}_2)}$. За теоремою ван Кампена ([6]) і лемами 3 і 1

$$\begin{aligned} \pi_1(\mathcal{A}) &= \pi_1(\mathcal{G}_{\Gamma(\mathcal{A})}, j) \simeq \\ &\simeq \pi_1(\mathcal{G}_{\Gamma(\mathcal{B}_2)} \amalg_{\mathcal{G}_{\Gamma(\mathcal{B}_1)}} \mathcal{G}_{\Gamma(\mathcal{A}_2)}, j) \simeq \\ &\simeq \pi_1(\mathcal{G}_{\Gamma(\mathcal{B}_2)}, j) \amalg_{\pi_1(\mathcal{G}_{\Gamma(\mathcal{B}_1)}, 1)} \pi_1(\mathcal{G}_{\Gamma(\mathcal{A}_2)}, j) = \\ &= \pi_1(\mathcal{B}_2) \amalg_{\pi_1(\mathcal{B}_1)} \pi_1(\mathcal{A}_2) \simeq \\ &\simeq \pi_1(\mathcal{A}_1) \amalg_{\{e\}} \pi_1(\mathcal{A}_2) \simeq \pi_1(\mathcal{A}_1) * \pi_1(\mathcal{A}_2). \quad \square \end{aligned}$$

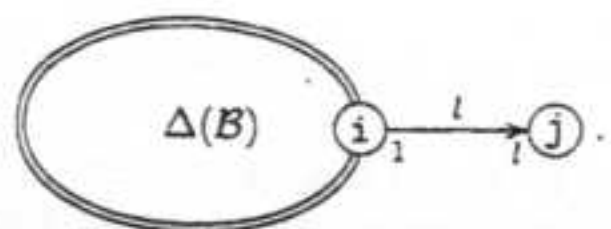
Лема 5. Нехай $\mathcal{A} \in Q_r$ — бімодульна задача, діаграма якої має вигляд



де $\mathcal{B}, \mathcal{A}_2$ — підзадачі $\mathcal{A}$. Тоді

$$\pi_1(\mathcal{A}) \simeq \pi_1(\mathcal{A}_1) * \pi_1(\mathcal{A}_2),$$

де $\mathcal{A}_1$ — підзадача $\mathcal{A}$ з діаграмою



Доведення. Розглянемо бімодульну задачу $\mathcal{B}_1$, розміченою діаграмою $\Delta(\mathcal{B}_1)$ якої є точка $\textcircled{1}$. Розглянемо функтори $F : \mathcal{G}_{\Gamma(\mathcal{B}_1)} \rightarrow \mathcal{G}_{\Gamma(\mathcal{A}_1)}$,

$G : \mathcal{G}_{\Gamma}(\mathcal{B}_1) \rightarrow \mathcal{G}_{\Gamma}(\mathcal{A}_2)$ такі, що $F(1) = j$, $G(1) = j$. Очевидно, $\mathcal{G}_{\Gamma}(\mathcal{A}) \simeq \mathcal{G}_{\Gamma}(\mathcal{A}_1) \amalg_{\mathcal{G}_{\Gamma}(\mathcal{B}_1)} \mathcal{G}_{\Gamma}(\mathcal{A}_2)$. Тоді згідно з теоремою ван Кампена ([6]) і лемою 1

$$\begin{aligned} \pi_1(\mathcal{A}) &= \pi_1(\mathcal{G}_{\Gamma}(\mathcal{A}), j) \simeq \\ &\simeq \pi_1(\mathcal{G}_{\Gamma}(\mathcal{A}_1) \amalg_{\mathcal{G}_{\Gamma}(\mathcal{B}_1)} \mathcal{G}_{\Gamma}(\mathcal{A}_2), j) \simeq \\ &\simeq \pi_1(\mathcal{G}_{\Gamma}(\mathcal{A}_1), j) \amalg_{\pi_1(\mathcal{G}_{\Gamma}(\mathcal{B}_1), 1)} \pi_1(\mathcal{G}_{\Gamma}(\mathcal{A}_2), j) = \\ &= \pi_1(\mathcal{A}_1) \amalg_{\pi_1(\mathcal{B}_1)} \pi_1(\mathcal{A}_2) \simeq \\ &\simeq \pi_1(\mathcal{A}_1) \amalg_{\{e\}} \pi_1(\mathcal{A}_2) \simeq \pi_1(\mathcal{A}_1) * \pi_1(\mathcal{A}_2). \quad \square \end{aligned}$$

Лема 6. Нехай $\mathcal{A} \in Q_r^*$ – бімодульна задача, діаграма якої має вигляд



де $l_i \geq 2$, $s_{ij} < l_i$, $j = 1, 2$, $i = 1, 2, \dots, n$. Тоді фундаментальна група бімодульної задачі $\mathcal{A}$ ізоморфна $\mathbb{Z}$.

Доведення. Аналогічне доведенню леми 2. $\square$

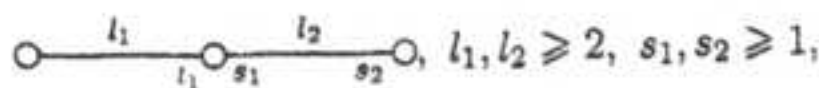
Теорема 1. Нехай $\mathcal{A}$ – бімодульна задача з класу Q_r^* . Тоді її група Пуанкаре є вільною.

Доведення. Випливає з лем 3, 4, 5, 2 і 6. $\square$

Твердження 1. Нехай $\mathcal{A}$ – бімодульна задача з класу Q_r^* . Тоді її група Пуанкаре є вільною групою рангу 1 тоді і тільки тоді, коли діаграма $\Delta(\mathcal{A})$ не містить наступні фрагменти:



$$l_1, l_2 \geq 2, s_{ij} \geq 1, i, j = 1, 2,$$



$$l_1, l_2 \geq 2, s_1, s_2 \geq 1,$$

причому перший фрагмент містить принаймні три ребра.

Доведення. Необхідність випливає з лем 4, 5 і 6.

Достатність. Згідно з лемою 3 достатньо обмежитись випадком лінійки. Оскільки $\Delta(\mathcal{A})$ не містить перший фрагмент, то ваги усіх ребер більші за 1, а через відсутність другого фрагмента всі початкові та кінцеві кроки строго менші за ваги відповідних ребер. Тому твердження випливає з леми 6. $\square$

Зауважимо, що для кожного цілого $n \geq 0$ існує бімодульна задача $\mathcal{A} \in Q_r$ така, що $\pi_1(\mathcal{A})$ є вільною групою рангу n .

Заключні зауваження

У даній роботі розроблено методи знаходження фундаментальних груп бімодульних задач з використанням техніки групоїдів. Ці методи можна ефективно використати і для інших класів бімодульних задач.

Список використаних джерел

1. V. M. Babych, N. S. Golovashchuk. An Application of Covering Techniques // Науковий вісник Ужгородського ун-ту. Сер. матем. і інформ. – 2003. – 8 – с. 4–14.
2. Bautista R. On algebras of strongly unbounded representation type // Comment. Math. Helv. – 1985. – 60, no. 3 – p. 392–399.
3. Bautista R., Gabriel P., Roiter A. V., Salmeron L. Representation-finite algebras and multiplicative bases // Invent. math. – 1985. – 81. – P. 217–286.
4. Bongartz K. A criterion for finite representation type // Math. Ann. – 1984. – 269, no. 1 – p. 1–12.
5. W. W. Crawley-Boevey. Matrix problems and Drozd's theorem // Topics in Algebra, Part 1 (Warsaw, 1988), Banach Center Publ. 26, Part 1, PWN, Warsaw, 1990. – p. 199–222.
6. Габриель П., Цисман. М. Категории частных и теория гомотопий. – М.: Мир, 1971. – 296 с.
7. Ю. А. Дрозд. Матричные задачи и категории матриц. Зап. Науч. Сем. Ленинград. Отдел. Мат. Инст. им. Стеклова (ЛОМИ), 1972. – 28 – С. 144–153.
8. Drozd Yu., Ovsienko S. Coverings of tame boxes and algebras. – Bonn, 2000. – 44 p. – Preprint of Max-Planck Institute, MPI-2000-26.
9. Gabriel P., Roiter A. V. Representations of finite-dimensional algebras. – Algebra VIII, Encyclopaedia of Math.Sc. – Springer, 1992. – Vol.73.
10. Головащук Н. С. Накрывающие некоторого класса свободных матричных задач // Доклады АН УССР, Сер. А. – 1987. – 10 – С. 3–5.

Надійшла до редколегії 10.12.2007

УДК 512.53

Оксана О. Безущак, к. ф.-м. н., доцент

Відношення Гріна інверсної напівгрупи частково визначених ко-скінченних ізометрій дискретної лінійки

Розглядається інверсна напівгрупа частково визначених ко-скінченних ізометрій дискретної лінійки, описуються її ідеали та відношення Гріна.

Ключові слова: інверсна напівгрупа, відношення Гріна.

E-mail: bezusch@univ.kiev.ua

Статтю представив д.ф.-м.н., професор В.В.Кириченко

Вступ

Об'єктом дослідження цієї статті є інверсна напівгрупа ID_∞ частково визначених ко-скінченних ізометрій дискретної лінійки. Деякі базові властивості цього моноїда (зокрема ріст) розглядалися в [1]. Ми продовжуємо вивчення напівгрупи ID_∞ — описуємо її головні ідеали та відношення Гріна.

Розглянемо нескінченну дієдральну групу D_∞ , яка породжується двома твірними a та b і визначальними співвідношеннями $a^2 = 1$, $aba = b^{-1}$. Кожний елемент цієї групи канонічно зображається словом в алфавіті $S = \{a, b\}$ виду $b^k a^\epsilon$, де $k \in \mathbb{N} \cup \{0\}$, $\epsilon \in \{0, 1\}$. Елементи D_∞ можна інтерпретувати як ізометрії дискретної лінійки $\mathbb{Z}$. Тоді твірні a та b матимуть вигляд: $a: x \rightarrow -x$, $x \in \mathbb{Z}$, $b: x \rightarrow x + 1$, $x \in \mathbb{Z}$. Приєднавши до цієї групи ідемпотент $f: x \rightarrow x$, для всіх $x \in \mathbb{Z} \setminus \{0\}$, і $f(0) = \emptyset$, одержимо інверсний моноїд ID_∞ усіх частково визначених ізометрій дискретної лінійки $\mathbb{Z}$, область невизначення яких є скінченною множиною.

Твердження 1 ([1]). Кожний елемент $g \in ID_\infty$ допускає єдине зображення у вигляді

$$g = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t} b^k a^\epsilon, \quad (1)$$

де $\alpha_1, \dots, \alpha_t \in \mathbb{Z}$ такі, що $\alpha_1 < \dots < \alpha_t$, $t \in \mathbb{N} \cup \{0\}$, $k \in \mathbb{Z}$, $\epsilon \in \{0, 1\}$.

Теорема 1 ([1]). Моноїд ID_∞ задається такою системою твірних елементів і визначаль-

Oksana O. Bezushchak, PhD, Associate Professor

Green's relations of the inverse semigroup of partially defined co-finite isometrics of discrete line

The inverse semigroup of partially defined co-finite isometrics of discrete line is considered and it is described ideals and Green's relations.

Key words: inverse semigroup, Green's relations.

них співвідношень:

$$\begin{aligned} ID_\infty = \langle a, b, b^{-1}, f \mid a^2 = Id, bb^{-1} = b^{-1}b = Id, \\ aba = b^{-1}, f^2 = f, af = fa, \\ fb^{-r}fb^r = b^{-r}fb^rf, r \in \mathbb{N} \rangle, \end{aligned}$$

причому система визначальних співвідношень є незвідною.

1 Головні ідеали моноїда ID_∞

Нехай $g \in ID_\infty$ — елемент виду (1). Символом $Dom\,g$ позначимо область визначення елемента g , символом $Ran\,g$ — область значень g . Ко-область визначення $\mathbb{Z} \setminus Dom\,g$ елемента g позначимо як $Codom\,g$, а ко-область значень $\mathbb{Z} \setminus Ran\,g$ елемента g позначимо як $Coran\,g$. Спільне значення $|Dom\,g| = |Ran\,g|$ називається рангом елемента g і позначається $rk(g)$. Спільне значення $|Codom\,g| = |Coran\,g|$ називається ко-рангом елемента g і позначається $cor\,rk(g)$. Також назвемо $b^{-\alpha_1}fb^{\alpha_1} \dots b^{-\alpha_t}fb^{\alpha_t}$ ідемпотентною частиною g і позначимо $idem(g)$, а $b^k a^\epsilon$ — груповою частиною g і позначимо $gr(g)$. Тоді $g = idem(g) \cdot gr(g)$, $Codom\,g = Codom\,idem(g)$,

$$Coran\,g = a^\epsilon(b^k(Coran\,idem(g))),$$

причому елементи g_1 та g_2 моноїда ID_∞ рівні тоді й тільки тоді, коли $gr(g_1) = gr(g_2)$, $Codom\,idem(g_1) = Codom\,idem(g_2)$.

Зауваження 1. Відмітимо, що коли $Dom\,g_1 = Dom\,g_2$ та $Ran\,g_1 = Ran\,g_2$, то, взагалі кажучи, $g_1 \neq g_2$.

Лема 1. (а) Нехай $g \in ID_\infty$ виду (1). Тоді

$$\text{Dom } g = \mathbb{Z} \setminus \{\alpha_1, \dots, \alpha_t\}, \text{ Ran } g = \mathbb{Z} \setminus \{(-1)^\epsilon(\alpha_1 + k), \dots, (-1)^\epsilon(\alpha_t + k)\}.$$

(б) Нехай $g_1 = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t} b^{\bar{k}} a^{\bar{\epsilon}}$,

$$g_2 = b^{-\beta_1} f b^{\beta_1} \dots b^{-\beta_s} f b^{\beta_s} b^k a^\epsilon \in ID_\infty$$

виду (1). Тоді $\text{Dom } g_1 g_2 = \mathbb{Z} \setminus \{\alpha_1, \dots, \alpha_t,$

$$-\bar{k} + (-1)^\epsilon \beta_1, \dots, -\bar{k} + (-1)^\epsilon \beta_s\},$$

$$\text{Ran } g_1 g_2 = \mathbb{Z} \setminus \{(-1)^{\epsilon+\bar{\epsilon}}(\alpha_1 + \bar{k} + (-1)^\epsilon k),$$

$$\dots, (-1)^{\epsilon+\bar{\epsilon}}(\alpha_t + \bar{k} + (-1)^\epsilon k),$$

$$(-1)^\epsilon(\beta_1 + k), \dots, (-1)^\epsilon(\beta_s + k)\}.$$

Доведення. (а) Елемент $h = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t}$ є ідемпотентом. І оскільки $h(x) = x$ для довільного $x \in \text{Dom } h = \mathbb{Z} \setminus \{\alpha_1, \dots, \alpha_t\}$, то $\text{Ran } h = \mathbb{Z} \setminus \{\alpha_1, \dots, \alpha_t\}$. Тому

$$\text{Ran } h b^k = \mathbb{Z} \setminus \{\alpha_1 + k, \dots, \alpha_t + k\}.$$

Якщо $\epsilon = 1$, то $ga = hb^k$, і оскільки $y \in \text{Ran } g \Leftrightarrow -y \in \text{Ran } ga$, то

$$\text{Ran } h b^k a = \mathbb{Z} \setminus \{-\alpha_1 - k, \dots, -\alpha_t - k\}.$$

(б) При $\bar{\epsilon} = 1$ маємо

$$g_1 g_2 = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t} b^{\bar{k}+\beta_1} f b^{-\bar{k}-\beta_1} \dots b^{\bar{k}+\beta_s} f b^{-\bar{k}-\beta_s} b^{\bar{k}-k} a^{\epsilon+1}.$$

Якщо $\bar{\epsilon} = 0$, то

$$g_1 g_2 = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t} b^{\bar{k}-\beta_1} f b^{-\bar{k}+\beta_1} \dots b^{\bar{k}-\beta_s} f b^{-\bar{k}+\beta_s} b^{\bar{k}+k} a^\epsilon.$$

Далі залишилось скористатися п.(а). $\square$

Нагадаємо, що множини виду $M \cdot d$, $d \in M$, називаються лівими головними ідеалами моноїда M , множини виду $d \cdot M$, $d \in M$, називаються правими головними ідеалами моноїда M , а множини виду $M \cdot d \cdot M$, $d \in M$, називаються двосторонніми головними ідеалами моноїда M .

Твердження 2. Нехай $d \in ID_\infty$.

(1) Лівий головний ідеал $I_l(d)$ моноїда ID_∞ має вигляд $\{g \in ID_\infty \mid \text{Ran } g \subseteq \text{Ran } d\}$.

(2) Правий головний ідеал $I_r(d)$ моноїда ID_∞ має вигляд $\{g \in ID_\infty \mid \text{Dom } g \subseteq \text{Dom } d\}$.

Доведення. Нехай $d = b^{-\beta_1} f b^{\beta_1} \dots b^{-\beta_s} f b^{\beta_s} b^k a^\epsilon$.

(1) Покажемо спочатку, що для довільного $g \in I_l(d)$ має місце включення $\text{Ran } g \subseteq \text{Ran } d$. Справді, $g = \bar{g} \cdot d$ для деякого $\bar{g} = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t} b^{\bar{k}} a^{\bar{\epsilon}}$. Тому згідно леми 1

$$\begin{aligned} \text{Ran } \bar{g} \cdot d = \mathbb{Z} \setminus \{ & (-1)^{\epsilon+\bar{\epsilon}}(\alpha_1 + \bar{k} + (-1)^\epsilon k), \dots, \\ & (-1)^{\epsilon+\bar{\epsilon}}(\alpha_t + \bar{k} + (-1)^\epsilon k), (-1)^\epsilon(\beta_1 + k), \dots, \\ & (-1)^\epsilon(\beta_s + k)\} \subseteq \text{Ran } d = \mathbb{Z} \setminus \{ & (-1)^\epsilon(\beta_1 + k), \\ & \dots, (-1)^\epsilon(\beta_s + k)\}. \end{aligned}$$

Навпаки, нехай $\text{Ran } g \subseteq \text{Ran } d$,

$$g = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t} b^{\bar{k}} a^{\bar{\epsilon}}.$$

Виберемо

$$\bar{g} = b^{-x_1} f b^{x_1} \dots b^{-x_r} f b^{x_r} b^x a^y$$

так, щоб

$$g = b^{-x_1} f b^{x_1} \dots b^{-x_r} f b^{x_r} b^{x-(-1)^y \beta_1} f b^{-x+(-1)^y \beta_1} \dots b^{x-(-1)^y \beta_s} f b^{-x+(-1)^y \beta_s} b^{x+(-1)^y k} a^{\epsilon+y}.$$

Тоді $b^{\bar{k}} a^{\bar{\epsilon}} = b^{x+(-1)^y k} a^{\epsilon+y}$, тобто $y = \bar{\epsilon} - \epsilon \pmod{2}$, $x = \bar{k} + (-1)^{\bar{\epsilon}-\epsilon+1} k$, і $\{\alpha_1, \dots, \alpha_t\} = \{x_1, \dots, x_r, -x + (-1)^y + \beta_1, \dots, -x + (-1)^y + \beta_s\}$. Але за припущенням $\{(-1)^\epsilon(\beta_1 + k), \dots, (-1)^\epsilon(\beta_s + k)\} \subseteq \{(-1)^\epsilon(\alpha_1 + \bar{k}), \dots, (-1)^\epsilon(\alpha_t + \bar{k})\}$. Тому $(-1)^\epsilon(\beta_1 + k) = (-1)^\epsilon(\alpha_{i_1} + \bar{k})$, $\dots$, $(-1)^\epsilon(\beta_s + k) = (-1)^\epsilon(\alpha_{i_s} + \bar{k})$ для деяких попарно різних $i_1, \dots, i_s \in \{1, \dots, t\}$. Таким чином, $\{\alpha_1, \dots, \alpha_t\} = \{x_1, \dots, x_r, -\bar{k} + (-1)^{\bar{\epsilon}-\epsilon}(k + \beta_1), \dots, -\bar{k} + (-1)^{\bar{\epsilon}-\epsilon}(k + \beta_s)\} = \{x_1, \dots, x_r, -\bar{k} + (-1)^\epsilon(-1)^{\bar{\epsilon}}(\bar{k} + \alpha_{i_1}), \dots, -\bar{k} + (-1)^\epsilon(-1)^{\bar{\epsilon}}(\bar{k} + \alpha_{i_s})\} = \{x_1, \dots, x_r, \alpha_{i_1}, \dots, \alpha_{i_s}\}$. Отже, виберемо $r = t - s$ і $x_1 < \dots < x_r$ так, щоб $\{x_1, \dots, x_r\} = \{\alpha_1, \dots, \alpha_t\} \setminus \{\alpha_{i_1}, \dots, \alpha_{i_s}\}$.

(2) Покажемо, що для довільного $g \in I_r(d)$ має місце включення $\text{Dom } g \subseteq \text{Dom } d$. Справді, $g = d \cdot \bar{g}$ для деякого

$$\bar{g} = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t} b^{\bar{k}} a^{\bar{\epsilon}}.$$

Тому згідно леми 1 $\text{Dom } g = \mathbb{Z} \setminus \{\beta_1, \dots, \beta_s, -k + (-1)^\epsilon \alpha_1, \dots, -k + (-1)^\epsilon \alpha_t\} \subseteq \text{Dom } d = \mathbb{Z} \setminus \{\beta_1, \dots, \beta_s\}$.

Навпаки, нехай $\text{Dom } g \subseteq \text{Dom } d$,

$$g = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t} b^{\bar{k}} a^{\bar{\epsilon}}.$$

Виберемо

$$\bar{g} = b^{-x_1} f b^{x_1} \dots b^{-x_r} f b^{x_r} b^x a^y$$

так, щоб

$$g = b^{-\beta_1} f b^{\beta_1} \dots b^{-\beta_s} f b^{\beta_s} b^{k-(-1)^{\varepsilon} x_1} f \\ b^{-k+(-1)^{\varepsilon} x_1} \dots b^{k-(-1)^{\varepsilon} x_r} f b^{-k+(-1)^{\varepsilon} x_r}$$

$b^{k+(-1)^{\varepsilon} x} a^{y+\varepsilon}$. Тоді $b^k a^{\varepsilon} = b^{k+(-1)^{\varepsilon} x} a^{y+\varepsilon}$, тобто $y = \bar{\varepsilon} - \varepsilon$, $x = (-1)^{\varepsilon}(\bar{k} - k)$, і $\{\alpha_1, \dots, \alpha_t\} = \{\beta_1, \dots, \beta_s, -k + (-1)^{\varepsilon} + x_1, \dots, -k + (-1)^{\varepsilon} + x_r\}$. Але за припущенням $\{\beta_1, \dots, \beta_s\} \subseteq \{\alpha_1, \dots, \alpha_t\}$. Тому виберемо $x_1 < \dots < x_r$ так, щоб $\{-k + (-1)^{\varepsilon} x_1, \dots, -k + (-1)^{\varepsilon} x_r\} = \{\alpha_1, \dots, \alpha_t\} \setminus \{\beta_1, \dots, \beta_s\}$, тобто $i_1, \dots, i_r \in \{1, \dots, t\}$ так, щоб $x_1 = (-1)^{\varepsilon}(\alpha_{i_1} + k)$, $\dots$, $x_r = (-1)^{\varepsilon}(\alpha_{i_r} + k)$, де або $i_1 < \dots < i_r$, або $i_r < \dots < i_1$. $\square$

Наслідок 1. (1) Для $d_1, d_2 \in ID_{\infty}$ має місце включення $I_l(d_1) \subseteq I_l(d_2)$ тоді й лише тоді, коли $\text{Ran } d_1 \subseteq \text{Ran } d_2$.

(2) Для $d_1, d_2 \in ID_{\infty}$ має місце включення $I_r(d_1) \subseteq I_r(d_2)$ тоді й лише тоді, коли $\text{Dom } d_1 \subseteq \text{Dom } d_2$.

Для довільної скінченної підмножини $A \subset \mathbb{Z}$ позначимо $I_l(A) := \{g \in ID_{\infty} \mid \text{Ran } g \subseteq \mathbb{Z} \setminus A\}$ та $I_r(A) := \{g \in ID_{\infty} \mid \text{Dom } g \subseteq \mathbb{Z} \setminus A\}$. Тоді

Твердження 3. (1) Множина $I_l(A)$ ($I_r(A)$) є лівим (правим) головним ідеалом моноїда ID_{∞} .

(2) Існує взаємно однозначна відповідність між скінченними підмножинами $\mathbb{Z}$ та лівими (правими) головними ідеалами моноїда ID_{∞} .

Доведення. (1) Нехай $A = \{\alpha_1, \dots, \alpha_t\}$. Тоді рівність $I_r(A) = I_r(d)$, де $d = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t}$, випливає з твердження 2 (2) і того, що $\text{Dom } d = \mathbb{Z} \setminus \{\alpha_1, \dots, \alpha_t\}$. Застосувавши перетворення $g \mapsto g^{-1}$ можна отримати і двоїсте твердження про ліві головні ідеали.

(2) легко випливає з (1). $\square$

Означення 1. Нехай g — довільний елемент напівгрупи ID_{∞} , його область невизначення $\text{Codom } g$ є деякою множиною $\{\alpha_1, \dots, \alpha_t\}$, де $t = |\text{Codom } g| \geq 1$. Шириною $w(g)$ елемента g будемо називати набір натуральних чисел $(n_1, \dots, n_{t-1})$, де $n_i = \alpha_{i+1} - \alpha_i$ для $i = 1, \dots, t-1$ при $t > 1$. Якщо $t = 1$, то вважатимемо шириною $w(g)$ елемента g рівною $\emptyset$.

Нехай тепер $g_1, g_2 \in ID_{\infty}$.

Означення 2. Нехай ширина елемента g_1 та g_2 не рівна $\emptyset$. Будемо казати, що ширина $w(g_1) = (n_1, \dots, n_{t-1})$ елемента g_1 не менша за ширину $w(g_2) = (m_1, \dots, m_{s-1})$ елемента g_2 (тобто $w(g_2) \leq w(g_1)$), якщо виконуються такі умови:

$$(1) \text{cork}(g_1) \leq \text{cork}(g_2);$$

$$(2) \text{існують такі } j_1, \dots, j_t \in \{1, \dots, s\}, j_1 < \dots < j_t, \text{ що або } n_1 = m_{j_1} + \dots + m_{j_2-1}, \dots, n_{t-1} = m_{j_{t-1}} + \dots + m_{j_t-1}, \text{ або } n_{t-1} = m_{j_1} + \dots + m_{j_2-1}, \dots, n_1 = m_{j_{t-1}} + \dots + m_{j_t-1}.$$

Окремо вважатимемо шириною $\emptyset$ елемента $g \in ID_{\infty}$ з $|\text{Codom } g| = 1$ не меншою за ширину довільного іншого елемента з ID_{∞} .

Умову (2) можна переписати в іншому вигляді. Нехай $g_1 = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t} b^{k_1} a^{\varepsilon_1}$, $g_2 = b^{-\mu_1} f b^{\mu_1} \dots b^{-\mu_s} f b^{\mu_s} b^{k_2} a^{\varepsilon_2}$. Тоді:

$$(2^*) \text{існують такі } j_1, \dots, j_t \in \{1, \dots, s\}, j_1 < \dots < j_t, \text{ та } \delta \in \mathbb{Z}, \text{ що або } b^{\delta} : \alpha_1 \mapsto \mu_{j_1}, \dots, b^{\delta} : \alpha_t \mapsto \mu_{j_t}, \text{ або } b^{\delta} : -\alpha_t \mapsto \mu_{j_1}, \dots, b^{\delta} : -\alpha_1 \mapsto \mu_{j_t}.$$

Зауваження 2. Відмітимо, що умова $w(g_1) = w(g_2)$ означає одночасно виконання $w(g_1) \leq w(g_2)$ та $w(g_2) \leq w(g_1)$, тобто згідно означення 2 або $w(g_1) = \emptyset$ та $w(g_2) = \emptyset$, або:

$$(1) \text{cork}(g_1) = \text{cork}(g_2);$$

$$(2) n_1 = m_1, \dots, n_{t-1} = m_{t-1} \text{ або } n_{t-1} = m_1, \dots, n_1 = m_{t-1}, \text{ де, як і раніше, } w(g_1) = (n_1, \dots, n_{t-1}), w(g_2) = (m_1, \dots, m_{s-1}).$$

Твердження 4. Нехай $d \in ID_{\infty}$. Двосторонній головний ідеал $I(d)$ моноїда ID_{∞} має вигляд $\{g \in ID_{\infty} \mid w(g) \leq w(d)\}$.

Доведення. Нехай $d = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t} b^q a^{\varepsilon}$. Також нехай $g_1 = b^{-\beta_1} f b^{\beta_1} \dots b^{-\beta_s} f b^{\beta_s} b^{k_1} a^{\varepsilon_1}$, $g_2 = b^{-\gamma_1} f b^{\gamma_1} \dots b^{-\gamma_r} f b^{\gamma_r} b^{k_2} a^{\varepsilon_2}$ — деякі елементи із ID_{∞} . Тоді згідно леми 1 (b)

$$\text{Codom } g_1 \cdot d \cdot g_2 = \mathbb{Z} \setminus \{\beta_1, \dots, \beta_s,$$

$$-k_1 + (-1)^{\varepsilon_1} \alpha_1, \dots, -k_1 + (-1)^{\varepsilon_1} \alpha_t,$$

$$-k_1 + (-1)^{\varepsilon_1+1}q + (-1)^{\varepsilon_1+\varepsilon}\gamma_1, \dots, -k_1 + (-1)^{\varepsilon_1+1}q + (-1)^{\varepsilon_1+\varepsilon}\gamma_r\}, \quad (2)$$

де $\beta_1 < \dots < \beta_s$, $-k_1 + (-1)^{\varepsilon_1}\alpha_1 < \dots < -k_1 + (-1)^{\varepsilon_1}\alpha_t$ або $-k_1 + (-1)^{\varepsilon_1}\alpha_t < \dots < -k_1 + (-1)^{\varepsilon_1}\alpha_1$, $-k_1 + (-1)^{\varepsilon_1+1}q + (-1)^{\varepsilon_1+\varepsilon}\gamma_1 < \dots < -k_1 + (-1)^{\varepsilon_1+1}q + (-1)^{\varepsilon_1+\varepsilon}\gamma_r$ або $-k_1 + (-1)^{\varepsilon_1+1}q + (-1)^{\varepsilon_1+\varepsilon}\gamma_r < \dots < -k_1 + (-1)^{\varepsilon_1+1}q + (-1)^{\varepsilon_1+\varepsilon}\gamma_1$.

Якщо тепер g_1, g_2 такі, що елемент $g \in I(d)$ дорівнює $g = g_1 \cdot d \cdot g_2$, то $\text{cork}(g_1 \cdot d \cdot g_2) \geq \max\{s, t, r\} \geq \text{cork}(d) = t$, тобто виконується умова (1) означення 2. Далі зауважимо, що елементи множини $\text{Codom } g_1 \cdot d \cdot g_2$ (див. (2)), взагалі кажучи, не впорядковані по зростанню, більш того, можуть збігатися. Тому залишимо по одному входженню однакових елементів і впорядкуємо їх строго по зростанню. Матимемо: $\text{Codom } g_1 \cdot d \cdot g_2 = \{\dots, -k_1 + \alpha_1, \dots, -k_1 + \alpha_2, \dots, -k_1 + \alpha_t, \dots\}$ при $\varepsilon_1 = 0$ та $\text{Codom } g_1 \cdot d \cdot g_2 = \{\dots, -k_1 - \alpha_t, \dots, -k_1 - \alpha_{t-1}, \dots, -k_1 - \alpha_1, \dots\}$ при $\varepsilon_1 = 1$. І оскільки $\text{Codom } d = \{\alpha_1, \dots, \alpha_t\}$, то при $\varepsilon_1 = 0$ одержимо $b^{-k_1} : \alpha_1 \mapsto -k_1 + \alpha_1, \dots, b^{-k_1} : \alpha_t \mapsto -k_1 + \alpha_t$, а при $\varepsilon_1 = 1$ одержимо $b^{-k_1} : -\alpha_1 \mapsto -k_1 - \alpha_t, \dots, b^{-k_1} : -\alpha_t \mapsto -k_1 - \alpha_1$. Тим самим, умова (2*) виконується. Тому $w(g) \leq w(d)$.

Навпаки, нехай $w(g) \leq w(d)$,

$$g = b^{-\mu_1} f b^{\mu_1} \dots b^{-\mu_m} f b^{\mu_m} b^{\bar{k}} a^{\bar{\varepsilon}}.$$

Знайдемо такі

$$g_1 = b^{-\beta_1} f b^{\beta_1} \dots b^{-\beta_s} f b^{\beta_s} b^{k_1} a^{\varepsilon_1}$$

та

$$g_2 = b^{-\gamma_1} f b^{\gamma_1} \dots b^{-\gamma_r} f b^{\gamma_r} b^{k_2} a^{\varepsilon_2},$$

що $g = g_1 \cdot d \cdot g_2$. Але $g = g_1 \cdot d \cdot g_2$ тоді й лише тоді, коли

$$gr(g) = gr(g_1 \cdot d \cdot g_2), \quad (3)$$

$$\text{Codom idem}(g) = \text{Codom idem}(g_1 \cdot d \cdot g_2). \quad (4)$$

Використовуючи лему 1 із рівності (3) отримаємо систему

$$\begin{cases} k_1 + (-1)^{\varepsilon_1}q + (-1)^{\varepsilon_1+\varepsilon}k_2 = \bar{k}, \\ \varepsilon_1 + \varepsilon_2 + \varepsilon = \bar{\varepsilon} \pmod{2}, \end{cases} \quad (5)$$

яка, легко видно, є сумісною відносно $k_1, k_2, \varepsilon_1, \varepsilon_2$. Права частина рівності (4) виду (2), а ліва $\text{Codom idem}(g) = \{\mu_1, \dots, \mu_m\}$.

Крім того, умова (1) означення 2 означає, що $\text{cork}(d) = |\{\alpha_1, \dots, \alpha_t\}| = t \leq \text{cork}(g) = |\{\mu_1, \dots, \mu_m\}| = m$, а умова (2*) означає існування таких $i_1, \dots, i_t \in \{1, \dots, m\}$, $i_1 < \dots < i_t$, та $\delta \in \mathbb{Z}$, що або

$$b^\delta : \alpha_1 \mapsto \mu_{i_1}, \dots, b^\delta : \alpha_t \mapsto \mu_{i_t}, \quad (6)$$

або

$$b^\delta : -\alpha_t \mapsto \mu_{i_1}, \dots, b^\delta : -\alpha_1 \mapsto \mu_{i_t}. \quad (7)$$

Тому виберемо $\beta_1, \dots, \beta_s, \gamma_1, \dots, \gamma_r$ так, щоб $\beta_1 < \dots < \beta_s, \gamma_1 < \dots < \gamma_r$ і $\{\beta_1, \dots, \beta_s, \gamma_1, \dots, \gamma_r\} = \{\mu_1, \dots, \mu_m\} \setminus \{\mu_{i_1}, \dots, \mu_{i_t}\}$. Більш того, якщо дія елемента b^δ задається співвідношеннями (6), то в якості розв'язків системи (5) виберемо $k_1 = -\delta, \varepsilon_1 = 0$, бо $-k_1 + (-1)^{\varepsilon_1}\alpha_1 = \mu_{i_1}, \dots, -k_1 + (-1)^{\varepsilon_1}\alpha_t = \mu_{i_t}$. Звідки, $k_2 = (-1)^\varepsilon(\bar{k} - q + \delta), \varepsilon_2 = \bar{\varepsilon} - \varepsilon \pmod{2}$. Якщо ж дія елемента b^δ задається співвідношеннями (7), то в якості розв'язків системи (5) виберемо $k_1 = -\delta, \varepsilon_1 = 1$, бо $-k_1 + (-1)^{\varepsilon_1}\alpha_t = \mu_{i_1}, \dots, -k_1 + (-1)^{\varepsilon_1}\alpha_1 = \mu_{i_t}$. Звідки, $k_2 = (-1)^{\varepsilon+1}(\bar{k} + q + \delta), \varepsilon_2 = \bar{\varepsilon} - \varepsilon - 1 \pmod{2}$. Отже, потрібні g_1 та g_2 існують і твердження доведено. $\square$

Наслідок 2. Для $d_1, d_2 \in ID_\infty$ має місце включення $I(d_1) \subseteq I(d_2)$ тоді й лише тоді, коли $w(d_1) \leq w(d_2)$.

Довільний скінченний набір натуральних чисел $(u_1, \dots, u_r), r \geq 1$, є шириною елементів деякого класу з ID_∞ . Ці елементи матимуть вигляд:

$$b^{-\alpha} f b^{\alpha} b^{-\alpha-u_1} f b^{\alpha+u_1} \dots b^{-\alpha-u_1-\dots-u_r} f b^{\alpha+u_1+\dots+u_r} b^k a^\varepsilon$$

для довільних $\alpha \in \mathbb{Z}, k \in \mathbb{Z}, \varepsilon \in \{0, 1\}$. Елементи ширини 0 мають вигляд: $b^{-\alpha} f b^{\alpha} b^k a^\varepsilon, \alpha \in \mathbb{Z}, k \in \mathbb{Z}, \varepsilon \in \{0, 1\}$.

Позначимо символом $I(w)$ множину всіх $g \in ID_\infty$, таких що $w(g) \leq w$, де w або 0, або деякий набір $(u_1, \dots, u_r), r \geq 1$, натуральних чисел.

Твердження 5. (1) Множина $I(w)$ є двостороннім ідеалом моноїда ID_∞ .

(2) Якщо $w = (u_1, \dots, u_r)$, то маємо ланцюг строгих включень:

$$I(\emptyset) \supset I((u_1)) \supset I((u_1, u_2)) \supset \dots \supset I((u_1, u_2, \dots, u_r)).$$

Доведення. (1) Для довільних $g \in I(w)$ та $g_1, g_2 \in ID_\infty$ має місце включення $g_1 \cdot g \cdot g_2 \in I(w)$, бо $Codom\ g_1 \cdot g \cdot g_2$ є ущільненням множини $Codom\ g_1 \cdot g$, тому $w(g_1 \cdot g \cdot g_2) \leq w(g_1 \cdot g)$. У свою чергу, нерівність $w(g_1 \cdot g) \leq w(g)$ випливає з вигляду $Codom\ g_1 \cdot g$ (лема 1 (b)) і умови (2*) означення 2.

(2) легко випливає із означення 2. $\square$

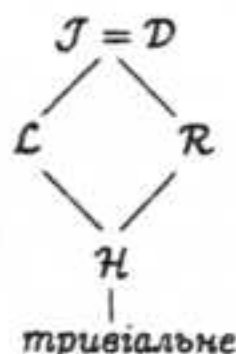
2 Відношення Гріна моноїда ID_∞

Теорема 2. Нехай g_1, g_2 — довільні елементи з ID_∞ . Тоді

- (1) $g_1 \mathcal{L} g_2$ тоді і тільки тоді, коли $Ran\ g_1 = Ran\ g_2$.
- (2) $g_1 \mathcal{R} g_2$ тоді і тільки тоді, коли $Dom\ g_1 = Dom\ g_2$.
- (3) $g_1 \mathcal{H} g_2$ тоді і тільки тоді, коли $Ran\ g_1 = Ran\ g_2$, $Dom\ g_1 = Dom\ g_2$.
- (4) $g_1 \mathcal{J} g_2$ тоді і тільки тоді, коли $w(g_1) = w(g_2)$.
- (5) $g_1 \mathcal{D} g_2$ тоді і тільки тоді, коли існує елемент $g \in ID_\infty$ такий, що $Ran\ g_1 = Ran\ g$, $Dom\ g = Dom\ g_2$.

Доведення. Пункти (1) та (2) випливають з твердження 2, (3) та (5) є наслідками пунктів (1) та (2). Пункт (4) виконується згідно твердження 4. $\square$

Наслідок 3. Діаграма відношень Гріна є такою:



Доведення. Строгі включення $\mathcal{H} \subset \mathcal{L}$, $\mathcal{H} \subset \mathcal{R}$, $\mathcal{L} \subset \mathcal{D}$, $\mathcal{R} \subset \mathcal{D}$ очевидні згідно теореми 2 та леми 1. Те, що $\mathcal{H}$ не є тривіальним відношенням "рівності" випливає із зауваження 1.

Покажемо, що $\mathcal{J} = \mathcal{D}$. Нехай $g_1, g_2 \in ID_\infty$,

$$g_1 = b^{-\alpha_1} f b^{\alpha_1} \dots b^{-\alpha_t} f b^{\alpha_t} b^{k_1} a^{\epsilon_1},$$

$$g_2 = b^{-\beta_1} f b^{\beta_1} \dots b^{-\beta_s} f b^{\beta_s} b^{k_2} a^{\epsilon_2},$$

такі що $g_1 \mathcal{D} g_2$, тобто згідно теореми 2 існує

$$g \in ID_\infty, \quad g = b^{-x_1} f b^{x_1} \dots b^{-x_r} f b^{x_r} b^k a^\epsilon,$$

таке, що

$$Ran\ g_1 = \mathbb{Z} \setminus \{(-1)^{\epsilon_1}(\alpha_1 + k_1), \dots, (-1)^{\epsilon_t}(\alpha_t + k_t)\}$$

$$= Ran\ g = \mathbb{Z} \setminus \{(-1)^\epsilon(x_1 + k), \dots, (-1)^\epsilon(x_r + k)\},$$

$$Dom\ g = \mathbb{Z} \setminus \{x_1, \dots, x_r\} =$$

$$Dom\ g_2 = \mathbb{Z} \setminus \{\beta_1, \dots, \beta_s\}.$$

Тоді $t = r = s$ і $cork(g_1) = t = s = cork(g_2)$. Також маємо, що $x_1 = \beta_1, \dots, x_t = \beta_t$ і або $(-1)^{\epsilon_1}(\alpha_1 + k_1) = (-1)^\epsilon(\beta_1 + k), \dots, (-1)^{\epsilon_t}(\alpha_t + k_t) = (-1)^\epsilon(\beta_t + k)$, або $(-1)^{\epsilon_1}(\alpha_1 + k_1) = (-1)^\epsilon(\beta_t + k), \dots, (-1)^{\epsilon_t}(\alpha_t + k_t) = (-1)^\epsilon(\beta_1 + k)$.

Звідси, якщо $\epsilon_1 = 0$ і $\alpha_1 + k_1 = (-1)^\epsilon(\beta_1 + k), \dots, \alpha_t + k_t = (-1)^\epsilon(\beta_t + k)$, то з $\alpha_1 < \dots < \alpha_t$ та $\beta_1 < \dots < \beta_t$ отримаємо $\epsilon = 0$ і $\alpha_2 - \alpha_1 = \beta_2 - \beta_1, \dots, \alpha_t - \alpha_{t-1} = \beta_t - \beta_{t-1}$. Тому $w(g_1) = w(g_2)$. Якщо $\epsilon_1 = 0$ і $\alpha_1 + k_1 = (-1)^\epsilon(\beta_t + k), \dots, \alpha_t + k_t = (-1)^\epsilon(\beta_1 + k)$, то $\epsilon = 1$ і $\alpha_2 - \alpha_1 = \beta_t - \beta_{t-1}, \dots, \alpha_t - \alpha_{t-1} = \beta_2 - \beta_1$. Тому умова (2) зауваження 2 виконується і $w(g_1) = w(g_2)$.

Аналогічно при $\epsilon_1 = 1$ і $-(\alpha_1 + k_1) = (-1)^\epsilon(\beta_1 + k), \dots, -(\alpha_t + k_t) = (-1)^\epsilon(\beta_t + k)$ маємо $\epsilon = 1$, $\alpha_2 - \alpha_1 = \beta_2 - \beta_1, \dots, \alpha_t - \alpha_{t-1} = \beta_t - \beta_{t-1}$ і $w(g_1) = w(g_2)$. Якщо ж $\epsilon_1 = 1$ і $-(\alpha_1 + k_1) = (-1)^\epsilon(\beta_t + k), \dots, -(\alpha_t + k_t) = (-1)^\epsilon(\beta_1 + k)$, то отримаємо $\epsilon = 0$, $\alpha_2 - \alpha_1 = \beta_t - \beta_{t-1}, \dots, \alpha_t - \alpha_{t-1} = \beta_2 - \beta_1$ і знову $w(g_1) = w(g_2)$. Отже, $g_1 \mathcal{J} g_2$. $\square$

Список використаних джерел

1. O. Bezushchak, On growth of the inverse semigroup of partially defined co-finite automorphisms of integers, Algebra and discrete mathematics (2004), no. 2, 5–11.
2. P. Catarino, P. Higgins, The monoid of orientation-preserving mappings on a chain, Semigroup Forum 58 (1999), no. 2, 190–206.
3. O. Ganyushkin, V. Mazorchuk, Introduction to classical finite transformation semigroups, Springer-Verlag (2008), 330p.

Надійшла до редколегії 05.09.2007

УДК 519.21

Світлана Я. Гончарова, к.ф.-м.н.

Рівняння Беллмана для імпульсних процесів переносу

Досліджується оптимальне стохастичне керування імпульсними процесами переносу. Виводиться рівняння Беллмана.

Ключові слова: імпульсний процес переносу, оптимальне стохастичне керування, рівняння Беллмана.

*E-mail: goncharov@kw.ukrtel.net

Статтю представив доктор фіз.-мат. наук Козаченко Ю.В.

Вступ

Провідною метою даної роботи є подальший розвиток теорії напівмарковських випадкових еволюцій.

Об'єктом дослідження роботи є модель імпульсного процесу переносу в напівмарковському випадковому середовищі, що має інтерпретацію напівмарковської випадкової еволюції.

Предмет дослідження – оптимальне стохастичне керування імпульсними процесами переносу.

Методами дослідження даної роботи є методи напівмарковських випадкових еволюцій [4].

Проблеми оптимального стохастичного керування розглядалась лише для марковських систем [1] – [3], [5] – [7], тому одержані результати відрізняються від раніше відомих, так як досліджуються напівмарковські процеси.

Основна частина

Нехай (Ω, F, P) імовірнісний простір, на якому розглядатимемо випадкові величини із значеннями у вимірному просторі (X, X) .

Означення 1. Напівмарковським процесом називається процес $x(t)$, що задається співвідношеннями:

$$x(t) = x_{v(t)}, \quad (1)$$

$$v(t) = \max \{n: \tau_n \leq t\},$$

$$\tau_n = \sum_{k=1}^n \theta_k, n \geq 0, \tau_0 = 0, \quad (2)$$

де τ_n – моменти відновлення,

θ_k – невід'ємні випадкові величини, що задають інтервали між моментами відновлення.

Моменти відновлення $\tau_n, n \geq 0$, є марковськими моментами для напівмарковського про-

S. Ya. Goncharova, Cand.Sci. (Phys.-Math.)

The Bellman equation for impulse processes of transfer

The optimal stochastic control for the impulse processes of transfer is investigated. The Bellman equation is proved.

Key Words: impulse process of transfer, optimal stochastic control, Bellman equation.

цесу $x(t)$.

Надалі розглядатимемо регулярні напівмарковські процеси, тобто такі, для яких з імовірністю 1 відбувається скінченне число відновлень за скінченний час:

$$P \{v(t) < +\infty\} = 1, \quad \forall t \in R_+.$$

Означення 2. Керований імпульсний процес переносу $z(t)$ в напівмарковському випадковому середовищі задається розв'язком інтегрального рівняння

$$z(t) = z + \int_T^t v(z(s), x(s), u) ds - \sum_{k=v(\dot{O})}^{v(t)} a(x_k), \quad (3)$$

$$t \geq T \geq 0,$$

де функція $v(z, x, u)$ неперервна;

$\dot{e}$ – параметр керування,

$a(x)$ – невід'ємна, вимірна, обмежена функція на X .

Припустимо, що параметр керування $\dot{e}$ в рівнянні (3) вибирається у вигляді функції від $z(t)$, $x(t)$ та $\gamma(t)$, тобто

$$\dot{e} = u(z(t), x(t), \gamma(t)),$$

де $x(t)$ – напівмарковський процес,

$$\gamma(t) = t - \tau_{v(t)}, t \geq \dot{O}, \text{ – дефект-процес.}$$

Таке керування називатимемо марковським.

Вважатимемо, що $\dot{e} = u(z, x, t)$ є дійсною скалярною або векторною функцією.

Функцію $\dot{e} = u(z, x, t)$ назвемо допустимою, якщо функція $v(z, x, u(z, x, t))$ в рівнянні (3) неперервна; неперервно диференційовна по z ; обмежена по x ; невід'ємна і така, що

$$v(0, x, u(0, x, t)) = 0, \quad \forall x \in X, t \in R_+,$$

та

$$0 \leq v'_z \leq K, K > 0.$$

Клас допустимих функцій позначимо U .

Кожній функції $u \in U$ (допустимому керуванню) відповідає імпульсний процес переносу $z(t)$, що є єдиним розв'язком рівняння (3).

Надалі писатимемо замість $u(z(t), x(t), \gamma(t))$ просто $u(t)$.

Розглянемо варіант задачі оптимального стохастичного керування, що полягає у виборі такого керування $u \in U$,

де U – компактна множина допустимих керувань у просторі дійсних неперервних функцій на $R_+ \times X \times R_+$, що задовольняють локальній умові Ліпшица по z ,

яке переводить трьохкомпонентний неперервний справа строго марковський процес $(z(t), x(t), \gamma(t))$,

де $z(t)$ – керований імпульсний процес переносу, визначений в (3),

з початкового стану $(z, x, \dot{O}) \in G_0 \times X \times R_+$ в множину $\bar{G}^0 \times X \times R_+$ з імовірністю 1,

де $G^0 = [0, z_0)$, $z_0 \in R_+$, $\bar{G}^0 = [z_0, +\infty)$, і щоб при цьому мінімізувався в порівнянні з іншими керуваннями із раніше визначеної множини допустимих керувань U функціонал якості (ціни)

$$C(z, x, \dot{O}, u) = E_{z, x, \dot{O}}^u \left[\int_0^{\tau_u} k(z(s), x(s), \gamma(s), u(s)) ds + b(z(\tau_u), x(\tau_u), \gamma(\tau_u)) \right], \quad (4)$$

де $\tau_u = \min\{t \geq \dot{O} : (z(t), x(t), \gamma(t)) \in G^0 \times X \times R_+\}$ – випадковий момент попадання в множину $\bar{G}^0 \times X \times R_+$;

$b(z, x, t)$ – невід'ємна, обмежена та неперервно диференційовна на $R_+ \times X \times R_+$ функція;

$k(z, x, t, u) \geq k_1 > 0$ – неперервна за сукупністю змінних функція.

Якщо $(z(t), x(t), \gamma(t)) \in G^0 \times X \times R_+$, $\forall t < +\infty$, то $\tau_u = +\infty$.

Визначимо

$$C_0(z, x, \dot{O}) = \inf_u C(z, x, \dot{O}, u), \quad (5)$$

де інфімум береться по всіх допустимих керуваннях.

Функцію $C_0(z, x, \dot{O})$ в (5) назвемо оптимальним функціоналом якості або оптимальним функціоналом ціни.

Наступна теорема є модифікацією принципу Беллмана стосовно задач оптимального керуван-

ня стохастичними системами.

Теорема. Нехай $z(t)$ є керованим імпульсним процесом переносу, визначеним в (3), $C(z, x, \dot{O}, u)$ – функціонал якості (ціни), визначений в (4), $\bar{G}^0 \times X \times R_+$ – цільова множина, де $\bar{G}^0 = [z_0, +\infty)$, $z_0 \in R_+$.

Допустиме керування $\bar{u}$, $\bar{u} \in U$, якому відповідає оптимальна якість (ціна) $C_0(z, x, \dot{O}) \in \text{Dom}(L_u)$, є оптимальним тоді і тільки тоді, коли $\bar{N}_0$ задовольняє рівняння

$$\begin{cases} \inf_{u \in U} [L_u C_0(z, x, \dot{O}) + k(z, x, \dot{O}, u)] = 0, \\ \forall (z, x, \dot{O}) \in G^0 \times X \times R_+; \\ C_0(z, x, \dot{O}) = b(z, x, \dot{O}), \\ \forall (z, x, \dot{O}) \in \partial G \times X \times R_+. \end{cases} \quad (6)$$

де $G^0 = [0, z_0)$,

$\partial G = \{0, z_0\}$,

оператор L_u – твірний оператор марковського процесу $(z(t), x(t), \gamma(t))$.

Доведення.

$$\begin{aligned} L_u f(z, x, t) &= \frac{d}{dt} f(z, x, t) + \\ &+ \lambda_x(t) [Pf(z, x, T) - f(z, x, t)] + \\ &+ v(z, x, u) \frac{d}{dz} f(z, x, t) + \\ &+ [Pf(z - a(x), x, t) - f(z, x, t)] \end{aligned} \quad (7)$$

– твірний оператор марковського процесу $(z(t), x(t), \gamma(t))$,

де $\lambda_x(t) = g_x(t) / \bar{G}_x(t)$;

$g_x(t) = dG_x(t) / dt$;

$\bar{G}_x(t) = 1 - G_x(t)$,

$P(x, A) = P\{x_{n+1} \in A / x_n = x\}$,

$G_x(t) = P\{\theta_{n+1} \leq t / x_n = x\}$,

$x_n \in X, A \in X, t \geq 0$;

$Pa(x) = \int_X P(x, dy) a(y)$;

$\forall f(z, x, t) \in C^1(R_+ \times X \times R_+)$, $\forall u \in U$.

Нехай $\bar{u}(t)$ є оптимальним керуванням, тоді

$$\begin{aligned} C_0(z, x, \dot{O}) &= E_{z, x, \dot{O}}^{\bar{u}} \left[\int_0^{\tau_{\bar{u}}} k(z(s), x(s), \gamma(s), \bar{u}(s)) ds + \right. \\ &\quad \left. + b(z(\tau_{\bar{u}}), x(\tau_{\bar{u}}), \gamma(\tau_{\bar{u}})) \right]. \end{aligned} \quad (8)$$

Визначимо процес

$$\bar{z}(t) = z + \int_T^t v(\bar{z}(s), x(s), \bar{u}) ds - \sum_{k=v(\bar{O})}^{v(t)} a(x_k), \quad (9)$$

де керування $\bar{u} = \bar{u}(t)$ таке, що

$$\bar{u}(t) = \begin{cases} \bar{u}(t), & t \geq T + \Delta, \\ u(t), & T < t < T + \Delta, \end{cases} \quad u \in U,$$

де $\Delta > 0$.

Процес $(\bar{z}(t), x(t), \gamma(t))$ є марковським.

Згідно принципу динамічного програмування [2] для оптимізації керування на проміжку $[\bar{O}, \tau_{\bar{u}}]$ при фінальній платі $b(z(\tau_{\bar{u}}), x(\tau_{\bar{u}}), \gamma(\tau_{\bar{u}}))$ можна спочатку оптимізувати керування на проміжку $[\bar{O} + \Delta, \tau_{\bar{u}}]$ для $\Delta > 0$ (з тією ж фінальною платою), а потім оптимізувати керування на проміжку $[\bar{O}, T + \Delta]$ при фінальній платі $C_0(z(T + \Delta), x(T + \Delta), \gamma(T + \Delta))$.

Тоді з оптимальності функціоналу якості C_0 випливає

$$C_0(z, x, \bar{O}) \leq E_{z, x, T}^{\bar{u}} \left[C_0(z(T + \Delta \wedge \tau_{\bar{u}}), x(T + \Delta \wedge \tau_{\bar{u}}), \gamma(T + \Delta \wedge \tau_{\bar{u}})) + \int_0^{T + \Delta \wedge \tau_{\bar{u}}} k(\bar{z}(s), x(s), \gamma(s), \bar{u}(s)) ds \right]. \quad (10)$$

В якості часового індексу використовується $\bar{O} + \Delta \wedge \tau_{\bar{u}}$, так як керування закінчується в момент $\tau_{\bar{u}}$.

Вважатимемо, що

$$\lim_{\Delta \rightarrow 0} P_{z, x, T}^{\bar{u}} \{ \tau_{\bar{u}} \leq T + \Delta \} = 0, \quad (11)$$

$$\forall (z, x, T) \in G^0 \times X \times R_+$$

— це справедливо для неперервних справа процесів [5] $(z(t), x(t), \gamma(t))$, що розглядаються в даній роботі.

Так як, враховуючи (11),

$$\begin{aligned} \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} [E_{z, x, T}^{\bar{u}} (C_0(z(T + \Delta \wedge \tau_{\bar{u}}), x(T + \Delta \wedge \tau_{\bar{u}}), \gamma(T + \Delta \wedge \tau_{\bar{u}})) - C_0(z, x, T))] &= L_{\bar{u}} \cdot C_0(z, x, \bar{O}); \\ \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} E_{z, x, T}^{\bar{u}} \int_T^{T + \Delta \wedge \tau_{\bar{u}}} k(z(s), x(s), \gamma(s), \bar{u}(s)) ds &= \\ &= k(z, x, T, \bar{u}), \end{aligned} \quad (12)$$

то, розділивши нерівність (10) на $\Delta > 0$, отримаємо

$$L_{\bar{u}} C_0(z, x, \bar{O}) + k(z, x, \bar{O}, \bar{u}) \geq 0. \quad (13)$$

Якщо в нерівності (13) замість керування $\bar{u}(t)$ взяти керування $\bar{u}(t)$, то отримаємо рівність. Тобто

$$L_{\bar{u}} C_0(z, x, \bar{O}) + k(z, x, \bar{O}, \bar{u}) = 0.$$

Підсумовуючи все вище доведене, отримаємо рівняння (6).

Доведемо обернений факт.

Нехай функція $C_0(z, x, \bar{O}) \in \text{Dom}(L_{\bar{u}})$ задовольняє рівняння (6).

Покажемо, що керування $\bar{u}(t)$, якому відповідає функція $C_0(z, x, \bar{O})$, є оптимальним керуванням і $C_0(z, x, \bar{O})$ дійсно є оптимальною якістю.

У випадку, коли функція

$$b(z, x, t) \neq 0,$$

введення в критерій якості члена

$$E_{z, x, T}^{\bar{u}} b(z(\tau_{\bar{u}}), x(\tau_{\bar{u}}), \gamma(\tau_{\bar{u}})),$$

що оцінює „кінцевий стан” процесу, не має змісту, якщо не виконується умова

$$E_{z, x, T}^{\bar{u}} \tau_{\bar{u}} < +\infty,$$

$$\forall (z, x, \bar{O}) \in G^0 \times X \times R_+.$$

За аналогом формули Динкіна для напівмарковських випадкових еволюцій [2], враховуючи умову (6) та властивості функції $k(z, x, t, u)$, отримаємо:

$$\begin{aligned} -E_{z, x, \bar{O}}^{\bar{u}} C_0(z(t \wedge \tau_{\bar{u}}), x(t \wedge \tau_{\bar{u}}), \gamma(t \wedge \tau_{\bar{u}})) + \\ + C_0(z, x, \bar{O}) &= -E_{z, x, \bar{O}}^{\bar{u}} \int_0^{t \wedge \tau_{\bar{u}}} L_{\bar{u}} C_0(z(s), x(s), \gamma(s)) ds = \\ &= E_{z, x, \bar{O}}^{\bar{u}} \int_T^{t \wedge \tau_{\bar{u}}} k(z(s), x(s), \gamma(s), \bar{u}(s)) ds \geq \\ &\geq \hat{e}_1 E_{z, x, T}^{\bar{u}} \int_T^{t \wedge \tau_{\bar{u}}} I_{z, x, T}(s, \omega) ds = \\ &= \hat{e}_1 (E_{z, x, T}^{\bar{u}}(t \wedge \tau_{\bar{u}}) - T), \quad T \leq t, \end{aligned}$$

де $I_{z, x, T}(s, \omega)$ — характеристична функція множини тих точок (s, ω) , для яких початковий стан $(z, x, T) \in G^0 \times X \times R_+$.

З невід'ємності функції $C_0(z, x, \bar{O})$ і останнього співвідношення випливає

$$E_{z, x, T}^{\bar{u}} \tau_{\bar{u}} \leq C_0(z, x, T) / k_1 + T,$$

де $E_{z, x, T}^{\bar{u}} \tau_{\bar{u}}$ — середній час перебування процесу $(z(t), x(t), \gamma(t))$ в множині $G^0 \times X \times R_+$ до моменту першого виходу.

Нехай функція $\bar{u}(t)$ — допустиме керування, якому відповідає процес $\bar{z}(t)$, визначений в рівнянні (9).

З рівняння (6) випливає:

$$\begin{aligned} L_{\bar{u}} C_0(z, x, \dot{O}) + k(z, x, \dot{O}, \bar{u}) &= 0, \\ L_{\bar{u}} C_0(z, x, \dot{O}) + k(z, x, \dot{O}, \bar{u}) &\geq 0, \\ \forall (z, x, \dot{O}) &\in G^0 \times X \times R_+. \end{aligned} \quad (14)$$

З аналогу формули Динкіна для напівмарковських випадкових еволюцій [2] і, враховуючи співвідношення (14), отримаємо:

$$\begin{aligned} -E_{z,x,\dot{O}}^{\bar{u}} C_0(z(t \wedge \tau_{\bar{u}}), x(t \wedge \tau_{\bar{u}}), \gamma(t \wedge \tau_{\bar{u}})) + \\ + C_0(z, x, \dot{O}) = -E_{z,x,\dot{O}}^{\bar{u}} \int_0^{t \wedge \tau_{\bar{u}}} L_{\bar{u}} C_0(z(s), x(s), \gamma(s)) ds = \\ = E_{z,x,\dot{O}}^{\bar{u}} \int_T^{t \wedge \tau_{\bar{u}}} k(z(s), x(s), \gamma(s), \bar{u}(s)) ds, \quad T \leq t, \end{aligned} \quad (15)$$

та

$$\begin{aligned} -E_{z,x,\dot{O}}^{\bar{u}} C_0(\bar{z}(t \wedge \tau_{\bar{u}}), x(t \wedge \tau_{\bar{u}}), \gamma(t \wedge \tau_{\bar{u}})) + \\ + C_0(z, x, \dot{O}) = -E_{z,x,\dot{O}}^{\bar{u}} \int_T^{t \wedge \tau_{\bar{u}}} L_{\bar{u}} C_0(\bar{z}(s), x(s), \gamma(s)) ds \leq \\ \leq E_{z,x,\dot{O}}^{\bar{u}} \int_T^{t \wedge \tau_{\bar{u}}} k(\bar{z}(s), x(s), \gamma(s), \bar{u}(s)) ds, \quad T \leq t. \end{aligned} \quad (16)$$

Відмітимо, що

$$\begin{aligned} E_{z,x,\dot{O}}^{\bar{u}} C_0(\bar{z}(t \wedge \tau_{\bar{u}}), x(t \wedge \tau_{\bar{u}}), \gamma(t \wedge \tau_{\bar{u}})) &\rightarrow \\ &\rightarrow E_{z,x,\dot{O}}^{\bar{u}} b(\bar{z}(\tau_{\bar{u}}), x(\tau_{\bar{u}}), \gamma(\tau_{\bar{u}})), \\ E_{z,x,\dot{O}}^{\bar{u}} C_0(z(t \wedge \tau_{\bar{u}}), x(t \wedge \tau_{\bar{u}}), \gamma(t \wedge \tau_{\bar{u}})) &\rightarrow \\ &\rightarrow E_{z,x,\dot{O}}^{\bar{u}} b(z(\tau_{\bar{u}}), x(\tau_{\bar{u}}), \gamma(\tau_{\bar{u}})), \end{aligned} \quad (17)$$

коли $t \rightarrow +\infty$.

З співвідношень (15), (16) і (17) випливає, що

$$\begin{aligned} C(z, x, T, \bar{u}) = E_{z,x,\dot{O}}^{\bar{u}} \left[\int_T^{\tau_{\bar{u}}} k(z(s), x(s), \gamma(s), \bar{u}(s)) ds + \right. \\ \left. + b(z(\tau_{\bar{u}}), x(\tau_{\bar{u}}), \gamma(\tau_{\bar{u}})) \right] = C_0(z, x, T) \leq \\ \leq E_{z,x,\dot{O}}^{\bar{u}} \left[\int_T^{\tau_{\bar{u}}} k(\bar{z}(s), x(s), \gamma(s), \bar{u}(s)) ds + \right. \\ \left. + b(\bar{z}(\tau_{\bar{u}}), x(\tau_{\bar{u}}), \gamma(\tau_{\bar{u}})) \right] = C(z, x, T, \bar{u}), \end{aligned}$$

що показує оптимальність ціни C_0 та керування $\bar{u}(t)$.

Теорему доведено.

Рівняння (6) назвемо рівнянням Беллмана для імпульсного процесу переносу $z(t)$, визначеного в рівнянні (3).

Висновок

Рівняння Беллмана є фундаментальним в теорії керування імпульсними процесами переносу.

Одержані результати доповнюють і узагальнюють відповідні дослідження напівмарковських випадкових еволюцій і співпадають з відомими результатами для марковських систем.

Практична цінність результатів дослідження визначається можливістю їх використання в теорії випадкових процесів.

Імпульсний процес переносу $z(t)$, визначений в рівнянні (3), є процесом накопичення [4], а тому може використовуватись як стохастична модель напівмарковського процесу ризику в страховій математиці.

Список використаних джерел

1. Гихман И.И., Скороход А.В. Управляемые случайные процессы. – К.: Наукова думка, 1977. – 252 с.
2. Дынкин Е.Б., Юшкевич А.А. Управляемые марковские процессы и их приложения. – М.: Наука, 1975. – 335 с.
3. Korolyuk V.S., Swishchuk A.V. Evolution of systems in random media // CRC Press. USA. – 1995. – 356 p.
4. Кушнер Г. Стохастическая устойчивость и управление. – Москва, Мир. – 1969. – 200 с.
5. Портенко Н.И., Скороход А.В. О существовании ε -оптимальных марковских стратегий для управляемых диффузионных процессов // Вопросы статистики и управления случайными процессами. – К.: Ин-т математики АН УССР, 1973. – С. 204–227.
6. Swishchuk A.V. Control of evolutionary stochastic systems // Proc. 12th Prague Conf. – Academic Publish. House. – 1994. – P. 285–288.
7. Хасьминский Р. Устойчивость систем дифференциальных уравнений при случайных возмущениях их параметров. – М.: Наука, 1969. – 368 с.

Надійшла до редколегії 25.05.2006

Валентина Дармосюк, аспірантка

Властивості квазіфробеніусових кілець

Ми розглядаємо зв'язки між властивістю кільця A бути квазіфробеніусовим та такою ж властивістю для кілець eAe , де e ненульовий ідемпотент та $e \neq 1$.

Ключові слова: квазіфробеніусове кільце, сагайдак, алгебра шляхів

V. Darmosiuk, PhD student

Properties of quasi-Frobenius rings

We consider relations between property of a ring A to be quasi-Frobenius and the same property for rings eAe , where e is nonzero idempotent and $e \neq 1$.

Key Words: quasi-Frobenius ring, quiver, path algebra

E-mail: darmosiuk@gmail.com

Статтю представив д.ф.-м.н., професор В.В.Кириченко

1 Вступ

Нехай A артинове справа і зліва кільце (артинове кільце), R – його радикал Джекобсона. Такі кільця є напівдосконалими [1, теорема 10.3.5].

Згідно [1, твердження 11.1.1] правий регулярний модуль A_A розкладається в пряму суму попарно неізоморфних нерозкладних проєктивних A -модулів (головних A -модулів):

$A_A = P_1^{n_1} \oplus \dots \oplus P_s^{n_s}$, де n_i – кратність модуля P_i ($i=1, \dots, s$).

Нехай $1=f_1+\dots+f_s$ – відповідний розклад $1 \in A$ в суму попарно ортогональних ідемпотентів, тобто, $f_i A = P_i^{n_i}$. Тоді $A f_i = Q_i^{n_i}$ ($i=1, \dots, s$) і ${}_A A = Q_1^{n_1} \oplus \dots \oplus Q_s^{n_s}$ – розклад його регулярного A -модуля ${}_A A$ в пряму суму попарно неізоморфних нерозкладних проєктивних лівих A -модулів (головних лівих A -модулів).

Ідемпотент кільця A називається тривіальним, якщо він дорівнює 0 або 1, в іншому випадку ідемпотент називається нетривіальним.

Нагадаємо, що цоклем модуля M називається сума всіх його простих підмодулів. Ця сума позначається $\text{soc} M$. Через $\text{top} M$ позначається фактормодуль M/MR . Якщо $P_1, \dots, P_s$ всі нерозкладні попарно неізоморфні головні A -модулі, то $U_1 = \text{top} P_1, \dots, U_s = \text{top} P_s$ всі попарно неізоморфні прості A -модулі, тобто кожен головний A -модуль P містить один максимальний підмодуль PR .

Всі результати про головні та прості модулі мають місце для будь-яких напівдосконалих

кілець. Ми будемо використовувати ті самі позначення, що і для артинових кілець.

Означення. Напівдосконале кільце A дозволяє підстановку Накаями $\nu(A): i \rightarrow \nu(i)$ множини $\{1, \dots, s\}$, якщо виконуються наступні умови:

(пр 1) $\text{soc} P_k = \text{top} P_{\nu(k)}$;

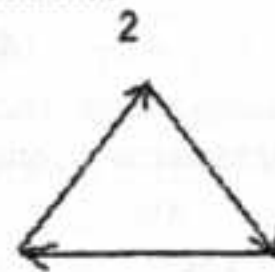
(пр 2) $\text{soc} Q_{\nu(k)} = \text{top} Q_k$.

Артинове кільце A називається квазіфробеніусовим кільцем (QF - кільцем), якщо воно дозволяє підстановку Накаями.

Квазіфробеніусове кільце A називається фробеніусовим кільцем, якщо $n_{\nu(i)} = n_i$ для всіх $i=1, \dots, s$.

Будемо казати, що артинове кільце A задовільняє умову F , якщо для довільного нетривіального ідемпотента $e \in A$ кільце eAe квазіфробеніусове.

2 Приклад



Нехай 1 2 3 - сагайдак, який є простим циклом C_3 з трьома вершинами і стрілками $\sigma_{12}, \sigma_{23}, \sigma_{31}$. Позначимо через kC_3 алгебру шляхів сагайдака C_3 над полем k [2, гл. III].

Позначимо через I - ідеал, який породжується усіма стрілками та розглянемо ідеал I^2 .

Нехай $B = kC_3 / I^2$. Зрозуміло, що $B = P_1 \oplus P_2 \oplus P_3$, причому $l(P_1) = l(P_2) = l(P_3) = 2$, де $l(M)$ - довжина композиційного ряду модуля M . Алгебра B є напівланцюговою (узагальненою однорядною). Згідно з теоремою X.4.1 [2] алгебра B фробеніусова.

Нехай $T_n(k)$ - алгебра всіх верхніх трикутних матриць над полем k . Сагайдак цієї узагальненої однорядної алгебри є ланцюг, який має n вершин. За теоремою X.4.1 [2] алгебра $T_n(k)$ не є квазіфробеніусовою.

Квазіфробеніусова алгебра B має три попарно ортогональних локальних ідемпотента e_1, e_2, e_3 і одиниця $1 \in A$ має такий розклад:

$$1 = e_1 + e_2 + e_3.$$

Розглянемо суму двох ідемпотентів $e_1 + e_2$. Тоді k - алгебра $(e_1 + e_2)B(e_1 + e_2)$ має базис e_1, e_2 та σ_{12} . Ця k - алгебра ізоморфна алгебрі $T_2(k)$.

Таким чином, ми побудували фробеніусову k -алгебру B , для якої існує нетривіальний ідемпотент $e \in B$ такий, що алгебра eBe не є квазіфробеніусовою.

3 Головна теорема

В цьому розділі ми будемо розглядати артінові кільця, сагайдаки яких мають не менше чотирьох вершин.

Нехай артінове кільце A задовольняє умові F і $A_A = P_1^{n_1} \oplus \dots \oplus P_r^{n_r}$ - розклад правого регулярного A - модуля A_A , такий як в пункті 1. Якщо $n_1 = n_2 = \dots = n_r = 1$, то кільце A зведене. Припустимо, що кільце A не є зведеним. Можна вважати, що $n_1 > 1$ та $f_1 = e_1 + \dots + e_{n_1}$ - розклад ідемпотента f_1 в суму попарно ортогональних ідемпотентів.

Тоді кільце $(e_1 + f_2 + \dots + f_r)A(e_1 + f_2 + \dots + f_r)$, за умовою F , є квазіфробеніусовим.

Це кільце еквівалентно в сенсі Моріти кільцю $\text{End}_A(P_1 \oplus \dots \oplus P_r)$, яке є зведеним і фробеніусовим. Тому і кільце A є квазіфробеніусовим.

Сформулюємо головну теорему.

Теорема. Нехай артінове кільце A задовольняє умову F . Тоді якщо кількість

вершин в сагайдаку $Q(A)$ кільця A не менше чотирьох, то кільце A квазіфробеніусове.

Доведення. Ми вже довели, що не зведене артінове кільце A , яке задовольняє умову F є квазіфробеніусовим. Тому можна вважати, що кільце A є зведеним і $A_A = P_1 \oplus \dots \oplus P_r$ є розкладом регулярного A - модуля в пряму суму попарно неізоморфних головних A - модулів.

Нехай $P = eA$ - головний модуль. Припустимо, що $\text{soc} P \supset U \oplus U$, де U - простий A -модуль. За лемою 11.1.2 [1] існує локальний ідемпотент f такий, що $U = Uf$.

Тому кільце $(e + f)A(e + f)$ не є квазіфробеніусовим за теоремою 4.5.2 [3]. Це суперечить тому, що кільце A задовольняє умові F .

Тому кождою будь-якого головного модуля P є простим A - модулем.

Припустимо, що неізоморфні головні модулі $P_i = e_i A$ та $P_j = e_j A$ містять простий A -модуль $U = Uf$.

В цьому випадку кільце $(e_i + e_j + f)A(e_i + e_j + f)$ не є квазіфробеніусовим за теоремою 4.5.2 [3]. Знову отримали протиріччя. Тому, якщо P_i та P_j не ізоморфні головні A - модулі, то $\text{soc} P_i \neq \text{soc} P_j$. Такі самі міркування мають місце для лівих A -модулів. Звідси, за теоремою 4.5.2 [3], випливає, що кільце A квазіфробеніусове. Теорему доведено.

Список використаних джерел

- 1) Hazewinkel M., Gubareni N. and Kirichenko V.V., Algebras, Rings and Modules, Vol. 1, Kluwer Academic Publishers, Mathematics and Its Applications, v. 575, 380 p.
- 2) Дрозд Ю.А., Кириченко В.В., Конечномерные алгебры, К., «Вища школа», 1980, 192 с.
- 3) Hazewinkel M., Gubareni N. and Kirichenko V.V., Algebras, Rings and Modules, Vol. 2, Springer, Mathematics and Its Applications, v. 586, 400 p.

Надійшла до редколегії 15.11.2007

УДК 512.64

Сергій М. Дяченко, аспірант

Ручні сагайдаки відносно поточково-проективних зображень

Ми вводимо поняття поточно-проективних зображень для сагайдака маркованого скінченновимірними зведеними алгебрами і даємо опис ручних сагайдаків відносно таких зображень.

Ключові слова: зображення, матрична задача, сагайдак.

sn_dyachenko@univ.kiev.ua

Статтю представив д.ф.-м.н., професор В.В.Кириченко

В теорії зображень орієнтований граф $Q = (Q_0, Q_1)$ з множиною вершин Q_0 та множиною стрілок Q_1 називають сагайдаком. Початок і кінець стрілки $x \in Q_1$ будемо позначати $\alpha(x)$ і $\beta(x)$ відповідно.

Нехай k — довільне фіксоване алгебраїчно замкнене поле. Ми розглядаємо лише (праві) скінченновимірні над k векторні простори; це стосується, зокрема, алгебр і модулів. Більш того, розглядаються лише базові алгебри, тобто такі, для яких фактор-алгебра по радикалу є прямою сумою полів (це по суті не є обмеженням, бо кожна алгебра Моріта-еквівалентна базовій алгебрі); добре відомо, що кожна така алгебра є алгеброю шляхів деякого орієнтованого графа із співвідношеннями довжини $d > 1$ (див., напр. [2]). Ми будемо надалі ототожнювати базову алгебру із відповідною алгеброю шляхів графа і замість вершина (шлях) графа, яким задана алгебра, будемо говорити вершина (шлях) алгебри. Клас всіх базових алгебр над k позначимо $\mathcal{A}$; через $\mathcal{A}_0$ будемо позначати клас всіх нерозкладних алгебр із $\mathcal{A}$. $\mathcal{A}$ -маркованим (відповідно $\mathcal{A}_0$ -маркованим) сагайдаком називаємо сагайдак Q разом з відображенням $\Delta : Q_0 \rightarrow \mathcal{A}$ (відповідно $\Delta : Q_0 \rightarrow \mathcal{A}_0$). Часто будемо говорити просто "маркований сагайдак" (у випадках, коли ясно іде мова про клас $\mathcal{A}$ чи клас $\mathcal{A}_0$, або коли це не суттєво). При цьому замість вершини з надписаною над нею алгеброю $\Delta(i)$ ми будемо писати просто $\Delta(i)$ (при геометричному зображенні сагайдаків).

Будемо говорити, що задано зображення $\mathcal{A}$ -маркованого (чи $\mathcal{A}_0$ -маркованого) сагайдака $\bar{Q} = (Q, \Delta) = (Q_0, Q_1, \Delta)$, якщо для кожної вершини задано $\Delta(i)$ -модуль P_i , а для кожної стрілки

S. M. Dyachenko, PhD student

Tame quivers relative to pointwise projective representations

We introduce a notion of pointwise projective representations of a quiver marked by finite dimensional reduced algebras and give a description of all quivers which are tame relative to such representations.

Key Words: representation, matrix problem, quiver.

ки $e : i \rightarrow j$ задано k -лінійне відображення $\gamma_e : P_i \rightarrow P_j$. Два зображення $\{P_i, \gamma_e\}$ і $\{P'_i, \gamma'_e\}$ назвемо еквівалентними, якщо існують ізоморфізми $\xi_i \in \text{Hom}_{\Delta(i)}(P_i, P'_i)$ такі, що для кожної стрілки $e : i \rightarrow j$ справедлива рівність $\gamma_e \xi_j = \xi_i \gamma'_e$ (ми перемножаємо відображення зліва направо). Прямою сумою зображень називається зображення $\{P_i \oplus P'_i, \gamma_e \oplus \gamma'_e\}$. Зображення називається розкладним, якщо воно еквівалентне прямій сумі ненульових зображень, і нерозкладним в іншому разі.

У випадку, коли кожний P_i є проективним модулем над алгеброю $\Delta(i)$, ми називаємо зображення $\{P, \gamma\}$ $\mathcal{A}$ -маркованого сагайдака $\bar{Q}$ поточно-проективним зображенням (скорочено *proj*-зображенням).

Маркований сагайдак назвемо *proj*-ручним (відповідно *proj*-диким), якщо задача про опис його *proj*-зображень є ручною (відповідно дикою) задачею; загальні означення ручних і диких задач див. в [3]. Зауважимо, що ручні задачі включають в себе задачі скінченного типу (тобто задачі із скінченим числом нееквівалентних нерозкладних зображень). В цій статті ми опишемо усі *proj*-ручні $\mathcal{A}$ -марковані сагайдаки.

Зауважимо, що *proj*-зображення є узагальненням звичайних зображень сагайдака [10] (треба покласти $\Delta(i) = k$ для всіх $i \in Q_0$). Зв'язні сагайдаки скінченного типу описані в [10] — це діаграми Динкіна A_n, D_n, E_6, E_7, E_8 , а зв'язні сагайдаки ручного нескінченного типу — в [4] і незалежно в [9] — це розширені діаграми Динкіна $\hat{A}_n, \hat{D}_n, \hat{E}_6, \hat{E}_7, \hat{E}_8$.

Очевидно, що достатньо розглянути випадок, коли сагайдак зв'язний. Оскільки проективні модулі прямої суми алгебр вичерпуються пря-

ними сумами проективних модулів (для алгебр, що є прямими доданками), то достатньо вважати, що всі алгебри нерозкладні, іншими словами, достатньо розглядати поточно-проективні зображення A_0 -маркованих сагайдаків. Далі, як легко бачити, задача про опис *proj*-зображень A_0 -маркованого сагайдака $\bar{Q} = (Q, \Delta)$ не може бути ручною, якщо ручною не є задача про опис звичайних зображень сагайдака Q , то будемо (окрім сказаного) вважати, що $\bar{Q} = (Q, \Delta)$ є звичайною або розширеною діаграмою Динкіна.

Сформулюємо задачу про *proj*-зображення маркованого сагайдака (Q, Δ) на матричній мові. Виберемо в кожній алгебрі $\Delta(l)$ базис $B_l = \{v_i^l | i = 1 \dots n_l\}$, який складається із шляхів в графі. Кожній стрілці $e : l \rightarrow s$ відповідає блочна матриця M_e . Вона розбита на горизонтальні смуги відповідні елементам базису B_l і на вертикальні смуги відповідні елементам базису B_s , ми позначаємо смуги елементами базису. Розмірності смуг, які відповідають шляхам, що починаються в тій самій вершині є рівними. Допустимими є наступні перетворення:

1) можна робити будь-яке елементарне перетворення із стовпцями смуги v_i^s (з рядками смуги v_i^l), якщо $\alpha(v_j^s) = \alpha(v_i^s)$ ($\alpha(v_j^l) = \alpha(v_i^l)$), то таке ж саме елементарне перетворення треба зробити зі стовпцями смуги v_j^s (з рядками смуги v_j^l);

2) якщо для трьох елементів базису виконується $v_m^s v_i^s = v_j^s$ ($v_m^l v_i^l = v_j^l$) то довільний стовпець смуги v_i^s (рядок смуги v_j^l), помножений на елемент поля, можна додавати до стовпця смуги v_j^s (до рядка смуги v_i^l); якщо крім цього виконується $v_m^s v_i^s = v_j^s$ ($v_m^l v_i^l = v_j^l$), то необхідно додати стовпець смуги v_i^s , відповідний стовпцю смуги v_i^l , помножений на той же елемент поля, до стовпця смуги v_j^s , відповідного стовпцю смуги v_j^l (додати рядок смуги v_j^l , відповідний рядку смуги v_j^l , помножений на той же елемент поля, до рядка смуги v_i^l , відповідного рядку смуги v_i^l).

Якщо дві матриці відповідають стрілкам сагайдака, що мають однаковий початок (кінець), то перетворення над рядками (стовпцями) цих матриць є одночасними; якщо стрілка відповідна першій матриці закінчується в тій вершині, в якій починається стрілка відповідна другій матриці, то перетворення над рядками другої матриці і стовпцями першої є взаємно оберненими.

Через Λ_i , $i = 1, \dots, 5$ будемо позначати алгебри, які задані наступними орієнтованими графами $\Gamma = (\Gamma_0, \Gamma_1)$ зі співвідношеннями:

$$\Lambda_1: \Gamma_0 = \{1\}, \Gamma_1 = \emptyset, \text{ тобто } \Lambda_1 = k;$$

$$\Lambda_2: \Gamma_0 = \{1\}, \Gamma_1 = \{a : 1 \rightarrow 1\}, a^2 = 0;$$

$$\Lambda_3: \Gamma_0 = \{1\}, \Gamma_1 = \{a : 1 \rightarrow 1\}, a^3 = 0;$$

$$\Lambda_4: \Gamma_0 = \{1, 2\}, \Gamma_1 = \{a : 1 \rightarrow 2\};$$

$$\Lambda_5: \Gamma_0 = \{1, 2\}, \Gamma_1 = \{a : 1 \rightarrow 2, b : 2 \rightarrow 1\}, ab = ba = 0.$$

Розглядаючи діаграми Динкіна (прості чи розширені), ми нумеруємо їх вершини натуральними числами природним чином (зліва направо). Також, ми будемо використовувати скорочення ч. в. множина для поняття частково впорядкованої множини, яке буде зустрічатися в майбутньому.

В цій статті ми доведемо наступну теорему.

Теорема 1. *Proj-ручними зв'язними A_0 -маркованими сагайдаками $\bar{Q} = (Q, \Delta)$ є наступні і лише вони:*

$$1a) Q = A_2, \Delta(2) \in \{\Lambda_1, \dots, \Lambda_5\}, \Delta(1) = \Lambda_1;$$

$$1b) Q = A_2, \Delta(1) \in \{\Lambda_1, \dots, \Lambda_5\}, \Delta(2) = \Lambda_1;$$

$$1c) Q = A_2, \Delta(1) = \Delta(2) = \Lambda_2;$$

$$2a) Q = A_3, \Delta(1) = \Delta(3) = \Lambda_1, \Delta(2) = \Lambda_2;$$

$$2b) Q = A_3, \Delta(1), \Delta(3) \in \{\Lambda_1, \Lambda_2\}, \Delta(2) = \Lambda_1;$$

$$3) Q = A_n, n \geq 4, \Delta(1), \Delta(n) \in \{\Lambda_1, \Lambda_2\}, \Delta(i) = \Lambda_1 \text{ для } i \neq 1, n;$$

$$4) Q = D_n, n \geq 4, \Delta(n) \in \{\Lambda_1, \Lambda_2\}, \Delta(i) = \Lambda_1 \text{ для } i \neq n;$$

$$5) Q = E_6, E_7, E_8, \hat{A}_n, \hat{D}_n, \hat{E}_6, \hat{E}_7, \hat{E}_8, \Delta(i) = \Lambda_1 \text{ для довільного } i.$$

Доведення. Для простоти, в кожному конкретному випадку ми будемо фіксувати напрямки стрілок сагайдака (для інших напрямків доведення аналогічне).

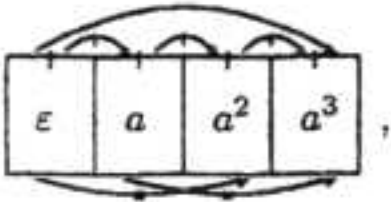
1-ий крок. Доведемо спочатку, що якщо сагайдак $\bar{Q}$ є *proj*-ручним, то відображення Δ набуває значення в множині $\{\Lambda_1, \dots, \Lambda_5\}$. Це твердження достатньо довести лише для сагайдаків $k \rightarrow \Gamma$ та $\Gamma \rightarrow k$. Дійсно, якщо $e : i \rightarrow j$, $e \in Q_1$, то задача про *proj*-зображення маркованого сагайдака (Q, Δ) містить в собі задачу про *proj*-зображення маркованого сагайдака $\Delta(i) \xrightarrow{e} \Delta(j)$, яка, в свою чергу, містить задачу про *proj*-зображення $k \rightarrow \Gamma$ та $\Gamma \rightarrow k$.

Ми розглянемо лише випадок сагайдака вигляду $\bar{Q} = k \rightarrow \Gamma$, для $\bar{Q} = \Gamma \rightarrow k$ доведення аналогічне.

Покажемо, що якщо $\bar{Q}$ є *proj*-ручним, то $\Gamma \in \{\Lambda_1, \dots, \Lambda_5\}$. Як і раніше, множину вершин і множину стрілок алгебри Γ позначаємо Γ_0 і Γ_1 .

Доведемо спочатку, що Γ не містить вершину, з якої виходить дві стрілки. Дійсно, припустимо, що це не так. Тоді 1) $a : 1 \rightarrow 1, b : 1 \rightarrow 1$, або 2) $a : 1 \rightarrow 1, b : 1 \rightarrow 2$, або 3) $a : 1 \rightarrow 2, b : 1 \rightarrow 3$ і очевидно, що задача про $rgoj$ -зображення $\bar{Q}$ буде містити задачу про зображення ч. в. множини $S = \{a, b, c\}$ з відношенням еквівалентності $a \sim b \sim c$, де $b < a, b < c$ у випадку 1), $b < a$ у випадку 2) і a, b, c попарно непорівняльні у випадку 3); згідно [7, 8] ці задачі дикі. До того ж, неможливий випадок 2) при іншому напрямку стрілки b ; доведення аналогічне (у цьому випадку $S = \{a, b, c, d\}$, $b < a, b < c, a \sim b, c \sim d$). Таким чином, Γ або складається з однієї петлі, або не містить петель.

Розглянемо спочатку перший випадок: $\Gamma_0 = \{1\}$, $\Gamma_1 = \{a : 1 \rightarrow 1\}$. Співвідношення для такої алгебри має вигляд $a^n = 0$. Випадкам $a^2 = 0$ і $a^3 = 0$ відповідають алгебри $\Gamma = \Lambda_2$ і $\Gamma = \Lambda_3$. Покажемо, що випадок $a^n = 0$ при $n = 4$ (а значить і при $n > 4$) неможливий. Згідно з приведеним вище описом матричної задачі в загальному випадку, маємо таку задачу:



де позначені одним символом додавання треба робити одночасно, а ϵ, a, a^2, a^3 позначають номери вертикальних смуг. Розглянемо четвірки матриць $T(B, C, D)$ такого вигляду:

	E																		
				E		E													
			E																
				E															
								E	E										
														0	B	C	D		

де на порожніх місцях стоять нульові матриці, довгі вертикальні лінії позначають границі матриць. Легко бачити, що $T(B, C, D)$ і $T(B', C', D')$ еквівалентні тоді і лише тоді, коли трійки матриць B, C, D і B', C', D' еквівалентні як зображення ч. в. множини з інволюцією $S = \{1, 2, 3\}$, де всі точки непорівнянні і $1^* = 2$, а це дика задача [7, 8]. Тут і в подальшому ми не вказуємо значення інволюції на елементах, які інволютивні самі собі, тобто ми пропускаємо $3^* = 3$.

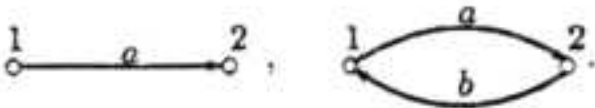
Перейдемо тепер до другого випадку, коли Γ не містить петель. Покажемо, що вершин може

бути не більше 2. Якщо це не так, то граф містить один із підграфів:



причому, можна вважати, що $ab = 0$ (інакше задача буде складнішою). Їм як легко бачити відповідають задачі про зображення наступних ч. в. множин з інволюцією: $S = \{\epsilon_1, \epsilon_2, \epsilon_3, a, b\}$, де $\epsilon_2 < a, \epsilon_2 < b, \epsilon_1^* = a, \epsilon_3^* = b$ та S , де $\epsilon_2 < a, \epsilon_3 < b, \epsilon_2^* = a, \epsilon_3^* = b$. Це дикі задачі [7, 8].

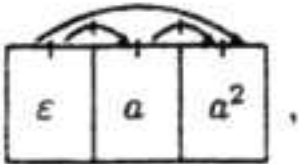
У випадку, коли Γ має одну вершину, $\Gamma = \Lambda_1$. Для двох вершин можливі такі два випадки:



У першому випадку $\Gamma = \Lambda_4$. Якщо у другому випадку немає співвідношення $ab = 0$, то задача про зображення $\bar{Q}$ буде містити задачу про зображення частково впорядкованої множини з відношенням еквівалентності $S = \{\epsilon_1, a, ab\}$, $\epsilon_1 < ab, \epsilon_1 \sim a \sim ab$, а це дика задача [7, 8]. Отже, $ab = 0$, аналогічно $ba = 0$, тоді $\Gamma = \Lambda_5$.

2-ий крок. Перейдемо безпосередньо до доведення теореми. Далі при доведенні теореми ми розглядаємо лише такі марковані сагайдаки (Q, Δ) , що $\Delta(j) \in \{\Lambda_1, \dots, \Lambda_5\}$ для всіх вершин сагайдака.

Маркований сагайдак Λ_2 має вигляд $\Gamma_1 \rightarrow \Gamma_2$. 1a) $\Gamma_1 = k$. Покажемо, що для довільного $\Gamma_2 \in \{\Lambda_1, \dots, \Lambda_5\}$ сагайдак $\bar{Q} = k \rightarrow \Gamma_2 \in rgoj$ -ручним. Якщо $\Gamma_2 = \Lambda_1$, то ми отримаємо задачу скінченного типу [10]. Якщо $\Gamma_2 = \Lambda_3$, то маємо таку задачу:

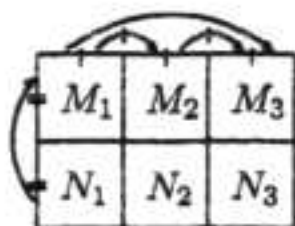


де позначені одним символом додавання треба робити одночасно. Це задача скінченного типу [5]. У випадках $\Gamma_2 = \Lambda_2, \Lambda_4, \Lambda_5$ задача про $rgoj$ -зображення $\bar{Q}$ є задачею про зображення частково впорядкованої множини з інволюцією $S = \{\epsilon < a\}, \epsilon^* = a; S = \{\epsilon_1, \epsilon_2 < a\}, \epsilon_1^* = a; S = \{\epsilon_1 < b, \epsilon_2 < a\}, \epsilon_1^* = a, \epsilon_2^* = b$ відповідно. Всі ці задачі є ручними [7, 8]. Випадок 1b) доводиться аналогічно.

Для подальшого доведення помітимо, що задача про $rgoj$ -зображення $\Gamma \rightarrow \Lambda_2$ є послабленням кожної із задач $\Gamma \rightarrow \Lambda_i, i = 2, \dots, 5$, те саме справедливо для $\Lambda_2 \rightarrow \Gamma$.

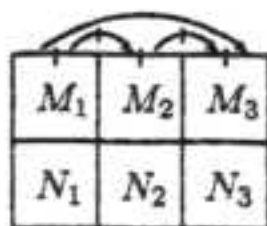
Розглянемо тепер випадок 1c): $\Gamma_1, \Gamma_2 \neq k$.

Легко бачити, що якщо $(\Gamma_1, \Gamma_2) = (\Lambda_i, \Lambda_j)$, де $i = 2, \dots, 5$, $j = 4, 5$, або симетрично $i = 4, 5$, $j = 2, \dots, 5$ то задача про зображення $\bar{Q}$ містить дику задачу про зображення пари ч. в. множин з інволюцією $S_1 = \{a_1, a_2\}$, точки непорівнянні, $a_1^* = a_2$, $S_2 = \{b_1 < b_2\}$, $b_1^* = b_2$ [6]. Якщо $(\Gamma_1, \Gamma_2) = (\Lambda_2, \Lambda_3)$, то ми отримаємо задачу



Якщо покласти $N_1 = E$, $M_1 = N_2 = N_3 = 0$, то для матриць M_2, M_3 отримаємо задачу про приведення пари матриць одночасними перетвореннями подібності, отже, задача дика. Аналогічно задача буде диною, якщо $(\Gamma_1, \Gamma_2) = (\Lambda_3, \Lambda_2)$. Таким чином, залишився єдиний випадок $(\Gamma_1, \Gamma_2) = (\Lambda_2, \Lambda_2)$. В цьому випадку ми отримаємо ручну задачу про зображення пари напівланцюгів $S_1 = \{\varepsilon_1 < a_1\}$, $S_2 = \{\varepsilon_2 < a_2\}$, $\varepsilon_1^* = a_1$, $\varepsilon_2^* = a_2$ [1].

$\bar{A}_3$ має вигляд $\Gamma_1 \leftarrow \Gamma_2 \rightarrow \Gamma_3$. Якщо $\Gamma_1 = \Gamma_3 = k$, а $\Gamma_2 \in \{\Lambda_4, \Lambda_5\}$, тоді задача про *proj*-зображення $\bar{A}_3$ містить дику підзадачу про зображення пари ч. в. множин з інволюцією $S_1 = \{a_1, a_2\}$, $S_2 = \{b_1, b_2\}$, точки непорівнянні, $a_1^* = a_2$ [6], а якщо $\Gamma_2 = \Lambda_3$, то задача має вигляд



Поклавши $N_1 = E$, $N_2 = N_3 = 0$, для матриць M_1, M_2, M_3 отримаємо задачу про зображення ч. в. множини з відношенням еквівалентності $S = \{a, b, c\}$ (точки непорівнянні) $a \sim b \sim c$; згідно [7, 8] ця задача дика, а отже, $\Gamma_2 \in \{\Lambda_1, \Lambda_2\}$.

Якщо $\Gamma_1 = \Gamma_2 = k$, а $\Gamma_3 \in \{\Lambda_4, \Lambda_5\}$, тоді задача про *proj*-зображення $\bar{A}_3$ містить дику підзадачу про зображення ч. в. множини з інволюцією $S_1 = \{\varepsilon, \varepsilon_1, a\}$, точки непорівнянні, $\varepsilon_1^* = a$ [7, 8], а якщо $\Gamma_3 = \Lambda_3$, то ми отримаємо задачу

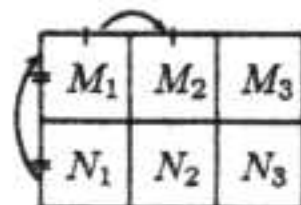


Розглянемо четвірки матриць $T(A, B, C, P, Q)$:

0	A	0	P	0	B	Q	C	0	E
E									
		E							
				E					
							E	E	

де на порожніх місцях стоять нульові матриці (довгі вертикальні лінії позначають границі матриць). Легко бачити, що $T(A, B, C, P, Q)$ і $T(A', B', C', P', Q')$ еквівалентні тоді і лише тоді, коли п'ятірки матриць A, B, C, P, Q і A', B', C', P', Q' еквівалентні як зображення ч. в. множини з відношенням еквівалентності $S = \{a, b, c, p, q\}$, $a < b < c$, $p < b$, $a < q$, $a \sim b \sim c$, $p \sim q$, а це дика задача [7, 8]. Таким чином, $\Gamma_3 \in \{\Lambda_1, \Lambda_2\}$. З міркувань симетричності $\Gamma_1 \in \{\Lambda_1, \Lambda_2\}$.

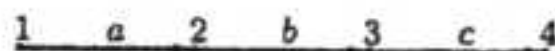
Покажемо, що якщо $\Gamma_2 = \Lambda_2$, то $\Gamma_1 = \Gamma_3 = k$. Якщо це не так, то задача про *proj*-зображення $\bar{A}_3$ міститиме наступну підзадачу:



Поклавши $N_1 = E$, $M_1 = N_2 = N_3 = 0$, для матриць M_2, M_3 отримаємо дику задачу про звичайне зображення сагайдака $Q = (Q_0, Q_1)$, $Q_0 = \{1, 2\}$, $Q_1 = \{a : 1 \rightarrow 1, b : 1 \rightarrow 2\}$ [4, 9]. Якщо $\Gamma_2 = \Lambda_2$, $\Gamma_1 = \Gamma_3 = k$, то ми отримаємо ручну задачу про зображення в'язки напівланцюгів (S_1, S_2) , де $S_1 = \{1, 2\}$, точки непорівнянні, $S_2 = \{3 < 4\}$, $3^* = 4$ [1]. Таким чином, ми довели пункт 2а).

Якщо $\Gamma_2 = k$, $\Gamma_1 = \Gamma_3 = \Lambda_2$, то ми отримаємо ручну задачу про зображення ч. в. множини з інволюцією $S = \{1 < 2, 3 < 4\}$, $1^* = 2$, $3^* = 4$ [7, 8], тому задача буде ручною і для випадків $(\Gamma_1, \Gamma_3) = (k, k), (k, \Lambda_2), (\Lambda_2, k)$. Таким чином, ми довели пункт 2б).

Для продовження доведення теореми помітимо, що якщо маркований сагайдак з трьох послідовних стрілок



proj-ручний, то $\Delta(2) = \Delta(3) = k$. Дійсно, якщо, наприклад, $\Delta(2) \neq k$, то задача про *proj*-зображення такого сагайдака містить дику задачу про зображення ч. в. множини з інволюцією

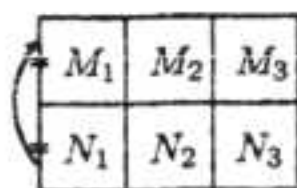
$S = \{1, 2, 3\}$, де всі точки непорівнянні і $1^* = 2$ [7, 8].

Розглянемо $\overline{A_n}$ $n \geq 4$: $1 \rightarrow 2 \rightarrow \dots \rightarrow n$, $\Gamma_j = \Delta(j)$. Згідно щойно доведеного, $\Gamma_j = k$, $j \neq 1, n$, а якщо ми покладемо матриці $(2, 3), \dots, (n-2, n-1)$ одиничними, то ми отримаємо задачу про *proj*-зображення $\Gamma_1 \rightarrow k \rightarrow \Gamma_n$, звідки $\Gamma_1, \Gamma_n \in \{\Lambda_1, \Lambda_2\}$. В цих чотирьох випадках, привівши по-черзі матриці $(1, 2), (2, 3), \dots, (n-2, n-1)$, для матриці $(n-1, n)$ отримаємо ручну задачу про зображення в'язки напівланцюгів [1].

Розглянемо D_n , $n \geq 4$:



Позначимо $\Delta(j) = \Gamma_j$. Нехай спочатку $n = 4$. Застосуємо результат для A_3 до $1 \rightarrow 2 \rightarrow 3$ та $1 \rightarrow 2 \rightarrow 1'$. Тоді $\Gamma_j \in \{k, \Lambda_2\}$, $j = 1, 1', 2, 3$. Якщо $\Gamma_2 = \Lambda_2$, $\Gamma_j = k$, $j = 1, 1', 3$, то задача має вигляд:



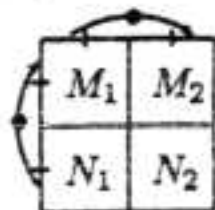
Поклавши $N_1 = N_2 = N_3 = E$, $M_1 = 0$, для матриць M_2, M_3 отримаємо задачу про приведення пари матриць одночасними перетвореннями подібності, отже, задача дика і $\Gamma_2 = k$. Якщо серед алгебр $\Gamma_1, \Gamma_{1'}, \Gamma_2$ більше однієї рівні Λ_2 , то задача про *proj*-зображення $\overline{D_4}$ містить дику задачу про зображення ч. в. множини з інволюцією $S = \{\varepsilon_1 < a_1, \varepsilon_2 < a_2, \varepsilon_3\}$, $\varepsilon_1^* = a_1$, $\varepsilon_2^* = a_2$ [7, 8]. Отже, лише одна з цих алгебр може бути рівною Λ_2 , інші дві рівні k .

Розглянемо тепер випадок D_n , $n > 4$. $\Gamma_j = k$, $j = 2, \dots, n-2$. $\Gamma_1, \Gamma_{1'}, \Gamma_{n-1} \in \{k, \Lambda_2\}$ згідно з результатом для A_n . Якщо ми покладемо одиничними матриці $(2, 3), \dots, (n-3, n-2)$, то отримаємо задачу про *proj*-зображення $\overline{D_4}$, де $\Delta(3) = \Gamma_{n-1}$. Згідно вищесказаного лише одна з алгебр $\Gamma_1, \Gamma_{1'}, \Gamma_{n-1}$ рівна Λ_2 . Не можливо $\Gamma_1 = \Lambda_2$, адже в цьому випадку задача про *proj*-зображення $\overline{D_n}$ містить (при $(4, 5) = \dots = (n-2, n-1) = 0$ і приведенні $(3, 4)$ до канонічного вигляду) дику задачу про зображення ч. в. множини з інволюцією $S = \{a_1 < a_2, b, c < d\}$, $a_1^* = a_2$ [7, 8]. Аналогічно не можливо $\Gamma_{1'} = \Lambda_2$. Таким чином, ми довели, що якщо $\overline{D_n}$ — *proj*-ручний, то його вершини помічені так, як сказано в пункті 4). Якщо $\Gamma_{n-1} = k$, то ми отримаємо задачу скінченного типу [10]. Якщо $\Gamma_{n-1} = \Lambda_2$, то привівши по-черзі матриці $(1, 2), (2, 1'), \dots, (n-3, n-2)$, для матриці $(n-2, n-1)$ отримаємо ручну задачу про зобра-

ження в'язки напівланцюгів [1].

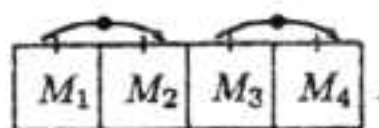
Неважко довести, що якщо сагайдаки $E_6, E_7, E_8, \hat{A}_n, \hat{D}_n, \hat{E}_6, \hat{E}_7, \hat{E}_8$ є *proj*-ручними, то всі їх вершини помічені полем k .

Розглянемо спочатку сагайдак $\hat{A}_0$, який складається з однієї вершини 0 і однієї петлі $e: 0 \rightarrow 0$. Якщо $\Delta(0) \neq k$, то задача про *proj*-зображення $(\hat{A}_0, \Delta)$ містить задачу:



Поклавши $N_1 = E$, $M_1 = 0$, для матриць M_2, N_2 отримаємо задачу про приведення пари матриць одночасними перетвореннями подібності, отже, задача дика і $\Delta(0) = k$.

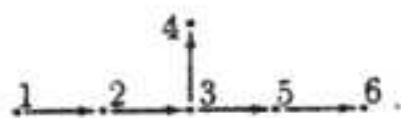
Для сагайдака $Q = \hat{A}_1$, $Q_0 = \{0, 1\}$, $Q_1 = \{a, b: 0 \rightarrow 1\}$ помітимо, що якщо $\Delta(1) \neq k$, то задача про його *proj*-зображення містить задачу



Поклавши $M_1 = E$, $M_2 = 0$, для матриць M_3, M_4 отримаємо задачу про приведення пари матриць одночасними перетвореннями подібності, отже, задача дика і $\Delta(1) = k$, аналогічно $\Delta(0) = k$.

Для $Q = \hat{D}_4$, $Q_0 = \{0, 1, 2, 3, 4\}$, $Q_1 = \{a_i: 0 \rightarrow i | i = 1, 2, 3, 4\}$, згідно доведеного для D_4 , $\Delta(0) = k$, а якщо для однієї з вершин $i = 1, 2, 3, 4$ $\Delta(i) \neq k$, то задача про його *proj*-зображення містить дику задачу про зображення ч. в. множини з інволюцією $S = \{a_1 < a_2, b, c, d\}$, $a_1^* = a_2$ [7, 8].

Для решти сагайдаків той факт, що їх вершини помічені полем доводиться за допомогою нескладних комбінаторних міркувань. Наприклад, для E_6



Якщо розглянемо підсагайдак типу A_5 : $1 \rightarrow \dots \rightarrow 6$, то маємо $\Delta(j) = k$, $j = 2, 3, 5$. Якщо покласти матрицю $(2, 3)$ одиничною, то ми отримаємо зображення сагайдака типу D_5 , який отримується з цього вилученням стрілки $2 \rightarrow 3$ і отождоженням вершин 2 і 3, тому $\Delta(1) = \Delta(4) = k$, аналогічно $\Delta(6) = k$.

Протилежне твердження, тобто, якщо всі вершини перелічених в пункті 5) сагайдаків помічені полем k , то задача про *proj*-зображення такого сагайдака буде ручною, випливає з результатів [10, 4, 9]. $\square$

Сформулюємо тепер теорему про опис A -маркованих proj -ручних сагайдаків, яка по суті еквівалентна вище доведеній теоремі 1. Введемо наступні алгебри з класу A :

$$\Lambda_6 = \Lambda_1 \oplus \Lambda_1,$$

$$\Lambda_7 = \Lambda_1 \oplus \Lambda_1 \oplus \Lambda_1,$$

$$\Lambda_8 = \Lambda_1 \oplus \Lambda_1 \oplus \Lambda_1 \oplus \Lambda_1,$$

$$\Lambda_9 = \Lambda_1 \oplus \Lambda_2,$$

$$\Lambda_{10} = \Lambda_1 \oplus \Lambda_1 \oplus \Lambda_2,$$

$$\Lambda_{11} = \Lambda_2 \oplus \Lambda_2.$$

Теорема 2. *Proj-ручними зв'язними A -маркованими сагайдаками $\bar{Q} = (Q, \Delta)$ є наступні і лише вони:*

$$1a) Q = A_2, \Delta(2) \in \{\Lambda_1, \dots, \Lambda_{11}\}, \Delta(1) = \Lambda_1;$$

$$1b) Q = A_2, \Delta(1) \in \{\Lambda_1, \dots, \Lambda_{11}\}, \Delta(2) = \Lambda_1;$$

$$1c) Q = A_2, \Delta(1), \Delta(2) \in \{\Lambda_2, \Lambda_6\};$$

$$2a) Q = A_3, \Delta(1) = \Delta(3) = \Lambda_1, \Delta(2) \in \{\Lambda_2, \Lambda_6\};$$

$$2b) Q = A_3, \Delta(1), \Delta(3) \in \{\Lambda_1, \Lambda_2, \Lambda_6\}, \Delta(2) = \Lambda_1;$$

$$3) Q = A_n, n \geq 4, \Delta(1), \Delta(n) \in \{\Lambda_1, \Lambda_2, \Lambda_6\}, \Delta(i) = \Lambda_1 \text{ для } i \neq 1, n;$$

$$4) Q = D_n, n \geq 4, \Delta(n) \in \{\Lambda_1, \Lambda_2, \Lambda_6\}, \Delta(i) = \Lambda_1 \text{ для } i \neq n;$$

$$5) Q = E_6, E_7, E_8, \hat{A}_n, \hat{D}_n, \hat{E}_6, \hat{E}_7, \hat{E}_8, \Delta(i) = \Lambda_1 \text{ для довільного } i.$$

Цю теорему неважко довести, використовуючи результат щойно доведеної теореми 1, тому що зображення сагайдака вершини якого помічені розкладними алгебрами еквівалентне зображенню іншого маркованого сагайдака, (цей сагайдак отримується розмноженням вершин, помічених розкладними алгебрами і дублюванням стрілок) вершини якого помічені вже нерозкладними алгебрами. Наприклад, задача про опис proj -зображень сагайдака $\Lambda_1 \rightarrow \Lambda_2 \oplus \Lambda_3$ рівносильна задачі про опис proj -зображень сагайдака $\Lambda_2 \leftarrow \Lambda_1 \rightarrow \Lambda_3$.

Висновки

Отже, ми розглянули поняття поточно-проективних зображень сагайдаків маркованих

скінченновимірними базовими алгебрами. Це поняття є узагальненням звичайного зображення сагайдака введеного Габріелем [10]. Ми дали опис усіх сагайдаків, для яких задача про опис поточно-проективних зображень є ручною задачею. Спочатку ми довели теорему про сагайдаки марковані нерозкладними алгебрами, а потім сформулювали еквівалентну цій теорему, яка дає опис сагайдаків маркованих алгебрами без умови нерозкладності.

Список використаних джерел

1. Бондаренко В. М. Представления связок полупростых множеств и их приложения // Алгебра и анализ — 1991. — Том 3 (вып. 5). — С. 38 — 67.
2. Дрозд Ю. А., Кириченко В. В. Конечномерные алгебры // К. "Вища школа", 1980. — 200 с.
3. Дрозд Ю. А. О ручных и диких матричных задачах // Матричные задачи, — К. Институт математики АН УССР. — 1977. — С. 104-114.
4. Назарова Л. А. Представления колчанов бесконечного типа. // Изв. АН СССР. — 1973. — 37, №4. — С. 752-791.
5. Назарова Л. А., Ройтер А. В. Категорные матричные задачи и проблема Брауэра-Трелла // препринт 73.9. Киев: Наукова думка, 1973. — 100 с.
6. Bondarenko V. M. On certain aspects of the representation theory of bundles of semichains // Алгебраїчні структури та їх застосування. Праці III-ї міжнародної алгебраїчної конференції в Україні — Київ, 2002. — С. 190 - 200.
7. Bondarenko V. M., Zavadskij A. G. Posets with an equivalence relation of tame type and of finite growth // Canad. Math. Soc. Conf. Proc. — 1991. — 11. — P. 67 — 88.
8. Bondarenko V. M., Zavadskij A. G. Tame posets with equivalence relation // Contem. Math. — 1992. — 131 (part 2). — P. 237 — 251.
9. Donovan P., Freislich M. R. The representation theory of finite graphs and associated algebras // Carleton Lecture Notes. — 1973. — №5. — P. 3-86.
10. Gabriel P. Unzerlegbare Darstellungen, I // Manus. Math. — 1972. — 6, №1. — P. 71-103.

Надійшла до редколегії 06.11.2007

УДК 512.53

Анатолій В. Жучок

Оболонка зсувів Кліффордової напівгрупи

Описано оболонку зсувів Кліффордової напівгрупи.

Ключові слова: Кліффордова напівгрупа, лівий зсув, зв'язані зсуви, оболонка зсувів.

E-mail: zhuchok_a@mail.ru

Статтю представив доктор фізико-математичних наук, професор Кириченко В.В.

Вступ. Однією із актуальних задач теорії напівгруп є задача описання структурних властивостей напівгруп в термінах їх оболонок зсувів. Одним із перших результатів цього напрямку залишається результат Петрича [1] про будову оболонки зсувів регулярної напівгрупи Ріса матричного типу над групою з нулем. Пізніше в роботі [2] було отримано модифікацію цього результату за допомогою конструкції вінцевого добутку та вивчено алгебраїчні властивості напівгруп односторонніх зсувів напівгруп Ріса матричного типу. Далі в [3] автори узагальнили результат Петрича, описавши оболонки зсувів квазірегулярних напівгруп Ріса матричного типу над моноїдом з нулем, використовуючи при цьому визначену ними конструкцію голоморфного добутку. Опираючись на згадані результати, природньо, отже, постає проблема побудови та вивчення оболонок зсувів теоретико-напівгрупових конструкцій таких, як наприклад, сполуки напівгруп.

Одним із прикладів сполуки напівгруп є конструкція Кліффордової напівгрупи (=сильної напіврешітки груп). Ендоморфізми Кліффордових напівгруп вивчалися в [4]. В [5] описуються зображення сильних напіврешіток напівгруп.

В цій роботі охарактеризовано оболонку зсувів Кліффордової напівгрупи.

Основні визначення та позначення. Нехай S – довільна напівгрупа. Перетворення $\lambda(\rho)$ множини S називається лівим (правим) зсувом напівгрупи S , якщо $(uv)\lambda = (u\lambda)v$ ($((uv)\rho = u(\nu\rho))$) для всіх $u, v \in S$. Множина всіх лівих зсувів напівгрупи S є напівгрупою відносно операції композиції перетворень. Цю напівгрупу називають напівгрупою лівих зсувів напівгрупи S . Аналогічно визначається напівгрупа правих зсувів напівгрупи S . Напівгрупу лівих

Anatolii V. Zhuchok

The translational hull of Clifford semigroup

The translational hull of Clifford semigroup is described.

Key Words: Clifford semigroup, left translation, linked translations, translational hull.

(правих) зсувів напівгрупи S будемо позначати через $T_l(S)$ ($T_r(S)$).

Кожному елементу a напівгрупи S зіставимо перетворення λ_a (ρ_a) напівгрупи S , поклавши $x\lambda_a = ax$ ($x\rho_a = xa$) для всіх $x \in S$. Перетворення λ_a (ρ_a) називається внутрішнім лівим (правим) зсувом напівгрупи S , який відповідає елементу $a \in S$. Множина всіх внутрішніх лівих (правих) зсувів напівгрупи S є піднапівгрупою напівгрупи лівих (правих) зсувів напівгрупи S (див. [6]).

Лівий зсув λ та правий зсув ρ напівгрупи S називаються зв'язаними, якщо $u(v\lambda) = (u\rho)v$ для всіх $u, v \in S$. На множині всіх пар (λ, ρ) , де λ і ρ – зв'язані лівий та правий зсуви напівгрупи S , визначимо операцію, поклавши

$$(\lambda_1, \rho_1)(\lambda_2, \rho_2) = (\lambda_2\lambda_1, \rho_1\rho_2).$$

Дана операція є асоціативною. Напівгрупу, отриману в такий спосіб, називають оболонкою зсувів напівгрупи S . Оболонку зсувів напівгрупи S будемо позначати через $TH(S)$.

Нехай $\overline{TH}(S) = \{(\lambda_a, \rho_a) \in TH(S) | a \in S\}$.

Легко перевірити, що множина $\overline{TH}(S)$ є піднапівгрупою напівгрупи $TH(S)$ (див. [6]). Напівгрупу $\overline{TH}(S)$ називають внутрішньою частиною оболонки зсувів напівгрупи S .

Нагадаємо визначення Кліффордової напівгрупи.

Нехай Λ – комутативна напівгрупа ідемпотентів, G_i ($i \in \Lambda$) – сімейство груп, які попарно не перетинаються. Нехай далі для будь-яких елементів $i, j \in \Lambda$ таких, що $j \leq i$

$\phi_{i,j} : G_i \rightarrow G_j$, які задовольняють умовам:

1. для кожного $i \in \Lambda$ перетворення $\phi_{i,i}$ є тотожнім;
2. для будь-яких $j, i, k \in \Lambda$ таких, що $j \leq i \leq k$, має місце умова $\phi_{k,i} \phi_{i,j} = \phi_{k,j}$.

В термінах операцій груп $G_i, i \in \Lambda$ на множині $S = \bigcup_{i \in \Lambda} G_i$ визначимо операцію за правилом:

$$xy = (x\phi_{i,j})(y\phi_{j,k})$$

для всіх $x \in G_i, y \in G_j, i, j \in \Lambda$.

Відносно цієї операції S є напівгрупою. Отриману напівгрупу будемо позначати через $S = S[\Lambda; G_i, \phi_{i,j}]$. Напівгрупу $S = S[\Lambda; G_i, \phi_{i,j}]$ називають Кліффордовою напівгрупою або сильною напіврешіткою груп.

Зауважимо, що через e_i ми будемо позначати одиницю групи $G_i, i \in \Lambda$.

Основні відомості про Кліффордові напівгрупи можна знайти, наприклад, в [7].

1. Зсуви Кліффордової напівгрупи. В цьому пункті описано ліві (праві) зсуви Кліффордової напівгрупи.

1.1. Нехай $S = S[\Lambda; G_i, \phi_{i,j}]$ – Кліффордова напівгрупа.

Лема. Перетворення λ Кліффордової напівгрупи $S = S[\Lambda; G_i, \phi_{i,j}]$ є лівим зсувом тоді й лише тоді, коли існують лівий зсув λ_0 напівгрупи Λ та сімейство відображень $\lambda_i : G_i \rightarrow G_{i\lambda_0}, i \in \Lambda$ такі, що $a\lambda = a\lambda_i$, де $a \in G_i, i \in \Lambda$, та виконуються умови:

$$\phi_{i,k}\lambda_k = \lambda_i\phi_{i\lambda_0, k\lambda_0} \quad (1)$$

для всіх $i, k \in \Lambda, k \leq i$,

$$(ab)\lambda_i = a\lambda_i b$$

для всіх $a, b \in G_i, i \in \Lambda$.

Доведення. Нехай λ – лівий зсув напівгрупи S . Для $e_i \in G_i, i \in \Lambda$ покладемо $e_i\lambda \in G_j$ для деякого $j \in \Lambda$. Маємо:

$$\begin{aligned} e_i\lambda &= (e_i e_i)\lambda = (e_i\lambda) e_i = \\ &= (e_i\lambda)\phi_{j,i} (e_i\phi_{i,i}) = \\ &= (e_i\lambda)\phi_{j,i} e_i = (e_i\lambda)\phi_{j,i}, \end{aligned}$$

звідки $e_i\lambda \in G_j$. Отже, $j \leq i$.

Крім цього, для довільного

$a_i \in G_i, i \in \Lambda$ отримуємо:

$$\begin{aligned} a_i\lambda &= (e_i a_i)\lambda = (e_i\lambda) a_i = \\ &= (e_i\lambda)\phi_{j,i} (a_i\phi_{i,i}) = \\ &= (e_i\lambda)\phi_{j,i} (a_i\phi_{i,j}) = (e_i\lambda)(a_i\phi_{i,j}), \end{aligned}$$

звідки $a_i\lambda \in G_j$. Отже, для кожного $i \in \Lambda$ існує $j \in \Lambda$ такий, що $G_i\lambda \subseteq G_j$. Таким чином, лівий зсув λ напівгрупи S визначається сімейством $\{\lambda_i | i \in \Lambda\}$ відображень $\lambda_i : G_i \rightarrow G_{i\lambda_0}$, де λ_0 – деяке перетворення напівгрупи Λ :

$$a\lambda = a\lambda_i, \text{ де } a \in G_i, i \in \Lambda.$$

Оскільки λ – лівий зсув, то $(a_i b_i)\lambda = (a_i\lambda) b_i$ для всіх $a_i \in G_i, b_i \in G_i, i \in \Lambda$, і отже, $(i\lambda_0)l = (il)\lambda_0, i, l \in \Lambda$, звідки λ_0 – лівий зсув напівгрупи Λ .

Нехай $i \geq l, i, l \in \Lambda$. Для елементів $a_i \in G_i, e_l \in G_l$ при цьому отримуємо:

$$\begin{aligned} (a_i e_l)\lambda &= (a_i\phi_{i,l})\lambda = (a_i\phi_{i,l})\lambda_l, \\ (a_i\lambda) e_l &= (a_i\lambda_l) e_l = \\ &= (a_i\lambda_l)\phi_{i\lambda_0, (il)\lambda_0} e_l\phi_{l, (il)\lambda_0} = \\ &= (a_i\lambda_l)\phi_{i\lambda_0, (il)\lambda_0} = (a_i\lambda_l)\phi_{i\lambda_0, l\lambda_0}, \end{aligned}$$

звідки $\phi_{i,l}\lambda_l = \lambda_i\phi_{i\lambda_0, l\lambda_0}$.

Очевидно, що $(ab)\lambda_i = a\lambda_i b$ для всіх $a, b \in G_i, i \in \Lambda$.

Навпаки, нехай тепер λ_0 – лівий зсув напівгрупи $\Lambda, \{\lambda_i | i \in \Lambda\}$ – сімейство відображень $\lambda_i : G_i \rightarrow G_{i\lambda_0}$, що задовольняють умовам $\phi_{i,l}\lambda_l = \lambda_i\phi_{i\lambda_0, l\lambda_0}$ для всіх $i \geq l, i, l \in \Lambda$, $(ab)\lambda_i = a\lambda_i b$ для всіх $a, b \in G_i, i \in \Lambda$, τ – перетворення напівгрупи S , яке визначається рівністю:

$$a\tau = a\lambda_i, \text{ де } a_i \in G_i, i \in \Lambda.$$

Для будь-яких $a_i \in G_i, b_l \in G_l, i, l \in \Lambda$ при цьому маємо:

$$\begin{aligned} (a_i b_l)\tau &= ((a_i\phi_{i,l})(b_l\phi_{l,i}))\tau = \\ &= ((a_i\phi_{i,l})(b_l\phi_{l,i}))\lambda_u = (a_i\phi_{i,l})\lambda_u (b_l\phi_{l,i}) = \\ &= ((a_i\phi_{i,l})\lambda_u)\phi_{(i)\lambda_0, (il)\lambda_0} (b_l\phi_{l,i})\phi_{(l)\lambda_0, (il)\lambda_0} = \\ &= ((a_i\phi_{i,l})\lambda_u)\phi_{(i)\lambda_0, (il)\lambda_0} (b_l\phi_{l, (il)\lambda_0}) = \\ &= (a_i\phi_{i,l})\lambda_u (b_l\phi_{l, (il)\lambda_0}), \end{aligned}$$

$$\begin{aligned}(a_i \tau) b_l &= (a_i \lambda_l) b_l = \\ &= ((a_i \lambda_l) \phi_{i\lambda_0, (i\lambda_0)l}) (b_l \phi_{l, (i\lambda_0)l}) = \\ &= ((a_i \lambda_l) \phi_{i\lambda_0, (il)\lambda_0}) (b_l \phi_{l, (il)\lambda_0}).\end{aligned}$$

Таким чином, $(a_i b_l) \tau = (a_i \tau) b_l$ для всіх $a_i \in G_i, b_l \in G_l, i, l \in \Lambda$, звідки τ – лівий зсув.

Лему доведено.

1.2. Позначимо для зручності лівий зсув λ Кліффордової напівгрупи S через $[\lambda_0, \lambda_i]_{i \in \Lambda}$.

З леми п.1.1 отримуємо

Наслідок. Нехай $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda}$ – лівий зсув Кліффордової напівгрупи $S[\Lambda; G_i, \phi_{i,j}]$. Мають місце рівності:

- (i) $a\lambda = (e_i \lambda_i)(a \phi_{i, i\lambda_0})$ для всіх $a \in G_i, i \in \Lambda$;
- (ii) $e_{ij} \lambda_{ij} = e_i \lambda_i \phi_{i\lambda_0, (ij)\lambda_0}$ для всіх $i, j \in \Lambda$.

1.3. З леми п.1.1 отримуємо

Наслідок. Нехай $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda}$ – лівий зсув Кліффордової напівгрупи $S[\Lambda; G_i, \phi_{i,j}]$. Якщо $i, j \in \Lambda$ і $i \geq j$, то $(\text{Im } \phi_{i,j}) \lambda_j \subseteq \text{Im } \phi_{i\lambda_0, j\lambda_0}$.

Доведення безпосередньо випливає з умови $\phi_{i,k} \lambda_k = \lambda_i \phi_{i\lambda_0, k\lambda_0}$ для всіх $i, k \in \Lambda, k \leq i$.

1.4. З леми п.1.1 отримуємо

Наслідок. Нехай $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda}$ – лівий зсув Кліффордової напівгрупи $S[\Lambda; G_i, \phi_{i,j}]$. Мають місце твердження:

- (i) якщо $\phi_{i,j}$ є епіморфізмом, то λ_j визначається через λ_i ;
- (ii) якщо $\phi_{i\lambda_0, j\lambda_0}$ є біекцією, то λ_i визначається через λ_j .

1.5. У двоїстий спосіб (див. п.1.1) доводиться наступна

Лема. Перетворення ρ Кліффордової напівгрупи $S = S[\Lambda; G_i, \phi_{i,j}]$ є правим зсувом тоді й лише тоді, коли існують правий зсув ρ_0 напівгрупи Λ та сімейство відображень $\rho_i: G_i \rightarrow G_{i\rho_0}, i \in \Lambda$ такі, що $a\rho = a\rho_i$, де $a \in G_i, i \in \Lambda$, та виконуються умови:

$$\phi_{i,k} \rho_k = \rho_i \phi_{i\rho_0, k\rho_0} \quad (2)$$

для всіх $i, k \in \Lambda, k \leq i$,

$$(ab)\rho_i = a(b\rho_i)$$

для всіх $a, b \in G_i, i \in \Lambda$.

1.6. Позначимо для зручності правий зсув

ρ напівгрупи S через $[\rho_0, \rho_i]_{i \in \Lambda}$.

З леми п.1.5 отримуємо

Наслідок. Нехай $[\rho_0, \rho_i]_{i \in \Lambda}$ – правий зсув Кліффордової напівгрупи $S = S[\Lambda; G_i, \phi_{i,j}]$. Мають місце рівності:

- (i) $a_i \rho = a_i \phi_{i, i\rho_0} (e_i \rho_i)$ для всіх $a \in G_i, i \in \Lambda$;
- (ii) $e_{ij} \rho_{ij} = e_j \rho_j \phi_{j\rho_0, (ij)\rho_0}$ для всіх $i, j \in \Lambda$.

1.7. З леми п.1.5 отримуємо

Наслідок. Нехай $[\rho_0, \rho_i]_{i \in \Lambda}$ – правий зсув Кліффордової напівгрупи $S = S[\Lambda; G_i, \phi_{i,j}]$. Якщо $i, j \in \Lambda$ і $i \geq j$, то $(\text{Im } \phi_{i,j}) \rho_j \subseteq \text{Im } \phi_{i\rho_0, j\rho_0}$.

1.8. З леми п.1.5 отримуємо

Наслідок. Нехай $[\rho_0, \rho_i]_{i \in \Lambda}$ – правий зсув Кліффордової напівгрупи $S = S[\Lambda; G_i, \phi_{i,j}]$. Мають місце твердження:

- (i) якщо $\phi_{i,j}$ є епіморфізмом, то ρ_j визначається через ρ_i ;
- (ii) якщо $\phi_{i\rho_0, j\rho_0}$ є біекцією, то ρ_i визначається через ρ_j .

2. Зв'язані зсуви Кліффордової напівгрупи. В цьому пункті сформульовано необхідні та достатні умови зв'язаності зсувів Кліффордової напівгрупи та описано її ліві (праві) внутрішні зсуви.

2.1. Наступна теорема описує необхідні та достатні умови зв'язаності лівого і правого зсувів Кліффордової напівгрупи.

Теорема. Лівий зсув $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda}$ і правий зсув $\rho = [\rho_0, \rho_i]_{i \in \Lambda}$ Кліффордової напівгрупи $S = S[\Lambda; G_i, \phi_{i,j}]$ є зв'язаними тоді і лише тоді, коли $(e_i \rho_i) \phi_{i\rho_0, (i\rho_0)j} = (e_j \lambda_j) \phi_{j\lambda_0, i(j\lambda_0)}$ для всіх $i, j \in \Lambda$.

Доведення. Нехай λ та ρ – зв'язані лівий і правий зсуви напівгрупи S . Тоді для будь-яких $a_i \in G_i, b_j \in G_j, i, j \in \Lambda$, використовуючи умову (i) наслідку п.1.2, маємо:

$$\begin{aligned}(a_i \rho) b_j &= ((a_i \phi_{i, i\rho_0}) (e_i \rho_i)) b_j = \\ &= ((a_i \phi_{i, i\rho_0}) (e_i \rho_i)) \phi_{i\rho_0, (i\rho_0)j} (b_j \phi_{j, (i\rho_0)j}) = \\ &= (a_i \phi_{i, (i\rho_0)j}) (e_i \rho_i) \phi_{i\rho_0, (i\rho_0)j} (b_j \phi_{j, (i\rho_0)j}), \\ a_i (b_j \lambda) &= a_i ((e_j \lambda_j) (b_j \phi_{j, j\lambda_0})) = \\ &= (a_i \phi_{i, i(j\lambda_0)}) ((e_j \lambda_j) (b_j \phi_{j, j\lambda_0})) \phi_{j\lambda_0, i(j\lambda_0)} = \\ &= (a_i \phi_{i, i(j\lambda_0)}) (e_j \lambda_j) \phi_{j\lambda_0, i(j\lambda_0)} (b_j \phi_{j, i(j\lambda_0)}).\end{aligned}$$

$$\begin{aligned} \text{Оскільки } (a_i \rho) b_j &= a_i (b_j \lambda), \text{ то} \\ (i \rho_0) j &= i (j \lambda_0) \end{aligned} \quad (3)$$

для всіх $i, j \in \Lambda$. Таким чином, з рівності

$$\begin{aligned} (a_i \phi_{i, (i \rho_0) j}) (e_i \rho_i) \phi_{i \rho_0, (i \rho_0) j} (b_j \phi_{j, (i \rho_0) j}) &= \\ = (a_i \phi_{i, (i \rho_0) j}) (e_j \lambda_j) \phi_{j \lambda_0, i (j \lambda_0)} (b_j \phi_{j, (i \rho_0) j}) \end{aligned}$$

випливає, що

$$(e_i \rho_i) \phi_{i \rho_0, (i \rho_0) j} = (e_j \lambda_j) \phi_{j \lambda_0, i (j \lambda_0)} \quad (4)$$

для всіх $i, j \in \Lambda$.

Нехай λ і ρ – лівий і правий зсуви напівгрупи S , для яких має місце рівність (4). З (4) безпосередньо випливає рівність (3).

Для довільних $a_i \in G_i, b_j \in G_j, i, j \in \Lambda$ маємо:

$$\begin{aligned} (a_i \rho) b_j &= ((a_i \phi_{i, i \rho_0}) (e_i \rho_i)) b_j = \\ &= ((a_i \phi_{i, i \rho_0}) (e_i \rho_i)) \phi_{i \rho_0, (i \rho_0) j} (b_j \phi_{j, (i \rho_0) j}) = \\ &= (a_i \phi_{i, (i \rho_0) j}) (e_i \rho_i) \phi_{i \rho_0, (i \rho_0) j} (b_j \phi_{j, (i \rho_0) j}), \\ a_i (b_j \lambda) &= a_i ((e_j \lambda_j) (b_j \phi_{j, j \lambda_0})) = \\ &= (a_i \phi_{i, i (j \lambda_0)}) ((e_j \lambda_j) (b_j \phi_{j, j \lambda_0})) \phi_{j \lambda_0, i (j \lambda_0)} = \\ &= (a_i \phi_{i, i (j \lambda_0)}) (e_j \lambda_j) \phi_{j \lambda_0, i (j \lambda_0)} (b_j \phi_{j, i (j \lambda_0)}). \end{aligned}$$

Використовуючи (3) і (4), дійдемо висновку, що $(a_i \rho) b_j = a_i (b_j \lambda)$ при будь-яких $a_i \in G_i, b_j \in G_j, i, j \in \Lambda$. Отже, (λ, ρ) – пара зв'язаних лівого λ і правого ρ зсувів напівгрупи S .

Теорему доведено.

2.2. Внутрішній лівий зсув λ_a Кліффордової напівгрупи S , що відповідає елементу $a_i \in G_i, i \in \Lambda$, будемо позначати через $(\lambda)_a$, щоб відрізнити його від відображення λ_i .

Наступна теорема дає необхідні і достатні умови, за якими лівий зсув Кліффордової напівгрупи є внутрішнім.

Теорема. Лівий зсув $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda}$ Кліффордової напівгрупи S є внутрішнім лівим зсувом, що відповідає елементу $a_i \in G_i, i \in \Lambda$, тоді й лише тоді, коли $e_j \lambda_j = a_i \phi_{i, j}$ для всіх $j \in \Lambda$.

Доведення. Якщо $\lambda = (\lambda)_a$ – внутрішній лівий зсув напівгрупи S , який відповідає елементу $a_i \in G_i, i \in \Lambda$, то, використовуючи умову (i) наслідку п.1.2, маємо:

$$e_j \lambda = (e_j \lambda_j) (e_j \phi_{j, j \lambda_0}) =$$

$$\begin{aligned} &= (e_j \lambda_j) e_{j \lambda_0} = e_j \lambda_j = \\ &= a_i \phi_{i, j} = (a_i \phi_{i, j}) e_j = \\ &= (a_i \phi_{i, j}) (e_j \phi_{j, j}) = a_i e_j = e_j (\lambda)_a \end{aligned}$$

для всіх $j \in \Lambda$. Звідси $e_j \lambda_j = a_i \phi_{i, j}$ для всіх $j \in \Lambda$.

Навпаки, нехай $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda}$ – лівий зсув напівгрупи S , для якого $e_j \lambda_j = a_i \phi_{i, j}$ для всіх $j \in \Lambda$. З останньої рівності безпосередньо випливає, що $j \lambda_0 = i j, i, j \in \Lambda$. Для довільного $b_j \in G_j, j \in \Lambda$ маємо:

$$\begin{aligned} b_j \lambda &= (e_j \lambda_j) (b_j \phi_{j, j \lambda_0}) = (a_i \phi_{i, j}) (b_j \phi_{j, j \lambda_0}) = \\ &= (a_i \phi_{i, j}) (b_j \phi_{j, j}) = a_i b_j = b_j (\lambda)_a, \end{aligned}$$

звідки $\lambda = (\lambda)_a$.

Теорему доведено.

2.3. Для внутрішніх правих зсувів у двоїстий спосіб доводиться наступна

Теорема. Правий зсув $\rho = [\rho_0, \rho_i]_{i \in \Lambda}$ Кліффордової напівгрупи S є внутрішнім правим зсувом, що відповідає елементу $a_i \in G_i, i \in \Lambda$, тоді й лише тоді, коли $e_j \rho_j = a_i \phi_{i, j}$ для всіх $j \in \Lambda$.

3. Оболонка зсувів Кліффордової напівгрупи. В цьому пункті охарактеризовано Кліффордову напівгрупу та її оболонку зсувів.

3.1. Нехай I – довільна непорожня множина, T^0 – напівгрупа з нулем. $I \times I$ -матрицю X' над T^0 будемо називають мономіальною по рядках, якщо кожний її рядок містить не більше одного ненульового елемента напівгрупи T^0 . Множина всіх $I \times I$ -матриць, мономіальних по рядках, є напівгрупою (див.[6]). Позначимо її через $A(I; T^0)$.

Нехай $H = \{S_i\}_{i \in X}$ – сімейство довільних напівгруп. Для кожної пари $(i, j) \in X \times X$ через $D^{(i, j)}$ позначимо сукупність всіх відображень $\varphi: S_i \rightarrow S_j$ та покладемо $B = \bigcup_{(i, j) \in X \times X} D^{(i, j)}$.

На множині B природньо визначеною є часткова операція $\bullet$ композиції відображень. Довизначимо на множині $B \cup \{0\}$ (0 – зовнішній ідемпотент) операцію наступним чином:

$$\begin{aligned} \varphi \bullet \psi &= \begin{cases} \varphi \bullet \psi, & j = m, \\ 0, & j \neq m, \end{cases} \\ \varphi \bullet 0 &= 0 \bullet \varphi = 0 \bullet 0 = 0 \end{aligned}$$

для всіх $\varphi \in D^{(i,j)}, \psi \in D^{(m,p)}$.

Дана операція є асоціативною. Щойно отриману напівгрупу позначимо через $B^0[H]$.

3.2. Нехай $S = S[\Lambda; G_i, \phi_{i,j}]$ – Кліффордова напівгрупа, $F = \{G_i\}_{i \in \Lambda}$. В позначеннях п. 3.1 отримуємо наступне описання напівгрупи $T_i(S)$ лівих зсувів Кліффордової напівгрупи S .

Твердження. Існує мономорфізм

$$T_i(S) \rightarrow A(\Lambda; B^0[F]).$$

Доведення. Безпосередньо перевіряється, що стверджуваний мономорфізм визначається відображенням

$$\gamma: T_i(S) \rightarrow A(\Lambda; B^0[F]):$$

$$\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda} \mapsto \lambda \gamma = (a_{in}), \quad \text{де}$$

$$(a_{in}) = \begin{cases} \lambda_i \Leftrightarrow n = i\lambda_0, \\ 0 \text{ в інших випадках} \end{cases} \quad (i, n \in \Lambda).$$

Твердження доведено.

3.3. Для напівгрупи $T_i(S)$ правих зсувів Кліффордової напівгрупи S аналогічно доводиться наступне

Твердження. Існує мономорфізм

$$T_i(S) \rightarrow A(\Lambda; B^0[F]).$$

3.4. Напівгрупу $(A(\Lambda; B^0[F]))^2$ з операцією

$$((a_{ij}), (b_{ij}))((c_{ij}), (d_{ij})) =$$

$$= ((c_{ij})(a_{ij}), (b_{ij})(d_{ij}))$$

для всіх $(a_{ij}), (b_{ij}), (c_{ij}), (d_{ij}) \in A(\Lambda; B^0[F])$ позначимо через $\bar{A}(\Lambda; B^0[F])$.

У відповідності до визначення оболонки зсувів як безпосередній наслідок теореми п.2.1 та тверджень п.п. 3.2, 3.3 отримуємо наступне описання оболонки зсувів $TH(S)$ Кліффордової напівгрупи $S = S[\Lambda; G_i, \phi_{i,j}]$.

Лема. Існує мономорфізм

$$TH(S) \rightarrow \bar{A}(\Lambda; B^0[F]).$$

3.5. Нагадаємо, що довільна напівгрупа T називається слабо редукованою, якщо для $a, b \in T$ з рівностей $\lambda_a = \lambda_b$ та $\rho_a = \rho_b$ випливає, що $a = b$. Відомо, що напівгрупа T ізоморфна внутрішній частині $\overline{TH}(T)$ своєї оболонки зсувів $TH(T)$ тоді й лише тоді коли T є слабо редукованою (див., наприклад [6]). Добре відомо також, що будь-яка Кліффордова напівгрупа S є слабо редукованою ([8]). Отже, $S \cong \overline{TH}(S)$ і як

безпосередній наслідок леми п. 3.4 отримуємо наступну характеристику Кліффордової напівгрупи $S = S[\Lambda; G_i, \phi_{i,j}]$.

Теорема. Існує мономорфізм

$$S[\Lambda; G_i, \phi_{i,j}] \rightarrow \bar{A}(\Lambda; B^0[F]).$$

3.6. Оскільки кожний лівий зсув $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda}$ Кліффордової напівгрупи $S = S[\Lambda; G_i, \phi_{i,j}]$ визначається зсувами напівгрупи Λ та відображеннями $\lambda_i, i \in \Lambda$ (див.п.1.1), то описання властивостей лівих зсувів Кліффордової напівгрупи зводиться до описання властивостей відображень $\lambda_i, i \in \Lambda$ та зсувів напівгрупи Λ .

В пунктах 3.6-3.11 ми розглянемо для лівого зсуву $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda}$ та правого зсуву $\rho = [\rho_0, \rho_i]_{i \in \Lambda}$ властивості відображень $\lambda_i, i \in \Lambda$ та $\rho_i, i \in \Lambda$, відповідно.

З умови (1) п.1.1 та наслідку п.1.3 безпосередньо випливає

Теорема. Нехай напівгрупа Λ має найменший елемент δ і нехай усі гомоморфізми $\phi_{i,j}, i, j \in \Lambda$ є ізоморфізмами. Тоді кожне відображення λ_i визначається через λ_δ за формулою

$$\lambda_i = \phi_{i,\delta} \lambda_\delta \phi_{i\lambda_0, \delta\lambda_0}^{-1}.$$

3.7. Має місце

Теорема. Нехай усі гомоморфізми $\phi_{i,j}, i, j \in \Lambda$ є ізоморфізмами. Тоді кожне відображення λ_i визначається через будь-яке відображення λ_j та гомоморфізми за формулою

$$\lambda_i = \phi_{i,\gamma} \phi_{j,\gamma}^{-1} \lambda_j \phi_{j\lambda_0, \gamma\lambda_0} \phi_{i\lambda_0, \gamma\lambda_0}^{-1}, \quad \text{де } \gamma = ij.$$

Доведення. Використовуючи умову (1) п.1.1, маємо:

$$\phi_{i,\gamma} \lambda_\gamma = \lambda_i \phi_{i\lambda_0, \gamma\lambda_0}, \quad \phi_{j,\gamma} \lambda_\gamma = \lambda_j \phi_{j\lambda_0, \gamma\lambda_0}$$

для всіх $i, j \in \Lambda$, де $\gamma = ij$. Звідси $\lambda_i = \phi_{i,\gamma} \lambda_\gamma \phi_{i\lambda_0, \gamma\lambda_0}^{-1}$, $\lambda_j = \phi_{j,\gamma}^{-1} \lambda_j \phi_{j\lambda_0, \gamma\lambda_0}$. Підставляючи другу рівність в першу, отримуємо шукане рівняння.

Теорему доведено.

3.8. З умови (1) п.1.1 безпосередньо випливає

Теорема. Нехай $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda}$ – лівий зсув Кліффордової напівгрупи S . Тоді $\text{Im } \lambda_j \supseteq (\text{Im } \lambda_i) \phi_{i\lambda_0, j\lambda_0}$.

3.9. Нехай $\rho = [\rho_0, \rho_i]_{i \in \Lambda}$ – правий зсув Кліффордової напівгрупи S .

У двоїстий спосіб (п.3.6) для правих зсувів Кліффордової напівгрупи отримуємо

Теорема. Нехай напівгрупа Λ має найменший елемент δ і нехай усі гомоморфізми $\phi_{i,j}$, $i, j \in \Lambda$ є ізоморфізмами. Тоді кожне відображення ρ_i визначається через ρ_δ за формулою

$$\rho_i = \phi_{i,\delta} \rho_\delta \phi_{i\rho_0, \delta\rho_0}^{-1}.$$

3.10. У двоїстий спосіб до п.3.7 доводиться наступна

Теорема. Нехай усі гомоморфізми $\phi_{i,j}$, $i, j \in \Lambda$ є ізоморфізмами. Тоді кожне відображення ρ_i визначається через будь-яке відображення ρ_j та гомоморфізми за формулою

$$\rho_i = \phi_{i,\gamma} \phi_{j,\gamma}^{-1} \rho_j \phi_{j\rho_0, \gamma\rho_0} \phi_{i\rho_0, \gamma\rho_0}^{-1}, \text{ де } \gamma = ij.$$

3.11. З умови (2) п. 1.5 безпосередньо випливає

Теорема. Нехай $\rho = [\rho_0, \rho_i]_{i \in \Lambda}$ – правий зсув Кліффордової напівгрупи S . Тоді $\text{Im } \rho_j \supseteq (\text{Im } \rho_i) \phi_{i\rho_0, j\rho_0}$.

3.12. В наступних пунктах роботи ми розглянемо зсуви комутативної напівгрупи ідемпотентів Λ .

Теорема. $TH(\Lambda) \cong T_l(\Lambda) = T_r(\Lambda)$.

Доведення. Нехай λ та ρ – лівий та правий зсуви напівгрупи Λ . Для будь-яких $a, b \in \Lambda$ маємо:

$$\begin{aligned} (ab)\lambda &= (ba)\lambda = (b\lambda)a = a(b\lambda), \\ (ab)\rho &= (ba)\rho = b(a\rho) = (a\rho)b, \end{aligned}$$

звідки $T_l(\Lambda) = T_r(\Lambda)$.

Якщо λ та ρ – зв'язані лівий та правий зсуви напівгрупи Λ , то для будь-якого $a \in \Lambda$ маємо:

$$\begin{aligned} a(a\lambda) &= (a\lambda)a = (aa)\lambda = a\lambda = \\ &= a\rho = (aa)\rho = a(a\rho) = (a\rho)a. \end{aligned}$$

Отже, $\lambda = \rho$. Навпаки, якщо $\lambda = \rho$, то для будь-яких $a, b \in \Lambda$ маємо:

$$\begin{aligned} a(b\lambda) &= a(b\rho) = (ab)\rho = \\ &= (ab)\lambda = (a\lambda)b = (a\rho)b, \end{aligned}$$

звідки λ та ρ – зв'язані. Отже, шуканий ізоморфізм визначається відображенням $(\tau; \tau) \mapsto \tau$.

Теорему доведено.

3.13. Має місце

Твердження. $T_l(\Lambda)$ – комутативна напівгрупа ідемпотентів.

Доведення. Для будь-якого зсуву τ напівгрупи Λ та для будь-якого $a \in \Lambda$ маємо:

$$a\tau^2 = ((aa)\tau)\tau = (a\tau a)\tau = a\tau a\tau = a\tau,$$

звідки $\tau^2 = \tau$.

Згідно теореми п.3.12 для будь-яких зсувів τ_1, τ_2 напівгрупи Λ також маємо:

$$\begin{aligned} a\tau_1\tau_2 &= (aa)\tau_1\tau_2 = (a(a\tau_1))\tau_2 = \\ &= a\tau_2 a\tau_1 = (a\tau_2 a)\tau_1 = \\ &= ((aa)\tau_2)\tau_1 = a\tau_2\tau_1, \end{aligned}$$

звідки $\tau_1\tau_2 = \tau_2\tau_1$. Твердження доведено.

3.14. Має місце

Твердження. Існує мономорфізм $T_l(\Lambda) \rightarrow \text{End}(\Lambda)$.

Доведення. Нагадаємо, що $\tau^2 = \tau$ для будь-якого зсуву $\tau \in T_l(\Lambda)$. Стверджуваний мономорфізм випливає з того, що для будь-яких $a, b \in \Lambda$ має місце рівність:

$$(ab)\tau = (ab)\tau^2 = (a\tau b)\tau = a\tau b\tau.$$

Твердження доведено.

3.15. Нехай τ – лівий зсув напівгрупи Λ . Через I_τ позначимо множину $\{y \in Y \mid y\tau = y\}$. Оскільки $\tau^2 = \tau$, то $I_\tau \neq \{\emptyset\}$ при будь-якому $\tau \in T_l(\Lambda)$.

Теорема. Нехай τ – лівий зсув напівгрупи Λ . Тоді

- 1) I_τ – ідеал напівгрупи Λ ;
- 2) $I_{\tau_1} I_{\tau_2} = I_{\tau_1} \cap I_{\tau_2} = I_{\tau_1 \tau_2}$ для будь-яких зсувів $\tau_1, \tau_2 \in T_l(\Lambda)$;
- 3) для будь-якого $b \in \Lambda$ існує $g \in \Lambda$ такий, що $I_\tau \cap b\Lambda \subseteq g\Lambda$.

Доведення. Нехай $x, y \in I_\tau$. Згідно твердження п.3.14

$$xy = (x\tau)(y\tau) = (xy)\tau,$$

звідки випливає, що $xy \in I_\tau$. Для будь-якого $z \in \Lambda$ маємо, далі $(xz)\tau = (x\tau)z = xz$, звідки $xz \in I_\tau$ і, аналогічно, $zx \in I_\tau$. Отже, I_τ – ідеал напівгрупи Λ .

Нехай $x \in I_{\tau_1} I_{\tau_2}$. Тоді

$$x = y_1 y_2, y_1 \in I_{\tau_1}, y_2 \in I_{\tau_2},$$

причому $y_1 \tau_1 = y_1, y_2 \tau_2 = y_2$. Оскільки

$$x\tau_1\tau_2 = (y_1y_2)\tau_1\tau_2 = (y_1\tau_1y_2)\tau_2 = \\ = y_1\tau_1y_2\tau_2 = y_1y_2 = x,$$

то $x \in I_{\tau_1\tau_2}$. Навпаки, нехай $x \in I_{\tau_1\tau_2}$. Тоді

$$x = x\tau_1\tau_2 = (xx)\tau_1\tau_2 = \\ = (x\tau_1x)\tau_2 = x\tau_1x\tau_2.$$

Оскільки $x\tau_1 \in I_{\tau_1}$ та $x\tau_2 \in I_{\tau_2}$, то $x \in I_{\tau_1}I_{\tau_2}$. Отже, $I_{\tau_1}I_{\tau_2} = I_{\tau_1\tau_2}$.

Нехай $x \in I_{\tau_1} \cap I_{\tau_2}$. Тоді

$$x\tau_1 = x, \quad x\tau_2 = x, \quad \text{звідки} \\ x = x\tau_1x\tau_2 = (x\tau_1x)\tau_2 = \\ = ((xx)\tau_1)\tau_2 = x\tau_1\tau_2,$$

тобто $x \in I_{\tau_1\tau_2}$. Навпаки, якщо $x \in I_{\tau_1\tau_2}$, то $x = x\tau_1\tau_2$, звідки

$$x\tau_1 = x\tau_1\tau_2\tau_1 = x\tau_1\tau_2 = x, \\ x\tau_2 = x\tau_1\tau_2\tau_2 = x\tau_1\tau_2 = x,$$

тобто $x \in I_{\tau_1} \cap I_{\tau_2}$.

Нехай $x \in I_{\tau} \cap b\Lambda$. Тоді $x = x\tau = bg$ для деякого $g \in \Lambda$. Звідси

$$x = (bg)\tau = (gb)\tau = g(b\tau), \quad \text{тому } x \in g\Lambda.$$

Теорему доведено.

4. Ідемпотентність напівгруп $T_l(S)$ та $T_r(S)$. В цьому пункті напівгрупа $T_l(S)$ ($T_r(S)$) лівих (правих) зсувів Кліффордової напівгрупи S досліджується на ідемпотентність.

4.1. Має місце

Твердження. Лівий зсув $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda}$ Кліффордової напівгрупи S є ідемпотентним тоді і тільки тоді, коли $e_{i\lambda_0}\lambda_{i\lambda_0} = e_{i\lambda_0}$ для всіх $i \in \Lambda$.

Доведення. Нагадаємо, що з твердження п.3.13 випливає, що $\lambda_0^2 = \lambda_0$ для будь-якого зсуву $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda} \in T_l(S)$. Нехай $\lambda^2 = \lambda$. Для довільного $a_i \in G_i, i \in \Lambda$, використовуючи умову (i) наслідку п.1.2, маємо:

$$a_i\lambda^2 = ((e_i\lambda_i)(a_i\phi_{i,\lambda_0}))\lambda = \\ = (e_{i\lambda_0}\lambda_{i\lambda_0})((e_i\lambda_i)(a_i\phi_{i,\lambda_0}))\phi_{i\lambda_0,i\lambda_0^2} = \\ = (e_{i\lambda_0}\lambda_{i\lambda_0})(e_i\lambda_i)(a_i\phi_{i,\lambda_0}) = \\ = (e_i\lambda_i)(a_i\phi_{i,\lambda_0}) = a_i\lambda,$$

звідки $e_{i\lambda_0}\lambda_{i\lambda_0} = e_{i\lambda_0}$ для всіх $i \in \Lambda$.

Навпаки. Нехай $\lambda = [\lambda_0, \lambda_i]_{i \in \Lambda} \in T_l(S)$,

причому $e_{i\lambda_0}\lambda_{i\lambda_0} = e_{i\lambda_0}$ для всіх $i \in \Lambda$. Тоді при будь-якому $a_i \in G_i, i \in \Lambda$ матимемо:

$$a_i\lambda^2 = ((e_i\lambda_i)(a_i\phi_{i,\lambda_0}))\lambda = \\ = (e_{i\lambda_0}\lambda_{i\lambda_0})((e_i\lambda_i)(a_i\phi_{i,\lambda_0}))\phi_{i\lambda_0,i\lambda_0^2} = \\ = (e_{i\lambda_0}\lambda_{i\lambda_0})(e_i\lambda_i)(a_i\phi_{i,\lambda_0}) = \\ = (e_{i\lambda_0})(e_i\lambda_i)(a_i\phi_{i,\lambda_0}) = \\ = (e_i\lambda_i)(a_i\phi_{i,\lambda_0}) = a_i\lambda.$$

Твердження доведено.

4.2. Аналогічно доводиться

Твердження. Правий зсув $\rho = [\rho_0, \rho_i]_{i \in \Lambda}$ Кліффордової напівгрупи S є ідемпотентним тоді і тільки тоді, коли $e_{i\rho_0}\rho_{i\rho_0} = e_{i\rho_0}$ для всіх $i \in \Lambda$.

Висновки

В роботі описано ліві (праві) зсуви Кліффордової напівгрупи, знайдено необхідні та достатні умови зв'язаності лівого та правого зсувів Кліффордової напівгрупи, охарактеризовано оболонку зсувів Кліффордової напівгрупи. Побудовано адекватну теоретико-напівгрупову конструкцію, в термінах якої описано Кліффордову напівгрупу.

Список використаних джерел

1. M.Petrich. The translational hull in semigroups and rings // Semigroup Forum. – 1970. – V.1. – P.283-360.
2. Жучок А.В. Зсуви напівгрупи Ріса матричного типу // Вісник Донецького ун-ту. Сер. А. Природничі науки, ч.1. – вип. 1. – 2005. – с.7-15.
3. А.И.Закусило, В.М.Усенко. Голоморфные сплетения и сдвиговые оболочки полугрупп // Труды ИПММ НАН Украины. 2005. Вып.11. – с.49-60.
4. M. Samman, J.Meldrum. On Endomorphisms of Semilattices of Groups // Algebra Colloquium 12:1 (2005), pp. 93-100.
5. G.Ayik, H. Ayik and Y.Unlu. Presentations and word problem for strong semilattices of semigroups // Algebra and Discrete Mathematics, No.4. (2005). pp. 28-35.
6. Клиффорд А., Престон Г. Алгебраическая теория полугрупп // М.:Мир. – 1972. – Т.1. – С.285.
7. J.M. Howie. An Introduction to Semigroup Theory, Acad. Press, London, 1976.
8. А.Артамонов, В.Н.Салий, Л.А. Скорняков и др. Общая алгебра. Т.2. – М.:Наука. – 1991. – 480с.

Надійшла до редколегії 17.04.2008 р.

УДК 512.53

Роман А. Заторський, к.ф.-м.н., доцент

Операції з формальними степеневими рядами з ненульовим вільним членом

Розглядаються операції з формальними степеневими рядами з ненульовим вільним членом.

Ключові слова: формальний степневий ряд, парадетермінант, парперманент.

E-mail: romazz@rambler.ru

Статтю представив д.ф.-м.н., професор А.П.Петравчук

1 Вступ

Центральним методом комбінаторного аналізу є метод генератрис, в якому важливу роль відіграють операції з формальними степеневими рядами. Однак, ряд формальних операцій із такими рядами пов'язані з певними труднощами.

В цій роботі, при допомозі апарату парадетермінантів та парперманентів трикутних матриць [1], побудовано ряд рекурсивних алгоритмів операцій із формальними степеневими рядами з ненульовим вільним членом. Позаяк для знаходження значення парафункцій трикутних матриць n -того порядку достатньо виконати всього $\frac{n(n-1)}{2}$ операцій множення і стільки ж операцій

додавання [2], то можна зробити висновок, що апарат парадетермінантів корисний при роботі з формальними степеневими рядами.

2. Операції із формальними степеневими рядами з ненульовим вільним членом

Нижче розглядатимемо степеневі ряди

$$A(x) = 1 + \sum_{i=1}^{\infty} a_i x^i, \quad (2.1)$$

де a_i , $i = 1, 2, \dots$ – деякі, взагалі кажучи, комплексні числа. Такі степеневі ряди називають формальними степеневими рядами, причому $A(x)$ у рівності (2.1) є лише позначенням цього ряду, але не завжди його сумою. Нижче ми будемо ігнорувати питання збіжності таких рядів.

Ми не наводитимемо загальновідомі означення суми та добутку формальних степеневих рядів, а розглянемо лише ті операції, які пов'язані з певними труднощами.

Справедлива наступна

Теорема 1. Нехай

Roman A. Zatorsky, Cand.Sci. (Phys.-Math.),
Associate Professor

Operations with formal power series which have nontrivial free numbers

Operations with formal power series which have nontrivial free coefficient are considered.

Key Words: formal power series, paraderterminant, parapermanent.

$$A(x) = 1 + \sum_{i=1}^{\infty} a_i x^i, B(x) = 1 + \sum_{i=1}^{\infty} b_i x^i, C(x) = 1 + \sum_{i=1}^{\infty} c_i x^i$$

причому виконується рівність

$$C(x) = \frac{A(x)}{B(x)}, \quad (1)$$

тоді справедливі рівності:

$$c_i = \sum_{j=0}^{i-1} (-1)^j (a_{i-j} - b_{i-j}) \cdot \left\langle \frac{b_{s-r+1}}{b_{s-r}} \right\rangle_{1 \leq r \leq s \leq j}, \quad (2)$$

де $i = 1, 2, \dots$ Прийнято вважати, що

$$\left\langle \frac{b_{s-r+1}}{b_{s-r}} \right\rangle_{1 \leq r \leq s \leq 0} = 1.$$

Доведення. Запишемо рівність (1) у вигляді системи рівнянь

$$b_i + b_{i-1}c_1 + b_{i-2}c_2 + \dots + b_1c_{i-1} + c_i = a_i, \quad i = 1, 2, \dots \quad (3)$$

і доведемо, що її розв'язком є вираз (2). Нижче ми користуємось позначенням

$$\diamond_j = \left\langle \frac{b_{s-r+1}}{b_{s-r}} \right\rangle_{1 \leq r \leq s \leq j}.$$

При $i = 1$ рівність (2), очевидно, справедлива. Доведемо її справедливість при $i = m+1$ із врахуванням того, що вона виконується для всіх $i = 1, 2, \dots, m$.

$$\begin{aligned} c_{m+1} &= a_{m+1} - b_{m+1} - \sum_{i=1}^m b_{m+1-i} \cdot c_i = a_{m+1} - b_{m+1} - \\ &- \sum_{i=1}^m b_{m+1-i} \cdot \sum_{j=0}^{i-1} (-1)^j (a_{i-j} - b_{i-j}) \cdot \diamond_j = \\ &= a_{m+1} - b_{m+1} - (a_m - b_m) \cdot b_1 \cdot \diamond_0 + \\ &+ (a_{m-1} - b_{m-1}) \cdot (b_1 \cdot \diamond_1 - b_2 \cdot \diamond_0) - \dots + \\ &+ (-1)^m \cdot (a_1 - b_1) \cdot (b_1 \cdot \diamond_{m-1} - b_2 \cdot \diamond_{m-2} + \dots \\ &+ (-1)^{m-1} b_m \cdot \diamond_0) = (a_{m+1} - b_{m+1}) \cdot \diamond_0 - (a_m - b_m) \cdot \diamond_1 + \end{aligned}$$

$$+ \dots + (-1)^m (a_1 - b_1) \cdot \diamond_m = \sum_{j=0}^m (-1)^j (a_{m+1-j} - b_{m+1-j}) \cdot \diamond_j = \left[(-1)^{\delta_{ij}} \cdot \frac{(i-j+1)p - (j-1)}{(i-j)p - j} \cdot \frac{a_{i-j+1}}{a_{i-j}} \right]_{1 \leq j \leq i \leq n} \quad (9)$$

Примітка. Система лінійних рівнянь

$$\sum_{i+j=k} b_i c_j = 0, \quad k = 1, 2, \dots, \quad (4)$$

де $0 \leq i, j \leq k$, причому $a_0 = c_0 = 1$, має розв'язок

$$c_i = (-1)^i \cdot \left\langle \frac{b_{s-r+1}}{b_{s-r}} \right\rangle_{1 \leq r \leq s \leq i} = (-1)^i \diamond_i, \quad i = 1, 2, \dots \quad (5)$$

Для доведення цього твердження достатньо у рівності (2) покласти $a_i = 0, i = 1, 2, \dots$, тоді матимемо:

$$c_i = \sum_{j=0}^i (-1)^j (-b_{i-j}) \cdot \diamond_j = (-1)^i \cdot \sum_{j=0}^i (-1)^{i+j+1} b_{i-j} \diamond_j = (-1)^i \cdot \diamond_i, \quad i = 1, 2, \dots$$

Наслідок 1. Справедлива рівність

$$(A(x))^{-1} = 1 - \diamond_1 x + \diamond_2 x^2 - \dots + (-1)^i \diamond_i x^i + \dots, \quad (6)$$

де

$$\diamond_i = \left\langle \frac{a_{s-r+1}}{a_{s-r}} \right\rangle_{1 \leq r \leq s \leq i}, \quad i = 1, 2, \dots \quad (7)$$

Знайдемо рекурсивний алгоритм (див. [2]) піднесення формального степеневому ряду $A(x)$ до степеня $p \in R$.

Нехай

$$\left(1 + \sum_{i=1}^{\infty} a_i x^i \right)^p = 1 + \sum_{i=1}^{\infty} b_i x^i.$$

Здиференціюємо обидві частини цієї рівності і помножимо на $A(x)$, тоді отримаємо рівність

$$pB(x)A'(x) = B'(x)A(x).$$

Застосовуючи правило згортки і зрівнюючи коефіцієнти при однакових степенях, отримаємо співвідношення

$$p \sum_{j=0}^{i-1} (i-j) b_j a_{i-j} = \sum_{j=0}^{i-1} (i-j) a_j b_{i-j}, \quad i = 1, 2, \dots$$

з яких при $i = n-1$ знаходимо шукане рекурентне співвідношення

$$b_n = \sum_{i=1}^n \frac{i(p+1)-n}{n} \cdot a_i b_{n-i}, \quad b_0 = 1, \quad n = 1, 2, \dots \quad (8)$$

Теорема 2. Якщо

$$B(x) = (A(x))^p,$$

де p – деяке дійсне число, то справедливі рівності:

$$b_n = (-1)^n \left\langle \frac{(i-j+1)p - (j-1)}{(i-j)p - j} \cdot \frac{a_{i-j+1}}{a_{i-j}} \right\rangle_{1 \leq j \leq i \leq n} =$$

Доведення. Рівність парадетермінанта і парперманента в рівностях (9) доводиться винесенням із кожного стовпчика парадетермінанта за його знак спільного множника, що дорівнює $(-1)^i$ і застосуванням теореми про зв'язок парперманента і парадетермінанта. Розкладемо парперманент із рівності (9) за елементами останнього рядка. При цьому отримаємо рівність (8).

Наслідок 2. Справедливі наступні комбінаторні тотожності:

$$(A(x))^n = 1 + \sum_{i=1}^{\infty} b_i x^i,$$

$$(A(x))^{-n} = 1 + \sum_{i=1}^{\infty} c_i x^i,$$

$$(A(x))^{\frac{1}{n}} = 1 + \sum_{i=1}^{\infty} d_i x^i,$$

$$(A(x))^{-\frac{1}{n}} = 1 + \sum_{i=1}^{\infty} e_i x^i,$$

де

$$b_i = (-1)^k \left\langle \frac{(i-j+1)n - (j-1)}{(i-j)n - j} \cdot \frac{a_{i-j+1}}{a_{i-j}} \right\rangle_{1 \leq j \leq i \leq k},$$

$$c_i = (-1)^k \left\langle \frac{(i-j+1)n + (j-1)}{(i-j)n + j} \cdot \frac{a_{i-j+1}}{a_{i-j}} \right\rangle_{1 \leq j \leq i \leq k},$$

$$d_i = (-1)^k \left\langle \frac{-(i-j+1) + (j-1)n}{-(i-j) + jn} \cdot \frac{a_{i-j+1}}{a_{i-j}} \right\rangle_{1 \leq j \leq i \leq k},$$

$$e_i = (-1)^k \left\langle \frac{(i-j+1) + (j-1)n}{(i-j) + jn} \cdot \frac{a_{i-j+1}}{a_{i-j}} \right\rangle_{1 \leq j \leq i \leq k}.$$

Доведення цих комбінаторних тотожностей одразу випливають із рівностей (9), якщо замінити в ній p відповідно на $n, -n, \frac{1}{n}, -\frac{1}{n}$.

Теорема 3. Нехай f і f^{-1} – взаємно обернені функції, тоді якщо справедлива рівність

$$f\left(1 + \sum_{i=1}^{\infty} a_i x^i\right) = 1 + \sum_{i=1}^{\infty} b_i x^i, \quad (10)$$

тут

$$b_i = \left\langle \tau_{x^p} \frac{a_{s-r+1}}{a_{s-r}} \right\rangle_{1 \leq r \leq s \leq i}, \quad i = 1, 2, \dots \quad (11)$$

то справедлива також рівність

$$f^{-1}\left(1 + \sum_{i=1}^{\infty} a_i x^i\right) = 1 + \sum_{i=1}^{\infty} c_i x^i, \quad (12)$$

де

$$c_i = \left\langle \tau_{s,s-r+1}^{-1} \frac{a_{s-r+1}}{a_{s-r}} \right\rangle_{1 \leq r \leq s \leq i}, i = 1, 2, \dots \quad (13)$$

Доведення. Перехід від рівності (10) до рівності

$$1 + \sum_{i=1}^{\infty} a_i x^i = f^{-1} \left(1 + \sum_{i=1}^{\infty} b_i x^i \right)$$

рівносильний розв'язанню системи рівнянь (11) відносно a_i . Покажемо, що розв'язком цієї системи рівнянь є

$$a_i = \left\langle \tau_{s,s-r+1}^{-1} \frac{b_{s-r+1}}{b_{s-r}} \right\rangle_{1 \leq r \leq s \leq i}, i = 1, 2, \dots \quad (14)$$

Випадок $i = 1$ перевіряється безпосередньо. Нехай рівність (14) справедлива для $i = 2, 3, \dots, k-1$. Доведемо справедливість цієї рівності для $i = k$. Розкладемо парадетермінант, у правій частині рівності (11), при $i = k$, за елементами k -того рядка

$$\tau_{kk} a_1 b_{k-1} - \tau_{kk} \tau_{k,k-1} a_2 b_{k-2} + \dots +$$

$$+ (-1)^{k-2} \tau_{kk} \dots \tau_{k2} a_{k-1} b_1 + (-1)^{k-1} \tau_{kk} \dots \tau_{k1} a_k = b_k$$

Замінімо в цій рівності змінні $a_1, a_2, \dots, a_{k-1}$ відповідними пара детермінантами із рівностей (14)

$$(-1)^{k-1} \tau_{kk} \dots \tau_{k1} a_k = b_k - \tau_{kk} b_{k-1} \left\langle \tau_{11}^{-1} b_1 \right\rangle +$$

$$+ \tau_{kk} \tau_{k,k-1} b_{k-2} \left\langle \tau_{11}^{-1} b_1 \right\rangle \left\langle \tau_{22}^{-1} \frac{b_2}{b_1} \tau_{21}^{-1} b_1 \right\rangle - \dots +$$

$$+ (-1)^{k-1} \tau_{kk} \dots \tau_{k2} b_1 \left\langle \tau_{s,s-r+1}^{-1} \frac{b_{s-r+1}}{b_{s-r}} \right\rangle_{1 \leq r \leq s \leq k-1}$$

Знайдемо із останньої рівності a_k :

$$a_k = (-1)^{k-1} \tau_{k1}^{-1} \dots \tau_{kk}^{-1} b_k + (-1)^{k-2} \tau_{k1}^{-1} \dots$$

$$\cdot \tau_{k,k-1}^{-1} b_{k-1} \left\langle \tau_{11}^{-1} b_1 \right\rangle + \dots + \tau_{k1}^{-1} b_1 \left\langle \tau_{s,s-r+1}^{-1} \frac{b_{s-r+1}}{b_{s-r}} \right\rangle_{1 \leq r \leq s \leq k-1}$$

Але права частина останньої рівності є розкладом парадетермінанта (14) при $i = k$ за елементами останнього рядка.

Тепер, для доведення теореми, у рівностях (14), залишається зробити заміну a_i на c_i і b_i на a_i , $i = 1, 2, \dots$

Приклад 1. Позаяк функції $f(x) = x^n$ і

$f^{-1}(x) = x^{\frac{1}{n}}$, при $x \geq 0$, взаємно обернені, то внаслідок теореми 3, за відомим формальним степеневим рядом $(A(x))^n$, можна знайти формальний степеневий ряд $(A(x))^{\frac{1}{n}}$. Для цього достатньо, за

коефіцієнтами

$$\tau_{ij} = \frac{(i-j+1)n - (j-1)}{(i-j)n - j}$$

елементів парадетермінантів із рівності (11) знайти відповідні коефіцієнти $\tau_{i,i-j+1}^{-1}$ елементів пара детермінантів із рівності (13):

$$\tau_{i,i-j+1}^{-1} = \left(\frac{(i - (i-j+1) + 1)n - (i-j+1-1)}{(i - (i-j+1))n - (i-j+1)} \right)^{-1} =$$

$$= \frac{-(i-j+1) + (j-1)n}{-(i-j) + jn}.$$

Теорема 4. Справедлива тотожність

$$(1 + a_1 x + a_2 x^2 + \dots)^n =$$

$$= 1 + \sum_{i=1}^{\infty} \left(\sum_{\lambda_1 + 2\lambda_2 + \dots + i\lambda_i = i} \frac{n^{\lambda_1 + \dots + \lambda_i}}{\lambda_1! \dots \lambda_i!} a_1^{\lambda_1} \dots a_i^{\lambda_i} \right) x^i,$$

тут n – натуральне число.

Доведення. Щоб знайти коефіцієнт біля степеня x^i , необхідно знайти суму всіх добутків

$$a_{1j_1} a_{2j_2} \dots a_{nj_n}, \quad (15)$$

для яких сума елементів мультимножини

$$\{j_1, j_2, \dots, j_n\} \quad (16)$$

дорівнює i , тобто виконується рівність

$$j_1 + j_2 + \dots + j_n = i. \quad (17)$$

Якщо мультимножину (16) записати в канонічному вигляді

$$\{0^{\lambda_0}, 1^{\lambda_1}, \dots, i^{\lambda_i}\}, \quad (18)$$

то елементи її первинної специфікації задовольняють рівняння

$$\lambda_1 + 2\lambda_2 + \dots + i\lambda_i = i. \quad (19)$$

Зафіксуємо мультимножину (18) і знайдемо число всіх мультимножин (16) з канонічним виглядом (18). Позаяк число ненульових елементів мультимножини (16) дорівнює $\lambda_1 + \lambda_2 + \dots + \lambda_i$, то існує

$$\binom{n}{\lambda_1 + \dots + \lambda_i} = \frac{n^{\lambda_1 + \dots + \lambda_i}}{(\lambda_1 + \dots + \lambda_i)!}$$

таких мультимножин. Якщо в останніх врахувати ще й порядок елементів, то всього добутків вигляду (15) з умовою (17) є

$$\frac{n^{\lambda_1 + \dots + \lambda_i}}{(\lambda_1 + \dots + \lambda_i)!} \cdot \frac{(\lambda_1 + \dots + \lambda_i)!}{\lambda_1! \dots \lambda_i!} = \frac{n^{\lambda_1 + \dots + \lambda_i}}{\lambda_1! \dots \lambda_i!}.$$

Отже існує

$$\frac{n^{\lambda_1 + \dots + \lambda_i}}{\lambda_1! \dots \lambda_i!}$$

добутків $a_1^{\lambda_1} \dots a_i^{\lambda_i}$ і коефіцієнт степеня x^i , з врахуванням умови (19), дорівнює

$$\sum_{\lambda_1+2\lambda_2+\dots+i\lambda_i=i} \frac{n^{\lambda_1+\dots+\lambda_i}}{\lambda_1! \dots \lambda_i!} \cdot a_1^{\lambda_1} \dots a_i^{\lambda_i}.$$

Твердження 1. Справедлива тотожність:

$$(1+x+x^2+\dots)^n = 1 + \frac{n^1}{1!}x + \frac{n^2}{2!}x^2 + \dots + \frac{n^i}{i!}x^i + \dots \quad (20)$$

Доведення. При $n=1$ рівність (20) очевидна. Нехай вона справедлива при $n=k$. Доведемо виконання індуктивного кроку.

$$(1+x+\dots+x^i+\dots) \cdot (1+\frac{k^1}{1!}x+\dots+\frac{k^i}{i!}x^i+\dots) =$$

$$= 1 + \sum_{i=1}^{\infty} (1+\frac{k^1}{1!}+\frac{k^2}{2!}+\dots+\frac{k^i}{i!}) \cdot x^i.$$

Але тотожність

$$\sum_{j=1}^i \frac{k^j}{j!} = \frac{(k+1)^i}{i!}$$

легко доводиться при допомозі тотожностей

$$\frac{(r+1)^i}{i!} = \frac{(i+1)^i}{r!}, \quad \sum_{j=1}^i \frac{j^{r-1}}{(r-1)!} = \frac{i^r}{r!}, \quad i, r \in N,$$

(див. [3]).

Твердження 2. Справедливі наступні тотожності:

$$\left[(-1)^{\delta_{ij}} \frac{(i-j+1)n+(j-1)}{(i-j)n+j} \right]_{1 \leq j \leq i \leq k} =$$

$$\sum_{\lambda_1+2\lambda_2+\dots+k\lambda_k=k} (-1)^s \frac{n^{\lambda_1+\dots+\lambda_k}}{\lambda_1! \dots \lambda_k!} = (-1)^k \frac{n^k}{k!},$$

Доведення. Це твердження одразу випливає із наслідку 2, теореми 4 та твердження 1.

Справедливі також наступні комбінаторні тотожності, які ми не доводимо. Зауважимо лише, що їх можна отримати, замінивши у рівностях твердження 2 величину n відповідно на $-n$, $\frac{1}{n}$

та $-\frac{1}{n}$.

$$\left[(-1)^{\delta_{ij}} \frac{(i-j+1)n-(j-1)}{(i-j)n-j} \right]_{1 \leq j \leq i \leq k} =$$

$$\sum_{\lambda_1+2\lambda_2+\dots+k\lambda_k=k} \frac{n^{\lambda_1+\dots+\lambda_k}}{\lambda_1! \dots \lambda_k!} = \frac{n^k}{k!}$$

$$\left[(-1)^{\delta_{ij}} \frac{-(i-j+1)+(j-1)n}{-(i-j)+jn} \right]_{1 \leq j \leq i \leq k} =$$

$$\sum_{\lambda_1+\dots+k\lambda_k=k} (-1)^s \frac{(n-1) \dots ((s-1)n-1)}{n^s \lambda_1! \dots \lambda_k!} =$$

$$= (-1)^k \frac{(n+1)(2n+1) \dots ((k-1)n+1)}{n^k k!},$$

$$\left[(-1)^{\delta_{ij}} \frac{(i-j+1)+(j-1)n}{(i-j)+jn} \right]_{1 \leq j \leq i \leq k} =$$

$$\sum_{\lambda_1+\dots+k\lambda_k=k} (-1)^s \frac{(n+1) \dots ((s-1)n+1)}{n^s \lambda_1! \dots \lambda_k!} =$$

$$= (-1)^k \frac{(n-1)(2n-1) \dots ((k-1)n-1)}{n^k k!},$$

в яких $s = \lambda_1 + \lambda_2 + \dots + \lambda_k$.

3 Висновки

Таким чином, операції ділення, піднесення до степеня та інші операції з формальними степеневими рядами з ненульовим вільним членом, можна виконувати при допомозі парафункцій трикутних матриць. При цьому загальний член формального ряду, що є результатом таких операцій, виражається через парафункції трикутних матриць, елементи яких виражаються через коефіцієнти вихідних рядів. Позаяк властивості парафункцій трикутних матриць вже достатньо добре вивчені, то метод дає змогу виявити ряд загальних властивостей та комбінаторних тотожностей з формальними степеневими рядами.

Список використаних джерел

1. Заторський Р.А. Про паравизначники та паракерманенти трикутних матриць// Математичні студії.- 2002.- Т.17, №1.- С.3-17.
2. Eulero L.. Institutiones calculi differentialis, 1755 № 41, p. 99-101.
3. Заторский Р.А. Определители треугольных матриц и траектории на диаграммах Ферре// Математические заметки.- 2002.- Т.72, Вып 6 - С. 834-852.

Надійшла до редколегії
15.10.2007

УДК 512.54

Д. І. Морозов

Спряженість автоморфізмів, що задаються лінійними функціями в групі скінченостанових автоморфізмів кореневого сферично однорідного дерева T_2 .

Дана стаття дає відповідь на питання, чи спряжені в $FAutT_2$ лінійні функції, що задають скінченостанові автоморфізми максимального про-порядку.

Ключові слова: централізатор, група автоморфізмів, скінчено-становий автоморфізм, сферично однорідне дерево.

Статтю представив д.ф.-м.н., професор А.П.Петравчук

1. Метою даної роботи є відповідь на питання, чи спряжені в $FAutT_2$ скінченостанові автомати максимального про-порядку, що задаються лінійними функціями. У статті використані наступні позначення:

T_2 - однорідне кореневе дерево валентності 2, координатизоване природнім чином,

$AutT_2$ - група автоморфізмів дерева T_2 ,

e - нейтральний елемент групи $AutT_2$,

$FAutT_2$ - підгрупа скінченостанових автоматних підстановок в $AutT_2$,

Z_2 - поле 2-адичних чисел.

Детальніше опишемо $FAutT_2$.

Дію автоморфізму $\alpha \in AutT_2$ на дереві T_2 можна закодувати наступним чином. Помітимо вершину n -того рівня 1-ю, якщо автоморфізм α переставляє вершини $(n+1)$ -го рівня, суміжні з нею, і 0-м, якщо залишає ці вершини на місці (вершиною n -того рівня називається вершина, що з'єднана з коренем n ребрами). Таке представлення автоморфізму назовемо портретом.

Далі розглядаються тільки піддерева, які ізоморфні T_2 з коренем в певній вершині, що не містять вершин попереднього для кореня рівня. Портрет автоморфізму індукує портрети на всіх таких піддеревах. Далі такі портрети будемо називати станами автоморфізму. Якщо для автоморфізму α маємо скінчену кількість станів, то назовемо α скінченостановим. Множина скінченостанових автоморфізмів утворює групу $FAutT_2$. Автоморфізми, що переводять квазіперіодичні кінці в квазіперіодичні утворюють

D. Morozov

Conjugation of automorphisms that are defined by linear functions in a group of finite state automorphisms of spherically homogeneous rooted tree T_2 .

This paper answers the question, if linear functions that define finite state automorphisms of maximal pro-order are conjugate in $FAutT_2$.

Key Words: centralizer, automorphism group, finite state automorphism, spherically homogeneous tree.

множину $QAutT_2$. Очевидно, $QAutT_2$ є напівгрупою.

Має місце наступне твердження

Теорема 1. $FAutT_2$ є власною підмножиною $QAutT_2$.

Доведення. Дійсно, якщо автоморфізм α переводить квазіперіодичний кінець x в неперіодичний кінець y , то портрет автоморфізму α містить нескінчену кількість станів, які починаються у вершинах кінця x , оскільки в портреті автоморфізму α вершини квазіперіодичного кінця x кодуються неперіодичною послідовністю двійкового 2-адичного розкладу кінця y .

Далі, розглянемо автоморфізм β , який має наступний портрет: всі вершини деякого неперіодичного кінця x кодуються послідовністю двійкового 2-адичного розкладу кінця x , всі інші вершини 0-ми. Очевидно β належить $QAutT_2$, але не є скінченостановим, оскільки містить нескінчену кількість станів, які починаються у вершинах кінця x . Тобто $FAutT_2$ є власною підмножиною $QAutT_2$.

$QAutT_2$ не є групою. Дійсно, обернений автоморфізм до автоморфізму β переводить квазіперіодичний кінець ...00000 в неперіодичний кінець x , тобто β належить $QAutT_2$, а β^{-1} не належить $QAutT_2$. $\square$

Розглянемо праву та ліву вершини 1-го рівня. Вони є коренями двох піддерев, які назовемо відповідно правим та лівим. Запис $\alpha = (\beta, \gamma)$

означає, що автоморфізм α на ліве піддерево діє автоморфізмом β , на ліве піддерево діє автоморфізмом γ , вершини 1-го рівня залишає на місці. Запис $\alpha = (\beta, \gamma) \circ \sigma$ означає, що автоморфізм α на ліве піддерево діє автоморфізмом β , на ліве піддерево діє автоморфізмом γ , вершини 1-го рівня міняє місцями. Автоморфізми β, γ є станами автоморфізму α .

Множення автоморфізмів відбувається наступним чином:

$$(\beta_1, \gamma_1) \circ (\beta_2, \gamma_2) = (\beta_1 \circ \beta_2, \gamma_1 \circ \gamma_2)$$

$$(\beta_1, \gamma_1) \circ \sigma \circ (\beta_2, \gamma_2) = (\beta_1 \circ \gamma_2, \gamma_1 \circ \beta_2) \circ \sigma$$

Поставимо у відповідність автоморфізмам з $AutT_2$ функції $Z_2 \rightarrow Z_2$. Нехай вершини T_2 занумеровано природним чином: для двох вершин $n+1$ -го рівня, що з'єднані з однією вершиною n -го рівня, права вершина помічена 1, ліва 0.

Поставимо у відповідність кінцю x з дерева T_2 2-адичне число $\bar{x}$, двійковий запис якого співпадає з кодуванням x в T_2 . Означимо функцію $f_\gamma : Z_2 \rightarrow Z_2$, що відповідає автоморфізму γ наступним чином $f_\gamma(\bar{x}) = x \circ \gamma$ (γ_f відповідає f). Далі будемо ототожнювати x та $\bar{x}$, f_γ та γ . Добуток автоморфізмів $\alpha \circ \beta$ будемо записувати як суперпозицію функцій $f_\alpha \circ f_\beta$ (першою діє ліва функція).

Наприклад, adding machine, що задається автоматом $\varepsilon = (\varepsilon, \varepsilon) \circ \sigma$ відповідає функції $f_\varepsilon(x) = x + 1$. Централізатор такої функції в $AutT_2$ складається з функцій $f(x) = x + p$, $p \in Z_2$ (див. [3]), тобто з 2-адичних степенів ε , оскільки

$$\varepsilon^p = \gamma_{x+1}^p = \gamma_{x+p} \quad (\forall p \in Z_2)$$

Множину обертовних елементів в кільці K позначимо, як K^* .

Множину $\{p = m/(2n+1) | m, n \in \mathbb{Z}\}$ позначимо, як Z_2^Q . Вона складається з 2-адичних чисел, що мають період в двійковому запису.

Лема 1. $(x+p) \in FAutT_2 \Leftrightarrow p \in Z_2^Q$.

Доведення. $\Rightarrow$ Автоморфізм $x+p : 0 \rightarrow p$. Скінченостановий автоморфізм переводить квазіперіодичні кінці в квазіперіодичні. Оскільки $\dots 000$ - квазіперіодичний, то і p - квазіперіодичний, тобто належить Z_2^Q . p як квазіперіодичне число має наступний двійковий запис:

$$p = \dots a_{k+m} \dots a_{k+1} \dots a_{k+m} \dots a_{k+1} a_k \dots a_2 a_1.$$

Натуральне число $a_{k+m} \dots a_{k+1}$ будемо називати періодом числа p .

p можна представити у вигляді:

$$p = \dots 0 \dots 01 \dots 0 \dots 01 \cdot 2^k \cdot q + a_k \dots a_2 a_1 \quad (1)$$

де період $0 \dots 01$ складається з m цифр, а $q = a_{k+m} \dots a_{k+1}$ - період числа p .

$\Leftarrow$ Нехай $\varepsilon = \gamma_{x+1}$, тобто ε - adding machine. Автоморфізм ε підноситься у 2-адичну степінь наступним чином:

$$\varepsilon^{2^p} = (\varepsilon^p, \varepsilon^p), \varepsilon^{2^{p+1}} = (\varepsilon^p, \varepsilon^{p+1}) \circ \sigma \quad (2)$$

Згідно з (1) $\varepsilon^p = (\varepsilon^{0 \dots 01 \frac{0 \dots 01}{m}})^{q \cdot 2^n} \circ \varepsilon^v$ де $q, m, v, n \in \mathbb{Z}^+$, v - початок довжини n двійкового запису числа p до початку періода q довжини m .

Згідно з (2) автоморфізм $\varepsilon^{0 \dots 01 \frac{0 \dots 01}{m}}$ містить m станів вигляду $(\varepsilon^{0 \dots 01 \frac{0 \dots 01}{m}})^{2^k}, k \in \{0, 1, \dots, m-1\}$, та $m-1$ станів вигляду $(\varepsilon^{0 \dots 01 \frac{0 \dots 01}{m}})^{2^k} \circ \varepsilon, k \in \{1, \dots, m-1\}$, тобто містить $2m-1$ станів. Оскільки $\varepsilon^{0 \dots 01 \frac{0 \dots 01}{m}}$ скінченостановий автоморфізм, то і ε^p - скінченостановий, так як $q, v, n \in \mathbb{Z}^+$. $\square$

Лема 2. Автоморфізм $px \in FAutT_2 \Leftrightarrow p \in (Z_2^Q)^*$.

Доведення. Автоморфізми $f_1(x) = (2k+1)x$ та $f_2(x) = (1/(2k+1))x$ скінченостанові одночасно, оскільки

$$f_1 = f_2^{-1}$$

$$((2m+1)/(2k+1))x = (2m+1)x \circ (1/(2k+1))x \\ -kx = -x \circ kx$$

то достатньо довести, що $-x$ та $(2t+1)x$ - скінченостанові $\forall t \in \mathbb{Z}^+$.

$(2t+1)x$ містить тільки стани вигляду $(2t+1)x + f$, де $f \in \{0, 1, \dots, (2t+1)\}$. Дійсно

$$(2t+1)x = ((2t+1)x, (2t+1)x + t) \\ (2t+1)x + 2f = ((2t+1)x + f, (2t+1)x + (k+f)) \\ (2t+1)x + (2f+1) = \\ = ((2t+1)x + f, (2t+1)x + (t+f+1)) \circ \sigma$$

Якщо $f \leq t$, то $t+f \leq 2k, t+f+1 \leq 2t+1$, тобто автоморфізм $(2t+1)x$ має не більше $2t+2$ станів.

Далі

$$-x = (-x, -x-1)$$

$$-x-1 = (-x-1, -x-1) \circ \sigma$$

тобто $-x$ - автоморфізм з 2-ма станами. $\square$

У рівнянні $\alpha^x = \beta$ розв'язок, який переводить ...000 в ...000 позначимо, як χ_0 .

Лема 3. Нехай ε та α спряжені в $\text{Aut}T_2$. Рівняння $\varepsilon^x = \alpha$ має розв'язок в $\chi \in F\text{Aut}T_2$ тоді і тільки тоді, коли $\exists \chi_0 \in F\text{Aut}T_2$, такий, що $\varepsilon^{\chi_0} = \alpha$ ($\chi_0 : \dots 000 \rightarrow \dots 000$).

Доведення. $\Leftarrow \chi = \chi_0$

$\Rightarrow$ Для розв'язку $\chi : 0 \rightarrow p, \chi \in F\text{Aut}T_2$ обов'язково $p \in Z_2^Q$, оскільки скінченостановий автоморфізм переводить квазіперіодичні кінці в квазіперіодичні. Тоді(за лем.1) ε^{-p} - скінченостановий автоморфізм і $\chi_0 = \varepsilon^{-p} \circ \chi : 0 \rightarrow 0$. $\square$

Лема 4. Всі $f(x) = ax + p, a (p \in (Z_2^Q)^*)$ спряжені з $f(x) = ax + 1$ в $F\text{Aut}T_2$.

Доведення. Для рівняння $\alpha_{ax+1}^x = ax + p, \chi_0 = px$. Дійсно

$$p^{-1}x \circ (ax + 1) \circ px =$$

$$= (a(p^{-1}x) + 1) \circ px = p(a(p^{-1}x) + 1) = ax + p$$

Згідно з лемою 2 $\chi_0 \in F\text{Aut}T_2$. $\square$

Лема 5. $f(x) = p_1x + p_2 \in F\text{Aut}T_2 \Leftrightarrow p_1, p_2 \in Z_2^Q$

Доведення.

$$p_1x + p_2 = p_1x \circ (x + p_2)(1)$$

$$\Leftrightarrow x + p_2 \in F\text{Aut}T_2, p_1x \in F\text{Aut}T_2 \text{ (леми 1,2).}$$

$\Rightarrow p_1x + p_2 : 0 \rightarrow p_2$. Оскільки $p_1x + p_2 \in F\text{Aut}T_2$, то $p_2 \in Z_2^Q$. За лемою 1 $x + p_2 \in F\text{Aut}T_2$. Згідно з (1) $p_1x \in F\text{Aut}T_2$. За лемою 2 $p_1 \in Z_2^Q$. $\square$

Теорема 2. Автоморфізми $f(x) = (4k+1)x + 1 (k \in Z_2)$ -автоморфізми максимального порядку.

Доведення.

$$(4k+1)x = ((4k+1)x, (4k+1)x + 2k) =$$

$$= ((4k+1)x, (4k+1)x) \circ (e, x + 2k)$$

Автоморфізм вигляду (a, a) належить коммутанту в $\text{Aut}T_2$. $x + 2k$, а відповідно і $(e, x + 2k)$ належать коммутанту в $\text{Aut}T_2$.

$$(4k+1)x + 1 = (4k+1)x \circ (x + 1)$$

$x + 1$ - максимального про-порядку, отже і $(4k+1)x + 1$ - максимального про-порядку. $\square$

Означимо $(4k+1)^{\hat{x}}$, як $\chi_0 : 0 \rightarrow 0$ в рівнянні $x + 1^x = (4k+1)x + 1$. Має місце рівність

$$(4k+1)^x = 4k(4k+1)^{\hat{x}} + 1 \quad (*)$$

де $(4k+1)^x$ -функція, що продовжується з $\mathbb{Z}$ на Z_2 по неперервності в топології на Z_2 , оскільки функція $\frac{(4k+1)^x - 1}{4k}$ є автоморфізмом з $\text{Aut}T_2$ і

$$\left(\frac{(4k+1)^x - 1}{4k}\right)^{-1} \circ (x + 1) \circ$$

$$\circ \left(\frac{(4k+1)^x - 1}{4k}\right) = (4k+1)x + 1$$

Звернемо увагу на те, що $(4k+1)^x$ не задається автоморфізмом з $\text{Aut}T_2$.

Теорема 3. Функції $f_1(x) = x + 1$ та $f_2(x) = 5x + 1$ не спряжені в $F\text{Aut}T_2$.

Доведення. Дійсно, згідно з (*) $5^x \in Z_2^q \Leftrightarrow 5^{\hat{x}} \in Z_2^q$. Далі $5^{\hat{x}} : 1/3 \rightarrow 5^{1/3}$. $\sqrt[3]{5}$ не належить $Z_2^q \Rightarrow 5^{1/3}$ не належить $Z_2^q \Rightarrow 5^{\hat{x}} : 1/3 \rightarrow 5^{1/3}$ не належить $F\text{Aut}T_2$. Згідно з лемою 3 $f_1(x) = x + 1$ та $f_2(x) = 5x + 1$ не спряжені в $F\text{Aut}T_2$. $\square$

Теорема 4. Автоморфізми $ax + 1$ та $a^{-1}x + 1 (a = 4k + 1, k \in Z_2^q, a \neq 1)$ не спряжені в $F\text{Aut}T_2$.

Доведення. Множина функцій

$$\{\chi_k(x) = \frac{a^k x + a^{\hat{k}}}{a^{k-1}((a-1)x + 1)} | k \in Z_2\} \quad (1)$$

містить всі розв'язки рівняння по невідомому χ

$$\chi^{-1} \circ (ax + 1) \circ \chi = \frac{1}{a}x + 1 \quad (2)$$

і тільки їх.

Покажемо, що $\chi_k(x)$ є розв'язком рівняння (2).

$$\chi^{-1}(x) = \frac{-a^{k-1}x + a^{\hat{k}}}{a^{k-1}((a-1)x - a)},$$

$$\chi_k^{-1}(x) \circ (ax + 1) \circ \chi_k(x) =$$

$$= \left(\frac{-a^{k-1}x + a^{\hat{k}}}{a^{k-1}((a-1)x - a)}\right) \circ$$

$$\begin{aligned} & \circ(ax+1) \circ \left(\frac{a^k x + a^{\widehat{k}}}{a^{k-1}((a-1)x+1)} \right) = \\ & = \left(a \left(\frac{-a^{k-1}x + a^{\widehat{k}}}{a^{k-1}((a-1)x-a)} \right) + 1 \right) \circ \\ & \circ \left(\frac{a^k x + a^{\widehat{k}}}{a^{k-1}((a-1)x+1)} \right) = \\ & = \left(\frac{-a^{k-2}x + a^{\widehat{k}-1}}{a^{k-2}((a-1)x-a)} \right) \circ \left(\frac{a^k x + a^{\widehat{k}}}{a^{k-1}((a-1)x+1)} \right) = \\ & \frac{(-x-a)/((a-1)x-a)}{-a/((a-1)x-a)} = \frac{-x-a}{-a} = \frac{1}{a}x+1. \end{aligned}$$

Далі, якщо χ є розв'язком рівняння (2) і переводить $\chi: 0 \rightarrow t, t \in Z_2$ (для рівняння (2) існує єдиний χ , такий, що $\chi: 0 \rightarrow t$, для фіксованого $t \in Z_2$. Див. [3]), то $\chi(x) = \chi_k(x)$, де $k = \log_a(a^{k-1}(a-1)t+1)$, тобто інших розв'язків, крім (1), немає.

Автоморфізм $\frac{a^k x + a^{\widehat{k}}}{a^{k-1}((a-1)x+1)}$ не є скінченостановим для $\forall k \in Z_2$.

Дійсно, має місце наступна рівність

$$\frac{a^k x + a^{\widehat{k}}}{a^{k-1}((a-1)x+1)} = (a^k x + a^{\widehat{k}}) \circ \left(\frac{x}{px+q} \right)$$

де $p = (a-1)/a, q = 1/a$. Далі, якщо автоморфізм $\frac{a^k x + a^{\widehat{k}}}{a^{k-1}((a-1)x+1)}$ є скінченостановим, то, за лемою 5, і автоморфізм $a^k x + a^{\widehat{k}}$ є скінченостановим, оскільки $\frac{a^k x + a^{\widehat{k}}}{a^{k-1}((a-1)x+1)}$ переводить кінець ...0000 в квазіперіодичний, за теоремою 1, кінець $a^{\widehat{k}}/a^{k-1}$, а кінці $a^k, a^{\widehat{k}}$ та $a^{\widehat{k}}/a^{k-1}$ є квазіперіодичними одночасно. А автоморфізм $\frac{x}{px+q}$ містить нескінчену кількість станів вигляду

$$\frac{x}{p2^t x + q}, t \in Z^+$$

тобто маємо протиріччя, і автоморфізм $\frac{a^k x + a^{\widehat{k}}}{a^{k-1}((a-1)x+1)}$ - нескінченостановий $\forall k \in Z_2$.

Як наслідок, рівняння (2) не має розв'язків в $FAutT_2$. $\square$

Звернемо увагу на те, що згідно з лемою 4 автоморфізми $a^{-1}x+1$ та $a^{-1}x-a^{-1}$ спряжені в $FAutT_2$. Тобто, згідно з теоремою 4 автоморфізми $ax+1$ та $a^{-1}x-a^{-1}$ не спряжені в $FAutT_2$, хоча і мають однаковий зріст, оскільки $ax+1^{-1} = a^{-1}x-a^{-1}$. Тобто має місце наступне твердження:

Теорема 5. Група $FAutT_2$ є амбівалентною.

Теорема 6. Функції $f_1(x) = (4k_1+1)x+1$ та $f_2(x) = (4k_2+1)x+1$ ($k_1, k_2 \in Z_2^Q$) спряжені в $FAutT_2 \Leftrightarrow 4k_1+1 = 4k_2+1$.

Доведення. $ax+1, bx+1$ ($a = 4k_1+1, b = 4k_2+1, k_1, k_2 \in Z_2^Q, (4k_1+1) \neq (4k_2+1)^{\pm 1}$) не спряжені в $FAutT_2$.

Дійсно, якщо автоморфізми $\alpha, \beta \in FAutT_2$ спряжені в $FAutT_2$, то для всіх $p \in Z_2$ має місце $(\alpha^p \in FAutT_2) \Leftrightarrow (\beta^p \in FAutT_2)$ (згідно з теоремою 4, ця умова не є достатньою). Далі, без обмеження загальності $b \neq a^n$ ($n \in Z, n \neq \pm 1$).

$\log_a((a-1)x+1)$ є біекцією $Z_2 \rightarrow Z_2$, як функція, обернена до $a^{\widehat{x}}$ ($a = 4k+1, k, x \in Z_2$).

$$(ax+1)^t = (a^t x + a^{\widehat{t}}) (t \in Z_2) (**).$$

За теоремою Діріхле про наявність простих чисел в арифметичній прогресії $\exists x \in Z$, таке, що $c = (a-1)x+1$ - просте число, якого немає в розкладі a та b . Тоді (згідно з лемою 5 та (**)) $(ax+1)^{\log_a c} \in FAutT_2$, оскільки $a^{\log_a c} = c$ є раціональним числом, а $(bx+1)^{\log_a c}$ не належить $FAutT_2$, оскільки $b^{\log_a c}$ не є раціональним числом. Додаючи теорему 4 отримаємо результат теореми. $\square$

Висновки

Сформулюємо відповідь на питання поставлене на початку статті.

Скінченостанові автоморфізми максимального про-порядку, що задаються лінійними функціями $ax+c, bx+d$ ($a, b, c, d \in Z_2$) спряжені в $FAutT_2$ тоді і тільки тоді, коли $a = b$.

Список використаних джерел

1. R.I. Grigorchuk, V.V. Nekrashevych, V.I. Sushchansky Automata, dynamical systems and groups. Proc. V.A. Steklov Inst. Math., Vol. 231 (2000), 134-215.
2. V. Nekrashevych Self-similar groups. AMS: Mathematical Surveys and Monographs, Vol. 117, 231 pp., 2005.
3. Боднарчук Ю.В., Морозов Д.І. Розширені 2-адичні числа як централізатори автоморфізмів регулярного кореневого дерева валентності 3.

Надійшла до редколегії 29.10.2007

УДК 512.53

Василь О. Пехтерев, к. ф.-м. н., асистент

Аutomорфізми напівгрупи OT_n

Вивчаються автоморфізми напівгрупи OT_n , яка є єдиним (з точністю до ізоморфізму) $\mathcal{R}$ -зрізом напівгрупи T_n .

Ключові слова: напівгрупи перетворень, зрізи, автоморфізм.

E-mail: vasiliiy@univ.kiev.ua

Статтю представив д.ф.-м.н., професор В.В.Кириченко

Вступ

Нехай $N = \{1, \dots, n\}$. На блоках кожного диз'юнктного розбиття $N = \dot{\cup}_{i=1}^k A_i$ множини N на k непорожніх блоків визначимо лінійний порядок $<$: $A_i < A_j$, якщо $\min(A_i) < \min(A_j)$, де $\min(A_i)$ позначає найменший елемент множини A_i відносно природного лінійного порядку на множині N . Напівгрупа OT_n визначається за таким правилом: для кожного розбиття $N = \dot{\cup}_{i=1}^k A_i$, де $A_1 < A_2 < \dots < A_k$ вона містить елемент a , котрий діє наступним чином: $a(A_i) = i$ для кожного $i = 1, \dots, k$. В роботі [?] показано, що $OT_n \in \mathcal{R}$ -зрізом T_n , тобто містить рівно по одному елементу з кожного $\mathcal{R}$ -класу T_n , а також доведено, що усі $\mathcal{R}$ -зрізи напівгрупи T_n можуть бути отримані з OT_n за допомогою спряження, тобто мають вигляд $\pi^{-1}OT_n\pi$ для деякої підстановки π симетричної групи S_n .

Дана робота присвячена вивченню автоморфізмів OT_n . Ми будемо дотримуватись стандартних позначень (див. [?]). Для кожного $a \in OT_n$ позначимо $\text{im}(a)$ та ρ_a відповідно образ елемента a і таке відношення еквівалентності на множині N , що $i\rho_a j$ тоді і лише тоді, коли $a(i) = a(j)$. Число $\text{rk}(a) = |\text{im}(a)|$ називається рангом перетворення a . Через 0 будемо позначати єдине перетворення множини N , ранг якого дорівнює одиниці, і яке є нулем напівгрупи OT_n . Через e будемо позначати тотожне перетворення множини N — одиницю даної напівгрупи. Легко бачити, що потужність напівгрупи OT_n збігається з кількістю диз'юнктивних розбиттів множини N на непорожні блоки, і дорівнює числу Бела B_n . Якщо a — елемент напівгрупи OT_n рангу k , то $\text{im}(a) = \{1, \dots, k\}$. Через a_i бу-

Vasyl O. Pyekhtyeryev, PhD, Assistant Professor

Automorphisms of the semigroup OT_n

We investigate automorphisms of the semigroup OT_n which is a unique (up to isomorphism) $\mathcal{R}$ -cross-section of T_n .

Key words: transformation semigroups, cross-section, automorphism.

демо позначати найменший елемент множини $a^{-1}(i)$ для кожного $i = 1, \dots, k$. Очевидно, що $a_i \geq i$ для всіх $i = 1, \dots, k$.

Основний результат

Нехай E — множина всіх ідемпотентів напівгрупи OT_n . Ясно, що кожен автоморфізм взаємно-однозначно відображає цю множину на себе.

Лема 1. Для кожного $a \in E$ множина тих $x \in E$, які задовольняють рівняння $xa = x$ має вигляд $M_a = \{x \in E \mid \text{rk}(x) \leq \text{rk}(a)\}$.

Доведення. Нехай $x \in E$ — розв'язок даного рівняння. Тоді маємо, що $\text{rk}(x) = \text{rk}(xa) \leq \text{rk}(a)$ і $x \in M_a$. Позаяк, для кожного $y \in M_a$ має місце включення $\text{im}(y) \subseteq \text{im}(a)$, то $ya = y$, бо кожен елемент $a \in E$ діє тотожно на своєму образі. $\square$

Лема 2. Для довільних автоморфізма f напівгрупи OT_n і елемента $a \in E$ має місце рівність $\text{rk}(f(a)) = \text{rk}(a)$.

Доведення випливає з того, що кожен автоморфізм індукує бієкцію між розв'язками в E рівнянь $xa = x$ та $yf(a) = y$. $\square$

Лема 3. Для кожного $a \in E$ множина розв'язків рівняння $xa = x$ має вигляд $M^a = \{x \in OT_n \mid \text{rk}(x) \leq \text{rk}(a)\}$.

Доведення. Аналогічне доведенню леми 1. $\square$

Лема 4. Для довільних автоморфізма f напівгрупи OT_n і елемента $a \in OT_n$ має місце рівність $\text{rk}(f(a)) = \text{rk}(a)$.

Доведення. Нехай $\text{rk}(a) = k$. Тоді елемент a буде розв'язком рівняння $xb = x$ для деякого

ідемпотента b рангу k . Тоді $f(a)$ буде розв'язком рівняння $yf(b) = y$. Позаяк за лемою 2 $\text{rk}(f(b)) = \text{rk}(b) = k$, то $\text{rk}(f(a)) \leq k = \text{rk}(a)$. Застосувавши аналогічні міркування до автоморфізма f^{-1} отримаємо потрібну рівність. $\square$

Для кожного натурального $k \leq n$ позначимо через E_k та R_k множини $\{x \in E \mid \text{rk}(x) \leq k\}$ і $\{x \in OT_n \mid \text{rk}(x) \leq k\}$ відповідно. З попередніх лем випливає, що кожен автоморфізм напівгрупи OT_n взаємно-однозначно відображає кожен з цих множин на себе.

Оскільки кожен елемент a множини E_{n-1} однозначно характеризується образом $a(n)$ елемента $n \in N$, то надалі через e_k будемо позначати ідемпотент рангу $n-1$, для якого $e_k(n) = k$.

Лема 5. Довільний автоморфізм напівгрупи OT_n індукує бієкцію на множині $\{e_1, e_2\}$, а на множині $\{e_3, \dots, e_{n-1}\}$ діє тотожно.

Доведення. Для кожного $a \in E_2$ знайдемо розв'язки рівняння $xa = a$ на множині E_{n-1} . Позаяк в цьому випадку $\text{im}(a) = \{1, 2\}$, то розбиття ρ_a складається з двох блоків $a^{-1}(1)$ та $a^{-1}(2)$. Очевидно, що e_k є розв'язком рівняння $xa = a$ тоді і лише тоді, коли елементи k і n множини N належать одному блоку ρ_a . Таким чином, кількість розв'язків даного рівняння на множині E_{n-1} залежить від елемента a . Більше того, ця кількість буде набувати свого максимального значення $n-2$ лише для таких двох елементів a_1 і a_2 множини E_2 , що $a_1^{-1}(1) = \{1\}$ і $a_2^{-1}(2) = \{2\}$. Тепер з того, що кожен автоморфізм f напівгрупи OT_n індукує бієкцію між розв'язками рівнянь $xa = a$ і $xf(a) = f(a)$ випливає, що f індукує бієкцію на двоелементній множині $\{a_1, a_2\}$. Звідки маємо, що f^2 діє тотожно на цій множині. Тому f^2 індукує бієкцію на множині розв'язків рівняння $xa_1 = a_1$, тобто на множині $E_{n-1} \setminus \{e_1\}$. Звідки $f^2(e_1) = e_1$. Аналогічно f^2 індукує бієкцію на множині $E_{n-1} \setminus \{e_2\}$ і $f^2(e_2) = e_2$. Припустимо, що $f(e_1) = e_k$, $k \neq 1$. Тоді $f(e_1)a_1 = e_ka_1 = a_1$. Звідки $e_1f(a_1) = f^2(e_1)f(a_1) = f(f(e_1)a_1) = f(a_1)$. Тому $f(a_1) = a_2$, бо e_1 не є розв'язком рівняння $xa_1 = a_1$. Звідси одразу випливає, що $f(e_1)$ не є розв'язком рівняння $xf(a_1) = f(a_1)$, тобто рівняння $xa_2 = a_2$. Звідки остаточно отримуємо, що $f(e_1) = e_2$. Таким чином ми показали, що f індукує бієкцію на двоелементній множині $\{e_1, e_2\}$.

Повторюючи міркування попереднього абзацу для рівнянь $xa = a$, де $a \in E_k$, $k \geq 3$ отри-

маємо, що $f(\{e_1, \dots, e_k\}) = \{e_1, \dots, e_k\}$ для всіх $k \geq 3$. Звідки $f(e_k) = e_k$ для всіх $k \geq 3$. $\square$

Лема 6. Довільний автоморфізм f напівгрупи OT_n повністю визначається своєю дією на множині E_{n-1} .

Доведення. Оскільки розбиття ρ_a кожного елемента $a \in R_{n-1}$ містить лише один двоелементний блок, то надалі a_{ij} будемо позначати той елемент множини R_{n-1} , двоелементний блок якого містить елементи i та j . Нехай π — перестановка на множині $\{1, \dots, n-1\}$, індукована дією автоморфізму f на множині E_{n-1} , тобто $f(e_i) = e_{\pi(i)}$. Очевидно, що $e_ia_{ij} = e_ja_{ij}$ для всіх i, j . Тому $e_{\pi(i)}f(a_{ij}) = e_{\pi(j)}f(a_{ij})$. Звідки $f(a_{ij})(\pi(i)) = f(a_{ij})(\pi(j))$. Оскільки $\text{rk}(f(a_{ij})) = n-1$ за лемою 4, то $f(a_{ij}) = a_{\pi(i)\pi(j)}$. Лишилось зауважити, що одиниця при кожному автоморфізмі переходить у одиницю, а множина R_{n-1} , поповнена одиницею, є незвідною системою твірних напівгрупи OT_n . $\square$

Наслідок. При $n \geq 4$ група автоморфізмів напівгрупи OT_n містить лише тотожний автоморфізм.

Доведення. З попередніх двох лем випливає, що досить розглядати лише дію автоморфізму на множині E_{n-1} , а також, що для такої дії можливі лише два випадки. Тому досить показати, що при $n \geq 4$ кожен автоморфізм f напівгрупи OT_n діє тотожно на множині $\{e_1, e_2\}$. Припустимо, що це не так. Тоді мають місце рівності $f(e_1) = e_2$, $f(e_2) = e_1$, $f(a_1) = a_2$, $f(a_2) = a_1$ і $f(a_{12}) = a_{12}$. Позначимо $c = a_{12}a_1$ і $d = f(a_{12})f(a_1)$. Тоді $f(c) = d$. Це суперечить тому, що f є автоморфізмом, бо блоки розбиттів ρ_c і ρ_d , які містять $n \in N$, мають різну потужність, а тому рівняння $xc = c$ та $xd = d$ мають різну кількість розв'язків на множині E_{n-1} . $\square$

Теорема. $|\text{Aut } OT_n| = 1$, при $n \neq 3$. $|\text{Aut } OT_3| = 2$.

Доведення. Для $n \geq 4$ це випливає з попереднього наслідку, а для $n \leq 3$ перевіряється безпосередньо. $\square$

Список використаних джерел

1. Pyekhtyeryev V. $\mathcal{H}$ - and $\mathcal{R}$ -cross-sections of the full finite semigroup T_n . // J. Algebra and Discrete Mathematics – 2003. – 3 – с. 82-88.
 2. Клиффорд А., Престон Г. Алгебраическая теория полугрупп. – М.: Мир, 1972. – 286 с.
- Надійшла до редколегії 05.09.2007

УДК 517.8

Марк Т. Тарашанський
Екстремальні продовження міри та гомоморфізми

Якщо μ – невід'ємна, скінчена σ – адитивна функція множин, визначена на σ – алгебрі підмножин B непустої множини Ω і λ – її екстремальне продовження на алгебру A , $B \subset A$, тоді існує такий гомоморфізм h алгебри A на алгебру класів μ – еквівалентності B/μ , що $\lambda(A) = \hat{\mu}(h(A))$ для будь-якого $A \in A$, де $\hat{\mu}$ – фактор-міра міри μ .

Ключові слова: алгебра підмножин, міра, екстремальне продовження міри, гомоморфізм.

E-mail: markt@net.lg.ua

Статтю представив д.ф.-м.н., професор В.В.Кириченко

1 Вступ

Нехай Ω – непуста множина, на якій виділена деяка алгебра підмножин A . Усюди далі $B \subset A$ – підалгебра алгебри A й μ – міра (тобто невід'ємна, скінченна скінченно-адитивна функція множин), визначена на алгебрі B . Позначимо через $S_\mu(A, B)$ множину всіх продовжень міри μ на алгебру A і нехай $\text{ex}S_\mu$ – множина його екстремальних точок, які називаються екстремальними продовженнями міри μ . Нехай далі, q_μ – канонічний гомоморфізм алгебри B на алгебру B/μ класів μ – еквівалентності. Позначимо через $\hat{\mu}$ фактор-міру, визначену на B/μ рівностями

$$\hat{\mu}(q_\mu(B)) = \mu(B), \quad B \in B.$$

Відомо, що $S_\mu(A, B) \neq \emptyset$ і якщо алгебра A породжена алгеброю B й деякою непустою підмножиною $A \subset \Omega$, $A \in B$, тоді й $\text{ex}S_\mu \neq \emptyset$ ([1], розділ 3). Останній результат узагальнювався багатьма авторами в різних напрямках. Зокрема, якщо алгебра B/μ повна, тоді $S_\mu(A, B)$ має непусту множину екстремальних точок ([2], теорема 1). Трохи інший підхід був запропонований в [3], де було показано, що $\lambda \in \text{ex}S_\mu$ тоді й тільки тоді, коли для кожного $A \in A$ й будь-якого $\varepsilon > 0$ знайдеться таке $B \in B$, що $\lambda(A \Delta B) < \varepsilon$.

Якщо $h: A \rightarrow B/\mu$ – гомоморфізм, що продовжує гомоморфізм q_μ , тоді рівності

$$\lambda(A) = \hat{\mu}(h(A)), \quad A \in A$$

Mark T. Tarashchanskii
Extreme extensions of measure and homomorphisms

If μ – non-negative, finite σ – additive function of set defined on σ – algebra of sets B of non-empty set Ω and λ – its extreme extension on algebra A , $B \subset A$, then exists such homomorphism h algebras A on algebra of equivalence classes B/μ , that $\lambda(A) = \hat{\mu}(h(A))$ for everyone $A \in A$, where $\hat{\mu}$ – the quotient-measure of measure μ .

Key Words: algebra of subsets, measure, extreme extension of a measure, homomorphism.

визначають деяке продовження міри μ на алгебру A . Позначимо через $H_\mu(A, B)$ множину всіх таких продовжень. З ([3], теорема 1) випливає, що $H_\mu(A, B) \subset \text{ex}S_\mu$, але ці множини не збігаються, як показує наступний приклад.

Приклад. Нехай Ω – злічена множина, на якій виділена деяка алгебра, але не σ – алгебра підмножин B , що містить всі одноточкові підмножини Ω , наприклад, алгебра, породжена всіма скінченими підмножинами Ω . Покладемо $\mu(\{\omega_n\}) = 2^{-n}$, $n = 1, 2, \dots$ для $\omega_n \in \Omega$. Таким чином, визначена міра на B . Нехай, далі, $C \subset \Omega$, $C \in B$. Визначимо A як алгебру, породжену B й $\tilde{N}$, і покладемо

$$\lambda(A) = \mu_*(A \cap C) + \mu^*(A \cap C')$$

для кожного $A \in A$.

Тоді λ – строго позитивна міра на A й екстремальне продовження міри μ ([3], приклад 1). Ясно, що не існує такого гомоморфізму $h: A \rightarrow B$, що $\lambda(A) = \mu(h(A))$, $A \in A$. Більше того, у цьому прикладі взагалі не існує жодного гомоморфізму $h: A \rightarrow B$, що залишає нерухомими елементи з алгебри B , тобто $H_\mu(A, B) = \emptyset$. $\square$

Проте, якщо алгебра B є σ – алгеброю, а міра μ σ – адитивна, тоді буде виконуватися рівність $H_\mu(A, B) = \text{ex}S_\mu$. Доказу цього твердження і його додатків присвячена ця замітка.

2 Екстремальні продовження

Для $A, B \in A$ символом $A \Delta B$ буде позначатися симетрична різниця елементів $A, B \in A$,

тобто $A \Delta B = (A \cap \bar{B}) \cup (\bar{A} \cap B)$, де $\bar{A}$ - доповнення елемента A в алгебрі A .

Нехай $B \subset A$ - підалгебра алгебри A . Тоді алгебра B^* , породжена алгеброю B й елементом $A \in A$, $A \notin B$ складається із всіх елементів $A^* \in A$, які мають вигляд

$$A^* = (B_1 \cap A) \cup (B_2 \cap \bar{A}), \quad B_1, B_2 \in B,$$

що випливає з того, що об'єднання й доповнення елементів такого вигляду є елементами того ж вигляду. Звідси отримаємо, що якщо алгебра $B \in \sigma$ -алгеброю, то така ж і алгебра B^* .

Нехай μ - міра, визначена на алгебрі B й λ - деяке її продовження на алгебру B^* . Якщо $B \in \sigma$ -алгеброю, а міра μ σ -адитивна, тоді й міра λ також σ -адитивна ([4], лема 4).

Теорема 1. Нехай $B \in \sigma$ -алгеброю, міра μ σ -адитивна й $\lambda \in S_\mu(A, B)$. Тоді наступні умови еквівалентні:

$$(1) \lambda \in \text{ex} S_\mu(A, B);$$

(2) для кожного $A \in A$ існує таке $B \in B$, що $\lambda(A \Delta B) = 0$;

(3) існує такий гомоморфізм $h: A \rightarrow B/\mu$, що $B \in h(B)$ для будь-якого $B \in B$ й $\lambda(A) = \hat{\mu}(h(A))$ для кожного $A \in A$.

Доведення. (1) $\Rightarrow$ (2). Як наслідок з ([3], теорема 1), для кожного $A \in A$ існує таке $B_n \in B$, що

$$\lambda(A \Delta B_n) < 2^{-n}, \quad n \in \mathbb{N}.$$

Визначимо $B'_n = \bigcap_{k=n}^\infty B_k$, $n \in \mathbb{N}$ і $B = \bigcup_{n \in \mathbb{N}} B'_n$. Оскільки алгебра $B \in \sigma$ -повною, то $B, B'_n \in B$, $n \in \mathbb{N}$. Маємо далі,

$$\lambda(A \Delta B'_n) \leq \lambda\left(\bigcup_{k=n}^\infty (A \cap \bar{B}_k)\right) + \lambda\left(\bigcap_{k=n}^\infty (B_k \cap \bar{A})\right).$$

Звуження міри λ на підалгебру B^* , породжену алгеброю B й елементом $A \in A$, є продовженням σ -міри μ на σ -алгебру B^* . Тому, як уже вказувалося, ця міра σ -адитивна. Тому для першого доданка в правій частині цієї нерівності, одержуємо

$$\begin{aligned} \lambda\left(\bigcup_{k=n}^\infty (A \cap \bar{B}_k)\right) &\leq \sum_{k=n}^\infty \lambda(A \cap \bar{B}_k) \leq \\ &\leq \sum_{k=n}^\infty \lambda(A \Delta B_k) \leq 2^{-n+1}. \end{aligned}$$

Оскільки

$$\lambda\left(\bigcap_{k=n}^\infty (B_k \cap \bar{A})\right) \leq \lambda(B_k \cap \bar{A}) \leq \lambda(B_k \Delta A) \leq 2^{-k}$$

для всіх $k = n, n+1, \dots$, то

$$\lambda\left(\bigcap_{k=n}^\infty (B_k \cap \bar{A})\right) = 0.$$

Звідси випливає, що $\lambda(A \Delta B'_n) \leq 2^{-n+1}$. Покажемо тепер, що $\lambda(A \Delta B) = 0$. Маємо

$$\lambda(A \Delta B) \leq \lambda(A \cap \bar{B}) + \lambda(B \cap \bar{A}).$$

По визначенню множини B , маємо

$$A \cap \bar{B} \subset \bigcup_{k=n}^\infty (A \cap \bar{B}_k)$$

для кожного $n \in \mathbb{N}$. Звідси

$$\begin{aligned} \lambda(A \cap \bar{B}) &\leq \lambda\left(\bigcup_{k=n}^\infty (A \cap \bar{B}_k)\right) \leq \sum_{k=n}^\infty \lambda(A \cap \bar{B}_k) \leq \\ &\leq \sum_{k=n}^\infty \lambda(A \Delta B_k) \leq 2^{-n+1} \end{aligned}$$

для кожного $n \in \mathbb{N}$, тобто $\lambda(A \cap \bar{B}) = 0$. Далі,

$$B \cap \bar{A} = \bigcup_{n=1}^\infty \bigcap_{k=n}^\infty (B_k \cap \bar{A}), \text{ послідовність}$$

$$C_n = \bigcap_{k=n}^\infty (B_k \cap \bar{A})$$

монотонно зростає й, як уже було показано, $\lambda(C_n) = 0$ для всіх $n \in \mathbb{N}$. Тому

$$\lambda(B \cap \bar{A}) = \lim_{n \rightarrow \infty} \lambda(C_n) = 0.$$

Звідси остаточно одержуємо $\lambda(A \Delta B) = 0$, що доводить необхідну імплікацію.

(2) $\Rightarrow$ (3). Для кожного $A \in A$ покладемо $h(A) = q_\mu(B)$, де множина $B \in B$ задовольняє умові $\lambda(A \Delta B) = 0$. Для перевірки коректності цього визначення необхідно показати, що елемент, який знаходиться в правій частині, не залежить від вибору множини $B \in B$. Інакше кажучи, необхідно показати, що якщо множини $B_1, B_2 \in B$ задовольняють умовам $\lambda(A \Delta B_i) = 0$, $i = 1, 2$, то $\mu(B_1 \Delta B_2) = 0$. Дійсно, нехай $\lambda(A \Delta B_i) = 0$, $i = 1, 2$. Справедливі співвідношення

$$\begin{aligned} (B_1 \cap \bar{A}) \cup (\bar{B}_1 \cap A) &\supseteq \\ &\supseteq (\bar{B}_2 \cap B_1 \cap \bar{A}) \cup (B_2 \cap \bar{B}_1 \cap A), \\ (B_2 \cap \bar{A}) \cup (\bar{B}_2 \cap A) &\supseteq \\ &\supseteq (B_2 \cap \bar{B}_1 \cap \bar{A}) \cup (\bar{B}_2 \cap B_1 \cap A). \end{aligned}$$

Поєднуючи ці співвідношення, одержуємо $(A \Delta B_1) \cup (A \Delta B_2) \supseteq (B_1 \Delta B_2)$. Звідси

$$\mu(B_1 \Delta B_2) = \lambda(B_1 \Delta B_2) \leq \lambda(A \Delta B_1) + \lambda(A \Delta B_2) = 0.$$

Таким чином, рівність $h(A) = q_\mu(B)$ однозначно визначає відображення алгебри A в алгебру B/μ , причому $h(B) = q_\mu(B)$ для $B \in B$. Зали-

шилося показати, що це відображення є гомоморфізмом.

Нехай $A \in \mathcal{A}$ і $B \in \mathcal{B}$ такі, що $\lambda(A \Delta B) = 0$. Тоді $\overline{A \Delta B} = \overline{A \Delta B}$ й, отже, $h(\overline{A}) = \overline{h(A)}$. Нехай, далі, $A_1, A_2 \in \mathcal{A}$ і $B_1, B_2 \in \mathcal{B}$ такі, що $\lambda(A_1 \Delta B_1) = 0$ й $\lambda(A_2 \Delta B_2) = 0$. Тоді

$$(A_1 \cup A_2) \Delta (B_1 \cup B_2) = ((A_1 \cup A_2) \cap (\overline{B_1} \cap \overline{B_2})) \cup ((B_1 \cup B_2) \cap (\overline{A_1} \cap \overline{A_2})).$$

Оскільки

$$(A_1 \cup A_2) \cap (\overline{B_1} \cap \overline{B_2}) \subseteq (A_1 \cap \overline{B_1}) \cup (A_2 \cap \overline{B_2}),$$

$$(B_1 \cup B_2) \cap (\overline{A_1} \cap \overline{A_2}) \subseteq (B_1 \cap \overline{A_1}) \cup (B_2 \cap \overline{A_2}),$$

тоді $(A_1 \cup A_2) \Delta (B_1 \cup B_2) \subseteq (A_1 \Delta B_1) \cup (A_2 \Delta B_2)$ й, тому,

$$\lambda((A_1 \cup A_2) \Delta (B_1 \cup B_2)) \leq \lambda(A_1 \Delta B_1) + \lambda(A_2 \Delta B_2) = 0.$$

З цього можна зробити висновок, що $h(A_1 \cup A_2) = h(A_1) \cup h(A_2)$.

(3) $\Rightarrow$ (1). Ця імплікація безпосередньо випливає з ([3], теорема 1) оскільки для кожного $A \in \mathcal{A}$ множина $B \in \mathcal{B}$ задовольняє співвідношенню $\lambda(A \Delta B) = 0$. $\square$

Розглянемо тепер один додаток попередньої теореми до зображення екстремальних продовжень міри на добуток алгебр. Нехай $(\Omega, \mathcal{A})$ і $(\Theta, \mathcal{B})$ – дві непустих множини з виділеними алгебрами підмножин; π_Ω і π_Θ – канонічні проекції декартова добутку $\Omega \times \Theta$ на Ω й Θ , відповідно; $\mathcal{A} \otimes \mathcal{B}$ – булів добуток алгебр $\mathcal{A}$ і $\mathcal{B}$, визначений як мінімальна алгебра на $\Omega \times \Theta$, що містить $\pi_\Omega^{-1}(A)$ і $\pi_\Theta^{-1}(B)$ (загальне визначення бульового добутку булевих алгебр див. в [5], стор. 66). Нехай, далі, μ – міра на алгебрі $\mathcal{A}$ й ν – міра на алгебрі $\mathcal{B}$. Існує й притім єдина міра $\mu \otimes \nu$, визначена на алгебрі $\mathcal{A} \otimes \mathcal{B}$ й задовольняюча для всіх $A \in \mathcal{A}$, $B \in \mathcal{B}$ співвідношенню

$$(\mu \otimes \nu)(\pi_\Omega^{-1}(A) \cap \pi_\Theta^{-1}(B)) = \mu(A) \cdot \nu(B).$$

Ототожнюючи алгебри $\mathcal{A}$ й $\mathcal{B}$ з їхніми ізоморфними образами в алгебрі $\mathcal{A} \otimes \mathcal{B}$, одержуємо, що для будь-якої міри ν на алгебрі $\mathcal{B}$, міра $\mu \otimes \nu$ є продовженням міри μ , але не всі продовження міри μ на $\mathcal{A} \otimes \mathcal{B}$ вичерпуються мірами такого вигляду як показує наступний простий приклад.

Приклад. Нехай $\Omega = \Theta = \{a, b\}$, $\mathcal{A} = \mathcal{B}$ і є алгебрами всіх підмножин, і $\mu(\{a\}) = \mu(\{b\}) = \frac{1}{2}$. Визначимо міру λ на алгебрі $\mathcal{A} \otimes \mathcal{B}$, поклавши

$$\lambda(\{a, a\}) = \lambda(\{b, b\}) = \frac{1}{2} \alpha,$$

$$\lambda(\{a, b\}) = \lambda(\{b, a\}) = \frac{1}{2}(1 - \alpha), \quad \alpha \in (0, 1).$$

Тоді λ продовжує μ , але не існує такої міри ν на $\mathcal{B}$, щоб $\lambda = \mu \otimes \nu$. $\square$

Нехай ν – двозначна (тобто приймаюча значення 0 і 1) міра на $\mathcal{B}$, $\lambda = \mu \otimes \nu$ і $I = \nu^{-1}(0)$. Для всякого $A \in \mathcal{A}$ покладемо

$$h(\pi_\Omega^{-1}(A) \cap \pi_\Theta^{-1}(B)) = \begin{cases} \emptyset, & B \in I; \\ q_\mu(A), & B \notin I. \end{cases}$$

Це відображення може бути продовжене до гомоморфізму $h: \mathcal{A} \otimes \mathcal{B} \rightarrow \mathcal{A}/\mu$ (див., наприклад, [5], теорема 12.2). Тоді для всякого $C \in \mathcal{A} \otimes \mathcal{B}$ справедливо співвідношення $\lambda(C) = \hat{\mu}(h(C))$. Відповідно до попередньої теореми це означає, що $\lambda \in \text{ex}S_\mu(\mathcal{A} \otimes \mathcal{B}, \mathcal{A})$.

Це відображення може бути продовжене до гомоморфізму $h: \mathcal{A} \otimes \mathcal{B} \rightarrow \mathcal{A}/\mu$ (див., наприклад, [5], теорема 12.2). Тоді для всякого $C \in \mathcal{A} \otimes \mathcal{B}$ справедливо співвідношення $\lambda(C) = \hat{\mu}(h(C))$. Відповідно до попередньої теореми це означає, що $\lambda \in \text{ex}S_\mu(\mathcal{A} \otimes \mathcal{B}, \mathcal{A})$.

Природно виникає питання чи зображується кожне екстремальне продовження міри μ у вигляді $\mu \otimes \nu$ для деякої двозначної міри ν на алгебрі $\mathcal{B}$?

Теорема 2. Нехай $\mathcal{A}$ є σ -алгеброю, міра μ σ -адитивна й $\lambda \in \text{ex}S_\mu(\mathcal{A} \otimes \mathcal{B}, \mathcal{A})$. Тоді існує така двозначна міра ν на алгебрі $\mathcal{B}$, що $\lambda = \mu \otimes \nu$.

Доведення. Будемо ототожнювати алгебри $\mathcal{A}$ й $\mathcal{B}$ з їхніми ізоморфними образами в $\mathcal{A} \otimes \mathcal{B}$. Оскільки міра λ є екстремальним продовженням, то в силу теореми 1 існує такий гомоморфізм $h: \mathcal{A} \otimes \mathcal{B} \rightarrow \mathcal{A}/\mu$, що $\lambda(C) = \hat{\mu}(h(C))$ для кожного $C \in \mathcal{A} \otimes \mathcal{B}$. Тому $\lambda^{-1}(0) = h^{-1}(\emptyset)$ і $h^{-1}(\emptyset) \cap \mathcal{A} = \mu^{-1}(0)$. Покладемо $I_0 = \lambda^{-1}(0) \cap \mathcal{B}$ й нехай I – деякий максимальний ідеал алгебри $\mathcal{B}$, такий що $I_0 \subset I$. Визначимо двозначну міру ν на алгебрі $\mathcal{B}$, що відповідає цьому ідеалу й покладемо $\lambda^* = \mu \otimes \nu$.

При такому визначенні міри λ^* справедливе відношення $\lambda^{-1}(0) \subset \lambda^{*-1}(0)$. Дійсно, оскільки $A \cap B \neq \emptyset$ для довільних $A \in \mathcal{A}$ і $B \in \mathcal{B}$, тоді

$h(A \cap B) = h(A) \cap h(B) = \emptyset$ тільки якщо або $h(A) = \emptyset$, або $h(B) = \emptyset$. Тому $\lambda(A \cap B) = 0$ для деяких $A \in \mathcal{A}$ і $B \in \mathcal{B}$, тільки тоді, коли або $\lambda(A) = 0$, або $\lambda(B) = 0$. Якщо $\lambda(A) = 0$, тоді й $\mu(A) = 0$. Тому $\lambda^*(A \cap B) = \mu(A)\nu(B) = 0$. Якщо $\lambda(B) = 0$, то в силу визначення міри ν , маємо $\nu(B) = 0$. Тому й у цьому випадку одержуємо $\lambda^*(A \cap B) = \mu(A)\nu(B) = 0$.

Покажемо, що справедливо й зворотне включення. Нехай $C \in \mathcal{A} \otimes \mathcal{B}$ і $\lambda^*(C) = 0$. Відповідно до попередньої теореми знайдеться таке $A \in \mathcal{A}$, що $\lambda(A \Delta C) = 0$, а тому, у силу попереднього міркування, і $\lambda^*(A \Delta C) = 0$. Тоді з рівності $0 = \lambda^*(A \Delta C) = \lambda^*(A \cap \bar{C}) + \lambda^*(\bar{A} \cap C)$ й співвідношення $\lambda^*(\bar{A} \cap C) \leq \lambda^*(C) = 0$ випливає, що $\lambda^*(A \cap \bar{C}) = 0$. Тепер з рівності $\lambda^*(A) = \lambda^*(A \cap C) + \lambda^*(A \cap \bar{C})$ й співвідношення $\lambda^*(A \cap C) \leq \lambda^*(C) = 0$ отримуємо, що $\lambda^*(A) = 0$.

Оскільки $\lambda^*(A) = \lambda(A) = \mu(A)$, тоді маємо $\lambda(A \cap C) \leq \lambda(A) = 0$. З рівності

$$0 = \lambda(A \Delta C) = \lambda(A \cap \bar{C}) + \lambda(\bar{A} \cap C)$$

одержуємо $\lambda(\bar{A} \cap C) = 0$. Тому

$$\lambda(C) = \lambda(A \cap C) + \lambda(\bar{A} \cap C) = 0.$$

Тобто $\Gamma_0 = \Gamma$.

Це означає, що звуження міри λ на алгебру $\mathcal{B}$ збігається з мірою ν й тому міри λ й λ^* збігаються на підалгебрах $\mathcal{A}$ і $\mathcal{B}$. Залишилося показати, що $\lambda = \lambda^*$.

Нехай $A \in \mathcal{A}$ і $B \in \mathcal{B}$. Якщо $\nu(B) = 0$, тоді $\lambda(A \cap B) \leq \lambda(B) = \nu(B) = 0$. Тому у цьому випадку маємо

$$\lambda(A \cap B) = \mu(A) \cdot \nu(B).$$

Якщо $\nu(B) = 1$, тоді

$$\mu(A) = \lambda(A) = \lambda(A \cap B) + \lambda(A \cap \bar{B}) = \lambda(A \cap B).$$

Таким чином, і в цьому випадку одержуємо, що $\lambda(A \cap B) = \mu(A) \cdot \nu(B)$. Отже, $\lambda = \lambda^*$. $\square$

3 Висновки

У роботі розглянуто множину $\text{ex}S_\mu(\mathcal{A}, \mathcal{B})$ екстремальних продовжень міри μ , визначеній на підалгебрі $\mathcal{B} \subset \mathcal{A}$, на алгебру $\mathcal{A}$. Доведено, що якщо міра μ σ -адитивна, а алгебра $\mathcal{B}$ є σ -алгеброю, то $\lambda \in \text{ex}S_\mu(\mathcal{A}, \mathcal{B})$ тоді і лише тоді, коли існує такий гомоморфізм $h: \mathcal{A} \rightarrow \mathcal{B}/\mu$, що $\lambda(A) = \hat{\mu}(h(A))$, де $\hat{\mu}$ – фактор-міра міри μ .

Список використаних джерел

1. Lipecki Z. On compactness and extreme points of some sets of quasi-measures and measures //Manuscr. Math. – 1995 – 86 – P.349-365
2. Lipecki Z. Extensions of positive operators and extreme points //Colloq. Math. – 1979 – 42 – P.285-289
3. Plachky D., Extremal and monogenic additive set functions //Proc. Amer. Math. Soc. – 1976 – 54 – P.193-196
4. Lipecki Z. On unique extensions of positive additive set functions //Arch. Math. (Basel) – 1983 – 41 – P. 71-79
5. Сикорский Р. Булевы алгебры. – М.: Мир - 1969 – 376с.

Надійшла до редколегії 10.11.2007

**ДИФЕРЕНЦІАЛЬНІ
РІВНЯННЯ,
МАТЕМАТИЧНА ФІЗИКА
ТА МЕХАНІКА**

УДК 531.36+629.78.001+62-755

Олег О. Горошко, д.ф.-м.н., професор
Ірина І. Філімоніхіна, асистент

**Достатні умови усунення
автобалансирами кута нутації
незрівноваженого обертового тіла в
ізолюваній системі**

За допомогою теореми Рауса одержані достатні умови усунення пасивними автобалансирами кута нутації незрівноваженого обертового тіла в ізолюваній системі, які із точністю до меж співпадають із необхідними. Умови не залежать від типу автобалансира і тому є узагальненими.

Ключові слова: обертове тіло, кут нутації, демпфер, усталений рух, стійкість.

*E-mail: fii@online.ua

Статтю представив проф. Мелешко В.В.

Вступ.

В роботах [1-4] визначалися умови, за яких пасивні автобалансири (АБ) зможуть усунути кут нутації незрівноваженого обертового несучого тіла, що входить до складу ізолюваної системи. Ця задача актуальна для штучних супутників Землі, космічних апаратів, положення яких у просторі стабілізується обертанням [5,6]. Слід відзначити, що АБ знайшли застосування у техніці для зрівноваження на ходу швидкообертових роторів – дисків ручних шліфувальних машин, барабанів побутових пральних машин, центрифуг, екстракторів, сепараторів, дисків в CD/DVD приводах тощо [7]. Застосування АБ у якості демпферів кута нутації це – єдиний пасивний метод усунення кута нутації обертових супутників, викликаного як неточним наданням супутнику початкового обертання, так і його незрівноваженістю відносно власної повздовжньої осі. В роботах [1-3] із застосуванням енергетичного підходу і емпіричного методу були знайдені узагальнені необхідні умови усунення АБ кута нутації, придатні для АБ будь-якого типу. В даній роботі, із застосуванням теореми Рауса, одержуються аналогічні достатні умови. При цьому також одержуються узагальнені рівняння усталених рухів вказаних вище систем, зокрема для випадку, коли система має елементи, здатні накопичувати потенціальну енергію,

Oleg O.Goroshko, Dt.Sci (Phys.-Math.), professor
I.I. Filimonikhina, assistant professor

**Sufficient conditions of removal by the
autobalancers of corner of notation of
unbalanced rotating body in the isolated
system**

By means of the Routh theorem are received the sufficient conditions of removal by passive autobalancers of corner of notation of unbalance spinning body in the isolated system, which coincide with necessary conditions within borders. Conditions do not depend on the type of autobalancer and so is generalized.

Key Words: rotating body, cone of nutation, damper, steady-state motion, stability.

обґрунтовується методика оцінки стійкості різних усталених рухів.

§1. Опис системи, основні допущення

Ізолювана матеріальна система складається з несучого обертового тіла (рис. 1, а) і приєднаних тіл, які можуть рухатися відносно несучого тіла. Відносному руху тіл, зокрема, перешкоджають сили в'язкого опору. В системі можуть бути елементи, здатні накопичувати потенціальну енергію.

Оскільки система ізолювана, то її центр мас – точка G – рухається рівномірно і прямолінійно. Не обмежуючи загальності, вважаємо, що він нерухомий. Також має місце закон збереження моменту кількості руху системи $K_G = \text{const}$. Будемо уявляти рух системи як складний [3]. За переносний рух вважатимемо обертання системи навколо центра мас із кутовою швидкістю обертання несучого тіла, а за відносний – рух елементів системи відносно її центра мас. Обертальну складову переміщень будемо задавати за допомогою кутів Ейлера-Крилова [8] (які ще називають кутами Кардана-Брайнта [9]). Введемо в розглядання осі $Gx_G y_G z_G$, в яких вісь z_G направлена уздовж вектора K_G . Нехай вони обертаються навколо осі z_G з кутовою швидкістю ω , яка відповідає першому повороту несучого тіла в просторі.

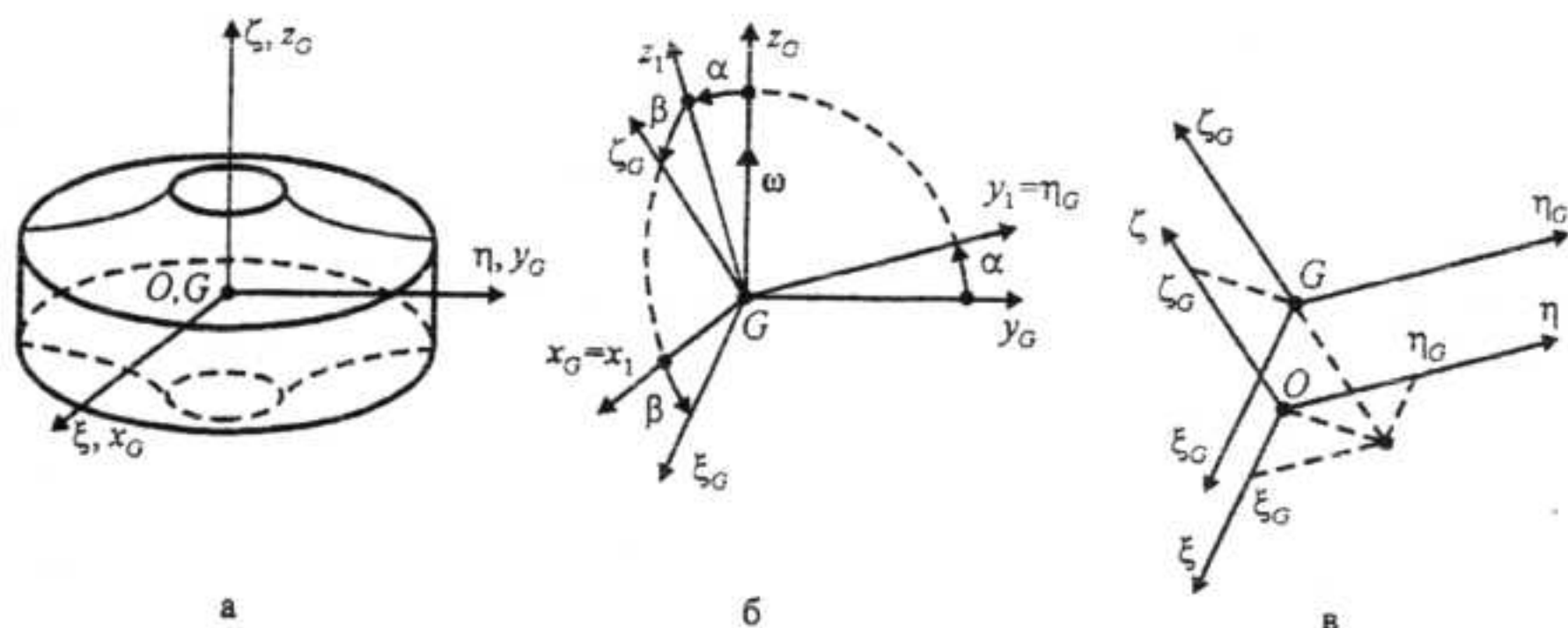


Рис. 1. Визначення руху несучого тіла

В процесі подальшого руху осі $Gx_Gy_Gz_G$ переходять в осі $O\xi\eta\zeta$, які визначають кінцеве положення несучого тіла, таким чином. Спочатку осі $Gx_Gy_Gz_G$ повертаються на кути Ейлера - Кривої α, β , як це показано на рис.1, б, внаслідок чого переходять в осі $G\xi_G\eta_G\zeta_G$. Потім осі $G\xi_G\eta_G\zeta_G$ переміщуються поступально на $-\xi_G, -\eta_G, -\zeta_G$, внаслідок чого переходять в осі $O\xi\eta\zeta$, як це показано на рис.1, в. Очевидно, що ξ_G, η_G, ζ_G - координати центра мас системи щодо осей $O\xi\eta\zeta$. Оскільки на систему діють дисипативні сили і під час перехідних процесів до системи не підводиться зовнішня і внутрішня енергія, то з часом система прийде до усталеного руху, у якому відносний рух приєднаних тіл припиниться і система буде обертатися як одне жорстке ціле навколо осі z_G .

Тензор інерції системи утворений несучим тілом і приєднаними тілами. Відносно осей $O\xi\eta\zeta$ його можна представити у вигляді

$$\mathbf{J}_O = \begin{pmatrix} J_\xi & -J_{\xi\eta} & -J_{\xi\zeta} \\ -J_{\xi\eta} & J_\eta & -J_{\eta\zeta} \\ -J_{\xi\zeta} & -J_{\eta\zeta} & J_\zeta \end{pmatrix}, \quad (1.1)$$

де відцентрові моменти інерції системи $J_{\xi\zeta}, J_{\eta\zeta}$ характеризують моментну незрівноваженість системи.

Тензор інерції системи щодо центральних осей системи $G\xi_G\eta_G\zeta_G$:

$$\mathbf{J}_G = \mathbf{J}_O - M_\Sigma \begin{pmatrix} (\eta_G^2 + \zeta_G^2) & -\xi_G\eta_G & -\xi_G\zeta_G \\ -\xi_G\eta_G & (\xi_G^2 + \zeta_G^2) & -\eta_G\zeta_G \\ -\xi_G\zeta_G & -\eta_G\zeta_G & (\xi_G^2 + \eta_G^2) \end{pmatrix}. \quad (1.2)$$

де M_Σ - маса всієї системи. Зауважимо, що координати ξ_G, η_G характеризують статичну незрівноваженість системи.

Нехай $\varphi_1, \dots, \varphi_n$ - узагальнені координати, які задають відносні положення приєднаних тіл відносно несучого тіла. Тоді

$$\begin{aligned} \xi_G &= \xi_G(\varphi_1, \dots, \varphi_n), & \eta_G &= \eta_G(\varphi_1, \dots, \varphi_n), \\ J_{\xi\zeta} &= J_{\xi\zeta}(\varphi_1, \dots, \varphi_n), & J_{\eta\zeta} &= J_{\eta\zeta}(\varphi_1, \dots, \varphi_n), \\ \mathbf{J}_G &= \mathbf{J}_G(\varphi_1, \dots, \varphi_n). \end{aligned} \quad (1.3)$$

Рух системи цілком визначають такі незалежні параметри

$$\alpha, \beta, \omega, \varphi_1, \dots, \varphi_n. \quad (1.4)$$

На усталених рухах системи вони - сталі. Параметр ω на цих рухах визначає кутову швидкість обертання системи навколо осі z_G .

§2. Диференціальні рівняння і рівняння усталених рухів системи

Кінетична і потенціальна енергії системи є функціями вигляду

$$\begin{aligned} T &= T(\omega, \alpha, \beta, \varphi_1, \dots, \varphi_n, \dot{\alpha}, \dot{\beta}, \dot{\varphi}_1, \dots, \dot{\varphi}_n), \\ \Pi &= \Pi(\varphi_1, \dots, \varphi_n). \end{aligned} \quad (2.1)$$

Тоді рівняння Лагранжа II роду дають такі диференціальні рівняння руху системи

$$\begin{aligned} \frac{d}{dt} \frac{\partial T}{\partial \omega} &= 0, & \frac{d}{dt} \frac{\partial T}{\partial \dot{\alpha}} - \frac{\partial T}{\partial \alpha} &= 0, & \frac{d}{dt} \frac{\partial T}{\partial \dot{\beta}} - \frac{\partial T}{\partial \beta} &= 0, \\ \frac{d}{dt} \frac{\partial T}{\partial \dot{\varphi}_j} - \frac{\partial T}{\partial \varphi_j} &= \frac{\partial \Pi}{\partial \varphi_j} + Q_j, & j &= \overline{1, n}, \end{aligned} \quad (2.2)$$

де Q_j - узагальнена сила в'язкого опору, яка діє на приєднане тіло j . Вона стає рівною нулю, коли відносний рух тіла припиняється ($\dot{\varphi}_j = 0$). Узагальнена координата γ , що відповідає кутовій

швидкості $\omega = \dot{\gamma}$, є циклічною координатою. Тому застосуємо до визначення умов стійкості усталеного руху теорему Рауса [10].

З рис. 1, а визначаємо вектор кутової швидкості обертання системи, знайдений в проекціях на осі $G\xi_G\eta_G\zeta_G$:

$$\Omega = \begin{bmatrix} \Omega_x \\ \Omega_y \\ \Omega_z \end{bmatrix} = \begin{bmatrix} \dot{\alpha} \cos \beta - \omega \cos \alpha \sin \beta \\ \dot{\beta} + \omega \sin \alpha \\ \dot{\alpha} \sin \beta + \omega \cos \alpha \cos \beta \end{bmatrix}. \quad (2.3)$$

Введемо у розглядання одиничні вектори $\mathbf{x}_G, \mathbf{y}_G, \mathbf{z}_G$, направлені по осях x_G, y_G, z_G . В проекціях на осі $G\xi_G\eta_G\zeta_G$:

$$\begin{aligned} \mathbf{x}_G &= (\cos \beta, 0, \sin \beta)^T, \\ \mathbf{y}_G &= (\sin \alpha \sin \beta, \cos \alpha, -\sin \alpha \cos \beta)^T, \\ \mathbf{z}_G &= (-\cos \alpha \sin \beta, \sin \alpha, \cos \alpha \cos \beta)^T. \end{aligned} \quad (2.4)$$

З (2.3) знаходимо, що

$$\Omega = \omega \mathbf{z}_G + \dot{\alpha} \mathbf{x}_G + \dot{\beta} (\mathbf{y}_G \cos \alpha + \mathbf{z}_G \sin \alpha). \quad (2.5)$$

Подамо рух системи як складний. За переносний рух приймемо її обертання разом з осями $G\xi_G\eta_G\zeta_G$ навколо центра мас із кутовою швидкістю Ω . За відносний рух приймемо рух системи відносно осей $G\xi_G\eta_G\zeta_G$. Тоді кінетичну енергію системи можна подати як суму трьох складових

$$T = T_e + T_r + T_{er}, \quad (2.6)$$

де T_e - кінетична енергія переносного руху, яка не залежить від відносних швидкостей, T_r - кінетична енергія відносного руху, яка залежить від квадратів відносних швидкостей, T_{er} - додаткові складові, які лінійні щодо відносних швидкостей.

Кінетична енергія переносного руху визначається, як

$$\begin{aligned} T_e &= \frac{1}{2} \Omega^T \mathbf{J}_G \Omega = \frac{1}{2} J_{z_G} \omega^2 - \omega [J_{x_G z_G} \dot{\alpha} + \\ &+ \dot{\beta} (J_{y_G z_G} \cos \alpha - J_{z_G} \sin \alpha)] + \frac{1}{2} [J_{x_G} \dot{\alpha}^2 + \\ &+ \dot{\beta}^2 (J_{y_G} \cos^2 \alpha + J_{z_G} \sin^2 \alpha - J_{y_G z_G} \sin 2\alpha) - \\ &- 2\dot{\alpha} \dot{\beta} (J_{x_G y_G} \cos \alpha + J_{x_G z_G} \sin \alpha)], \end{aligned} \quad (2.7)$$

де

$$\begin{aligned} J_{x_G} &= \mathbf{x}_G^T \mathbf{J}_G \mathbf{x}_G, \quad J_{y_G} = \mathbf{y}_G^T \mathbf{J}_G \mathbf{y}_G, \quad J_{z_G} = \mathbf{z}_G^T \mathbf{J}_G \mathbf{z}_G, \\ J_{x_G z_G} &= -\mathbf{x}_G^T \mathbf{J}_G \mathbf{z}_G, \quad J_{y_G z_G} = -\mathbf{y}_G^T \mathbf{J}_G \mathbf{z}_G, \\ J_{x_G y_G} &= -\mathbf{x}_G^T \mathbf{J}_G \mathbf{y}_G \end{aligned} \quad (2.8)$$

- осьові і відцентрові моменти інерції системи відносно осей $Gx_G y_G z_G$.

Перше рівняння в (2.2) дає такий перший інтеграл для визначення ω :

$$\begin{aligned} \frac{\partial T}{\partial \omega} &= J_{z_G} \omega - \\ &- [J_{x_G z_G} \dot{\alpha} + \dot{\beta} (J_{y_G z_G} \cos \alpha - J_{z_G} \sin \alpha)] + \\ &+ \tilde{F}(\dot{\phi}_i) = K, \end{aligned} \quad (2.9)$$

де $\tilde{F}(\dot{\phi}_i)$ - лінійні щодо відносних узагальнених швидкостей $\dot{\phi}_i$ члени, K - стала величина, що дорівнює модулю момента кількості руху системи $\mathbf{K}_G$.

З рівняння (2.9) знаходимо

$$\begin{aligned} \omega^* &= \frac{J_{x_G z_G} \dot{\alpha} + \dot{\beta} (J_{y_G z_G} \cos \alpha - J_{z_G} \sin \alpha) + K}{J_{z_G}} + \\ &+ F(\dot{\phi}_i). \end{aligned} \quad (2.10)$$

де $F(\dot{\phi}_i) = -\tilde{F}(\dot{\phi}_i) / J_{z_G}$.

Підставляючи (2.10) в (2.6), (2.7) знайдемо кінетичну енергію системи у вигляді

$$T^* = T_0^* + T_1^* + T_2^*, \quad (2.11)$$

де

$$T_0^* = K^2 / (2J_{z_G}) \quad (2.12)$$

- складова, що не залежить від узагальнених швидкостей, T_1^*, T_2^* - складові, що є відповідно лінійними і квадратичними формами щодо узагальнених швидкостей $\dot{\alpha}, \dot{\beta}, \dot{\phi}_1, \dots, \dot{\phi}_n$.

Функція Рауса матиме вигляд:

$$R = T^* - K\omega^* = R_0 + R_1 + R_2, \quad (2.13)$$

де

$$R_0 = -K^2 / (2J_{z_G}) \quad (2.14)$$

- складова, що не залежить від узагальнених швидкостей, R_1, R_2 - складові, що є відповідно лінійними і квадратичними формами щодо узагальнених швидкостей $\dot{\alpha}, \dot{\beta}, \dot{\phi}_1, \dots, \dot{\phi}_n$.

Рівняння руху Рауса для нециклічних координат приймуть вигляд

$$\frac{d}{dt} \frac{\partial R}{\partial \dot{\alpha}} - \frac{\partial R}{\partial \alpha} = 0, \quad \frac{d}{dt} \frac{\partial R}{\partial \dot{\beta}} - \frac{\partial R}{\partial \beta} = 0,$$

$$\frac{d}{dt} \frac{\partial R}{\partial \dot{\varphi}_j} - \frac{\partial R}{\partial \varphi_j} = \frac{\partial \Pi}{\partial \varphi_j} + Q_j, \quad / j = \overline{1, n} /. \quad (2.15)$$

Потенціальна енергія приведеної системи дорівнює

$$W(\alpha, \beta, \varphi_1, \dots, \varphi_n) = \Pi - R_0 = \Pi + K^2 / (2J_{z_G}). \quad (2.16)$$

Рівняння усталених рухів системи мають вигляд

$$\frac{\partial W}{\partial \alpha} = 0, \quad \frac{\partial W}{\partial \beta} = 0, \quad \frac{\partial W}{\partial \varphi_j} = 0 \quad / j = \overline{1, n} / , \quad (2.17)$$

або, після перетворень

$$\frac{\partial R_0}{\partial \alpha} = 0, \quad \frac{\partial R_0}{\partial \beta} = 0, \quad \frac{\partial R_0}{\partial \varphi_j} = \frac{\partial \Pi}{\partial \varphi_j} \quad / j = \overline{1, n} / .$$

Як було показано в роботі [3], перші два рівняння цієї системи, з врахуванням рівності (2.14), еквівалентні таким двом рівнянням: $J_{x_G z_G} = 0$,

$J_{y_G z_G} = 0$. Тому рівняння усталених рухів для розглядуваних систем приймають такий вигляд:

$$J_{x_G z_G} = 0, \quad J_{y_G z_G} = 0, \quad \frac{\partial R_0}{\partial \varphi_j} = \frac{\partial \Pi}{\partial \varphi_j} \quad / j = \overline{1, n} / . \quad (2.18)$$

З цих рівнянь випливає, що на будь-якому усталеному русі система обертається навколо головної центральної осі інерції z_G і осьовий момент інерції J_{z_G} є головним центральним.

Система рівнянь (2.18) є системою нелінійних алгебраїчних рівнянь. Така система допускає декілька істотно відмінних розв'язків, а у випадку АБ із багатьма коригувальними вантажами навіть сім'ї усталених рухів, що залежать від одного чи декількох параметрів [7].

§3. Оцінка стійкості усталених рухів

З рівнянь (2.17) випливає, що на усталеному русі приведена потенціальна енергія системи W приймає екстремальне, або критичне значення. Відповідно до теореми Рауса для стійкості усталеного руху достатньо, щоб на ньому приведена потенціальна енергія системи мала мінімум [10].

Введемо позначення

$$a_{ij} = \partial^2 W / \partial q_i \partial q_j,$$

$$q = (\alpha, \beta, \varphi_1, \dots, \varphi_n), \quad / i, j = \overline{1, n} / . \quad (3.1)$$

Нехай на певному усталеному русі

$$q = (\tilde{\alpha}, \tilde{\beta}, \tilde{\varphi}_1, \dots, \tilde{\varphi}_n), \quad a_{ij} = \tilde{a}_{ij}, \quad / i, j = \overline{1, n} / . \quad (3.2)$$

Тоді, відповідно до критерію Сильвестра,

умовами мінімуму функції W на цьому усталеному русі будуть [10]:

$$\tilde{a}_{ii} > 0, \quad / i = \overline{1, n} / ,$$

$$\Delta_j = \begin{vmatrix} \tilde{a}_{11} & \dots & \tilde{a}_{1j} \\ \vdots & & \vdots \\ \tilde{a}_{j1} & \dots & \tilde{a}_{jj} \end{vmatrix} > 0, \quad / j = \overline{2, n} / . \quad (3.3)$$

Якщо система не має елементів, здатних накопичувати потенціальну енергію, як це у випадку, коли приєднані тіла утворюють АБ, то $\Pi = 0$ і

$$W(\alpha, \beta, \varphi_1, \dots, \varphi_n) = -R_0 = K^2 / (2J_{z_G}). \quad (3.4)$$

З рівнянь (2.17) з врахуванням (3.4) випливає, що на усталеному русі осьовий момент інерції J_{z_G} приймає екстремальне, або критичне значення. Відповідно до теореми Рауса для стійкості усталеного руху достатньо, щоб на ньому осьовий момент інерції приймав максимальне значення.

В загальному випадку теорема Рауса не оборотна [10]. Але в даному випадку до розглядуваних систем не підводиться енергія, і під час перехідних процесів відбувається відносний рух приєднаних тіл і як наслідок – розсіюється енергія системи. Якщо б розсіювання енергії не було, то мав би місце інтеграл енергії вигляду $R_2 + W = \text{const}$. Якщо на систему діють сили в'язкого опору, лінійні щодо узагальнених швидкостей, то замість інтеграла енергії має місце такий баланс енергій

$$d(R_2 + W)dt = -2\Phi, \quad (3.5)$$

де Φ – дисипативна функція Релея. Тому можна стверджувати (аналог теореми Четаєва про оборотність теореми Лагранжа [10]), що якщо на усталеному русі потенціальна енергія приведеної системи не має мінімуму, то цей рух нестійкий. Твердження ілюструє рис. 2, на якому зображена важка кулька у полі сил ваги.

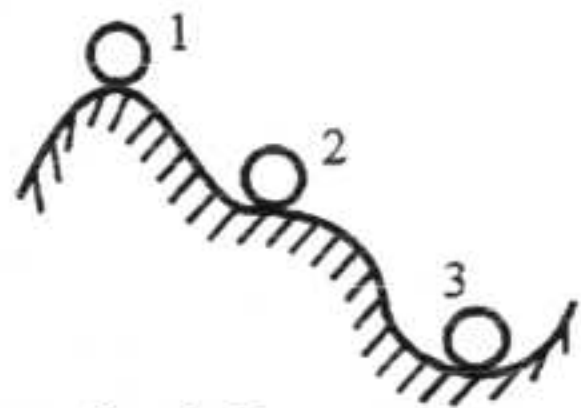


Рис. 2. Важка кулька

Її потенціальна енергія якісно характеризує різні енергетичні рівні потенціальної енергії

приведеної системи W . У точці максимуму – 1, або перегину – 2 поверхні, на якій знаходиться кулька, рівновага нестійка. У точці мінімуму – 3 рівновага – стійка. Умови розсіювання і не підведення енергії обов'язкові, оскільки завдяки ним точка може опускатися з верхніх енергетичних рівнів $(R_2 + W)$ на нижні, і покинувши певний верхній енергетичний рівень, вже не може до нього піднятися.

§4. Застосування теореми Рауса до незрівноваженого обертового тіла з АБ

1. Динамічне зрівноваження несучого тіла.

Нехай тіла АБ можуть усунути динамічну незрівноваженість несучого тіла. Нехай система здійснює основний рух – у якому обертове тіло зрівноважене і система обертається як одне ціле навколо осі z_G . Нехай на основному русі осі $O\xi\eta\zeta$ і $Gx_Gy_Gz_G$ співпадають і відносно них система має головні центральні осьові моменти інерції A, B, C . Нехай тіла АБ дещо відхилилися від положення, в якому зрівноважують несуче тіло. Тоді тензор інерції та координати центра мас системи щодо осей $O\xi\eta\zeta$, можна представити у вигляді

$$J_O = \begin{pmatrix} A + \varepsilon \tilde{J}_\xi & -\varepsilon \tilde{J}_{\xi\eta} & -\varepsilon \tilde{J}_{\xi\zeta} \\ -\varepsilon \tilde{J}_{\xi\eta} & B + \varepsilon \tilde{J}_\eta & -\varepsilon \tilde{J}_{\eta\zeta} \\ -\varepsilon \tilde{J}_{\xi\zeta} & -\varepsilon \tilde{J}_{\eta\zeta} & C \end{pmatrix},$$

$$\xi_G = \varepsilon \tilde{\xi}_G, \quad \eta_G = \varepsilon \tilde{\eta}_G, \quad \zeta_G = \varepsilon \tilde{\zeta}_G, \quad (4.1)$$

де $|\varepsilon| \ll 1$, і врахована властивість тіл АБ не змінювати осьовий момент інерції C [7].

Шукаємо α, β у вигляді

$$\alpha \approx \alpha_1 \varepsilon, \quad \beta \approx \beta_1 \varepsilon. \quad (4.2)$$

Тоді з врахуванням (1.2), (2.4) із (2.8), з точністю до величин другого порядку малості включно, знаходимо

$$J_{z_G} \approx C - \varepsilon^2 \{ [(C - B)\alpha_1^2 + (C - A)\beta_1^2 + M_\Sigma (\tilde{\xi}_G^2 + \tilde{\eta}_G^2) + 2(\alpha_1 \tilde{J}_{\eta\zeta} - \beta_1 \tilde{J}_{\xi\zeta})] \}. \quad (4.3)$$

Розглядатимемо параметри незрівноваженості $\tilde{\xi}_G, \tilde{\eta}_G, \tilde{J}_{\xi\zeta}, \tilde{J}_{\eta\zeta}$ як узагальнені координати, що характеризують рух системи. На основному русі вони дорівнюють нулю. Відмітимо, що вони незалежні (бо АБ може усунути динамічну незрівноваженість несучого тіла) і виражаються через узагальнені координати $\varphi_1, \dots, \varphi_n$, які задають положення тіл АБ щодо несучого тіла.

Введемо позначення

$$a_{11} = C - B, \quad a_{22} = C - A, \quad a_{33} = a_{44} = M_\Sigma, \\ a_{55} = a_{66} = 0, \quad a_{16} = 1, \quad a_{25} = -1. \quad (4.4)$$

Тоді, відповідно до критерію Сильвестра, з перших умов (3.3) знаходимо, що для того, щоб осьовий момент інерції J_{z_G} мав максимум, необхідно, щоб

$$C > A, B, \quad (4.5)$$

тобто щоб складене тіло, утворене несучим тілом і АБ, було сплюснуте. Нехай умови (4.5) виконуються. Тоді

$$\Delta_2 = (C - B)(C - A) > 0, \quad \Delta_3 = \Delta_2 M_\Sigma > 0, \\ \Delta_4 = \Delta_3 M_\Sigma > 0, \quad \Delta_5 = -M_\Sigma (C - B) < 0, \\ \Delta_6 = M_\Sigma^2 > 0. \quad (4.6)$$

Звідки видно, що достатні умови стійкості не виконуються. Оскільки на основному русі у системи немає мінімуму потенціальної енергії приведеної системи, і під час перехідних процесів в системі відбувається розсіювання енергії, то цей рух – нестійкий.

2. Статичне зрівноваження несучого тіла.

Нехай статичну незрівноваженість несучого тіла усуває один АБ. Нехай площина зрівноваження (і статичної незрівноваженості) паралельна площині $O\xi\eta$ і зміщена на координату b по осі ζ . Тоді

$$\tilde{J}_{\xi\zeta} = M_\Sigma \tilde{\xi}_G b, \quad \tilde{J}_{\eta\zeta} = M_\Sigma \tilde{\eta}_G b. \quad (4.7)$$

Тоді рівняння (4.3) прийме вигляд

$$J_{z_G} \approx C - \varepsilon^2 \{ [(C - B)\alpha_1^2 + (C - A)\beta_1^2 + M_\Sigma (\tilde{\xi}_G^2 + \tilde{\eta}_G^2) + 2M_\Sigma b(\alpha_1 \tilde{\eta}_G - \beta_1 \tilde{\xi}_G)] \}. \quad (4.8)$$

Розглядатимемо параметри незрівноваженості $\tilde{\xi}_G, \tilde{\eta}_G$ як узагальнені координати, що характеризують рух системи. На основному русі вони дорівнюють нулю. Відмітимо, що вони незалежні (бо АБ може усунути статичну незрівноваженість несучого тіла) і виражаються через узагальнені координати $\varphi_1, \dots, \varphi_n$, які задають положення тіл АБ щодо несучого тіла.

Введемо позначення

$$a_{11} = C - B, \quad a_{22} = C - A, \quad a_{33} = a_{44} = M_\Sigma, \\ a_{14} = bM_\Sigma, \quad a_{23} = -bM_\Sigma. \quad (4.9)$$

Тоді перші умови в (3.3) ($n = 4$) дають необхідні

умови стійкості (4.5). Нехай ці умови виконуються, тоді решта умов приймає вигляд:

$$\begin{aligned}\Delta_2 &= (C - B)(C - A) > 0, \\ \Delta_3 &= M_z(C - A)(C - B - b^2 M_z) > 0, \\ \Delta_4 &= M_z^2(C - A - b^2 M_z)(C - B - b^2 M_z) > 0. \quad (4.10)\end{aligned}$$

З врахуванням (4.5) і (4.10), достатніми умовами, що J_{zG} приймає на основному русі максимальне значення, будуть

$$2) C > A + b^2 M_z, \quad C > B + b^2 M_z. \quad (4.11)$$

У цьому випадку складене тіло – сплюснуте, і площина зрівноваження знаходиться поблизу центра мас системи.

Зауважимо, що у проведених дослідженнях тип АБ не обумовлювався. Тому одержані умови усунення кута нутації застосовні для АБ будь-якого типу, тобто є узагальненими.

Умови (4.11) є достатніми умовами стійкості основного руху (усунення кута нутації), причому вони із точністю до меж співпадають із необхідними умовами, знайденими в роботах [1-3].

Проведені дослідження дозволяють стверджувати, що повне усунення кута нутації обертового незрівноваженого супутника можливо тільки у випадку сплюснутого (так званого стійкого) статично незрівноваженого супутника за умови, що АБ зрівноважує супутник у площині статичного дисбалансу і ця площина віддалена від центра мас системи так, що виконуються умови (4.11). Ці результати були покладені в розробку способів застосування АБ у якості демпферів кута нутації сплюснутих обертових космічних апаратів [4].

Список використаних джерел

1. Філімоніхіна І.І. Умови зменшення автобалансирами кута нутації обертового супутника Землі // Всеукраїнський н.-т. журнал „Вібрації у техніці та технологіях”. – 2007. №1 (46), С. 34-37.

2. Філімоніхіна І.І., Філімоніхін Г.Б. Усталені рухи і умови самозрівноваження одного типу ізольованої системи // Вісник Київського ун-ту. Серія: фізико-матем. науки. 2007. - №3. - с. 103-109.
3. Филимоных И.И., Филимоных Г.Б. Условия уравновешивания автобалансирами вращающегося тела в изолированной системе // Прикладная механика, т.43, №11, 2007. – С.113-120.
4. Застосування пасивного автобалансира як демпфера кута нутації сплюснутого обертового космічного апарата: Пат. на корисну модель № 28407 Україна, МПК В64G 1/00 / І. І. Філімоніхіна, Г.Б. Філімоніхін (Україна); КНТУ - № 200708020; Заявл. 16.07.2007; Опубл. 10.12.2007, Бюл.№20.
5. Артюхин Ю.П., Каргу Л.И. Симаев В.Л. Системы управления космических аппаратов, стабилизируемых вращением. –М.: Наука, 1979. -295 с.
6. Попов В.И. Системы ориентации и стабилизации космических аппаратов. - М.: Машиностроение, 1986. - 184 с.
7. Філімоніхін Г.Б. Зрівноваження і віброзахист роторів автобалансирами з твердими коригувальними вантажами. - Кіровоград: КНТУ, 2004. – 352 с.
8. Павловський М.А. Теоретична механіка: Підручник. – К.: Техніка, 2002. – 512 с.
9. Докучаев Л.В., Рабинович Б.И. Анализ возмущенного движения вблизи границы устойчивости вращающегося КА типа Авроральный зонд проекта Интербол // Космические исследования, 1999, том. 37, №6, с. 589-597.
10. Меркин Д.Р. Введение в теорию устойчивости движения. – М.: Наука, 1987. – 304 с.

Надійшла до редколегії 23.04.08

УДК 517.925.4

*Ігор І. Король, к.ф.-м.н., доцент

Дослідження періодичних розв'язків автономних диференціальних систем з імпульсною дією у критичному випадку

Вивчаються питання існування періодичних розв'язків систем автономних диференціальних рівнянь, що піддаються імпульсному впливу. Для критичного випадку запропоновано новий чисельно-аналітичний алгоритм дослідження існування і побудови періодичних розв'язків, період яких співпадає або близький до періоду розв'язків породжуючої лінійної однорідної імпульсної системи.

Ключові слова: імпульсні системи, чисельно-аналітичний метод.

*E-mail: korol_igor@ukr.net

Статтю представив д.ф.-м.н, проф. Перестюк М.О.

У даній роботі розроблено новий чисельно-аналітичний алгоритм дослідження та наближеної побудови періодичних розв'язків нелінійних автономних систем диференціальних рівнянь, підпорядкованих імпульсному впливу. Встановлено необхідні і достатні умови існування розв'язків, знайдено оцінки похибки послідовних наближень. Аналогічні питання розглядалися в роботах [1,2,5]. Дана публікація є продовженням робіт [3,4].

Розглянемо задачу встановлення умов існування та знаходження - періодичних розв'язків автономної системи диференціальних рівнянь

$$\frac{dx}{dt} = Ax + g(x), \quad t \neq t_i, \quad (1)$$

з імпульсною дією

$$\Delta x|_{t=t_i} = B_i x(t_i - 0) + b_i(x(t_i - 0)), \quad (2)$$

де A , B_i - сталі $(n \times n)$ - матриці, $g(x)$, $b_i(x)$ - вектор-функції, причому $\det(E_n + B_i) \neq 0$, I_n - одинична матриця порядку n , $B_{i+p} = B_i$, $b_{i+p}(x) = b_i(x)$, $t_{i+p} = t_i + T$, $0 < t_1 < \dots < t_p < T$.

При цьому будемо розглядати критичний випадок - коли

А) відповідна породжуюча лінійна однорідна імпульсна система

*Ihor I. Korol, Cand.Sci (Phys.-Math.), Associate Professor,

Investigating for periodic solutions of autonomous differential systems with impulse action in a critical case

Problems of existence of periodic solutions of systems of the autonomous differential equations, exposed to impulse influence are studied. In a critical case the new numerical-analytical algorithm for investigating of existence and approximate constructing of periodic solutions which period coincides or is close to the period of solutions of corresponding generating linear homogeneous impulsive system is suggested.

Key words: impulsive systems, numerical-analytic method.

$$\frac{dx}{dt} = Ax, \quad t \neq t_i, \quad \Delta x|_{t=t_i} = B_i x(t_i - 0), \quad (3)$$

має k - параметричну сім'ю нетривіальних розв'язків зі спільним періодом T .

Відомо, що якщо виконується умова А, то спряжена до (3) лінійна однорідна система

$$\frac{dx}{dt} = -A^T x, \quad t \neq t_i, \quad \Delta x|_{t=t_i} = -(I_n + B_i^T) B_i^T x(t_i - 0) \quad (4)$$

теж має k лінійно незалежних T - періодичних розв'язків $\psi_1(t), \dots, \psi_k(t)$. Позначимо через $\Psi(t)$ $(n \times k)$ - вимірну матрицю, стовпцями якої є $\psi_1(t), \dots, \psi_k(t)$, через $\Theta(t)$ - $(n \times (n-k))$ - вимірну матрицю, стовпцями якої є ті $(n-k)$ лінійно незалежні розв'язки системи (4), які не є T - періодичними. Нехай $Y(t)$ - така фундаментальна матриця системи (4), у якої перші k стовпців утворені матрицею $\Psi(t)$, а інші $(n-k)$ стовпців - матрицею $\Theta(t)$. Фундаментальну матрицю $X(t)$ виберемо так, щоб $Y^T(t)X(t) = I_n$. Матрицант $X(t, 0)$ системи (3) можемо записати у вигляді

$$X(t, 0) = (Y^T(t))^{-1} Y^T(0) = (Y^T(t))^{-1} \begin{pmatrix} \Psi(0) \\ \Theta(0) \end{pmatrix}$$

Крім того, введемо наступні позначення

$$R_1(t) = \int_0^t \Psi^T(s) \Psi(s) ds, \quad R_2(t) = \int_0^t \Theta^T(s) \Psi(s) ds,$$

$$R_1 = R_1(T), \quad R(t) = \int_0^t Y^T(s) \Psi(s) ds = \begin{pmatrix} R_1(t) \\ R_2(t) \end{pmatrix},$$

$$U_1(t, s) = \Psi^T(s) - R_1(t) R_1^{-1} \Psi^T(s)$$

$$U_2(t, s) = \Theta^T(s) - R_2(t) R_1^{-1} \Psi^T(s)$$

$$U(t, s) = Y^T(s) - R(t) R_1^{-1} \Psi^T(s) = \begin{pmatrix} U_1(t, s) \\ U_2(t, s) \end{pmatrix}$$

$$V(t, s) = R(t) R_1^{-1} \Psi^T(s).$$

Сформулюємо необхідні та достатні умови існування розв'язку імпульсної системи (1), (2).

Лема 1. Нехай виконується умова А і імпульсна диференціальна система (1), (2) має періодичний розв'язок $\varphi(t)$ періоду T . Тоді його початкове значення визначається за формулою

$$\varphi(0) = \varphi(0, \xi^*) = F\xi^* + G^* \left[\int_0^T \Theta^T(s) g(\varphi(s)) ds + \sum_{i=1}^p \Theta^T(t_i + 0) b_i(\varphi(t_i - 0)) \right] \quad (5)$$

де $G = \Theta^T(T) - \Theta^T(0)$ – прямокутна $((n-k) \times n)$ - матриця рангу $(n-k)$; G^* – єдина $(n \times (n-k))$ - матриця, псевдообернена по Пенроузу до матриці G ; F – фундаментальна $(n \times k)$ - матриця розв'язків однорідної системи $Gx_0 = 0$; $\xi^* \in R^k$ є розв'язком визначального рівняння

$$\int_0^T \Psi^T(s) g(\varphi(s, \xi^*)) ds + \sum_{i=1}^p \Psi^T(t_i + 0) b_i(\varphi(t_i - 0), \xi^*) = 0. \quad (6)$$

Доведення. Нехай $\varphi(t)$, $\varphi(0) = x_0$ є розв'язком системи (1), (2), тоді при $t \neq t_i$ має місце тотожність

$$\varphi(t) = X(t, 0)x_0 + \int_0^t X(t, s)g(\varphi(s))ds + \sum_{0 \leq t_i < t} X(t, t_i + 0)b_i(\varphi(t_i - 0)),$$

де $X(t, t_i + 0) = X(t, t_i - 0)(I_n + B_i)^{-1}$. З T - періодичності $\varphi(t)$ одержуємо лінійну алгебраїчну систему

$$(I_n - X(T, 0))x_0 = X(T) \left\{ \int_0^T Y^T(s)g(\varphi(s))ds + \sum_{i=1}^p Y^T(t_i + 0)b_i(\varphi(t_i - 0)) \right\}.$$

Домножуючи її зліва на невироджену матрицю $X^{-1}(T)$ і, беручи до уваги структуру матриць $X(t)$ і $Y(t)$, одержимо наступну систему для знаходження x_0 :

$$\begin{pmatrix} 0 \\ G \end{pmatrix} x_0 = \begin{pmatrix} \int_0^T \Psi^T(s)g(\varphi(s))ds + \sum_{i=1}^p \Psi^T(t_i + 0)b_i(\varphi(t_i - 0)) \\ \int_0^T \Theta^T(s)g(\varphi(s))ds + \sum_{i=1}^p \Theta^T(t_i + 0)b_i(\varphi(t_i - 0)) \end{pmatrix}$$

Зрозуміло, що ця система є сумісною тоді і тільки тоді, коли виконується умова (6) і при цьому x_0 має вигляд (5). Лему доведено.

Розглянемо k - параметричне відображення $L_\xi x: C(R, R^n) \rightarrow C(R, R^n)$ і функціонал $\mu(x): C(R) \rightarrow R^k$:

$$\begin{aligned} (L_\xi x)(t) &= X(t, 0)F\xi^* + X(t, 0)G^* \left(\int_0^T U_2(T, s)g(x(s))ds + \sum_{i=1}^p U_2(T, t_i + 0)b_i(x(t_i - 0)) \right) + \\ &+ \int_0^t X(t, s)g(x(s))ds - \int_0^t X(t)U(t, s)g(x(s))ds + \sum_{0 \leq t_i < t} X(t)U(t, t_i + 0)b_i(x(t_i - 0)) - \sum_{t \leq t_i < T} X(t)U(t, t_i + 0)b_i(x(t_i - 0)), \\ \mu(x) &= \int_0^T \Psi^T(s)g(x(s))ds + \sum_{i=1}^p \Psi^T(t_i + 0)b_i(x(t_i - 0)). \end{aligned}$$

Зауважимо, що

$$\begin{aligned} (L_\xi x)(t) &= X(t, 0) \left\{ F\xi^* + G^* \left(\int_0^T \Theta^T(s)g(x(s))ds + \sum_{i=1}^p \Theta^T(t_i + 0)b_i(x(t_i - 0)) \right) \right\} + \\ &+ \int_0^t X(t, s)g(x(s))ds + \sum_{0 \leq t_i < t} X(t, t_i + 0)b_i(x(t_i - 0)) - X(t)Y^T(0)G^*R_1(T) + R(t)R_1^{-1}\mu(x) \end{aligned} \quad (7)$$

Достатні умови існування періодичних розв'язків досліджуваної системи містить наступне твердження.

Лема 2. Нехай виконується умова А. Якщо при цьому параметр $\xi = \xi^* \in R^k$ і функція $\varphi(t) = \varphi(t, \xi^*)$ такі, що виконується система рівнянь

$$\varphi = L_t \varphi, \quad (8)$$

$$\mu(\varphi) = 0, \quad (9)$$

то $\varphi(t) \in T$ - періодичним розв'язком імпульсної диференціальної системи (1), (2) і приймає початкове значення (5).

Доведення. Нехай кусково-неперервна з розривами першого роду при $t = t_i$ функція $\varphi(t)$ і векторний параметр $\xi \in R^k$ задовольняють рівняння (8) і (9). Тоді з (7) випливає, що $\varphi(t)$ є розв'язком інтегрального рівняння

$$x(t) = X(t, 0) \left\{ F\xi + G \cdot \left(\int_0^t \Theta^T(s) g(x(s)) ds + \sum_{i=1}^p \Theta^T(t_i + 0) b_i(x(t_i, -0)) \right) \right\} + \int_0^t X(t, s) g(x(s)) ds + \sum_{0 \leq t_i < t} X(t, t_i + 0) b_i(x(t_i, -0)), \quad (10)$$

Підставляючи $t = 0$, переконуємося, що виконується умова (5), а диференціюючи (10) по t , бачимо, що $\varphi(t)$ є розв'язком імпульсної диференціальної системи (1), (2). Покажемо періодич-

ність функції $\varphi(t)$. Підставляючи $t = T$ в (10) отримуємо

$$\varphi(t) = X(T, 0) \left\{ F\xi + G \cdot \left(\int_0^T \Theta^T(s) g(\varphi(s)) ds + \sum_{i=1}^p \Theta^T(t_i + 0) b_i(\varphi(t_i, -0)) \right) \right\} + \int_0^T X(T, s) g(\varphi(s)) ds + \sum_{i=1}^p X(T, t_i + 0) b_i(\varphi(t_i, -0)).$$

звідки, враховуючи (5) і властивості псевдооберненої матриці, одержимо, що

$$\begin{aligned} Y^*(T)(\varphi(0) - \varphi(T)) &= Y^*(T)(I_n - X(T)) \left\{ F\xi + G \cdot \left(\int_0^T \Theta^T(s) g(\varphi(s)) ds + \sum_{i=1}^p \Theta^T(t_i + 0) b_i(\varphi(t_i, -0)) \right) \right\} - \\ &\quad - \int_0^T X(T, s) g(\varphi(s)) ds - \sum_{i=1}^p X(T, t_i + 0) b_i(\varphi(t_i, -0)) = \\ &= Y^*(T)(I_n - X(T)) \left\{ F\xi + G \cdot \left(\int_0^T \Theta^T(s) g(\varphi(s)) ds + \sum_{i=1}^p \Theta^T(t_i + 0) b_i(\varphi(t_i, -0)) \right) \right\} - \\ &\quad - \left(\int_0^T \Psi^*(s) g(\varphi(s)) ds + \sum_{i=1}^p \Psi^*(t_i + 0) b_i(\varphi(t_i, -0)) \right) = \\ &= \left(\int_0^T \Theta^*(s) g(\varphi(s)) ds + \sum_{i=1}^p \Theta^*(t_i + 0) b_i(\varphi(t_i, -0)) \right) = \\ &= \begin{pmatrix} 0 \\ G \end{pmatrix} \left(F\xi + G \cdot \left(\int_0^T \Theta^T(s) g(\varphi(s)) ds + \sum_{i=1}^p \Theta^T(t_i + 0) b_i(\varphi(t_i, -0)) \right) \right) - \begin{pmatrix} \int_0^T \Theta^*(s) g(\varphi(s)) ds + \sum_{i=1}^p \Theta^*(t_i + 0) b_i(\varphi(t_i, -0)) \end{pmatrix} = 0. \end{aligned}$$

Оскільки $\det Y(t) \neq 0$, то $\varphi(0) = \varphi(T)$.

Лему доведено.

Для знаходження T - періодичного розв'язку системи (1), (2) розглянемо послідовність T - пері-

одичних кусково-неперервних з розривами першого роду при $t = t_i$ вектор-функцій (11):

$$\begin{aligned}
 x_{m+1}(t, \xi) = & x_0(t, \xi) + X(t, 0)G^+ \left(\int_0^T U_2(T, s)g(x_m(s, \xi))ds + \sum_{i=1}^p U_2(T, t_i + 0)b_i(x_m(t_i, -0, \xi)) \right) + \\
 & + \int_0^t X(t)U(t, s)g(x_m(s, \xi))ds - \int_t^T X(t)V(t, s)g(x_m(s, \xi))ds + \sum_{0 \leq t_i < t} X(t)U(t, t_i + 0)b_i(x_m(t_i, -0, \xi)) - \\
 & - \sum_{t \leq t_i < T} X(t)V(t, t_i + 0)b_i(x_m(t_i, -0, \xi)), \quad x_0(t, \xi) = X(t, 0)F\xi, \quad m=1, 2, \dots
 \end{aligned} \quad (11)$$

Обґрунтування алгоритму відшукування T -періодичних розв'язків імпульсної диференціальної системи (1), (2) і конструктивні достатні умови їх існування містять теореми 2-4 [3].

Зауваження 1. Принцип побудови послідовних наближень (11) ґрунтується на "збурення" правої частини системи (1) [3]. Нескладно показати, що аналогічний ефект можна одержати шляхом введення в будь-якій точці $\sigma \in [0, T]$ додаткового спеціально підбраного імпульсу

$$\Delta x|_{t=\sigma} = -b_\sigma,$$

$$\begin{aligned}
 b_\sigma = & \Psi(\sigma)P_\sigma^{-1} \left\{ \int_0^T \Psi^T(s)g(x(s))ds + \sum_{i=1}^p \Psi^T(t_i + 0)b_i(x(t_i, -0)) \right\} \\
 P_\sigma = & \Psi^T(\sigma)\Psi(\sigma).
 \end{aligned}$$

При цьому послідовні наближення до точного розв'язку будуть мати вигляд

$$\begin{aligned}
 x_{m+1}(t, \xi) = & x_0(t, \xi) + X(t, 0)G^+ \left(\int_0^T U_2(T, s)g(x_m(s, \xi))ds + \sum_{i=1}^p U_2(T, t_i + 0)b_i(x_m(t_i, -0, \xi)) \right) + \int_0^t X(t, s)g(x_m(s, \xi))ds + \\
 & + \sum_{0 \leq t_i < t} X(t, t_i + 0)b_i(x_m(t_i, -0, \xi)) - \delta_\sigma X(t, \sigma)\Psi(\sigma)P_\sigma^{-1} \left\{ \int_0^T \Psi^T(s)g(x(s))ds + \sum_{i=1}^p \Psi^T(t_i + 0)b_i(x(t_i, -0)) \right\}, \\
 x_0(t, \xi) = & X(t, 0)F\xi, \quad m=1, 2, \dots, \quad \delta_\sigma = \begin{cases} 1, & t > \sigma, \\ 0, & t \leq \sigma. \end{cases}
 \end{aligned}$$

Розглянемо питання існування і наближеної побудови періодичних розв'язків нелінійної імпульсної системи (1), (2), період T , яких є близьким до періоду T розв'язків породжуючої лінійної системи (3). Так як період T системи (1), (2) невідомий, то будемо шукати його у вигляді $T_1 = (1 + \lambda)T$, де λ – невідомий малий параметр, який будемо шукати при побудові розв'язку.

За допомогою заміни незалежної змінної $t = (1 + \lambda)\tau$ приходимо до задачі знаходження періодичного з відомим фіксованим періодом T розв'язку автономної системи

Особливістю автономних імпульсних систем є те, що, взагалі кажучи, їх розв'язки не є інваріантними щодо зсуву по t , а тому при побудові розв'язків потрібно знаходити як значення параметра $\xi \in R^k$, так і $\lambda \in R$.

Введемо до розгляду $(k+1)$ -параметричне відображення $\tilde{L}_T x: C(R, R^n) \rightarrow C(R, R^n)$, $\tilde{\xi} = (\xi, \lambda)$ і функціонал $\tilde{\mu}(x): C(R) \rightarrow R^k$:

$$\frac{dx}{d\tau} = Ax + f(x, \lambda), \quad \tau \neq \tau_i,$$

$$f(x, \lambda) = g(x) + \lambda(Ax + g(x)), \quad \tau = \tau_i(\lambda) = \frac{t_i}{1 + \lambda}, \quad (12)$$

$$\Delta x|_{\tau=\tau_i} = B_i x(\tau_i, -0) + b_i(x(\tau_i, -0))$$

$$\begin{aligned} (\tilde{L}_{\tilde{\xi}} x)(\tau) = & X(\tau, 0)F\xi + X(\tau, 0)G^* \left(\int_0^T U_2(T, s)f(x(s, \tilde{\xi}), \lambda)ds + \sum_{i=1}^p U_2(T, \tau_i + 0)b_i(x(\tau_i - 0, \tilde{\xi})) \right) + \\ & + \int_0^{\tau} X(\tau)U(\tau, s)f(x(s, \tilde{\xi}), \lambda)ds - \int_{\tau}^T X(\tau)V(\tau, s)f(x(s, \tilde{\xi}), \lambda)ds + \\ & + \sum_{0 \leq \tau_i < \tau} X(\tau)U(\tau, \tau_i + 0)b_i(x(\tau_i - 0, \tilde{\xi})) - \sum_{\tau \leq \tau_i < T} X(\tau)V(\tau, \tau_i + 0)b_i(x(\tau_i - 0, \tilde{\xi})) \\ \tilde{\mu}(x) = & \int_0^T \Psi^T(s)f(x(s, \tilde{\xi}), \lambda)ds + \sum_{i=1}^p \Psi^T(\tau_i + 0)b_i(x(\tau_i - 0, \tilde{\xi})) = 0. \end{aligned} \quad (16)$$

Згідно викладеного вище, мають місце наступні твердження.

Лема 3. Нехай виконується умова А і імпульсна диференціальна система (12) має періодич-

ний розв'язок $\varphi(\tau) = \varphi(\tau, \tilde{\xi}^*)$ періоду T . Тоді його початкове значення визначається за формулою

$$\varphi(0) = \varphi(0, \tilde{\xi}^*) = F\xi^* + G^* \left[\int_0^T \Theta^T(s)f(\varphi(s, \tilde{\xi}^*), \lambda^*)ds + \sum_{i=1}^p \Theta^T(\tau_i + 0)b_i(\varphi(\tau_i - 0, \tilde{\xi}^*)) \right], \quad (13)$$

де $\tilde{\xi}^* = (\xi^*, \lambda^*) \in R^{k+l}$ є розв'язком визначального рівняння

$$\tilde{\mu}(\varphi) = 0. \quad (14)$$

Лема 4. Нехай виконується умова А. Якщо при цьому параметр $\tilde{\xi} = \tilde{\xi}^* \in R^{k+l}$ і функція $\varphi(\tau) = \varphi(\tau, \tilde{\xi}^*)$ такі, що виконується система рівнянь (14), (15):

$$\varphi = \tilde{L}_{\tilde{\xi}} \varphi, \quad (15)$$

то $\varphi(\tau) \in T$ - періодичним розв'язком імпульсної

диференціальної системи (12) і приймає початкове значення (13). Відповідно функція $\varphi(t)$ з початковим значенням (13) є $T_i = (1 + \lambda^*)T$ - періодичним розв'язком системи (1), (2).

У цьому випадку T - періодичний розв'язок системи (12) будемо шукати як границю рекурентної послідовності

$$\begin{aligned} \tilde{x}_{m+1}(\tau, \tilde{\xi}) = & \tilde{x}_0(\tau, \tilde{\xi}) + X(\tau, 0)G^* \left(\int_0^T U_2(T, s)f(\tilde{x}_m(s, \tilde{\xi}), \lambda)ds + \sum_{i=1}^p U_2(T, \tau_i + 0)b_i(\tilde{x}_m(\tau_i - 0, \tilde{\xi})) \right) + \\ & + \int_0^{\tau} X(\tau)U(\tau, s)f(\tilde{x}_m(s, \tilde{\xi}), \lambda)ds - \int_{\tau}^T X(\tau)V(\tau, s)f(\tilde{x}_m(s, \tilde{\xi}), \lambda)ds + \sum_{0 \leq \tau_i < \tau} X(\tau)U(\tau, \tau_i + 0)b_i(\tilde{x}_m(\tau_i - 0, \tilde{\xi})) - \\ & - \sum_{\tau \leq \tau_i < T} X(\tau)V(\tau, \tau_i + 0)b_i(\tilde{x}_m(\tau_i - 0, \tilde{\xi})), \quad \tilde{x}_0(\tau, \tilde{\xi}) = X(\tau, 0)F\xi, \quad m=1, 2, \dots \end{aligned} \quad (16)$$

Припускаємо, що в області $(\tau, x, \lambda) \in R \times D \times I_{\lambda}$, де $D = \{x \in R^n : |x| \leq R\}$, $I_{\lambda} = [-\lambda, \lambda]$ виконуються наступні умови:

В) вектор-функція $g(x)$ визначена, неперервна і виконуються умови обмеженості і Ліпшиця з невід'ємними сталими векторами $M, M_i \in R^n$ і сталими матрицями K, K_i :

$$\begin{aligned} |g(x)| & \leq M, \quad |g(x') - g(x'')| \leq K|x' - x''|, \\ |b_i(x)| & \leq M_i, \quad |b_i(x') - b_i(x'')| \leq K_i|x' - x''|. \end{aligned}$$

де $i = \overline{1, p}$, $|x| = (|x_1|, \dots, |x_n|)$ і всі нерівності розуміємо покомпонентно;

С) існує не порожня множина $D_0 \times I_{\lambda} \subset R^{k+l}$ точок $\tilde{\xi} = (\xi, \lambda)$ така, що вектор-функція $x_0(\tau, \xi)$ лежить в області D разом із своїм β - околom, де $\tilde{M} = M + \lambda_0(M + |A|R)$,

$$\beta = \max_{\tau \in [0, T]} \left\{ \left(\int_0^T |X(\tau, 0)G^+U_2(T, s)| ds \tilde{M} + \sum_{i=1}^p |X(\tau, 0)G^+U_2(T, \tau_i + 0)| M_i \right) + \int_0^T |X(\tau)U(\tau, s)| ds \tilde{M} + \right. \\ \left. + \int_\tau^T |X(\tau)V(\tau, s)| ds \tilde{M} + \sum_{0 \leq \tau_i < \tau} |X(\tau)U(\tau, \tau_i + 0)| M_i + \sum_{\tau \leq \tau_i < T} |X(\tau)V(\tau, \tau_i + 0)| M_i \right\};$$

D) $r(\tilde{Q}) < 1$, де $r(\tilde{Q})$ – спектральний радіус оператора $\tilde{Q}x$:

$$(\tilde{Q}x)(\tau) = \int_0^T |X(\tau, 0)G^+U_2(T, s)| ds \tilde{K} + \sum_{i=1}^p |X(\tau, 0)G^+U_2(T, \tau_i + 0)| K_i + \\ + \int_0^T |X(\tau)U(\tau, s)| ds \tilde{K} + \int_\tau^T |X(\tau)V(\tau, s)| ds \tilde{K} + \\ + \sum_{0 \leq \tau_i < \tau} |X(\tau)U(\tau, \tau_i + 0)| K_i + \sum_{\tau \leq \tau_i < T} |X(\tau)V(\tau, \tau_i + 0)| K_i, \quad \tilde{K} = K + \lambda_0(K + |A|);$$

Зауважимо, що $r(\tilde{Q})$ рівний найбільшому з додатних власних значень оператора $\tilde{Q}$ і не перевищує найбільшого власного значення матриці $\tilde{Q}_0$:

$$\tilde{Q}_0 = \max_{\tau \in [0, T]} \left\{ \int_0^T |X(\tau, 0)G^+U_2(T, s)| ds \tilde{K} + \sum_{i=1}^p |X(\tau, 0)G^+U_2(T, \tau_i + 0)| K_i + \int_0^T |X(\tau)U(\tau, s)| ds \tilde{K} + \right. \\ \left. + \int_\tau^T |X(\tau)V(\tau, s)| ds \tilde{K} + \sum_{0 \leq \tau_i < \tau} |X(\tau)U(\tau, \tau_i + 0)| K_i + \sum_{\tau \leq \tau_i < T} |X(\tau)V(\tau, \tau_i + 0)| K_i \right\};$$

Теорема 1. Нехай автономна система диференціальних рівнянь з імпульсною дією (12) задовольняє умови A–D. Тоді:

1) в області $(\tau, \tilde{\xi}) \in R \times D_0 \times I_\lambda$ послідовність T -періодичних функцій $\tilde{x}(\tau, \tilde{\xi})$ вигляду (16) рівномірно збігається при $m \rightarrow \infty$ до деякої T -періодичної функції $\tilde{x}^*(\tau, \tilde{\xi})$ і при всіх натуральних m справджуються оцінки збіжності

$$|\tilde{x}^*(\tau, \tilde{\xi}) - \tilde{x}_m(\tau, \tilde{\xi})| \leq \tilde{Q}^m (I_n - \tilde{Q})^{-1} \beta;$$

2) для того, щоб функція $\tilde{x}^*(\tau) = \tilde{x}^*(\tau, \tilde{\xi}^*)$ була T -періодичним розв'язком системи диференціальних рівнянь з імпульсною дією (12) необ-

хідно і досить, щоб точка $\tilde{\xi} = \tilde{\xi}^*$ була розв'язком визначального рівняння $\tilde{\Delta}(\tilde{\xi}) = 0$, де

$$\tilde{\Delta}(\tilde{\xi}) = \int_0^T \Psi^T(s) f(\tilde{x}^*(s, \tilde{\xi}), \lambda) ds + \\ + \sum_{i=1}^p \Psi^*(t_i + 0) b_i(\tilde{x}^*(t_i - 0, \tilde{\xi})) = 0.$$

3) при $\tau = 0$ T -періодичний розв'язок $\tilde{x}^*(\tau)$ приймає початкове значення

$$\tilde{x}^*(0) = \tilde{x}^*(0, \tilde{\xi}^*) = F\tilde{\xi}^* + G^* \left[\int_0^T \Theta^T(s) f(\tilde{x}^*(s, \tilde{\xi}^*), \lambda^*) ds + \sum_{i=1}^p \Theta^T(\tau_i + 0) b_i(\tilde{x}^*(\tau_i - 0, \tilde{\xi}^*)) \right]. \quad (23)$$

Теорема 2. Нехай для системи (12) справедливі припущення A–D і, крім того:

1) існує опукла, замкнена область $\tilde{D}' = D'_0 \times I'_\lambda$, $D'_0 \subset D_0$, $I'_\lambda \subset I_\lambda$ така, що при деякому фіксованому натуральному m наближене

визначальне рівняння $\tilde{\Delta}_m(\tilde{\xi}) = 0$ має в області $\tilde{D}'$ єдиний розв'язок $\tilde{\xi} = \tilde{\xi}_m$ ненульового індексу, де

$$\tilde{\Delta}_m(\tilde{\xi}) \equiv \int_0^T \Psi^T(s) f(\tilde{x}_m(s, \tilde{\xi}), \lambda) ds + \sum_{i=1}^p \Psi^T(\tau_i + 0) b_i(\tilde{x}_m(\tau_i - 0, \tilde{\xi})),$$

2) на границі $\partial \tilde{D}'$ області $\tilde{D}'$ виконується нерівність

$$\inf_{\tilde{\xi} \in \partial \tilde{D}'} |\tilde{\Delta}_m(\tilde{\xi})| > \tilde{Q}^{m+1} (I_n - \tilde{Q})^T T \tilde{M}.$$

Тоді система (12) має T -періодичний розв'язок $\tilde{x}^*(\tau) = \tilde{x}^*(\tau, \tilde{\xi}^*)$ з початковим значенням $\tilde{x}^*(0) = \tilde{\xi}^*$, $\tilde{\xi}^* \in D'_0$. Відповідно, задача (1), (2) має розв'язок $x^*(t)$, $x^*(0) = \xi^*$ періоду $T_1 = (1 + \lambda^*)T$.

Доведення теорем 1 і 2 проводиться аналогічно до доведення теорем 2, 3 [3].

Зауваження 2. Нехай $B_i = B$ і матриці A і B комутують: $AB = BA$. При цьому якщо функція $x(t) \in T$ -періодичним розв'язком системи (12), то при довільному v функція $x(t+v)$ теж буде розв'язком цієї системи [2,6]. Константу v можемо підібрати так, щоб одна з координат $\xi_1, \dots, \xi_k$ була рівна нулю. Без обмеження загальності можемо вважати, що $\xi_k = 0$. У цьому випадку всі вищенаведені твердження будуть справедливими з тим уточненням, що при побудові періодичного розв'язку системи (12) нам потрібно буде знайти не $(k+1)$ -вимірний невідомий параметр $\tilde{\xi}$, а k -вимірний параметр ξ_0 : $\xi_0 \in R^k$, $\xi_0 = (\xi_1, \dots, \xi_{k-1}, \lambda)$.

2. Квазілінійні автономні системи

Розглянемо квазілінійну автономну систему звичайних диференціальних рівнянь

$$\begin{aligned} \frac{dx}{dt} &= Ax + \varepsilon g(x, \varepsilon), \quad t \neq t_i, \\ \Delta x|_{t=t_i} &= B_i x(t_i - 0) + \varepsilon b_i(x(t_i - 0), \varepsilon), \end{aligned} \quad (17)$$

де ε - малий додатний параметр, $\varepsilon \in I_\varepsilon = [0, \varepsilon_0]$, $\varepsilon_0 \ll 1$, $t \in R$, $x, g \in R^n$; у випадку, коли виконується умова А і функції $g(x, \varepsilon)$, $b_i(x(t_i - 0), \varepsilon)$ в області $(x, \varepsilon) \in D \times I_\varepsilon$ задовольняють умови обмеженості і Ліпшиця В.

Розглянемо питання існування і наближеної побудови $T_1 = T_1(\varepsilon)$ -періодичного розв'язку системи (17). Виконуючи заміну незалежної змінної $t = (1 + \varepsilon \alpha(\varepsilon))\tau$, $\alpha(0) = \alpha''$, $\alpha \in I_\alpha = [-\bar{\alpha}, \bar{\alpha}]$ перейдемо до задачі знаходження періодичного з відомим фіксованим періодом T розв'язку автономної системи

$$\begin{aligned} \frac{dx}{d\tau} &= Ax + \varepsilon h(x, \varepsilon), \quad \tau \neq \tau_i, \quad \tau_i = \tau_i(\varepsilon) = \frac{t_i}{1 + \varepsilon \alpha(\varepsilon)}, \\ \Delta x|_{\tau=\tau_i} &= B_i x(\tau_i - 0) + \varepsilon b_i(x(\tau_i - 0), \varepsilon), \\ h(x, \varepsilon) &= \alpha(\varepsilon)Ax + (1 + \varepsilon \alpha(\varepsilon))g(x, \varepsilon), \end{aligned} \quad (18)$$

який при $\varepsilon = 0$ перетворюється на один з T -періодичних розв'язків породжуючої для (18) лінійної однорідної системи

$$\frac{dx}{d\tau} = Ax, \quad \tau \neq \tau_i, \quad \Delta x|_{\tau=\tau_i} = B_i x(\tau_i - 0).$$

З метою відшукування T -періодичного розв'язку системи (18) побудуємо рекурентну послідовність

$$\begin{aligned} \bar{x}_{m+1}(\tau, \bar{\xi}) &= \bar{x}_0(\tau, \bar{\xi}) + \varepsilon \left\{ X(\tau, 0)G^* \left(\int_0^T U_2(T, s) h(\bar{x}_m(s, \bar{\xi}), \varepsilon) ds + \sum_{i=1}^p U_2(T, \tau_i + 0) b_i(\bar{x}_m(\tau_i - 0, \bar{\xi}), \varepsilon) \right) + \right. \\ &\quad + \int_0^\tau X(\tau)U(\tau, s) h(\bar{x}_m(s, \bar{\xi}), \varepsilon) ds - \int_\tau^T X(\tau)V(\tau, s) h(\bar{x}_m(s, \bar{\xi}), \varepsilon) ds + \\ &\quad \left. + \sum_{0 \leq \tau_i < \tau} X(\tau)U(\tau, \tau_i + 0) b_i(\bar{x}_m(\tau_i - 0, \bar{\xi}), \varepsilon) - \sum_{\tau \leq \tau_i < T} X(\tau)V(\tau, \tau_i + 0) b_i(\bar{x}_m(\tau_i - 0, \bar{\xi}), \varepsilon) \right\}, \\ \bar{x}_0(\tau, \bar{\xi}) &= X(\tau, 0)F\xi, \quad \bar{\xi} = (\xi, \varepsilon) \in R^{n+1}, \quad m = 1, 2, \dots \end{aligned}$$

Зрозуміло, що в області $(\tau, x, \alpha, \varepsilon) \in R \times D \times I_\alpha \times I_\varepsilon$ при достатньо малому ε_0 виконуються умови

C₁) існує не порожня множина $\bar{D} = D_1 \times I_\alpha \subset R^{k+1}$ точок $\bar{\xi} = (\xi, \alpha)$ така, що век-

тор-функція $\bar{x}_0(\tau, \bar{\xi})$ лежить в області $\bar{D}$ разом із своїм $\bar{\beta}$ -околом, де $\bar{M} = \bar{\alpha}|A|R + (1 + \varepsilon_0 \bar{\alpha})M$,

$$\bar{\beta} = \max_{\tau \in [0, T]} \left\{ \left(\int_0^T |X(\tau, 0)G^*U_2(T, s)| ds \bar{M} + \sum_{i=1}^p |X(\tau, 0)G^*U_2(T, \tau_i + 0)| M_i \right) + \int_0^T |X(\tau)U(\tau, s)| ds \bar{M} + \right. \\ \left. + \int_{\tau}^T |X(\tau)V(\tau, s)| ds \bar{M} + \sum_{0 \leq \tau_i < \tau} |X(\tau)U(\tau, \tau_i + 0)| M_i + \sum_{\tau \leq \tau_i < T} |X(\tau)V(\tau, \tau_i + 0)| M_i \right\};$$

D₁) $r(\bar{Q}) < 1$, де $r(\bar{Q})$ – спектральний радіус оператора $\bar{Q}x$, $\bar{K} = \bar{\alpha}|A| + (I + \varepsilon_0 \bar{\alpha})K$:

$$(\bar{Q}x)(\tau) = \int_0^T |X(\tau, 0)G^*U_2(T, s)| ds \bar{K} + \sum_{i=1}^p |X(\tau, 0)G^*U_2(T, \tau_i + 0)| K_i + \int_0^T |X(\tau)U(\tau, s)| ds \bar{K} + \\ + \int_{\tau}^T |X(\tau)V(\tau, s)| ds \bar{K} + \sum_{0 \leq \tau_i < \tau} |X(\tau)U(\tau, \tau_i + 0)| K_i + \sum_{\tau \leq \tau_i < T} |X(\tau)V(\tau, \tau_i + 0)| K_i,$$

а тому твердження, аналогічні теоремам 1 і 2 справедливі також для системи (18), а відповідні оцінки при цьому мають наступний вигляд:

$$|\bar{x}^*(\tau, \bar{\xi}, \varepsilon) - \bar{x}_m(\tau, \bar{\xi}, \varepsilon)| \leq \varepsilon_0 (\varepsilon_0 \bar{Q})^m (I - \varepsilon_0 \bar{Q})^{-1} \bar{M} r_1(\tau),$$

$$\inf_{\xi \in \partial(\bar{D})} |\bar{\Delta}_m(\bar{\xi})| > (\varepsilon_0 \bar{Q})^{m+1} (I - \varepsilon_0 \bar{Q})^{-1} T \bar{M},$$

$$\text{де } \bar{x}^*(\tau, \bar{\xi}, \varepsilon) = \lim_{m \rightarrow \infty} \bar{x}_m(\tau, \bar{\xi}, \varepsilon),$$

$$\bar{\Delta}(\bar{\xi}) \equiv \int_0^T \Psi^T(s) h(\bar{x}^*(s, \bar{\xi}), \varepsilon) ds +$$

$$+ \sum_{i=1}^p \Psi^T(\tau_i + 0) b_i(\bar{x}^*(\tau_i - 0, \bar{\xi}), \varepsilon),$$

$$\bar{\Delta}_m(\bar{\xi}) \equiv \int_0^T \Psi^T(s) h(\bar{x}_m(s, \bar{\xi}), \varepsilon) ds +$$

$$+ \sum_{i=1}^p \Psi^T(\tau_i + 0) b_i(\bar{x}_m(\tau_i - 0, \bar{\xi}), \varepsilon), \quad m = 0, 1, 2, \dots$$

Наступне твердження дозволяє робити висновки про існування періодичного розв'язку з аналізу наближеного визначального рівняння вже при $m = 0$.

Теорема 3. Нехай система (18) задовольняє умову A і відображення

$$\bar{\Delta}_0(\bar{\xi}) \equiv \int_0^T \Psi^T(s) h(\bar{x}_0(s, \bar{\xi}), \varepsilon) ds +$$

$$+ \sum_{i=1}^p \Psi^T(\tau_i + 0) b_i(\bar{x}_0(\tau_i - 0, \bar{\xi}), \varepsilon)$$

має в області $\bar{D}' \subset \bar{D}$, $\bar{D}' = D'_1 \times I'_\alpha$, $D'_1 \subset D_1$, $I'_\alpha \subset I_\alpha$ ізольовану особливу точку $\bar{\xi} = \bar{\xi}_0$ ненульового індексу.

Тоді існує ε_1 таке, що при всіх $0 < \varepsilon < \varepsilon_1$ система (17) має $T_1(\varepsilon) = (1 + \varepsilon \alpha^*)T$ - періодичний

розв'язок, який при $\varepsilon = 0$ перетворюється на один з T - періодичних розв'язків породжуючої лінійної однорідної системи (3).

Доведення аналогічне доведенню теореми 3 [4].

Одержані в даній роботі результати можуть бути використані при дослідженні існування і наближеній побудові періодичних розв'язків автономних систем диференціальних рівнянь з імпульсним впливом у критичному випадку.

Список використаних джерел

1. Бойчук А.А., Журавлев В.Ф., Чуйко С.М. Периодические решения нелинейных автономных систем в критических случаях// Укр. мат. журн. – 2005. – 52, № 9. – С. 1180–1186.
2. Бойчук А.А., Чуйко Е.В., Чуйко С.М. Периодические решения автономных систем с импульсным воздействием в критических случаях// Укр. мат. журн. – 1995. – 47, № 11. – С. 1478–1484.
3. Король І.І. Чисельно-аналітичний метод дослідження періодичних розв'язків диференціальних систем з імпульсною дією// Вісник КНУ, сер: Матем., механ.- 2004. Вип.13-14, с. 9-18.
4. Король І.І., Перестюк М.О. Ще раз про чисельно-аналітичний метод послідовних періодичних наближень А.М.Самойленка// Укр. мат. журн. – 2006. – 58, №4. – С. 472–489.
5. Boichuk A.A., Samoilenko A.M. Generalized Inver Бойчук А.А., Журавлев В.Ф., Чуйко С.М. Периодические решения нелинейных автономных систем в критических случаях// Укр. мат. журн. – 2005. – 52, № 9. – С. 1180–1186.se Operators and Fredholm Boundary Value Problems. VSP Utrecht Boston, 2004. – 318 p.
6. Samoilenko A.M., Perestyuk N.A. Impulsive differential equations. – Singapore, 1995. – 455 p.

Надійшла до редколегії 11.03.2008.

УДК 519.9

Євген Ю. Таран*, д.ф.-м.н., професор,
Роман Я. Кондрат, аспірант

РЕОЛОГИЯ РАЗВЕДЕННОГО РАСТВОРА ДЕФОРМИРУЕМЫХ МАКРОМОЛЕКУЛ У ТЕЧЕИ ПРОСТОГО ЗСУВУ

Вивчається реологічна поведінка розведеного розчину деформованих ланцюгових макромолекул Зімма у течії простого зсуву. Одержані результати використовуються для перевірки загальної реологічної моделі розведених розчинів деформованих ланцюгових макромолекул, побудованої авторами раніше.

Ключові слова: реологія, розведений розчин ланцюгових макромолекул, проста зсувна течія.

*E-mail: taran@univ.kiev.ua

Статтю представив проф. Мелешко В.В.

1 Вступ

У роботі вивчається реологічна поведінка у течії простого зсуву розведених розчинів деформованих ланцюгових макромолекул полімерів з низькомолекулярними розчинниками. Підвищений інтерес до таких розчинів виник тоді, коли структуру деформованих ланцюгових макромолекул почали вивчати реологічними методами [2, 5]. Суть методики таких досліджень полягає у порівнянні між собою значень ефективної в'язкості розведеного розчину макромолекул полімеру, які отримуються експериментально за допомогою віскозиметрів, зі значеннями, які обчислюються за допомогою відповідної математичної моделі. Така математична модель повинна визначати ефективну в'язкість розведеного розчину макромолекул полімеру, яка є макрореологічною характеристикою розчину, через мікрохарактеристики зв'язаних макромолекул.

Для моделювання деформованих ланцюгових макромолекул у даній роботі використовується перлинне намисто Роуза-Зімма [6, 7]. Реологічні характеристики розведених розчинів таких модельних макромолекул у течії простого зсуву досліджуються за допомогою визначальних рівнянь, отриманих у роботі [3].

2 Загальна реологічна модель розведених розчинів деформованих ланцюгових макромолекул

Співвідношення між характерними розмірами макромолекул полімерів з великою молекулярною вагою і молекул низькомолекулярних роз-

Ye.Yu. Taran*, D.Sci. (Phys.-Math.), Full Professor,
R.Ya. Kondrat, PhD student

RHEOLOGY OF DILUTE SOLUTION OF DEFORMABLE MACROMOLECULES IN A SIMPLE SHEARING FLOW

Rheological behaviour of dilute solution of Zimm deformable chain macromolecules in a simple shearing flow is studied. The obtained results are used for checking the general rheological model of deformable chain macromolecules constructed by authors earlier.

Key Words: rheology, dilute solution, chain macromolecules, simple shearing flow.

чинників дозволяє розглядати взаємодію розчинника зі зв'язаними у ньому макромолекулами як гідродинамічну [2, 5, 8]. Розчинник моделюється класичним континуумом з ньютонівським реологічним рівнянням. Найбільш досконалою гідродинамічною моделлю деформованих ланцюгових макромолекул є перлинне намисто Роуза-Зімма [6, 7] – ланцюг, який складається з n однакових послідовно з'єднаних вільнозчеплених сегментів. Кожний сегмент є гнучким субланцюгом, відстані між кінцями якого розподіляються за гауссовим законом. Гідродинамічна взаємодія модельної макромолекули з розчинником здійснюється бусинками (перлинками), які розташовані на кінцях сегментів модельного ланцюга і мають коефіцієнт ξ поступального тертя у ньютонівському розчиннику. Загальна кількість бусин такого ланцюга дорівнює $n+1$, вони нумеруються від 0 до n .

За таких припущень у роботі [3] побудована загальна реологічна модель розведених розчинів деформованих ланцюгових макромолекул з низькомолекулярними розчинниками. Напружений стан у довільних градієнтних течіях таких розчинів визначається рівнянням

$$T_{ij} = -p\delta_{ij} + 2\mu d_{ij} + N_0 \xi (d_{ik} \langle r_k^a r_j^a \rangle - \langle \dot{r}_i^a r_j^a \rangle). \quad (2.1)$$

У (2.1) T_{ij} – тензор напружень; p – тиск; δ_{ij} – символ Кронекера; μ – динамічна в'язкість розчинника; N_0 – кількість макромолекул в

одиниці об'єму розчину; d_{ij} – тензор швидкостей деформації, $d_{ij} = (1/2)(v_{i,j} + v_{j,i})$, де $v_{i,j}$ – тензор градієнта швидкості v_i розчинника; r_i^α – радіус-вектор, який визначає положення бусинки з номером α ; $\dot{r}_i^\alpha$ – яманівська похідна вектора r_i^α , $\dot{r}_i^\alpha = \dot{r}_i^\alpha - \omega_{ik} r_k^\alpha$, де крапка над r_i^α означає диференціювання за часом t ; ω_{ik} – тензор вихору швидкості, $\omega_{ik} = (1/2)(v_{i,k} - v_{k,i})$. Кутові дужки $\langle \rangle$ у рівнянні (2.1) означають осереднення у фазовому просторі координат бусинок модельної макромолекули за допомогою функції розподілу F , яка визначається рівнянням

$$\frac{\partial F}{\partial t} + \frac{\partial}{\partial r_i^\alpha} (\dot{r}_i^\alpha F) = 0. \quad (2.2)$$

У роботі [3] припускається, що конформація молекулярного ланцюга у градієнтних течіях розчину у загальному випадку може формуватися не тільки під дією гідродинамічних сил

$$g_i^\alpha = \xi(d_{ij} r_j^\alpha - \dot{r}_i^\alpha), \quad (2.3)$$

а й зовнішніх і/або внутрішньомолекулярних негідродинамічних сил. Загальне рівняння динаміки бусинок модельної макромолекули при цьому має вигляд [3]

$$\dot{r}_i^\alpha = (d_{ij} + \omega_{ij}) r_j^\alpha + \frac{1}{\xi} f_i^\alpha. \quad (2.4)$$

Для негідродинамічних сил при виведенні рівняння (2.4) враховується сумарно за допомогою результуючих векторів f_i^α .

У тензорних позначеннях, які використовуються у роботі, латинськими літерами $i, j, k, \dots$ позначаються нижні індекси. Вони можуть набувати значень 1, 2, 3, які відповідають осям Ox_1, Ox_2, Ox_3 прямокутної системи координат, початок якої збігається з центром мас макромолекули. Грецькими літерами $\alpha, \beta, \gamma, \dots$ позначаються верхні індекси, які можуть набувати значень від 0 до n , вони відповідають номерам бусин модельного молекулярного ланцюга. Кома у нижніх індексах означає диференціювання у напрямі координатної осі, яка позначена індексом, що йде за комою; за двома однаковими індексами одного доданку відбувається підсумовування у межах зміни індексу, що повторюється.

3 Стационарна течія простого зсуву розведеного розчину макромолекул Зімма

У цьому розділі визначальні рівняння (2.1), (2.2) і (2.4) використовуються для дослідження реологічної поведінки розведеного розчину макромолекул Зімма [1] у стаціонарній течії простого зсуву

$$v_x = Ky, \quad v_y = 0, \quad v_z = 0; \quad K = \text{const}. \quad (3.1)$$

Для цього врахуємо у загальному рівнянні динаміки молекулярного намиста (2.4) явний вигляд ефективних сил броунівської дифузії

$$\bar{f}_i^\alpha = -\frac{kT}{F} \frac{\partial F}{\partial r_i^\alpha}, \quad (3.2)$$

і внутрішньомолекулярних пружних сил

$$\bar{f}_i^\alpha = -2kT \mathfrak{A} A_{\alpha\beta} r_i^\beta, \quad (3.3)$$

які разом з гідродинамічними силами (2.3), за припущенням Зімма [1], впливають на конформацію і динаміку зв'язаних макромолекул і таким чином формують напружений стан у розчині. У (3.2), (3.3) k – стала Больцмана; T – абсолютна температура; $\mathfrak{A}$ – коефіцієнт внутрішньої пружності зв'язаних макромолекул, феноменологічний параметр, який характеризує макромолекули, що моделюються перлинним намистом; $A_{\alpha\beta}$ – матриця такого виду:

$$A_{\alpha\beta} = \begin{bmatrix} 1 & -1 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 \\ 0 & 0 & -1 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

Рівняння динаміки молекулярного намиста (2.4) при цьому стає

$$\dot{r}_i^\alpha = (d_{ij} + \omega_{ij}) r_j^\alpha - \frac{2kT}{\xi} \mathfrak{A} A_{\alpha\beta} r_i^\beta - \frac{kT}{\xi F} \frac{\partial F}{\partial r_i^\alpha}. \quad (3.4)$$

Врахування у (2.1) виразу для $\dot{r}_i^\alpha$, який отримується з (3.4), дозволяє знайти явний вигляд тензора напруження у довільних градієнтних течіях розведеного розчину пружних ланцюгових макромолекул Зімма

$$T_{ij} = -p\delta_{ij} + 2\mu d_{ij} + N_0 kT (2\mathfrak{A} A_{\alpha\beta} \langle r_i^\alpha r_j^\beta \rangle - n\delta_{ij}). \quad (3.5)$$

Осереднення у (3.5) здійснюється за допомогою функції розподілу, яка згідно з (2.2) і (3.4), є розв'язком рівняння

$$\frac{\partial F}{\partial t} - \left\{ \frac{kT}{\xi} \frac{\partial^2 F}{\partial r_i^\alpha \partial r_i^\alpha} - [(d_{lk} + \omega_{lk}) r_k^\alpha - \frac{2kT}{\xi} \mathfrak{A} A_{\alpha\beta} r_l^\beta] \frac{\partial F}{\partial r_i^\alpha} + \frac{12kTn}{\xi} \mathfrak{A} F \right\} = 0. \quad (3.6)$$

Дослідження рівнянь (3.5) і (3.6) показує, що їх використання значно спрощується після ортогонального перетворення координат r_i^α в η_i^α , при якому матриця $A_{\alpha\beta}$ приводиться до діагонального виду з головними значеннями λ_α .

При цьому $\lambda_0 = 0$, оскільки визначник матриці $A_{\alpha\beta}$ дорівнює нулю; $\sum_{\alpha=1}^n \lambda_\alpha = 2n$, як лінійний інваріант матриці $A_{\alpha\beta}$; $r_i^\alpha r_j^\alpha = \eta_i^\alpha \eta_j^\alpha$ із-за ортогональності перетворення координат. Після такого ортогонального перетворення координат рівняння (3.5), (3.6) набирають вигляду

$$T_{ij} = -p\delta_{ij} + 2\mu d_{ij} + N_0 kT (2\mathfrak{A} \sum_{\alpha=0}^n \lambda_\alpha \langle \eta_i^\alpha \eta_j^\alpha \rangle - n\delta_{ij}), \quad (3.7)$$

$$\frac{\partial F}{\partial t} - \sum_{\alpha=0}^n \left\{ \frac{kT}{\xi} \frac{\partial^2 F}{\partial \eta_i^\alpha \partial \eta_i^\alpha} - [(d_{lk} + \omega_{lk}) \eta_k^\alpha - \frac{2kT}{\xi} \mathfrak{A} \lambda_\alpha \eta_l^\alpha] \frac{\partial F}{\partial \eta_i^\alpha} + \frac{6kT}{\xi} \lambda_\alpha \mathfrak{A} F \right\} = 0. \quad (3.8)$$

Рівняння (3.8) є складним для розв'язання – його розв'язки можуть бути знайденими лише при малих градієнтах швидкості для окремих найпростіших течій у вигляді асимптотичних розвинень за малими градієнтами швидкості. Використання таким чином знайдених функцій розподілу при осередненні компонентів тензора напружень дозволяє знайти тільки їх наближені значення.

Тому у даній роботі рівняння (3.8) використовується для знаходження звичайних диференціальних рівнянь для моментів функції розподілу F

$$\frac{d\langle \eta_i^\alpha \eta_j^\alpha \rangle}{dt} = 2 \frac{kT}{\xi} \delta_{ij} + (d_{ik} + \omega_{ik}) \langle \eta_k^\alpha \eta_j^\alpha \rangle + (d_{jk} + \omega_{jk}) \langle \eta_k^\alpha \eta_i^\alpha \rangle - \frac{4kT}{\xi} \mathfrak{A} \lambda_\alpha \langle \eta_i^\alpha \eta_j^\alpha \rangle. \quad (3.9)$$

Рівняння (3.7) і (3.9) складають систему визначальних рівнянь для дослідження у розділах 3-5 реологічної поведінки розведених розчинів, що розглядаються, у течіях простого зсуву.

Згідно з (3.7) обчислення компонентів тензора напружень T_{ij} , що виникають у течії (3.1) розчину, потребує знання моментів $\langle \eta_i^\alpha \eta_j^\alpha \rangle$ функції розподілу F , яка визначається рівнянням (3.8).

З урахуванням умов симетрії

$$\langle \eta_i \eta_j \rangle = \langle \eta_j \eta_i \rangle, \\ (d_{ik} + \omega_{ik}) \langle \eta_k^\alpha \eta_j^\alpha \rangle = (d_{jk} + \omega_{jk}) \langle \eta_k^\alpha \eta_i^\alpha \rangle,$$

визначальне рівняння (3.9) дозволяє отримати систему звичайних диференціальних рівнянь для знаходження моментів $\langle \eta_i^\alpha \eta_j^\alpha \rangle$ функції розподілу F для течії (3.1) розчину

$$\begin{aligned} \frac{d\langle \eta_x^\alpha \eta_x^\alpha \rangle}{dt} &= -4 \frac{kT}{\xi} \mathfrak{A} \lambda_\alpha \langle \eta_x^\alpha \eta_x^\alpha \rangle + 2K \langle \eta_x^\alpha \eta_y^\alpha \rangle + 2 \frac{kT}{\xi}, \\ \frac{d\langle \eta_x^\alpha \eta_y^\alpha \rangle}{dt} &= -4 \frac{kT}{\xi} \mathfrak{A} \lambda_\alpha \langle \eta_x^\alpha \eta_y^\alpha \rangle + K \langle \eta_y^\alpha \eta_y^\alpha \rangle, \\ \frac{d\langle \eta_x^\alpha \eta_z^\alpha \rangle}{dt} &= -4 \frac{kT}{\xi} \mathfrak{A} \lambda_\alpha \langle \eta_x^\alpha \eta_z^\alpha \rangle + K \langle \eta_y^\alpha \eta_z^\alpha \rangle, \\ \frac{d\langle \eta_y^\alpha \eta_y^\alpha \rangle}{dt} &= -4 \frac{kT}{\xi} \mathfrak{A} \lambda_\alpha \langle \eta_y^\alpha \eta_y^\alpha \rangle + 2 \frac{kT}{\xi}, \\ \frac{d\langle \eta_y^\alpha \eta_z^\alpha \rangle}{dt} &= -4 \frac{kT}{\xi} \mathfrak{A} \lambda_\alpha \langle \eta_y^\alpha \eta_z^\alpha \rangle, \\ \frac{d\langle \eta_z^\alpha \eta_z^\alpha \rangle}{dt} &= -4 \frac{kT}{\xi} \mathfrak{A} \lambda_\alpha \langle \eta_z^\alpha \eta_z^\alpha \rangle + 2 \frac{kT}{\xi}. \end{aligned} \quad (3.10)$$

Для стаціонарної течії простого зсуву (3.1) похідні за часом у (3.10) дорівнюють нулю, тому система диференціальних рівнянь (3.10) обертається на систему лінійних алгебраїчних рівнянь, розв'язки якої мають вигляд

$$\langle \eta_x^\alpha \eta_z^\alpha \rangle = \langle \eta_y^\alpha \eta_z^\alpha \rangle = 0, \\ \langle \eta_x^\alpha \eta_y^\alpha \rangle = \langle \eta_z^\alpha \eta_z^\alpha \rangle = \frac{1}{2\mathfrak{A} \lambda_\alpha},$$

$$\langle \eta_x^a \eta_y^a \rangle = \frac{\xi}{kT} \cdot \frac{K}{8\alpha^2 \lambda_a^2}, \quad (3.11)$$

$$\langle \eta_x^a \eta_x^a \rangle = \frac{1}{2\alpha \lambda_a} + \left(\frac{\xi}{kT} \right)^2 \frac{K^2}{16\alpha^3 \lambda_a^3}.$$

Використання (3.11) у (3.7) з врахуванням того, що $\lambda_0 = 0$ і

$$\sum_{\alpha=1}^n \lambda_{\alpha} = 2n, \quad \sum_{\alpha=1}^n \frac{1}{\lambda_{\alpha}} = \frac{n^2}{6}, \quad \sum_{\alpha=1}^n \frac{1}{\lambda_{\alpha}^2} = \frac{n^4}{90},$$

дозволяє отримати компоненти тензора напружень T_{ij} у течії (3.1) розчину, що розглядається,

$$T_{xy} = \left(\mu + \frac{N_0 \xi n^2}{24\alpha} \right) K, \quad (3.12)$$

$$T_{xx} = T_{yy} = 0,$$

$$T_{zz} = -p + \frac{N_0 \xi^2 n^4}{720\alpha^2 kT} K^2, \quad (3.13)$$

$$T_{yy} = T_{zz} = -p. \quad (3.14)$$

Із співвідношення (3.12) випливає, що ефективна в'язкість розчину деформівних ланцюгових макромолекул Зімма

$$\mu_s^{-1} \equiv \frac{T_{xy}}{K} = \mu + \frac{N_0 \xi n^2}{24\alpha} \quad (3.15)$$

у простій зсувній течії (3.1) не залежить від швидкості зсуву K . Одержаний вираз (3.15) для ефективної зсувної в'язкості μ_s^{-1} співпадає з відповідним виразом, отриманим Зіммом [1] іншим шляхом.

Згідно з (3.13) і (3.14) розведений розчин деформівних ланцюгових макромолекул, який розглядається, виявляє ефект Вайссенберга – ненульову різницю нормальних напружень

$$T_{xx} - T_{zz} = \frac{N_0 \xi n^4}{720\alpha^2 kT} K^2. \quad (3.16)$$

Вона збільшується зі збільшенням швидкості зсуву K , тобто розведений розчин деформівних ланцюгових макромолекул у простій зсувній течії (3.1) веде себе як пружнов'язка рідина.

4 Реологія розчину після миттєвої зупинки течії простого зсуву

У цьому розділі досліджується реологічна поведінка розведеного розчину ланцюгових макромолекул Зімма після миттєвої зупинки простої зсувної течії (3.1). Припускається, що

при $t \leq 0$ розчин тече зі швидкістю (3.1), а при $t > 0$ – він знаходиться у стані спокою.

Після такої зупинки $d_{ij} + \omega_{ij} = 0$ для будь-яких i та j , а визначальне рівняння (3.9) приводить до наступної системи звичайних диференціальних рівнянь

$$\begin{aligned} \frac{d\langle \eta_x^a \eta_x^a \rangle}{dt} &= -4 \frac{kT}{\xi} \alpha \lambda_a \langle \eta_x^a \eta_x^a \rangle + 2 \frac{kT}{\xi}, \\ \frac{d\langle \eta_x^a \eta_y^a \rangle}{dt} &= -4 \frac{kT}{\xi} \alpha \lambda_a \langle \eta_x^a \eta_y^a \rangle, \\ \frac{d\langle \eta_x^a \eta_z^a \rangle}{dt} &= -4 \frac{kT}{\xi} \alpha \lambda_a \langle \eta_x^a \eta_z^a \rangle, \\ \frac{d\langle \eta_y^a \eta_y^a \rangle}{dt} &= -4 \frac{kT}{\xi} \alpha \lambda_a \langle \eta_y^a \eta_y^a \rangle + 2 \frac{kT}{\xi}, \\ \frac{d\langle \eta_y^a \eta_z^a \rangle}{dt} &= -4 \frac{kT}{\xi} \alpha \lambda_a \langle \eta_y^a \eta_z^a \rangle, \\ \frac{d\langle \eta_z^a \eta_z^a \rangle}{dt} &= -4 \frac{kT}{\xi} \alpha \lambda_a \langle \eta_z^a \eta_z^a \rangle + 2 \frac{kT}{\xi}. \end{aligned} \quad (4.1)$$

Течія розчину до її зупинки була стаціонарною течією простого зсуву (3.1), тому початкові умови для рівнянь (4.1) згідно з (3.11) матимуть вигляд

$$\begin{aligned} \langle \eta_x^a \eta_z^a \rangle_{t=0} &= \langle \eta_y^a \eta_z^a \rangle_{t=0} = 0, \\ \langle \eta_y^a \eta_y^a \rangle_{t=0} &= \langle \eta_z^a \eta_z^a \rangle_{t=0} = \frac{1}{2\alpha \lambda_a}, \\ \langle \eta_x^a \eta_y^a \rangle_{t=0} &= \frac{\xi}{kT} \cdot \frac{K}{8\alpha^2 \lambda_a^2}, \\ \langle \eta_x^a \eta_x^a \rangle_{t=0} &= \frac{1}{2\alpha \lambda_a} + \left(\frac{\xi}{kT} \right)^2 \frac{K^2}{16\alpha^3 \lambda_a^3}. \end{aligned}$$

Розв'язання відповідних задач Коші для рівнянь (4.1) дозволяє отримати такі моменти функції розподілу F

$$\begin{aligned} \langle \eta_x^a \eta_z^a \rangle &= \langle \eta_y^a \eta_z^a \rangle = 0, \\ \langle \eta_y^a \eta_y^a \rangle &= \langle \eta_z^a \eta_z^a \rangle = \frac{1}{2\alpha \lambda_a}, \\ \langle \eta_x^a \eta_y^a \rangle &= \frac{\xi}{kT} \cdot \frac{K}{8\alpha^2 \lambda_a^2} e^{-4 \frac{kT}{\xi} \alpha \lambda_a t}, \\ \langle \eta_x^a \eta_x^a \rangle &= \frac{1}{2\alpha \lambda_a} + \left(\frac{\xi}{kT} \right)^2 \frac{K^2}{16\alpha^3 \lambda_a^3} e^{-4 \frac{kT}{\xi} \alpha \lambda_a t}. \end{aligned} \quad (4.2)$$

З (4.2) випливає, що при $t \rightarrow \infty$ моменти функції розподілу F стають

$$\langle \eta_i^a \eta_j^a \rangle = 0 \text{ при } i \neq j, \\ \langle \eta_x^a \eta_x^a \rangle = \langle \eta_y^a \eta_y^a \rangle = \langle \eta_z^a \eta_z^a \rangle = \frac{1}{2\alpha\lambda_a}. \quad (4.3)$$

Використання (4.2) у (3.7) дозволяє отримати компоненти тензора напружень T_{ij} у розведеному розчині деформівних ланцюгових макромолекул Зімма після миттєвої зупинки його простої зсувної течії (3.1)

$$T_{xy} = \left(\mu + \frac{N_0 \xi}{4\alpha} \sum_{a=1}^n \frac{1}{\lambda_a} e^{-\frac{4kT}{\xi} \alpha \lambda_a t} \right) K, \\ T_{xz} = T_{yz} = 0, \\ T_{xx} = -p + \frac{N_0 \xi^2}{8\alpha^2 kT} \sum_{a=1}^n \frac{1}{\lambda_a^2} e^{-\frac{4kT}{\xi} \alpha \lambda_a t} \cdot K^2, \quad (4.4) \\ T_{yy} = T_{zz} = -p.$$

Із (4.4) випливає, що зі збільшення часу t (при $t \rightarrow \infty$) ефективна зсувна в'язкість

$$\mu_a^\perp \equiv \frac{T_{xy}}{K} = \mu + \frac{N_0 \xi}{4\alpha} \sum_{a=1}^n \frac{1}{\lambda_a} e^{-\frac{4kT}{\xi} \alpha \lambda_a t}$$

розведеного розчину деформівних ланцюгових макромолекул Зімма після миттєвої зупинки течії простого зсуву (3.1) зменшується, прямує до в'язкості розчинника, а різниця нормальних напружень

$$T_{xx} - T_{zz} = \frac{N_0 \xi^2}{8\alpha^2 kT} \sum_{a=1}^n \frac{1}{\lambda_a^2} e^{-\frac{4kT}{\xi} \alpha \lambda_a t} K^2,$$

яка є проявом ефекту Вайссенберга, прямує до нуля. При цьому спостерігається цілий спектр часів релаксації

$$t_a = \frac{\xi}{4\alpha kT \lambda_a}.$$

5 Реологія розчину після миттєвого увімкнення течії простого зсуву

У даному розділі досліджується реологічна поведінка розведеного розчину ланцюгових макромолекул Зімма після миттєвого увімкнення течії простого зсуву (3.1). Припускається, що при

$t \leq 0$ розчин знаходиться у стані спокою, а при $t > 0$ – він тече зі швидкістю (3.1).

Для дослідження такої течії слід розв'язати систему диференціальних рівнянь (3.10) з початковими умовами, які згідно з (4.3) мають вигляд

$$\langle \eta_x^a \eta_x^a \rangle_{t=0} = \langle \eta_y^a \eta_y^a \rangle_{t=0} = \langle \eta_z^a \eta_z^a \rangle_{t=0} = \frac{1}{2\alpha\lambda_a}, \\ \langle \eta_i^a \eta_j^a \rangle_{t=0} = 0 \text{ при } i \neq j.$$

Розв'язання відповідних задач Коші для диференціальних рівнянь (3.10) дозволяє отримати такі моменти функції розподілу F

$$\langle \eta_x^a \eta_z^a \rangle = \langle \eta_y^a \eta_z^a \rangle = 0, \\ \langle \eta_y^a \eta_y^a \rangle = \langle \eta_z^a \eta_z^a \rangle = \frac{1}{2\alpha\lambda_a}, \\ \langle \eta_x^a \eta_y^a \rangle = \frac{\xi}{kT} \cdot \frac{K}{8\alpha^2 \lambda_a^2} \left(1 - e^{-\frac{4kT}{\xi} \alpha \lambda_a t} \right), \quad (5.1) \\ \langle \eta_x^a \eta_x^a \rangle = \frac{1}{2\alpha\lambda_a} + \left(\frac{\xi}{kT} \right)^2 \frac{K^2}{16\alpha^3 \lambda_a^3} \left(1 - e^{-\frac{4kT}{\xi} \alpha \lambda_a t} \right).$$

Використання (5.1) у (3.7) дозволяє отримати компоненти тензора напружень T_{ij} у течії розчину деформівних ланцюгових макромолекул Зімма, яка виникає після миттєвого увімкнення течії простого зсуву (3.1)

$$T_{xy} = \left(\mu + \frac{N_0 \xi n^2}{24\alpha} - \frac{N_0 \xi}{4\alpha} \sum_{a=1}^n \frac{1}{\lambda_a} e^{-\frac{4kT}{\xi} \alpha \lambda_a t} \right) K, \\ T_{xz} = T_{yz} = 0, \quad (5.2) \\ T_{xx} = -p + \left(\frac{N_0 \xi^2 n^4}{720\alpha^2 kT} - \frac{N_0 \xi^2}{8\alpha^2 kT} \sum_{a=1}^n \frac{1}{\lambda_a^2} e^{-\frac{4kT}{\xi} \alpha \lambda_a t} \right) K^2, \\ T_{yy} = T_{zz} = -p.$$

Слід зазначити, що ефективна зсувна в'язкість суспензії $\mu_a^\perp$

$$\mu_a^\perp \equiv \frac{T_{xy}}{K} = \mu + \frac{N_0 \xi n^2}{24\alpha} - \frac{N_0 \xi}{4\alpha} \sum_{a=1}^n \frac{1}{\lambda_a} e^{-\frac{4kT}{\xi} \alpha \lambda_a t}$$

течії, яка виникає після миттєвого увімкнення течії простого зсуву, збільшується зі збільшенням часу t . При $t \rightarrow \infty$ $\mu_a^\perp$ досягає свого стаціонарного значення (3.15).

Аналогічно різниця нормальних напружень

$$T_{xx} - T_{zz} = \frac{N_0 \xi^2 n^4}{720 \alpha^2 kT} - \frac{N_0 \xi^2}{8 \alpha^2 kT} \sum_{\alpha=1}^n \frac{1}{\lambda_{\alpha}^2} e^{-4 \frac{kT}{\xi} \alpha \lambda_{\alpha} t}$$

збільшується зі збільшенням часу t і досягає при $t \rightarrow \infty$ свого найбільшого значення (3.16).

6 Висновки

1. У даній роботі здійснено перевірку побудованої у [3] загальної реологічної моделі розведених розчинів деформівних ланцюгових макромолекул. Так порівняння виразу (3.15)

ефективної в'язкості μ_{σ} розведеного розчину деформівних ланцюгових макромолекул Зімма у простій зсувній течії (3.1), знайденим за допомогою визначальних рівнянь (2.1), (2.2) і (2.4) цієї моделі, з відповідним виразом, одержаним Зіммом [1] іншим шляхом, засвідчує правильність цих рівнянь.

2. Тензорний вид визначального рівняння (2.1) дозволив отримати у даній роботі новий, порівняно з [1], результат, а саме – виявлено ефект Вайссенберга, який є проявом пружної поведінки розведеного розчину деформівних ланцюгових макромолекул Зімма у простій зсувній течії. Це дозволяє говорити про розведений розчин деформівних ланцюгових макромолекул Зімма, як про пружнов'язке рідке середовище.

3. Тензорний вид визначального рівняння (2.1) також дає можливість дослідження реологічної поведінки розведеного розчину деформівних ланцюгових макромолекул у довільних градієнтних течіях. Зокрема це стосується дослідження ефективної в'язкості такого розчину. Згідно з [4] ефективна в'язкість розведеного розчину деформівних ланцюгових макромолекул у його довільних градієнтних течіях може бути визначена за формулою

$$\mu_{\sigma} = \frac{T_y d_{ij}}{2 d_{km} d_{km}}.$$

4. Загальний вид рівняння (2.2) для функції розподілу елементів ланцюгової макромолекули під дією гідродинамічних і внутрішньомолекулярних сил, а також ефективних сил броунівської дифузії дозволяє досліджувати за допомогою рівнянь (2.1), (2.2), (2.4) нестационарну реологічну поведінку розчину. Так дослідження у розділах 4 і 5 даної роботи нестационарних реологічних характеристик розведеного розчину макромолекул Зімма

показує внутрішню узгодженість отриманих результатів. А саме, підтверджується асимптотична рівність (при $t \rightarrow \infty$) нестационарних реологічних характеристик цього розчину з відповідними стаціонарними характеристиками, отриманими у розділі 3. Цей факт також підтверджує правильність визначальних рівнянь (2.1), (2.2), (2.4), які використовувались для досліджень у даній роботі.

Список використаних джерел

1. Зимм Б. Динамика полимерных молекул в разбавленном растворе: вязкоупругие свойства, двойное лучепреломление и диэлектрические потери // Физика полимеров: Сборник / Под. ред. М.В. Волькенштейна. – М.: ИЛ, 1960. – С. 379-411.
2. Моравец Г. Макромолекулы в растворе. – М.: Мир, 1967. – 399с.
3. Придатченко Ю.В., Таран Є.Ю., Кондрат Р.Я. Структурно-феноменологічна теорія напруженого стану у градієнтних течіях розведених розчинів полімерів з деформівними ланцюговими макромолекулами // Укр. фіз. журн. – 2006. – 51, №4. – С. 374-380.
4. Таран Є.Ю. Структурно-феноменологіческая реология разбавленных суспензий: Дис. ... доктора физ.-мат. наук: 01.02.05. – Харьков, 1994. – 346с.
5. Цветков В.Н., Эскин В.Е., Френкель С.Я. Структура макромолекул в растворах. – М.: Наука, 1964. – 720с.
6. Beris A.N., Edwards B.J. Thermodynamics of Flowing Systems with Internal Microstructure. – New York: Oxford University Press, 1994. – 683p.
7. Bird R.B., Curtiss C.F., Armstrong R.C., Hassager O. Dynamics of Polymeric Fluids. V.2. – New York: Wiley, 1987. – 437p.
8. Kirkwood J.G. Macromolecules. – New York: Gordon and Breach, 1967. – 294p.

Надійшла до редколегії 15.03.2008

КОМП'ЮТЕРНІ НАУКИ ТА ІНФОРМАТИКА

УДК 518.9

Grigoryeva L.V., postgraduate student

Maple-exploring of a Free Flywheel Suspended by Superconductive Bearing

The software Maple as a powerful tool to analyze complicated non-linear dynamic systems is used to explore parametrical conditions guaranteeing the stability of resting or rotating free magnetically levitated flywheel's rotor, to solve the Cauchy problem and to build the phase portraits. The dynamical model is derived on the basis of analytical electromechanics with taking into account six degrees of freedom of a free rotor, the constancy of full magnetic fluxes coupled with immobile and rotor's superconducting rings, Lyapunov stability theorems, and Euler equations of a free body dynamics.

Key words: superconducting levitation, free flywheel, dynamics.

Григор'єва Л.В., аспірантка

Maple- дослідження вільного ротора на надпровідних підшипниках

Система комп'ютерної математики Maple як засіб аналізу складних нелінійних динамічних систем застосована для дослідження параметричних умов, що гарантують стійкість вільного магнітно левітуючого ротора у стані рівноваги або при обертанні, для розв'язання задачі Коші і для побудови фазових портретів. Динамічна модель отримана на основі аналітичної електромеханіки з врахуванням шести степенів свободи вільного ротора, сталості повного магнітного потоку, зв'язаного з нерухомими надпровідними кільцями та кільцями ротора, теорем Ляпунова про стійкість та динамічних рівнянь Ейлера для вільного тіла.

Ключові слова: надпровідна левітація, вільний маховик, динаміка.

E-mail: l.v.grigoryeva@gmail.com

Статтю представив член-кор. НАНУ, д. фіз.-мат. н., проф. Ляшко С.І.

I. INTRODUCTION

Magnetic bearings become a very important element for renewable energy sources, end-use energy efficiency, environmentally preferred advanced generation, and flywheel energy storage. As it is expected, devices using magnetic bearings can be widely implemented, from automobile or helicopter engines to power reserve flywheels to level peak energy consumptions.

We developed a new type of magnetic bearings based on the "Magnetic Potential Well" (MPW) phenomenon [1, 2]. For two constantly oriented closed superconducting loops, the MPW-manifestation signifies that with nearing of these spaced loops, their magnetic attractive force does not increase as usually but decreases, becomes zero and changes into the repulsive magnetic force before the spacing between loops equals zero. A similar picture can be observed in a permanent magnet-closed zero resistant loop pair. The MPW can be also realized in a system with many magnets.

Magnetic bearings have many advantages over ball, gas, and hydro bearings: practically unlimited operating time, absence of a lubricant, simplicity and reliability of operation, etc. But many applications of

magnetic bearings require solving some problems. One of them is the stability of the free resting or rotating rotor. Trials to suspend a body in the free equilibrium under action of magnetic forces are fruitless on the basis of Earnshaw's theorem [3]. There are only three exceptions not covered by this theorem. One of them is automation allowing transforming an unstable dynamic system into the stable system by e.g. feed-back control. The second is diamagnetism of a substance particularly bulk superconductors [4 - 6]. And finally, the MPW based on zero resistance of a closed current carrying loop [1, 2]. It should be noted that any of these exceptions assists but doesn't guarantee the free body rest or rotation stability.

Not only stability but also centring force levels are important in bearing applications. The stiffness of magnetic bearing must be on the level of commercial bearings operating with radial stiffness no less than 10^6 - 10^7 N/mm. Possibilities to satisfy these levels by known magnetic bearings do not look optimistic. Really, the top pressure of "warm" magnetic bearings with automation is restricted from above app. by $2 \cdot 10^6$ N/m² determined by the magnetic saturation induction of 2-2.5 Tesla. For characteristic bearing

size of 0.1 m it gives magnetic bearing stiffness of $2 \cdot 10^5$ N/m that corresponds with published data [7].

The top pressure of 10^3 - 10^4 N/m² for passive superconducting bearings based on the ideal diamagnetism is restricted by the first critical magnetic field, which is less than 0.2 Tesla for all known superconductors. This is too low value to be estimated as practicable.

In contrast to known magnetic bearings, instead of relying on the superconductor diamagnetism or automatic control to keep the air gap within required limits, we propose to use the MPW phenomenon. Manifestation of this phenomenon requires ideal electric conductivity in a closed loop of any shape. This allows using existing high current density superconductors e.g. niobium-titanium, niobium-germanium, and niobium-tin joining. Superconducting magnets using these superconductors and operating in the persistent current mode are in abundance. They are capable of generating constant magnetic field no less than 10-20 Tesla without electric losses that is unreachable in all other magnetic bearings. Before MPW these magnets were used as high magnetic field sources only. Now their applications can be MPW-bearings with pressure of 10^8 N/m² and stiffness of 10^7 N/m.

II. FLYWHEEL MODEL

A. Coordinate Systems

A flywheel is modeled by two stator's coaxial immobile superconductive rings magnetically interacting with two sets of N small planar superconductive circular loops (dipoles) equally shifted in relation to the rotor's axis (see Figure 1). The MPW is realized in a set of N dipoles and the nearer superconductive immobile ring. It means that at some rotor/stator coaxiality position, the electric current in each dipole is zero because its full magnetic flux (magnetic linkage) that is constant on the basis of the Faraday's electromagnetic induction law is created by the nearer immobile ring, yet at any another rotor/stator geometrical configuration this current is non-zero to satisfy the magnetic linkage constancy.

To determine the potential energy of magnetic interaction, it is convenient to introduce some trihedrons. The origin O of the immobile trihedron $Oxyz$ with unit vectors $\vec{i}_1, \vec{i}_2, \vec{i}_3$ is the symmetry center of two coaxial superconductive rings of radius a , and the vertical rings' symmetry axis Oz is parallel to the gravity force.

The origin O_1 of the second trihedron $O_1x_1y_1z_1$ with unit vectors $\vec{i}'_1, \vec{i}'_2, \vec{i}'_3$ coincides with a free rotor mass center so that axis O_1z_1 is directed along the rotor's symmetry axis. As an example, the case $N=4$ is shown in Figure 1. The free rotor is described by six degrees of freedom. The first three are Cartesian $Oxyz$ -coordinates x, y, z of the point O_1 . Non-dimensional linear coordinates X, Y , and Z are derived by division of a corresponding Cartesian $Oxyz$ -coordinate by value a . The other three degrees of freedom are the Euler-Krylov angles x_4, x_5, x_6 determining the rotor space orientation and x_4 is rotation relatively axis O_1x_1 , x_5 is rotation relatively axis O_1y_1 , and x_6 is rotation relatively axis O_1z_1 at that.

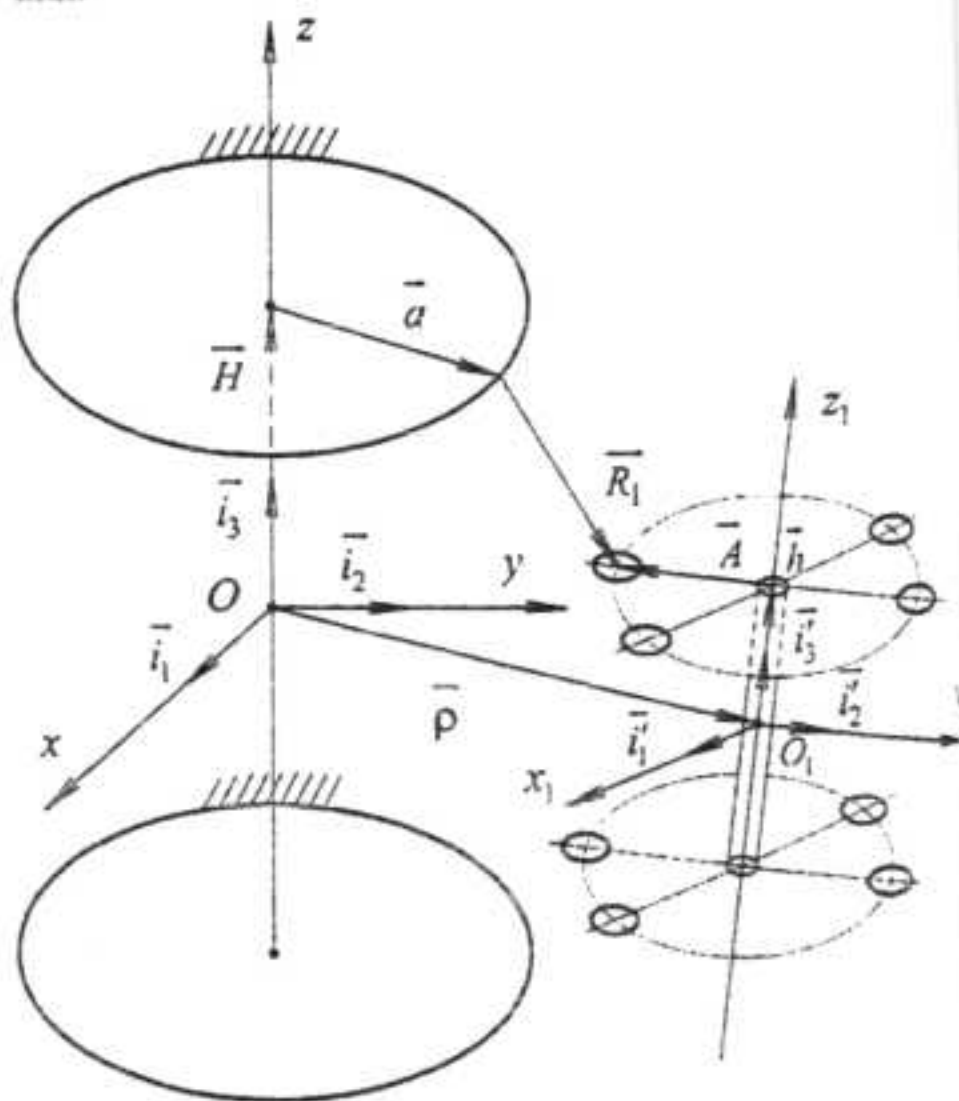


Fig. 1. Sketch of the flywheel.

B. Potential Energy of Gravity and Magnetic Interactions

The potential energy U of the flywheel dynamic system consists of two parts. One is the gravity represented by the rotor's constant weight force. The second is the sum of magnetic interactions of each j -immobile superconducting ring ($j=1,2$) and N dipoles of upper or lower part of the rotor. Magnetic interactions between dipoles are ignored as infinitesimal. Dipoles are supposed to be zero resistance closed loops with taking into account the constancy of their magnetic linkages as a consequence of zero resistance of each dipole's loop.

The MPW-position that is at our disposal corresponds to the coincidence of axes Oz and O_1z_1 . At this coaxiality position, the rotor's weight $G = mg$ where m is its mass and g is the gravity acceleration can be balanced at $Z = 0$ by the difference in attractive forces of the upper part of superconducting loops (one immobile and N on the rotor top) directed above and the lower similar part directed downward. Such coaxial equilibrium can be accomplished e.g. by the adjusted parameter $k = \Psi_2 \Psi_1^{-1} = const > 1$ that is ratio between the upper immobile superconducting ring magnetic linkage Ψ_2 and lower ring linkage Ψ_1 .

Below we use the matrix M determining a relative space orientation of the trihedrons $Oxyz$ and $O_1x_1y_1z_1$ (see e.g. [8], p.119) in the expanded form

$$M = \begin{pmatrix} c_5c_6 & c_4s_6 + s_4s_5c_6 & s_4s_6 - c_4s_5c_6 \\ -c_5s_6 & c_4c_6 - s_4s_5s_6 & s_4c_6 + c_4s_5s_6 \\ s_5 & -s_4c_5 & c_4c_5 \end{pmatrix} \quad (1)$$

where $c_r = \cos x_r$, $s_r = \sin x_r$, and elements of the first row present projections of the unit vector on axis O_1x_1 etc.

Introducing radius vector $\vec{\rho}$ of the point O_1 , vectors $\vec{h}, \vec{A}, \vec{R}_1, \vec{a}, \vec{H}$ forming the closed polygon for the upper part of magnetically interacting loops (see Figure 1) and similar vector polygon for the lower part of magnetically interacting loops (not depicted in Figure 1), and using (1) for the components of the vectors connecting arc's elements of rings and dipole centers after transformations, one can obtain

$$\begin{aligned} R_{j1} &= X - (-1)^j h a_{31} + A(a_{11} \cos t_{ij} + a_{21} \sin t_{ij}), \\ R_{j2} &= Y - (-1)^j h a_{32} + A(a_{12} \cos t_{ij} + a_{22} \sin t_{ij}), \\ R_{j3} &= Z - (-1)^j h a_{33} + A(a_{13} \cos t_{ij} + a_{23} \sin t_{ij}) + (-1)^j H \end{aligned} \quad (2)$$

where index $j = 1$ is applied to the upper and $j = 2$ to the lower parts of loops, parameter h is a non-dimensional half-distance between rotor's loops, H is the same for the immobile rings, and the angle t_{ij} together with non-dimensional shift A determine a dipole-place on the rotor.

The mutual inductance between a ring and a dipole can be derived as the magnetic flux generated by the unit electric current in the corresponding ring and piercing this dipole area. Assuming that each dipole plane is perpendicular to the rotor axis and its area equals S , taking into account known formulae for the ring magnetic induction (see e.g. [9]) and (2),

for the non-dimensional ring-dipole mutual inductance the following formula takes place

$$L_j = (b_{j1} (a_{31} R_{j1} + a_{32} R_{j2}) + b_{j2} a_{33}) \quad (3)$$

where

$$\begin{aligned} b_{j1} &= \frac{R_{j3}}{r_{j1}} \left(-K(k_j) + \frac{1 + r_{j1}^2 + R_{j3}^2}{(1 - r_{j1})^2 + R_{j3}^2} E(k_j) \right) \\ b_{j2} &= K(k_j) + \frac{1 - r_{j1}^2 - R_{j3}^2}{(1 - r_{j1})^2 + R_{j3}^2} E(k_j) \end{aligned} \quad (4)$$

$$r_{j1} = \sqrt{R_{j1}^2 + R_{j2}^2}$$

where $K(k_j)$, $E(k_j)$ are full elliptic integrals of the modulus

$$k_j = 2\sqrt{r_{j1}((1 + r_{j1})^2 + R_{j3}^2)^{-1}}. \quad (5)$$

Any non-dimensional ring-dipole inductance L_0 corresponding to the MPW-position is determined by zero values of all non-dimensional degrees of freedom ($X = Y = Z = x_4 = x_5 = x_6 = 0$) in (3).

On the basis of analytical electromechanics [10], the magnetic energy represented as a function of the magnetic linkages and mechanical coordinates (degrees of freedom) is the potential energy of an electromechanical dynamic system. Therefore, with taking into account the said suppositions and smallness in sizes of dipoles loops, one can derive the formula for the potential energy u in the non-dimensional form

$$u = UU_0^{-1} = \sum_{j=1}^2 \frac{\Psi_j^2}{\Psi_2^2} \sum_{i=1}^N (L_0 - L_{ij})^2 + U_1 Z \quad (6)$$

where characteristic magnetic energy U_0 and non-dimensional gravity energy U_1 are respectively

$$U_0 = \frac{1}{L} \left(\frac{\mu_0 S}{2a\pi} \frac{\Psi_2}{L_0} \right)^2, \quad U_1 = mg a L \left(\frac{\mu_0 S}{2a\pi} \frac{\Psi_2}{L_0} \right)^{-2}. \quad (7)$$

III. THE STABILITY PROBLEM

One of essential problems is the stability problem for either equilibrium or spinning of a levitated rotor. This problem can be investigated without corresponding dynamic equations analysis on the basis of the Lyapunov's theorem about partial stability [11]. This problem is reduced to finding the positivity conditions for the potential energy expanded into the Taylor series in corresponding variables. In our case using Maple [12], the Taylor series for $N=4$ dipole can be written as

$$u = B_1(X^2 + Y^2) + B_2Z^2 + B_3(x_4^2 + x_5^2). \quad (8)$$

where expressions for values B_i ($i=1, \dots, 3$) are too cumbersome functions of the geometrical parameters and here are omitted. In accordance with Lyapunov's theorem, the sufficient conditions of the free equilibrium stability are equivalent to positive all parameters B_i . As simulations show, this requirement can be satisfied by a relevant choice of the system parameters.

IV. THE DYNAMICS MODEL

The potential energy determined above and Euler equations of a free body dynamics [8] allow deriving the starting equations of the rotor motion in the non-dimensional form (function u is determined by (6))

$$\begin{aligned} \frac{d^2 X}{dt^2} &= -A_1 \frac{\partial u}{\partial X}, \\ \frac{d^2 Y}{dt^2} &= -A_1 \frac{\partial u}{\partial Y}, \\ \frac{d^2 Z}{dt^2} &= -A_1 \frac{\partial u}{\partial Z}, \\ \frac{dn_1}{dt} + k_1 n_3 n_2 &= -A_2 \frac{\partial u}{\partial x_4}, \\ \frac{dn_2}{dt} - k_1 n_1 n_3 &= -A_2 \frac{\partial u}{\partial x_5}, \\ \frac{dn_3}{dt} &= 0, \\ n_1 &= \cos x_5 \cos x_6 \frac{dx_4}{dt} + \sin x_6 \frac{dx_5}{dt}, \\ n_2 &= -\cos x_5 \sin x_6 \frac{dx_4}{dt} + \cos x_6 \frac{dx_5}{dt}, \\ n_3 &= \sin x_5 \frac{dx_4}{dt} + \frac{dx_6}{dt}. \end{aligned} \quad (9)$$

Here non-dimensional parameter $A_1 = U_0 (ma^2 w^2 L)^{-1}$ is ratio between magnetic characteristic energy U_0 and characteristic kinetic energy (w is the characteristic angular velocity); non-dimensional parameter $A_2 = U_0 (I_1 w^2 L)^{-1}$ determines ratio between magnetic characteristic energy U_0 and rotating kinetic energy with inertia moment I_1 relatively axis $O_1 x_1$ or $O_1 y_1$; non-dimensional parameter $k_1 = (I_3 - I_1) I_1^{-1}$ is ratio of the rotor main central inertia moments (I_3 is the inertia moment relatively axis $O_1 z_1$); n_1, n_2, n_3 are non-dimensional angular velocity components on

axis $O_1 x_1, O_1 y_1$, and $O_1 z_1$ respectively, and t is the non-dimensional time.

Expressions for partial derivatives of the non-dimensional potential energy u in (9) are too cumbersome and therefore omitted.

The conservative 12th-order system (9) with potential energy determined above is too complicated to be analyzed by known analytical methods. In this case the software Maple11 [12] is more relevant.

We have developed Maple-based tools that give full realization of the free rotor dynamics. As an example, we represent numerical solutions for the Cauchy problem and phase portraits building. The plots of solutions and phase portraits shown in Figures 2-5 depict the case of assumed $H - h = 0.1$, $A_1 = 1$, $A_2 = 1$, $k_1 = 0$, $U_1 = 0$ and the following initial conditions: $X(0) = 0.05$, $\dot{X}(0) = 0$, $Y(0) = 0$, $\dot{Y}(0) = 0.01$, $Z(0) = 0.01$, $\dot{Z}(0) = 0$, $x_4(0) = 0$, $x_5(0) = 0$, $x_6(0) = 0$, $n_1(0) = 0$, $n_2(0) = 0$, $n_3(0) = 100$.

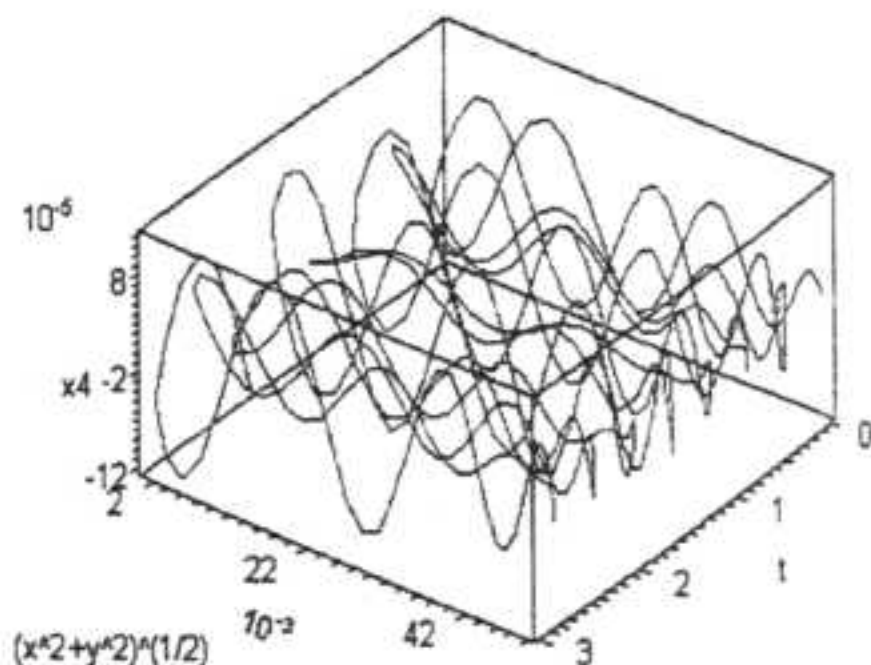


Fig. 2. Phase portrait $t, \sqrt{X(t)^2 + Y(t)^2}, x_4(t)$.

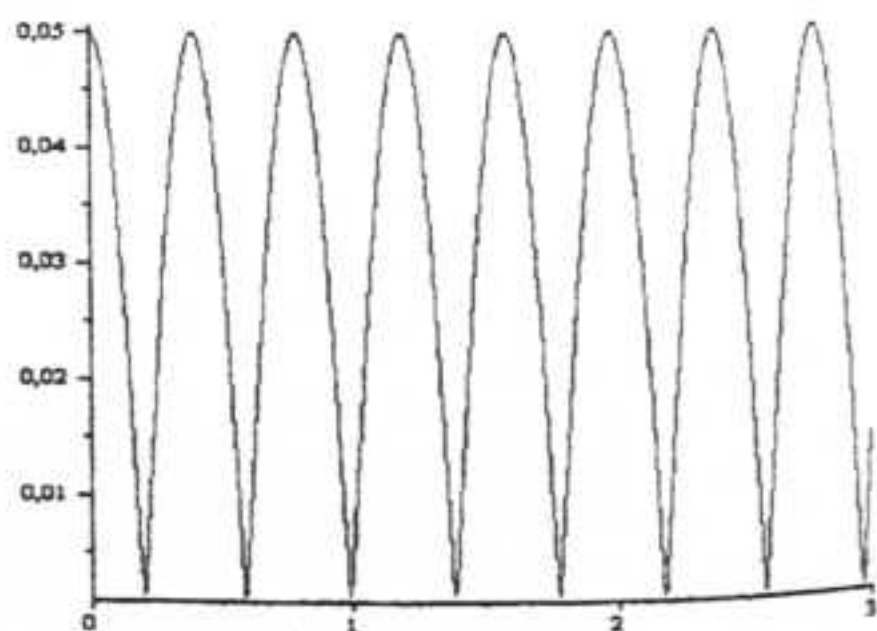


Fig. 3. Radial separation versus time.

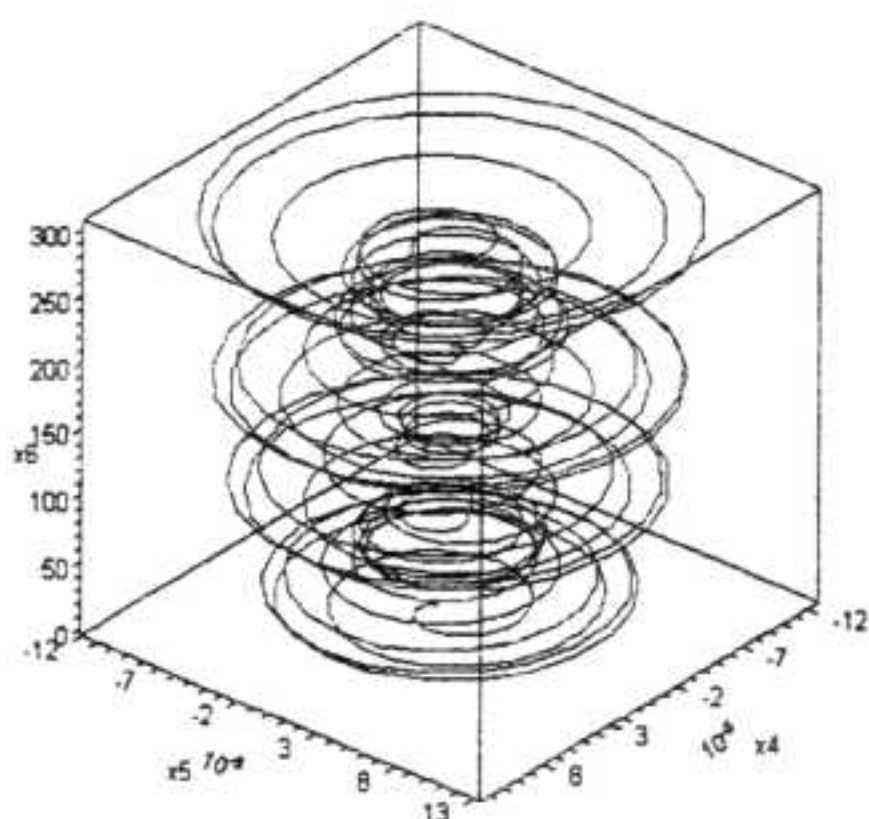


Fig. 4. Phase portrait of x_4, x_5, x_6 .

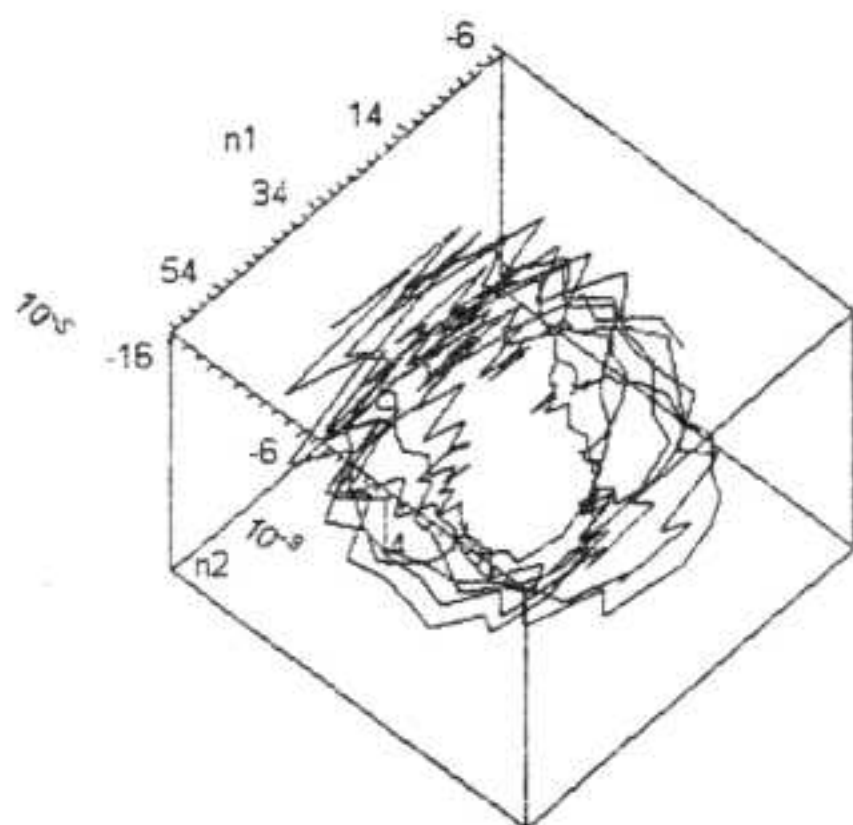


Fig. 5. Phase portrait of n_1, n_2, n_3 .

V. CONCLUSIONS

The free rotor dynamics problem is considered in the parts of mathematical model construction, substantiation of the free equilibrium stability (i.e. levitation), and dynamic analysis on the basis of the modern computer tools. The mathematical model taking into account six degrees of freedom proved to be complicated non-linear conservative dynamic system of the 12th-order, an analytically unsolvable problem.

The free equilibrium stability i.e. levitation derived on the Lyapunov's theorem is possible by the MPW-manifestation and relevant choice of the magnetic and geometrical parameters.

As examples demonstrate, the software Maple proved to be the effective computer tool to analyze

the free rotor dynamics. This approach can be useful to develop superconducting bearings [13-16].

VI. ACKNOWLEDGEMENT

I would like to express acknowledgment to prof. Vasyl' Kozoriz. The preparation of this paper would not have been possible without his support and advices and patents by his authority.

References

- [1] Kozoriz, V.V.: Dynamic Systems of Free Magnetically Interacting Bodies, Naukova Dumka, - Kyiv, 1981. (in Russian).
- [2] Kozoriz, V.V. www.maglev2006.de
- [3] Earnshaw, S.: On the Nature of Molecular Forces which Regulate the Constitution of Luminiferous Ether, Trans. Cambridge Phil. Soc. 7, 1842, pp. 97-112.
- [4] Braunbeck, W.: Freischwebende Körper in Elektrischen und Magnetischen Feld, Zeitschrift für Physics, 112 (11-12), 1939, pp.753-763.
- [5] Braunbeck, W.: Freies Schweben Diamagnetischen Körper in Magnetfeld, Zeitschrift für Physics, 112, 1939, pp.764-769.
- [6] Moon, F.C.: Superconducting Levitation Applications to Bearings and Magnetic Transportation, John Willey & Sons - New York, 1994.
- [7] Schweitzer, G., ed.: Magnetic Bearings, Springer-Verlag - Berlin, 1988.
- [8] Pars, L.A.: A Treatise on Analytical Dynamics, Heinemann - London, 1968.
- [9] Smyth, N.R.: Static and Dynamic Electricity, 3rd ed., McGraw-Hill - New York, 1968.
- [10] White, D.C.; Woodson, H.H.: Electromagnetic Energy Conversion, John Willey & Sons, Inc. - New-York, 1959.
- [11] Rumyantsev, V.V.: On Motion Stability Relatively the Part of Variables. Bulletin of Moscow University, Math., 4, 1957, pp. 9-16.
- [12] www.maplesoft.com
- [13] Kozoriz, V.: Super Conductive Bearing. The US-patent #6,608,417, B1; 8/2003.
- [14] Kozoriz, V.: Super Conductive Bearing. The US-patent #6,703,737, B2; 3/2004.
- [15] Kozoriz, V.: Super Conductive Bearing. The US-patent #6,770,994, B2; 8/2004.
- [16] Kozoriz, V.: Super Conductive Bearing. The US-patent # 6,856,060, B2; 2/2005.

Надійшла до редколегії 12.06.08

УДК 004.7

Антонюк М. А.

Програмні рішення Microsoft для навчальних закладів

В роботі розглянуто можливості навчальної програмної оболонки Microsoft Learning Gateway. На основі проведеного аналізу формулюються рекомендації що до впровадження розглянутої системи електронного навчання

Ключові слова: Class Server, SCORM, навчальний процес

Статтю представив д.т.н., проф. Гаращенко Ф.Г.

Загальновідомим є той факт, що процес навчання – це робота. Однак якщо для підвищення ефективності праці менеджера або бухгалтера ринок пропонує велику кількість рішень, то аналогічних завершених комплексних рішень для автоматизації робочих місць викладачів і учнів практично немає.

На ринку програмного забезпечення для освітніх установ домінують найпростіші системи керування навчанням (СКН), що забезпечують наступним функціоналом:

- створення й зберігання навчальних матеріалів;
- видача учням завдань для ознайомлення з навчальними матеріалами й для проходження тестів;
- контроль успішності учнів шляхом складання звітів по отриманим учнями оцінкам.

Основними недоліками таких систем є неможливість ефективної роботи з великою кількістю користувачів і їх абсолютно закрита архітектура. Навчальні матеріали зберігаються у форматах, найчастіше не документованих виробником СКН. Однак навіть якщо формат і документований, нерідко він не є стандартом і не сумісний з форматами зберігання навчальних матеріалів в інших системах. Отже, якщо користувач захоче змінити систему керування навчанням, йому доведеться витратити чимало часу й праці на конвертацію навчальних матеріалів з одного формату в інший. Також досить складно, а в багатьох випадках просто неможливо розширити функціонал СКН або інтегрувати його з іншим програмним забезпеченням, наприклад з бухгалтерією. Причина у відсутності відкритих і документованих інтерфейсів програмування.

M. A. Antoniuk

Microsoft software for educational institutions

Facilities of Microsoft Learning Gateway programm shell were considered in the work. Based on carried out analysis recommendations concerning implementation of considered electronic education system is formulated.

Key words: Class Server, SCORM, education process

У кожному навчальному процесі беруть участь три види користувачів: викладачі, учні й наставники. Викладачі створюють і вдосконалюють навчальні матеріали, видають і перевіряють завдання, керують процесом навчання. Учні виконують завдання, переглядаючи навчальні матеріали або відповідаючи на питання тестів. Їм необхідно створити найбільш комфортні для навчання умови – можливість навчатися де завгодно (у навчальному закладі, вдома, в Інтернет-клубі і т.д.), швидко одержувати консультації у викладачів по незрозумілим для них питанням, мати доступ до електронних бібліотек з додатковою літературою. Наставниками можуть бути як батьки (у випадку вищих і середніх навчальних закладів), так і менеджери, що направляють своїх підлеглих на навчання (у випадку комерційних навчальних центрів і навчальних відділів у комерційних компаніях). Наставник завжди може контролювати, які завдання видаються його підопічному і які оцінки були поставлені за виконані завдання (рис. 1).

Також не можна забувати про найважливіший фактор – комунікації між всіма учасниками освітнього процесу. Викладач спілкується з учнями, з одного боку, оперативно відповідаючи на їхні питання, а з іншого боку – задаючи свої власні питання для додаткового контролю розуміння учнями навчального матеріалу. Також спрощується контакт між викладачем і наставником, тому що тепер учителю немає необхідності вимагати від школяра з'явитися завтра з батьком – всю інформацію про погане поведіння недбайливого учня можна з легкістю донести до батьків.

Існують три основних види комунікації:

- класична електронна пошта.
- дискусійні форуми, які дозволяють організувати обговорення між багатьма

учасниками, що може бути корисно при ухваленні рішення про використання тих або інших методик у навчанні.

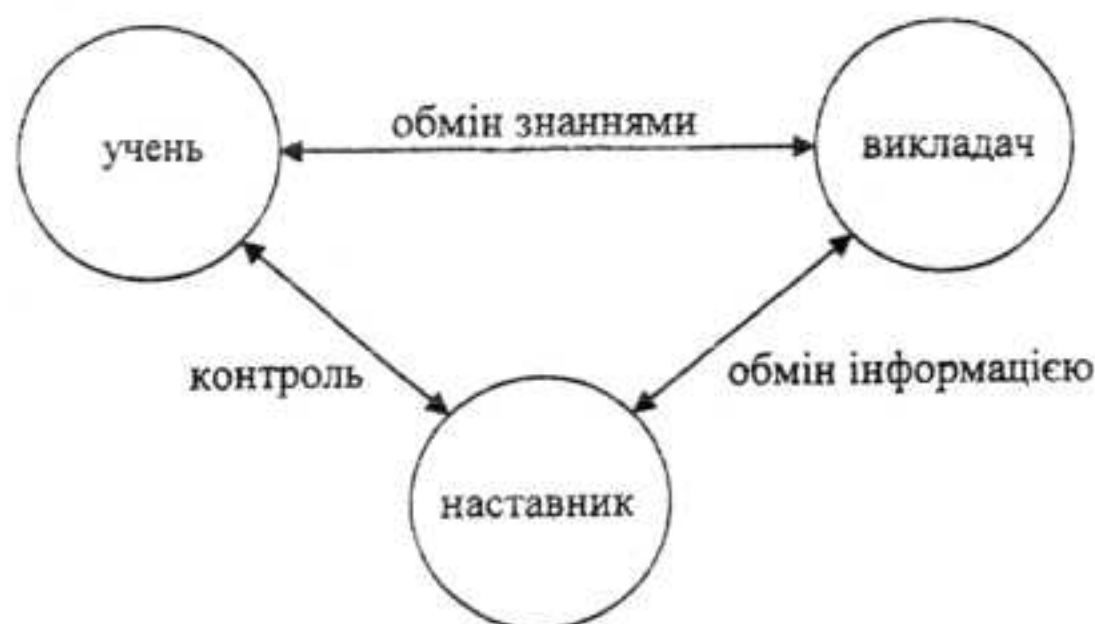


Рис. 1 Схема взаємодії учасників навчального процесу

• самий оперативний спосіб зв'язку – служби миттєвих повідомлень (Instant Messaging). Спілкування може здійснюватися за допомогою як текстових, так і голосових та відеоповідомлень.

Для вирішення цих задач, компанія Microsoft розробила Microsoft Learning Gateway (MLG) – засновану на стандартах, модульну архітектуру інформаційних систем для освітніх організацій. Ця архітектура заснована на використанні ряду стандартних серверних продуктів Microsoft. MLG забезпечує інфраструктуру створення порталів для доступу учнів, викладачів і батьків до освітньої інформації й сервісів.

Основні завдання, які ставились при створенні Microsoft Learning Gateway:

- Висока масштабованість (до 1,5 млн. користувачів)
- Надання всіх сервісів через Web-браузер
- Єдина реєстрація (single sign-on)
- Робота з порталом на основі ролей
- Максимальне використання готової функціональності продуктів Microsoft
- Можливість легкої інтеграції інших компонентів від Microsoft і інших постачальників рішень (відкрита інфраструктура)
- Можливість як централізованого, так і децентралізованого розгортання

Логічна архітектура Microsoft Learning Gateway містить у собі набір продуктів і технологій Microsoft, кожний з яких забезпечує частину загальної функціональності. При цьому розгортання MLG можна починати «з нуля» або

використати вже наявну інфраструктуру Microsoft.

Рішення Learning Gateway поєднує ряд ключових технологій корпорації Microsoft у єдину операційну структуру:

- База даних Microsoft SQL Server™ – забезпечує функції бази даних [4];
- Поштова служба Microsoft Exchange Server забезпечує – функції обміну повідомленнями й ведення календаря [5];
- Система створення і публікації навчальних курсів Microsoft Class Server – надає можливості керування навчальним курсом;
- Операційна система Microsoft Windows Server – служить базовою мережевою операційною системою [1];
- Універсальний мережевий екран Microsoft ISA Server – забезпечує функції безпеки й веб-кешування;
- Портал Microsoft SharePoint™ Portal Server – відповідає за подання інформації [2];
- Медіа сервер Windows Media Server – пропонує набір служб керування мультимедіа контентом.

Всі ці компоненти можна об'єднати в наступну схему (рис. 2)

Так як система є гнучкою і легко масштабованою, в залежності від потреб навчального закладу а також масштабності проекту можливі різні комбінації використання вказаних компонентів. Розглянемо три, найбільш вірогідних варіанти розгортання MLG.

1. Використання MLG поширюється на один предмет або один навчальний клас. В цьому

випадку мінімальним набором інструментів для побудови такого рішення буде – Class Server, Windows Server 2003, SQL Server. Останні два компоненти є невід'ємними при будь-якому масштабуванні проекту, так як потрібна базова операційна система, а також необхідне програмне

забезпечення для каталогізації та упорядкування в єдину базу даних навчального контенту. Викладач зможе локально створювати та редагувати навчальні матеріали, роздавати їх учням. Але в даному випадку обмін інформацією

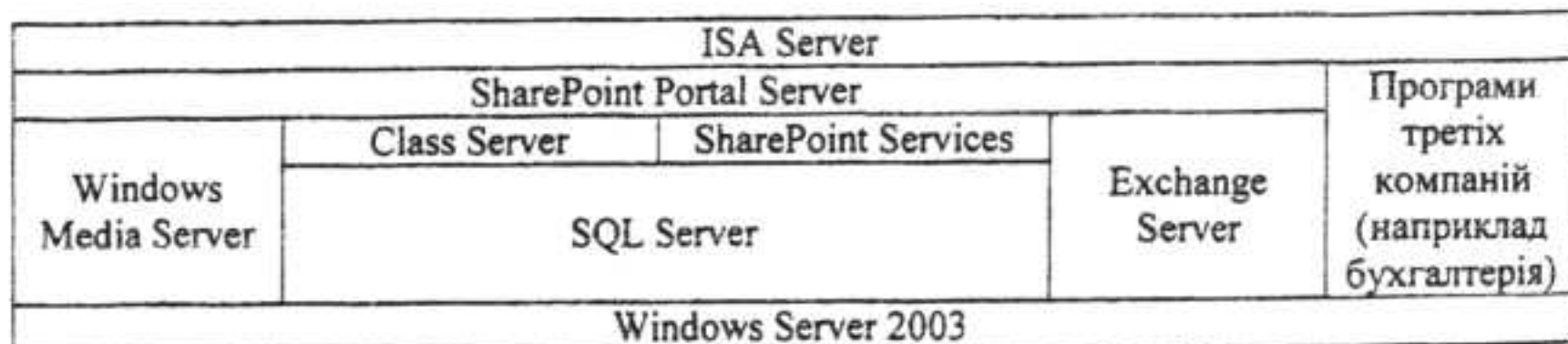


Рис. 2 Структура Microsoft Learning Gateway

буде відбуватися статично. Учні зможуть опрацьовувати навчальний матеріал та виконувати тестові завдання тільки у класі обмеженому однією робочою групою комп'ютерної мережі.

2. Полем використання, є певний структурний підрозділ (кафедра, факультет). До попереднього набору ПЗ необхідно додати Exchange Server та Share Point Services. Ці компоненти необхідні для обміну контентом між учасниками навчального процесу в незалежності від часу, проте є обмеження по місцю – навчальний корпус. При цьому викладачі додатково можуть вести моніторинг успішності груп студентів, які в свою чергу можуть працювати з навчальним матеріалом в режимі самопідготовки.

3. Навчання виходить за рамки навчального закладу – необхідно навчатись будь-де і будь-коли. ISA Server забезпечує надійність та безпеку передачі контенту. Викладачі також можуть працювати не обмежуючи себе територіально. Також при такому масштабуванні з'являється ще один учасник навчального процесу – батьки, вони можуть перевіряти успішність своїх дітей.

Невід'ємною складовою частиною архітектури Microsoft Learning Gateway є система керування навчальним процесом Microsoft Class Server. Цей серверний продукт Microsoft забезпечує функції планування навчального процесу, підготовки навчальних матеріалів і контрольних завдань, розповсюдження навчальних матеріалів і завдань, контроль знань і моніторинг успішності. Його функціонал дозволяє викладачеві створювати навчальний контент а також тестові завдання для студентів. Таким чином, навчальний заклад для вирішення перерахованих завдань може скористатися

функціональними можливостями Microsoft Class Server, використати розробки інших компаній або свої власні.

Ключові компоненти Class Server надають можливість для створення, візуалізації а також зберігання навчальних матеріалів (рис. 3)

- Class Server – визначає ролі користувачів, та розподіляє завдання між компонентами;
- Teacher Client – надає інтерфейс викладача для створення та редагування контенту;
- File Server та SQL Server – збереження та індексації навчальних матеріалів;
- Browser – надає інтерфейс для студента, та пропонує базу знань для навчання.

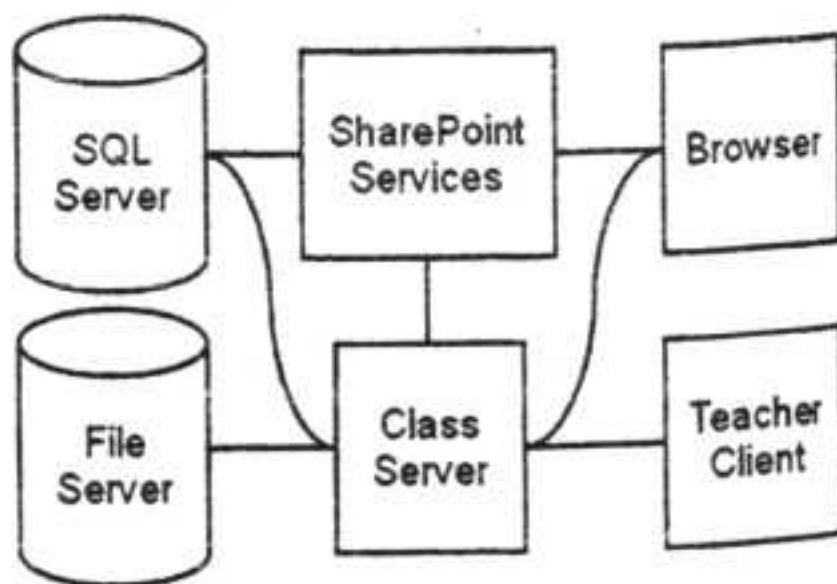


Рис. 3 Структура Class Server

MLG припускає можливість створення на основі єдиної технології ієрархії сайтів для різних факультетів (шкіл), сайти для студентів, викладачів, батьків, завдань і навчальних проектів, навчальних груп (або класів), кафедр і сайти по окремих предметних або тематичних областях.

Батьки, як правило, хотіли б активніше брати участь у навчанні своєї дитини, але часто

виявляють, що їм важко одержати актуальні й достовірні відомості. Microsoft Class Server дає можливість батькам зайти по паролю на спеціальну веб-сторінку, де вони зможуть відслідковувати успішність своєї дитини.

Microsoft Class Server пропонує багато можливостей, що збагачують процес навчання. Учні можуть працювати з навчальними матеріалами й виконувати домашні й перевірочні завдання (тести) з будь-якого комп'ютера, що має доступ в Інтернет, виходячи на захищену паролем веб-сторінку. Використання новітніх технологій перетворює процес навчання в захоплююче заняття.

Відкрита архітектура системи дає можливість використовувати Class Server для взаємодії методистів і викладачів, викладачів і учнів, а також їхніх батьків. Система може використовуватися як у школах, так і у вищих навчальних закладах.

Class Server – це система керування навчальним процесом, що поєднує п'ять функціональних підсистем:

- Керування навчальними планами.
- Керування навчальними матеріалами.
- Тестування й оцінка.
- Звітність по успішності учнів.
- Робота в Інтернеті.

Завдяки активному використанню мережі Інтернет, система керування навчальним процесом Microsoft Class Server дозволяє автоматизувати виконання багатьох адміністративних і професійних завдань, які викладачі звичайно виконують на папері.

Class Server дозволяє легко й зручно створювати навчальні матеріали й завдання, розсилати їх по мережі учням, одержувати по мережі виконані завдання й допомагає виконувати їхню перевірку.

Автоматизація цих функцій вивільняє час учителів, завдяки чому викладачі зможуть більш творчо підходити до викладання й приділяти більше уваги кожному учневі.

Як вже згадувалося, одним з недоліків більшості існуючих систем керування навчанням є використання закритого формату навчальних матеріалів. Навчальному закладу, що застосовує таке програмне забезпечення, доводиться витрачати досить великі ресурси для конвертації курсів у закритий формат. Крім того, не вдасться якщо буде потреба просто й дешево змінити програмне забезпечення, оскільки доведеться конвертувати велику кількість навчальних матеріалів у новий формат.

Наявність стандартів важливе для будь-якого користувача інформаційних технологій, тому що саме завдяки стандартизації кожний користувач може комбінувати програми різних виробників у відповідності зі своїми індивідуальними потребами. Якщо єдиний стандарт відсутній, то користувач повинен обмежуватися програмами лише одного виробника. Стандартизації підлягають як пристрої, так і програмне забезпечення, зокрема, програми, що використовуються в електронному навчанні.

Microsoft Learning Gateway працює з навчальними матеріалами, створеними відповідно до відкритих стандартів SIF (School Interoperability Framework), IMS (Instructional Management Systems) та SCORM (Sharable Content Object Reference Model). Стандарт SIF визначає правила, що дозволяють створювати навчальний контент не залежно від платформ та постачальників ПО.

Створена IMS інформаційна модель зберігання контенту, описує структури даних, покликані забезпечити сумісність матеріалів, створених за допомогою Інтернету, з інструментальними засобами розробки контенту, системами організації навчання і так званими робочими середовищами, або оперативними засобами керування виконанням програм. Модель зберігання контенту IMS створена для визначення стандартного набору структур, які можна використати для обміну навчальними матеріалами. Специфікація сумісності питань і систем тестування IMS описує структури даних, що забезпечують сумісність питань і систем тестування, створених на основі використання Інтернету. Головна мета цієї специфікації – дати користувачам можливість імпортувати й експортувати матеріали з питаннями й тестами, а також забезпечити сумісність змісту навчальних програм із системами оцінки.

В свою чергу SCORM містить вимоги до організації навчального матеріалу й всієї системи дистанційного навчання. SCORM дозволяє забезпечити сумісність компонентів і можливість їхнього багаторазового використання: навчальний матеріал представлений окремими невеликими блоками, які можуть включатися в різні навчальні курси й використовуватись системою дистанційного навчання незалежно від того, ким, де й за допомогою яких засобів були складені. Таким чином, для навчального закладу, що використовує MLG, стає доступний великий ринок курсів, створюваних великою кількістю незалежних видавців. До того ж навчальний

заклад зможе виступати як продавець своїх навчальних матеріалів.

Для Class Server існує чотири можливих настройки мережі.

- Зворотне настроювання проксі
- Настроювання, що забезпечує доступ тільки з Інтернету
- Настроювання з подвійною прив'язкою
- Настроювання, що забезпечує доступ тільки з внутрішньої мережі

Зворотне настроювання проксі

При такому настроюванні Class Server є частиною внутрішньої мережі й одержує пакети, адресовані йому з Інтернету, через зворотний проксі.

Для використання такого настроювання потрібні наступні складові.

- Підключення до локальної мережі
- Підключення до Інтернету
- Статична IP-адреса
- Зареєстроване ім'я домена й запису DNS
- Постачальник ISP, що дозволяє відправлення даних по протоколу HTTP на сервер навчального закладу

- Проксі-сервер, настроєний як зворотний проксі

Переваги такого настроювання

Настроювання зі зворотним проксі має наступні переваги.

- Викладачі можуть надавати доступ до навчальних матеріалів іншим викладачам у навчальному закладі й в Інтернеті.

- Викладачі можуть шукати навчальні матеріали в Інтернеті й у мережі.

Недоліки такого настроювання

Настроювання зі зворотним проксі має наступні недоліки.

- Для подібного настроювання потрібен проксі-сервер, що підтримує служби зворотного проксі. Щоб дозволити з'єднання між Class Server і Інтернетом, необхідно відкрити порт.

- Доступ з дому залежить від швидкості підключення навчального закладу до Інтернету.

Настроювання, що забезпечує доступ тільки з Інтернету

При такому настроюванні користувачі можуть одержати доступ до Class Server через Інтернет. Якщо існує локальна мережа, підключена до Інтернету через проксі-сервер, користувачі у мережі можуть одержати доступ до Class Server через проксі-сервер.

Для використання такого настроювання потрібні наступні складові.

- Підключення до Інтернету

- Статична IP-адреса
- Зареєстроване ім'я домена й запису DNS
- Постачальник ISP, що дозволяє відправлення даних по протоколі HTTP на свій сервер

Переваги такого настроювання

Настроювання, що забезпечує доступ тільки з Інтернету, має наступні переваги.

- Викладачі, учні й батьки можуть одержати доступ до системи з-за меж навчального закладу. Наприклад, вони можуть одержати доступ з дому, з бібліотеки, а також з переносного комп'ютера.

- Викладачі можуть надавати доступ до навчальних матеріалів іншим викладачам у навчальному закладі й в Інтернеті.

- Викладачі можуть шукати навчальні матеріали в Інтернеті й у локальній мережі.

Недоліки такого настроювання

Настроювання, що забезпечує доступ тільки з Інтернету, має наступні недоліки.

- Продуктивність при доступі з-за меж навчального закладу залежить від швидкості підключення користувача до Інтернету.

- Продуктивність користувача, що використовує локальну мережу, залежить від швидкості мережі, проксі-сервера й з'єднання між проксі-сервером і Class Server. Якщо один із цих елементів має низьку швидкість, продуктивність користувачів буде знижена.

Настроювання з подвійною прив'язкою

Подвійна прив'язка означає наявність двох мережних плат. При такому настроюванні доступ до Class Server можливий через локальну мережу й через Інтернет. Для використання такого настроювання потрібні наступні складові.

- Підключення до локальної мережі
- Підключення до Інтернету
- Дві мережні плати
- Статична IP-адреса
- Зареєстроване ім'я домена й запису DNS
- Постачальник ISP, що не накладає обмеження на трафік для статичної IP-адреси

Переваги такого настроювання

Настроювання з подвійною прив'язкою мають наступні переваги.

- Викладачі, учні й батьки можуть одержати доступ до системи з-за меж навчального закладу. Наприклад, вони можуть одержати доступ з дому, з бібліотеки, а також з переносного комп'ютера.

- Викладачі можуть надавати доступ до навчальних матеріалів іншим викладачам у школі й в Інтернеті.

- Викладачі можуть шукати навчальні матеріали в Інтернеті й у шкільній мережі.

Недоліки такого налаштування

Налаштування з подвійною прив'язкою мають наступні недоліки.

- Доступ у навчальному закладі залежить від швидкості локальної мережі.

- Доступ з дому залежить від швидкості підключення навчального закладу до Інтернету.

Налаштування, що забезпечує доступ тільки з локальної мережі

При такому налаштуванні Class Server підключений тільки до локальної мережі. Доступ до системи з Інтернету неможливий. Проте, використовуючи Class Server, викладачі можуть шукати навчальні матеріали в локальній мережі або в Інтернеті.

Переваги такого налаштування

Налаштування, що забезпечує доступ тільки з локальної мережі, має наступні переваги.

- Не потрібно налаштовувати підключення до Інтернету.

Недоліки такого налаштування

Налаштування, що забезпечує доступ тільки з локальної мережі, має наступні недоліки.

- Викладачі, учні й батьки не можуть одержати доступ до системи з-за меж навчального закладу.

- Викладачі можуть надавати доступ до навчальних матеріалів іншим викладачам у межах навчального закладу, однак не можуть надавати подібний доступ викладачам в Інтернеті.

Висновки

Отже, Microsoft Learning Gateway надає нові можливості в організації навчального процесу як для викладачів і учнів, так і для батьків або керівників, що направили на навчання своїх співробітників. Microsoft Learning Gateway збирає воедино можливості продуктів корпорації Microsoft для комплексної автоматизації процесу навчання, надаючи доступ до інформації й сервісів для учнів, викладачів і батьків або керівників. Тому розпочинати розгортання системи можна починати з малого (автоматизація певного предмету) і поступово підключаючи компоненти MLG розширювати область застосування та збільшувати інформаційну базу навчального контенту, що в подальшому дозволить автоматизувати весь процес навчання.

Зокрема, засоби Microsoft Learning Gateway дозволяють викладачеві досить просто створювати нові освітні ресурси й

використовувати їх при навчанні різних груп учнів. Інформація з кожної теми може зберігатися централізовано на сервері навчального закладу, підтримуючи, з одного боку, загальну базу знань й з іншого боку, дозволяючи викладачеві структурувати свої знання й навчальні матеріали. Викладач може підготувати завдання, та й ціле заняття вдома, а по приходу на робоче місце синхронізувати свої матеріали зі збереженими копіями в базі навчального закладу і видати студентам необхідний навчальний матеріал в структурованому вигляді для оптимального сприйняття останніми. Засоби автоматичної перевірки завдань дозволяють більше часу приділяти безпосередньо роботі зі студентами, а не перевірці виконаних у паперовому варіанті тестових і інших завдань.

В роботі дано практичні рекомендації по впровадженню архітектури MLG в навчальних закладах. В залежності від масштабів проекту можна попередньо оцінити набір необхідних інструментів та сервісів, що в подальшому дасть можливість розширювати та додавати компоненти архітектури. Також важливо враховувати рекомендації щодо взаємодії системи та користувачів, що зробить роботу всіх учасників навчального процесу максимально комфортно та корисною.

Список використаних джерел

1. Microsoft Windows Server 2003. Полное руководство / Моримото Р., Гардиньер К., Нозл М., Драуби О. - 2-е изд. - М.: "Вильямс", 2006. - 1312 с.
2. Нозл М., Спенс К. Microsoft SharePoint 2007. Полное руководство. - М.: "Вильямс", 2007. - 832с.
3. Андреев А.А., Солдаткин В.И. Дистанционное обучение: сущность, технология, организация. - М.: Издательство МЕСИ, 1999. - 196 с.
4. Гандерлой М., Джорден Д., Чанц Д. Освоение Microsoft SQL Server 2005 - М.: "Диалектика", 2007. - 1104 с.
5. Microsoft Exchange Server 2003. Полное руководство / Моримото Р., Гардиньер К., Нозл М., Кока Д. - М.: "Вильямс", 2006. - 1024 с.

Надійшла до редколегії 29.04.2007

УДК 519.876.5:624.131.542

Власюк А.П., д. т. н., проф., Мартинюк* П.М.,
к.ф.-м.н., доц.

Математичне моделювання контактного розмиву неоднорідного масиву ґрунту фільтраційним потоком

Побудовано математичну модель контактного фільтраційного розмиву масиву двошарового ґрунту. Для знаходження наближеного числового розв'язку відповідної двовимірної крайової задачі застосовано метод радіальних базисних функцій. Проведено ряд числових експериментів, які засвідчили ефективність запропонованого підходу і можливість фільтраційного руйнування масиву ґрунту шляхом контактного розмиву.

Ключові слова: контактний розмив, гранична швидкість фільтрації, метод радіальних базисних функцій.

*E-mail: Martinyuk@ukr.net

Статтю представив д.т.н., проф., Гаращенко Ф. Г.

1. Вступ. Руйнування гідротехнічних ґрунтових споруд та масивів ґрунтів в їх основах фільтраційними потоками ґрунтових вод є складною проблемою, вирішення та прогнозування якої пов'язане з безпекою життєдіяльності людей, екологічною безпекою та зменшенням економічних затрат. В роботі [1] наведено детальний аналіз причин та форм руйнувань ґрунтових гідротехнічних споруд. Є цікавим огляд фактичного матеріалу щодо руйнування таких споруд, виявлення їх причин та наслідків.

В роботі [2] наведено класифікацію видів фільтраційного руйнування ґрунтів, які, за твердженням автора, на той час найкраще вивчені в лабораторних умовах. Сюди відносяться: 1) випирання ґрунту під дією висхідного фільтраційного потоку; 2) суфозія; 3) контактний випір ґрунту; 4) контактний розмив. В даній роботі наведено результати багатьох експериментальних досліджень з вивчення всіх чотирьох видів фільтраційного руйнування незв'язних та зв'язних ґрунтів. Експериментально встановлено критичні градієнти та швидкості фільтрації, при яких розпочинається процес руйнування ґрунту, та їх залежність від типів ґрунтів та їх параметрів.

A.P.Vlasyuk, Doctor of technical sciences,
P. M. Martinyuk*, Candidate of physics and
mathematics sciences

The mathematical modeling of inhomogeneous soil wash-out

The mathematical model of stratified soil wash-out has been formulated. The numerical solution of the correspondent two-dimensional boundary-value problem has been found by the radial basis function method with discontinuous radial basis function. The numerical experiments have been lead. There has been shown that proposed method can be effective one.

Key Words: filtration erosion; critical filtration speed; radial basis function method.

Більш детально види та причини фільтраційного руйнування глинистих ґрунтів експериментально досліджено в роботі [3].

Однак, математичних моделей розглядуваних процесів у двох вищенаведених книгах не побудовано.

Також до видів фільтраційного руйнування ґрунтів потрібно віднести зосереджену фільтрацію. Так, в роботі [4] відмічається, що з 9248 гідротехнічних споруд, зареєстрованих в 1965 році Міжнародною комісією по великих греблях, на 535 греблях мали місце аварії та інциденти. В результаті аварій на 10 греблях загинуло близько 10 тисяч людей і матеріальні збитки склали кілька мільярдів доларів. При цьому, через зосереджену фільтрацію в основі та тілі греблі сталося 38% всіх аварій ґрунтових гребель із місцевих матеріалів. В роботі [5] відмічається, що основними причинами аварій ґрунтових гребель з водоводами є виникнення контактної зосередженої фільтрації вздовж водоводу, що супроводжується розмивом ґрунту тіла та основи греблі, а також суфозія ґрунту тіла греблі при пошкодженні напірного водоводу.

В роботі [6] досліджені практичні питання суфозії. Зокрема, відмічено, що винесення частинок розпочинається, коли швидкість фільтрації перевищує деяке критичне значення. В роботі проаналізовано формули, отримані для

критичних значень швидкості фільтрації іншими авторами. Самим автором встановлено, що швидкість винесення частинок прямо пропорційна діаметру частинки і коефіцієнту фільтрації ґрунту. Також прямо пропорційними є швидкість фільтраційного потоку та швидкість винесення частинки.

В роботі [7] наведено дослідження дифузійного вилуговування глин. Дане питання є важливим, оскільки дифузійне вилуговування впливає на інженерно-геологічні властивості глин, а отже і на умови фільтраційного руйнування цих ґрунтів.

В роботі [8] досліджено процеси суфозії та кольматації в незв'язних ґрунтах. Вивчено важливість даних процесів для роботи дренажних систем. Математична модель побудована в осесиметричному випадку для ґрунту, тверда фаза якого складається з двох фракцій, здійснено оцінки зміни коефіцієнта фільтрації в суфозійній та акумулюючій зонах.

Питанням математичного моделювання фільтрації ґрунтових вод з урахуванням суфозії та кольматації присвячено також роботу [9].

В монографії [10] наведені результати експериментальних досліджень радіальної фільтрації в незв'язних ґрунтах. На основі результатів проведених експериментів побудовано математичну модель фільтрації з урахуванням суфозійних процесів в придренній зоні.

Наведений короткий огляд літературних джерел показує, що дослідження фільтраційного руйнування ґрунтів є актуальним щодо практики гідротехнічного будівництва і математичне моделювання цих процесів знаходиться в постійному розвитку. В даній статті дослідимо питання математичного моделювання нестационарного процесу контактного розмиву ґрунту фільтраційним потоком. В стаціонарному випадку дана задача досліджена в роботі [11]. Мотивом є те, що цей процес фільтраційного руйнування ґрунтів може бути передумовою утворення шляхів зосередженої фільтрації [12]. А для ґрунтових споруд, як згадувалося вище, зосереджена фільтрація є основною причиною їх руйнування [4]. Також в даному випадку потрібно відмітити роботу [13], в якій вказується на небезпеку контактного розмиву при зведенні земляних споруд (укладанні ґрунту) на жорстких нерівних поверхнях.

Для чисельного розв'язування відповідної крайової задачі використаємо один із безсіткових методів [14]. Це мотивується тим, що в процесі розмиву з часом область розв'язку задачі буде

змінюватись. А в сіткових числових методах (наприклад, в методі скінченних різниць або в методі скінченних елементів) переформатувати геометричну сітку набагато складніше, ніж вузлову сітку в безсіткових методах. Тому для чисельний розв'язування відповідної крайової задачі, якою описується побудована математична модель, використаємо метод радіальних базисних функцій (РБФ) [15].

2. Постановка задачі та її математична модель. Нехай в областях Ω_1 та Ω_2 , відповідно, знаходяться два типи ґрунту, які контактують по рухомій межі Γ_{12} (рис.1). Область Ω_3 утворюється в результаті розмиву ґрунту в області Ω_1 і її займає порова рідина. Для спрощення форму області Ω_2 будемо вважати незмінною.

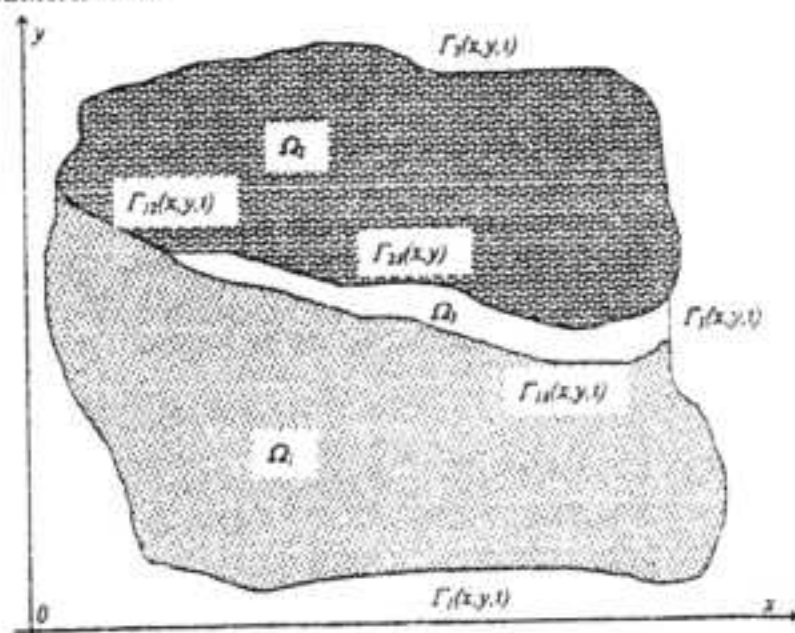


Рис.1. Контактний розмив ґрунту

Математична модель даної задачі контактного розмиву в двовимірному випадку буде описуватись наступною крайовою задачею:

$$\eta \frac{\partial h}{\partial t} = \nabla \cdot (K(e) \nabla h(x, y, t)), \quad (x, y) \in \Omega, \quad t > 0, \quad (1)$$

$$(u, n)|_{\Gamma^{(u)}} = 0, \quad h|_{\Gamma^{(h)}} = H(x, y, t), \quad (x, y) \in \Gamma^{(h)}, \quad (2)$$

$$[h]|_{\Gamma} = 0, \quad [u]|_{\Gamma} = 0, \quad \Gamma \in \{\Gamma_{12}; \Gamma_{13}; \Gamma_{23}\}, \quad (3)$$

$$h(x, y, 0) = h_0(x, y), \quad (x, y) \in \Omega, \quad (4)$$

$$c(x, y, t)(v, n) + |(v, d)| = -m\rho_T \frac{d\Gamma_{12}}{dt}, \quad (x, y) \in \Gamma_{12}, \quad (5)$$

$$u = -K(e) \nabla h, \quad (6)$$

$$v = F(x, y, t, u), \quad (x, y) \in \Gamma_{12}, \quad (7)$$

де $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$, $\Gamma^{(u)} \cup \Gamma^{(h)} = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$;

$$h(x, y, t) = \{h_i(x, y, t), \quad (x, y) \in \Omega_i,$$

$$K(e) = \{K_i(e), \quad (x, y) \in \Omega_i,$$

$$u(x, y, t) = \{u_i(x, y, t), \quad (x, y) \in \Omega_i,$$

$$\eta = \{\eta_i, \quad (x, y) \in \Omega_i, \quad i = \overline{1,3}, \quad \eta_3 = 0.$$

Рівняння (1) описує нестационарний режим фільтрації в деформівних масивах ґрунту Ω_1 та Ω_2 , і стаціонарний – в області Ω_3 . Припускається, що на межах контакту областей виконуються умови спряження ідеального контакту (3) [16]. Функція $\mathbf{v} = \mathbf{F}(x, y, t, \mathbf{u})$ пов'язує вектор швидкості фільтрації $\mathbf{u}$ та вектор швидкості руху твердих частинок ґрунту $\mathbf{v}$ на рухомій межі Γ_{13} . Тут: $h(x, y, t)$ – напір; e – коефіцієнт пористості ґрунту; ρ_T – густина твердих частинок ґрунту; m – відносний об'єм твердих частинок в одиниці об'єму ґрунту; $c(x, y, t)$ – концентрація твердих частинок ґрунту; η – коефіцієнт пружної ємкості ґрунту; $\mathbf{K}(e)$ – тензор коефіцієнта фільтрації; $\mathbf{n}$ – вектор напрямних косинусів зовнішньої нормалі; $\mathbf{d}$ – вектор напрямних косинусів дотичної. Вектор дотичної вважається додатньо направленим, якщо його поворот до вектора зовнішньої нормалі здійснюється проти руху годинникової стрілки.

Умова (5) задає кінематичну граничну умову на межі можливого контактного розмиву. Вивести її можна наступним чином. Виділимо на поверхні $\Gamma_{13}(x, y, t)$ при фіксованому t область, площею S . Виберемо в даній області точку (x, y) і осереднимо характеристики ґрунту даної області. Тобто, всі характеристики середовища у виділеній області будемо вважати рівними характеристикам у вибраній точці в заданий момент часу t . Тоді, за час dt з виділеної площі за рахунок розмиву відірветься певна кількість ґрунтових частинок, загальною масою

$$c(x, y, t)(\langle \mathbf{v}, \mathbf{n} \rangle + |\langle \mathbf{v}, \mathbf{d} \rangle|)S dt.$$

Якщо відбувається розмив ґрунту, то $c(x, y, t) = m\rho_T$. Якщо ж відбувається нанесення твердих частинок на поверхню ґрунту, то $c(x, y, t)$ – концентрація суспензії, яка фільтрується в ґрунт. Також ми беремо модуль від скалярного добутку $(\mathbf{v}, \mathbf{d})$, оскільки швидкість по дотичній завжди буде розвиваючою, незалежно від напрямку.

З іншого боку, маса твердих частинок, яка винесеться з виділеної області площею S поверхні $\Gamma_{13}(x, y, t)$ за час dt рівна

$$-m\rho_T(\Gamma_{13}(x, y, t+dt) - \Gamma_{13}(x, y, t))S.$$

Отже, маємо

$$-m\rho_T(\Gamma_{13}(x, y, t+dt) - \Gamma_{13}(x, y, t))S = c(x, y, t)(\langle \mathbf{v}, \mathbf{n} \rangle + |\langle \mathbf{v}, \mathbf{d} \rangle|)S dt.$$

Переходячи до границі при $dt \rightarrow 0$ отримуємо умову (5). Оскільки координати x, y межі $\Gamma_{13}(x, y, t)$ теж залежать від часу, то в (5) присутня повна похідна по часу.

Відмітимо, що умова (5) виведена, коли в області Ω_3 знаходиться суспензія концентрацією $c(x, y, t)$ і тверді частинки можуть накопичуватись на межі $\Gamma_{13}(x, y, t)$, а не лише покидати її.

3. Числове розв'язання крайової задачі (1)-(7) методом РБФ. Введемо позначення $X = (x, y)$. Покриємо замикання області $\overline{\Omega} = \overline{\Omega}_1 \cup \overline{\Omega}_2 \cup \overline{\Omega}_3$ вузловими точками $X_i = (x_i, y_i)$, $i = \overline{1, N}$, де: N^{Ω_i} вузлів лежать в кожній з областей Ω_i , $i = \overline{1, 3}$; $N^{(u)}$, $N^{(h)}$ – на відповідних частинах $\Gamma^{(u)}$ та $\Gamma^{(h)}$ границі $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3$; $2N^{\Gamma_{12}}$ – на межі спряження Γ_{12} ; $2N^{\Gamma_{13}}$ – на межі спряження Γ_{13} ; $2N^{\Gamma_{23}}$ – на межі спряження Γ_{23} . На межах спряження використовується подвійна нумерація вузлів.

Отже, $N = \sum_{i=1}^3 N^{\Omega_i} + N^{(u)} + N^{(h)} + 2N^{\Gamma_{12}} + 2N^{\Gamma_{13}} + 2N^{\Gamma_{23}}$. Не зменшуючи загальності, будемо вважати, що вузли з номерами $i = \overline{1, N^{\Omega}}$ лежать в області Ω ; вузли з номерами $i = N^{\Omega} + 1, N^{\Omega} + N^{\Gamma^{(h)}}$ – на межі $\Gamma^{(h)}$; вузли з номерами $i = N^{\Omega} + N^{\Gamma^{(h)}} + 1, N^{\Omega} + N^{\Gamma^{(h)}} + N^{\Gamma^{(u)}}$ – на межі $\Gamma^{(u)}$.

Наближений розв'язок задачі (1)-(7) шукаємо у вигляді

$$h(X, t) \approx h_N(X, t) = \sum_{i=1}^N a_i(t) \varphi_i(r_i, \varepsilon),$$

де $a_i(t)$, $i = \overline{1, N}$, – деякі невідомі коефіцієнти, які залежать лише від часу; $\varphi(r_i, \varepsilon)$, $i = \overline{1, N}$, – деякі розривні РБФ [15, 17], причому $r_i = |X - X_i| = \sqrt{(x - x_i)^2 + (y - y_i)^2}$. Тоді з рівняння (1), умов (2)-(4) отримаємо наступну задачу Коші для системи звичайних диференціальних рівнянь відносно вектора невідомих $\mathbf{A}(t) = \{a_i(t)\}_{i=1}^N$:

$$\mathbf{M} \cdot \frac{d\mathbf{A}}{dt} + \mathbf{L} \cdot \mathbf{A}(t) = \mathbf{F}(t), \quad (8)$$

$$\mathbf{M}^{(1)} \cdot \mathbf{A}^{(0)} = \mathbf{F}^{(1)}, \quad (9)$$

де $\mathbf{A}^{(0)} = \{a_i(0)\}_{i=1}^N$, $\mathbf{M} = \{m_{ij}\}_{i,j=1}^N$, $\mathbf{M}^{(1)} = \{m_{ij}^{(1)}\}_{i,j=1}^N$,

$$\mathbf{L} = \{l_{ij}\}_{i,j=1}^N, \quad \mathbf{F} = \{f_i\}_{i=1}^N, \quad \mathbf{F}^{(1)} = \{f_i^{(1)}\}_{i=1}^N,$$

$$m_{ij} = \begin{cases} \eta \varphi_j(r_{ij}, \varepsilon), i = \overline{1, N^2}; \\ 0, i = N^2 + 1, N; \end{cases}$$

$$m_{ij}^{(1)} = \varphi_j(r_{ij}, \varepsilon), i = \overline{1, N}; f_i^{(1)} = h_0(X_i), i = \overline{1, N};$$

$$f_i = \begin{cases} 0, i = \overline{1, N^2}; \\ H(X_i, t), i = N^2 + 1, N^2 + N^{r^{(h)}}; \\ 0, i = N^2 + N^{r^{(h)}} + 1, N; \end{cases}$$

$$r_{ij} = |X_j - X_i| = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}.$$

Спеціального визначення потребують елементи матриці $\mathbf{L}$. Так, $l_{ij} = -\nabla \cdot (\mathbf{K} \nabla \varphi_j(r_{ij}, \varepsilon))$ при $i = \overline{1, N^2}$. У випадку $i = N^2 + 1, N^2 + N^{r^{(h)}}$: $l_{ij} = \varphi_j(r_{ij}, \varepsilon)$. І $l_{ij} = (-\mathbf{K} \nabla \varphi_j(r_{ij}, \varepsilon), \mathbf{n})$, якщо $i = N^2 + N^{r^{(h)}} + 1, N^2 + N^{r^{(h)}} + N^{r^{(u)}}$. Якщо ж вузол X_j лежить на межі спряження з боку області Ω^* , а X_k - відповідний йому вузол на тій же межі спряження з боку іншої області Ω_* , то

$$l_{ij} = \begin{cases} \varphi_j(r_{ij}, \varepsilon), X_j \in \Omega^*; \\ -\varphi_j(r_{kj}, \varepsilon), X_j \in \Omega_*; \\ 0, \text{в інших випадках}; \end{cases}$$

$$l_{kj} = \begin{cases} (-\mathbf{K} \nabla \varphi_j(r_{ij}, \varepsilon), \mathbf{n}), X_j \in \Omega^*; \\ (\mathbf{K} \nabla \varphi_j(r_{kj}, \varepsilon), \mathbf{n}), X_j \in \Omega_*; \\ 0, \text{в інших випадках}. \end{cases}$$

Для знаходження чисельного розв'язку задачі Коші (8), (9) поділимо часовий відрізок $[0, t_0]$ на

M однакових частин з кроком $\tau = \frac{t_0}{M}$.

Наближений розв'язок системи звичайних диференціальних рівнянь (7) можна знайти за допомогою чисто неявної різницевої схеми або схеми Кранка-Ніколсона [16].

Кінематична гранична умова (5) використовується для апроксимації положення межі спряження $\Gamma_{13}(x, y, t)$. Якщо на часовому шарі під номером (s) крайова задача (1)-(7) розв'язана, то, наприклад, у випадку розмиву масиву ґрунту положення $\Gamma_{13}^{(s+1)}$ межі $\Gamma_{13}(x, y, t)$ на часовому шарі $(s+1)$ визначається з умови

$$\frac{\Gamma_{13}^{(s+1)}(x, y) - \Gamma_{13}^{(s)}(x, y)}{\tau} = (\mathbf{v}^{(s)}, \mathbf{n}) + |(\mathbf{v}^{(s)}, \mathbf{d})|, s = \overline{0, M-1}.$$

При $t=0$ межа $\Gamma_{13}(x, y, t)$ складається з одного вузла, який розміщений в точці G .

4. Результати числових експериментів. Вияснення можливості та величини контактного розмиву ґрунтів фільтраційним потоком проведено на тестовому прикладі. В зв'язку з цим розглянуто задачу фільтрації через прямокутний масив ґрунту товщиною s м та довжиною l м, який складається з двох типів ґрунтів (рис. 2). Початкова межа контакту паралельна осі Ox і задається рівнянням $y = \frac{s}{2}$. Ґрунти вважались ізотропними за своїми характеристиками. Коефіцієнти фільтрації покладались рівними $k_1 = 0,05 \text{ см/с}$, $k_2 = 1 \text{ см/с}$, що відповідає чистому піску [18]. Коефіцієнти пружної ємкості для пісків згідно [19] покладались рівними $\eta_1 = \eta_2 = 0,007 \text{ м}^{-1}$. Для області розмиву Ω_2 приймалось $\eta_3 = 0$. Критична швидкість фільтрації покладалась рівною $u_{kr} = 1,5 \text{ см/с}$ [2]. Між швидкістю фільтрації та швидкістю руху твердих частинок приймалась лінійна залежність

$$\mathbf{v} = \alpha^* \mathbf{u}, \text{ де } \alpha^* = \begin{cases} \alpha, u \geq u_{kr}; \\ 0, u < u_{kr}. \end{cases}$$

Оскільки ґрунт, можливість розмиву якого досліджувалась, брався як мілкозернистий пісок, то покладалось $\alpha = 1$. Причому припускалось, що піщана частинка, яка відірвалася від масиву ґрунту, миттєво виноситься фільтраційним потоком.

Граничні умови приймались наступні:

$$h(x, y, t)|_{OC} = H_1, h(x, y, t)|_{AB} = H_2,$$

$$\frac{\partial h}{\partial y}|_{CB} = \frac{\partial h}{\partial y}|_{OA} = 0,$$

де $H_1 = \text{const}$, $H_2 = \text{const}$. Тобто, припускалось, що верхня CB та нижня OA межі ґрунту непроникні. Початкова умова для напору задавалась у вигляді

$$h_0(x, y) = x \cdot H_2 + \frac{l-x}{l} H_1.$$

Для спрощення вважалось, що масив ґрунту в області Ω_2 не змінює своєї форми, навіть у випадку утворення зони розмиву.

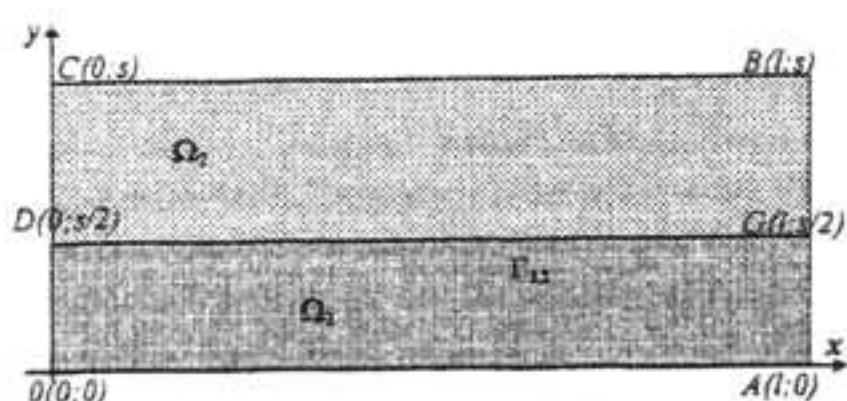


Рис.2. Масив ґрунту

Якщо зону розмиву Ω_3 змодельовати, як тріщину з паралельними стінками, то для швидкості руху рідини в такій тріщині відома формула Буссінеска, яка підтверджена в роботі [20] для $\delta \leq 1$ см

$$u_i = -\frac{g}{12\nu} \delta^2 \frac{\partial h}{\partial \bar{l}},$$

де ν - коефіцієнт кінематичної в'язкості порової рідини; g - прискорення вільного падіння; δ - величина тріщини (відстань між паралельними стінками); вектор $\bar{l}$ визначається напрямом тріщини. Для води [19] $\nu = 10^{-2} \text{ см}^2/\text{с}$. Якщо припустити, що стінки тріщини паралельні до осі

Ox , то в області Ω_3 маємо $u_x = -\frac{g}{12\nu} \delta^2 \frac{\partial h}{\partial x}$,

$u_y \equiv 0$, і $(k_3)_x = 8,32 \cdot 10^3 \delta^2 \text{ л/см} \cdot \text{с}$, $(k_3)_y = 0$.

Тоді при $\delta = 0,5$ см маємо $(k_3)_x = 2,08 \cdot 10^3 \text{ см/с}$, що при одиничному градієнті напору дає швидкість руху рідини величиною $20,8 \text{ м/с}$.

Урахування ефектів шороховатості стінок тріщини та турбулентності фільтраційного потоку [20] вносить певні зміни в числові значення, але не значні.

Числові експерименти проводились при $H_1 = 12$ м, $H_2 = 6$ м, $s = 2$ м, $l = 3$ м і загальною кількістю вузлів 683, причому 101 вузол був розміщений на межі Γ , а 52 на межі спряження Γ_{12} (рис. 3). В околі межі спряження використовувалось згущення вузлової сітки. Для дискретизації в часі використовувалась чисто неявна різницева схема з кроком $\tau = 10$ с.

Для проведення числових експериментів та побудови розривних РБФ [15, 17] використовувались наступні неперервні РБФ [14]:

мультиквадратична РБФ $\varphi(r, \epsilon) = \sqrt{1 + (r\epsilon)^2}$;

обернена мультиквадратична РБФ

$\varphi(r, \epsilon) = \frac{1}{\sqrt{1 + (r\epsilon)^2}}$; обернена квадратична РБФ

$\varphi(r, \epsilon) = \frac{1}{1 + (r\epsilon)^2}$; гаусівська РБФ $\varphi(r, \epsilon) = e^{-(r\epsilon)^2}$. У

вищенаведених функцій $\epsilon > 0$ - так званий параметр форми.

При проведенні чисельних експериментів використовувались різні значення "коефіцієнта фільтрації" зони розмиву.

При $k_3 = k_2$ прорив настає через 160 с (рис. 5). При цьому ширина зони розмиву постійно збільшується (рис. 4, 5).

При $k_3 = 0,5 \cdot k_2$ прорив настає за той же час, але ширина зони розмиву з часом не збільшується, а залишається сталою. При цьому зона розмиву утворюється лише в околі межі контакту областей Ω_1 та Ω_2 .

При збільшенні величини k_3 прорив настає значно раніше. Наприклад, при $k_3 = 2k_2$ прорив настає через 90 с. При швидкостях фільтрації в зоні розмиву, які відповідають формулі Буссінеска, прорив настає за лічені секунди.

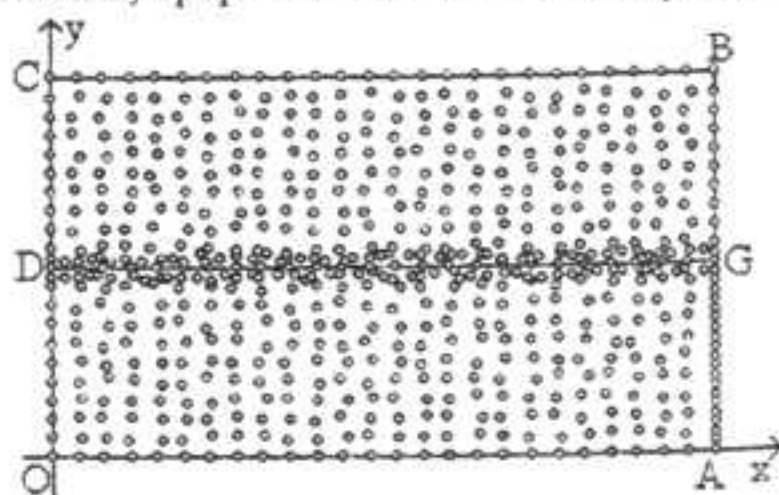


Рис.3. Вузлова сітка

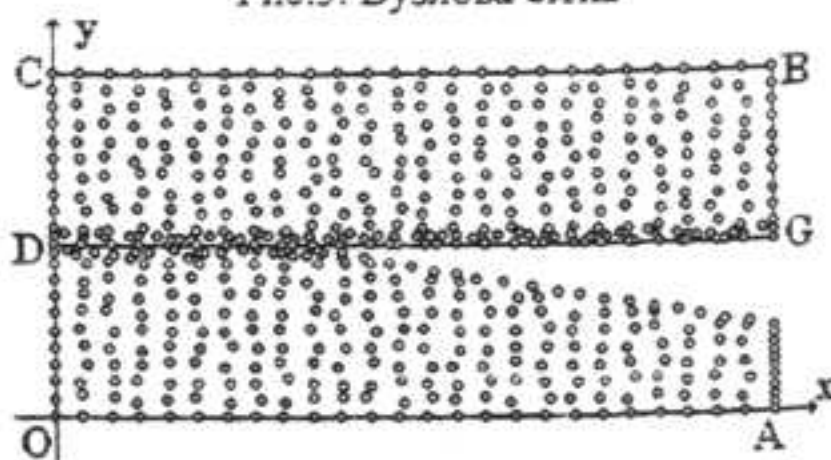


Рис. 4. Зона розмиву при $k_3 = k_2$, $t = 100$ с

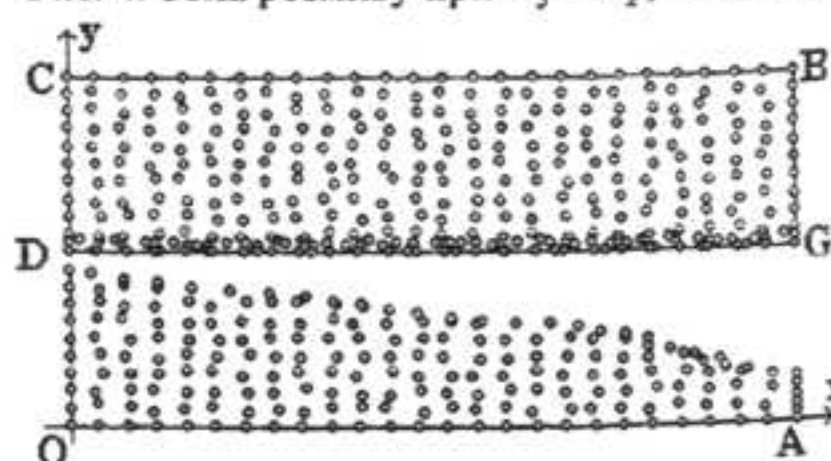


Рис. 5. Зона розмиву при $k_3 = k_2$, $t = 160$ с

5. Висновки. Отже, в даній роботі поставлена і чисельно розв'язана нестационарна задача про контактний розмив ґрунтів фільтраційним потоком. Аналіз результатів проведених чисельних експериментів дозволив зробити наступні висновки:

1) у випадку досягнення критичної швидкості фільтрації зона розмиву зростає катастрофічно швидко, що приводить до прориву;

2) час настання прориву та величина зони розмиву значно залежать від швидкості руху рідини в самій зоні розмиву;

3) безсіткові методи при їх застосуванні до досліджуваного класу задач не вимагають переформатування геометричної сітки, що значно спрощує їх використання в задачах з рухомими межами в порівнянні із сітковими методами.

Варто відмітити, що дана робота ставить і проблемні питання. По-перше, дослідження руху рідини в зоні розмиву вимагає розв'язування відповідної крайової задачі Нав'є-Стокса в даній області. По-друге, в математичній моделі потрібно врахувати фактор, що в зоні розмиву рухається не сама рідина, а суспензія. Ці питання будуть предметом наших подальших досліджень та публікацій.

Список використаних джерел

1. Гельфер А.А. Причины и формы разрушения гидротехнических сооружений. – Москва: Госстройиздат, 1936. – 320 с.
2. Истомина В.С. Фильтрационная устойчивость грунтов. – Москва: Госстройиздат, 1957. – 295 с.
3. Истомина В.С., Буренкова В.В., Мишурова Г.В. Фильтрационная прочность глинистых грунтов. – Москва: Стройиздат, 1975. – 219 с.
4. Гогоберидзе М.И., Какауридзе Р.Г., Микашвили Ю.Н., Мирцхулава Д.Ц. Обобщение данных статистического анализа аварий и инцидентов в аспекте надежности плотин // Сообщения Акад. наук Груз. ССР. – 1977. – №3. – С.681-684.
5. Анахаев К.Н., Гегиев К.А., Амшонов Б.Х. Об авариях и повреждениях земляных плотин с водоводами: причины и способы совершенствования противофильтрационной защиты // Гидротехническое строительство. – 2004. – №3. – С. 30-36.
6. Патрашев А.Н. Напорное движение грунтового потока, сопровождающееся выносом мелких частиц грунта // Изв. Науч.-иссл. ин-та гидротехн. – 1938. – Т. XXII. – С. 5-50.
7. Затенацкая Н.П., Сафохина И.А. Диффузионное выщелачивание глин. – Москва: Наука, 1968. – 88 с.
8. Поляков В.Л. О фильтрационных деформациях грунта с образованием аккумулярующих зон // Прикладна гідромеханіка. – 2003. – Т.5, №2. – С. 45-56.
9. Кремез В.С. Актуальные вопросы моделирования подтопления территорий и других опасных процессов, связанных с изменением режима грунтовых вод // Екологія довкілля та безпека життєдіяльності. – 2003. – №6. – С.56-64.
10. Дмитриев А.Ф., Хлапук Н.Н., Дмитриев Д.А. Деформационные процессы в несвязных грунтах и придренной зоне и их влияние на работу осушительно-увлажнительных систем. – Ровно: Изд-во РГТУ, 2002. – 145 с.
11. Власюк А.П., Мартинюк П.М. До питання про утворення зосереджених фільтраційних потоків в неоднорідному масиві ґрунту // Abstracts of International Workshop "Problems of decision making under uncertainties", Skhidnytsia, Ukraine, May 21-25, 2006. – С. 81-83.
12. Sellmeijer J.B., Koenders M. A mathematical model for erosion under a dam // Appl. Math. Model. – 1991. – 15. – Pp. 646-651.
13. Тартаковский Д.М., Васильев И.М. О неравномерных деформациях земляных плотин в областях примыкания грунтов к жестким неровным поверхностям // Вопросы проектирования высоких плотин: Труды ЛПИ. – 1968. – №292. – С.81-88.
14. Iske A. Radial basis functions: basics, advanced topics and meshfree methods for transport problems // Rend. Sem. Mat. Univ. Pol. Torino. – 2003. – Vol. 61, 3. – Pp.247-284.
15. Мартинюк П.М. Метод радіальних базисних функцій в крайових задачах для параболічних рівнянь з умовами спряження // Вісник Київського ун-ту. Серія фіз.-матем. науки. – 2007. – Вип.4. – С.178-181.
16. Сергиенко И.В., Скопецкий В.В., Дейнека В.С. Математическое моделирование и исследование процессов в неоднородных средах. – Киев: Наук. думка, 1991. – 432с.
17. Мартинюк П.М. Числове розв'язання крайових задач з умовами спряження методом радіальних базисних функцій // Вісник Нац. ун-ту водн. госп-ва та природокористув. – 2007. – Вип. 2(38). – С. 324-331.
18. Полубаринова-Кочина П.Я. Теория движения грунтовых вод. – Москва: Наука, 1977. – 664с.
19. Шестаков В.М. Динамика подземных вод. – Москва: Изд-во Моск. ун-та, 1979. – 368 с.
20. Ламизе Г.М. Фильтрация в трещиноватых породах. – Москва: Госэнергоиздат, 1951. – 127 с.

Надійшла до редколегії 19.01.2008

УДК 517.44: 004

Волинець* В.І., к.т.н., доцент

Аналіз точності рекурентних методів обчислення багатовимірних дискретних перетворень Фур'є та Хартлі в арифметиці з плаваючою комою

На основі статистичного методу проведено аналіз точності рекурентних методів обчислення звичайних і модифікованих багатовимірних дискретних перетворень Фур'є та Хартлі на стрибкових і ковзних інтервалах за всіма та одним виміром в арифметиці з плаваючою комою. Отримано аналітичні вирази для визначення точності обчислення перетворень в залежності від розрядності даних та кількості ітерацій обчислення, на підставі яких проведено порівняльний аналіз точності рекурентних методів обчислення.

Ключові слова: аналіз точності, рекурентні методи, дискретне перетворення Фур'є, дискретне перетворення Хартлі, арифметика з плаваючою комою.

*E-mail: victvol@mail.ru

Статтю представив д.ф.-м.н. Ніколюк П.К.

В основі динамічного спектрального аналізу багатовимірних сигналів, який проводиться на ковзних або стрибкових інтервалах вхідного сигналу, тобто коли черговий фрагмент вхідного сигналу відрізняється від попереднього фрагменту відповідно на одну або декілька груп відліків за одним або всіма вимірами, лежить використання рекурентних методів обчислення багатовимірних дискретних перетворень Фур'є (ДПФ) та Хартлі (ДПХ) [1-3], арифметична складність яких значно нижча за складність прямих та швидких методів обчислення дискретних перетворень, оскільки рекурентні методи враховують результати обчислення на попередніх інтервалах вхідного сигналу.

Важливим критерієм вибору методу обчислення є точність обчислення, яку він забезпечує. У праці [4] проведений аналіз точності рекурентних методів обчислення одновимірних звичайних і модифікованих ДПФ і ДПХ в арифметиці з плаваючою комою, в основу якого покладений статистичний метод аналізу, при якому кожному джерелу елементарної похибки, що виникає внаслідок округлення або

V.I. Volynets*, Ph. D.

Exactness analysis of recurrent methods of calculation of multidimensional discrete Fourier and Hartley transforms in floating point arithmetic

On the basis of statistical method the exactness analysis of recurrent methods of calculation of ordinary and modified multidimensional discrete Fourier and Hartley transforms on skipping and sliding intervals on all and one measuring in floating point arithmetic is executed. Analytical expressions for determination of exactness of calculation of transforms depending on the bit capacity of data and amount of iterations of calculation are obtained, on the basis of which the comparative exactness analysis of recurrent methods of calculation is executed.

Key Words: exactness analysis, recurrent methods, discrete Fourier transform, discrete Hartley transform, floating point arithmetic.

утинання результату операції додавання чи множення, ставиться у відповідність генератор випадкової похибки з рівномірним законом розподілу та робиться припущення, що всі джерела елементарних похибок не корелюють між собою та з вхідним сигналом, який є білим шумом. В якості кількісної оцінки точності обчислення приймаються середньоквадратичне значення (СКЗ) похибки обчислення перетворення та відношення СКЗ похибки обчислення до СКЗ перетворення, яке фізично інтерпретується через відношення потужності шуму до потужності сигналу.

У цій роботі ставиться завдання провести аналіз точності рекурентних методів обчислення звичайних і модифікованих багатовимірних ДПФ і ДПХ на стрибкових і ковзних інтервалах за всіма і одним виміром в арифметиці з плаваючою комою.

Рекурентні методи обчислення звичайних і модифікованих багатовимірних ДПФ на стрибкових інтервалах за всіма вимірами базуються на таких рекурентних виразах [3]:

$$F_1''(k_1, k_2, \dots, k_r) = F_0''(k_1, k_2, \dots, k_r) + \Delta F_1(k_1, k_2, \dots, k_r) \cdot W^{\sum_{i=1}^r \frac{M_i k_i}{N_i}}, \quad (2)$$

де $F_1(k_1, k_2, \dots, k_r)$, $F_1''(k_1, k_2, \dots, k_r)$ та $F_0(k_1, k_2, \dots, k_r)$, $F_0''(k_1, k_2, \dots, k_r)$ – звичайне і модифіковане r -вимірне ДПФ r -вимірного вхідного сигналу $x(n_1, n_2, \dots, n_r)$ розміром $\prod_{i=1}^r N_i$ на черговому (першому) та попередньому

(нульовому) інтервалах вхідного сигналу, зсунутих відносно один одного на m_i відліків за i -м виміром, де $i = \overline{1, r}$; M_i – зсув вхідного сигналу на попередньому інтервалі відносно початку координат за i -м виміром; $W = \exp(-j2\pi)$, де $j = \sqrt{-1}$;

$$\Delta F_1(k_1, k_2, \dots, k_r) = \sum_{i=0}^{2^r-2} \sum_{n_1=a_1}^{b_1-1} \sum_{n_2=a_2}^{b_2-1} \dots \sum_{n_r=a_r}^{b_r-1} \Delta x(n_1, n_2, \dots, n_r) \cdot W^{\sum_{i=1}^r \frac{n_i k_i}{N_i}},$$

де $a_i = m_i s_i$ та $b_i = m_i q_i + N_i s_i$, де s_i – $(r-i)$ -й розряд двійкового представлення числа s , а $q_i = 1 - s_i$;

$$\Delta x(n_1, n_2, \dots, n_r) = x(N_1 q_1 + n_1 + M_1, N_2 q_2 + n_2 + M_2, \dots, N_r q_r + n_r + M_r) - x(n_1 + M_1, n_2 + M_2, \dots, n_r + M_r).$$

Рекурентні методи обчислення звичайних і стрибкових інтервалах за всіма вимірами модифікованих багатовимірних ДПХ на базуються на таких рекурентних виразах [3]:

$$H_1(k_1, k_2, \dots, k_r) = [H_0(k_1, k_2, \dots, k_r) + \Delta H_1(k_1, k_2, \dots, k_r)] \cdot \cos\left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i}\right) - [H_0(-k_1, -k_2, \dots, -k_r) + \Delta H_1(-k_1, -k_2, \dots, -k_r)] \cdot \sin\left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i}\right), \quad (3)$$

$$H_1''(k_1, k_2, \dots, k_r) = H_0''(k_1, k_2, \dots, k_r) + \Delta H_1(k_1, k_2, \dots, k_r) \cdot \cos\left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i}\right) + \Delta H_1(-k_1, -k_2, \dots, -k_r) \cdot \sin\left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i}\right), \quad (4)$$

де $H_1(k_1, k_2, \dots, k_r)$, $H_1''(k_1, k_2, \dots, k_r)$ та $H_0(k_1, k_2, \dots, k_r)$, $H_0''(k_1, k_2, \dots, k_r)$ – звичайне і модифіковане r -вимірне ДПХ r -вимірного

вхідного сигналу $x(n_1, n_2, \dots, n_r)$ на черговому та попередньому інтервалах вхідного сигналу;

$$\Delta H_1(k_1, k_2, \dots, k_r) = \sum_{i=0}^{2^r-2} \sum_{n_1=a_1}^{b_1-1} \sum_{n_2=a_2}^{b_2-1} \dots \sum_{n_r=a_r}^{b_r-1} \Delta x(n_1, n_2, \dots, n_r) \cdot \cos\left(2\pi \sum_{i=1}^r \frac{n_i k_i}{N_i}\right),$$

де $\text{cas}(X) = \cos(X) + \sin(X)$.

Рекурентні вирази для обчислення звичайних і модифікованих багатовимірних ДПФ і ДПХ на стрибкових інтервалах за одним виміром отримують з виразів (1) – (4) для $m_i = 0$ та $M_i = 0$ у коефіцієнтах W , $\cos$ та $\sin$ для $i = \overline{1, r-1}$. Рекурентні вирази для обчислення звичайних і модифікованих багатовимірних ДПФ

і ДПХ на ковзних інтервалах отримують з виразів (1) – (4) для $m_i = 1$, де $i = \overline{1, r}$.

Проведемо аналіз точності рекурентних методів обчислення звичайного багатовимірного ДПФ. Обчислення дійсної (Re) та уявної (Im) частин ДПФ за виразом (1) для дійсної послідовності $x(n_1, n_2, \dots, n_r)$ здійснюється за такими виразами:

$$Q(\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r)) = \sum_{s=0}^{2^r-2} \left[\sum_{n_1=a_1}^{b_1-1} \sum_{n_2=a_2}^{b_2-1} \dots \sum_{n_r=a_r}^{b_r-1} \Delta x(n_1, n_2, \dots, n_r) \cdot \left(1 + t_1 \varepsilon_{\delta_{s, n_1, n_2, \dots, n_r}} \right) \times \right. \\ \left. \times \left(-\sin \left(2\pi \sum_{i=1}^r \frac{n_i k_i}{N_i} \right) \right) \cdot \left(1 + t_2 \varepsilon_{\delta_{s, n_1, n_2, \dots, n_r}} \right) \cdot \prod_{i=1}^r \left(1 + \varepsilon_{\delta_{s_i}} \right) \cdot \prod_{j=\begin{cases} 1, s=0 \\ s, \text{інакше} \end{cases}}^{2^r-2} \left(1 + t_3 \varepsilon_{\delta_{s_j}} \right) \right]$$

де

$$t_1 = \begin{cases} 0, M_i + m_i < N_i \\ 1, \text{інакше} \end{cases}, \quad t_2 = \begin{cases} 0, (m_i = 1) \wedge (s = 0) \\ 1, \text{інакше} \end{cases} \quad \text{для } i = \overline{1, r}; \quad t_3 = \begin{cases} 0, m_i = 0 \\ 1, \text{інакше} \end{cases} \quad \text{для } i = \overline{1, r-1}.$$

Беручи до уваги, що результат обчислення значення X з врахуванням абсолютної похибки обчислення $E(X)$ визначається як $Q(X) = X + E(X)$, на підставі (7) та (8) можна отримати рекурентні вирази для визначення похибок обчислення дійсної та уявної частин ДПФ, які без врахування похибок обчислення вищих порядків мають такий вигляд:

$$E(\operatorname{Re} F_1(k_1, k_2, \dots, k_r)) = \operatorname{Re} F_1(k_1, k_2, \dots, k_r) \cdot \varepsilon_{\delta_1} + [\operatorname{Re} F_0(k_1, k_2, \dots, k_r) + \operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r)] \times \\ \times \cos \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) \cdot (\varepsilon_{\delta_1} + \varepsilon_{\delta_{11}}) - [\operatorname{Im} F_0(k_1, k_2, \dots, k_r) + \operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r)] \cdot \sin \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) \times \\ \times (\varepsilon_{\delta_2} + \varepsilon_{\delta_{12}}) + [E(\operatorname{Re} F_0(k_1, k_2, \dots, k_r)) + E(\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r))] \cdot \cos \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) - \\ - [E(\operatorname{Im} F_0(k_1, k_2, \dots, k_r)) + E(\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r))] \cdot \sin \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right), \quad (9)$$

$$E(\operatorname{Im} F_1(k_1, k_2, \dots, k_r)) = \operatorname{Im} F_1(k_1, k_2, \dots, k_r) \cdot \varepsilon_{\delta_4} + [\operatorname{Re} F_0(k_1, k_2, \dots, k_r) + \operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r)] \times \\ \times \sin \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) \cdot (\varepsilon_{\delta_1} + \varepsilon_{\delta_{13}}) + [\operatorname{Im} F_0(k_1, k_2, \dots, k_r) + \operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r)] \cdot \cos \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) \times \\ \times (\varepsilon_{\delta_2} + \varepsilon_{\delta_{14}}) + [E(\operatorname{Re} F_0(k_1, k_2, \dots, k_r)) + E(\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r))] \cdot \sin \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) + \\ + [E(\operatorname{Im} F_0(k_1, k_2, \dots, k_r)) + E(\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r))] \cdot \cos \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right), \quad (10)$$

де

$$E(\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r)) = \sum_{s=0}^{2^r-2} \sum_{n_1=a_1}^{b_1-1} \sum_{n_2=a_2}^{b_2-1} \dots \sum_{n_r=a_r}^{b_r-1} \Delta x(n_1, n_2, \dots, n_r) \cdot \cos \left(2\pi \sum_{i=1}^r \frac{n_i k_i}{N_i} \right) \times \\ \times \left(t_1 \varepsilon_{\delta_{s, n_1, n_2, \dots, n_r}} + t_2 \varepsilon_{\delta_{s, n_1, n_2, \dots, n_r}} + \sum_{i=\begin{cases} 1, n_i=a_i \text{ для } i=\overline{1, r} \\ \sum_{j=1}^r (n_j - a_j) \prod_{j=i+1}^r (b_j - a_j), \text{інакше} \end{cases}}^{\prod_{i=1}^r (b_i - a_i) - 1} \varepsilon_{\delta_{s_i}} + t_3 \cdot \sum_{j=\begin{cases} 1, s=0 \\ s, \text{інакше} \end{cases}}^{2^r-2} \varepsilon_{\delta_{s_j}} \right),$$

$$E(\text{Im } \Delta F_1(k_1, k_2, \dots, k_r)) = \sum_{s=0}^{2^r-2} \sum_{n_1=a_1}^{b_1-1} \sum_{n_2=a_2}^{b_2-1} \dots \sum_{n_r=a_r}^{b_r-1} \Delta x(n_1, n_2, \dots, n_r) \cdot \left(-\sin \left(2\pi \sum_{i=1}^r \frac{n_i k_i}{N_i} \right) \right) \times$$

$$\times \left(t_1 \varepsilon_{\delta_{s,n_1,n_2,\dots,n_r}} + t_2 \varepsilon_{\delta_{s,n_1,n_2,\dots,n_r}} + \sum_{i=1}^r \prod_{j=i+1}^r (b_j - a_j) - 1 \varepsilon_{\delta_{s,n_1,n_2,\dots,n_r}} + t_3 \cdot \sum_{l=\begin{cases} 1, s=0 \\ s, \text{інакше} \end{cases}}^{2^r-2} \varepsilon_{\delta_{s,n_1,n_2,\dots,n_r}} \right).$$

Враховуючи припущення, покладені в значення, внаслідок чого СКЗ похибок основу статистичного методу аналізу, що вхідний обчислення визначаються значеннями дисперсій сигнал є білим шумом і відповідно значення похибок обчислення (D), котрі отримуємо на перетворення також є білим шумом, математичне підставі (9) та (10): очікування похибок обчислення має нульове

$$D[E(\text{Re } F_1(k_1, k_2, \dots, k_r))] = D[\text{Re } F_1(k_1, k_2, \dots, k_r)] \cdot D[\varepsilon_0] +$$

$$+ (D[\text{Re } F_0(k_1, k_2, \dots, k_r)] + D[\text{Re } \Delta F_1(k_1, k_2, \dots, k_r)]) \cdot \cos^2 \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) \cdot (D[\varepsilon_0] + D[\varepsilon_u]) +$$

$$+ (D[\text{Im } F_0(k_1, k_2, \dots, k_r)] + D[\text{Im } \Delta F_1(k_1, k_2, \dots, k_r)]) \cdot \sin^2 \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) \cdot (D[\varepsilon_0] + D[\varepsilon_u]) +$$

$$+ (D[E(\text{Re } F_0(k_1, k_2, \dots, k_r))] + D[E(\text{Re } \Delta F_1(k_1, k_2, \dots, k_r))]) \cdot \cos^2 \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) +$$

$$+ (D[E(\text{Im } F_0(k_1, k_2, \dots, k_r))] + D[E(\text{Im } \Delta F_1(k_1, k_2, \dots, k_r))]) \cdot \sin^2 \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right), \quad (11)$$

$$D[E(\text{Im } F_1(k_1, k_2, \dots, k_r))] = D[\text{Im } F_1(k_1, k_2, \dots, k_r)] \cdot D[\varepsilon_0] +$$

$$+ (D[\text{Re } F_0(k_1, k_2, \dots, k_r)] + D[\text{Re } \Delta F_1(k_1, k_2, \dots, k_r)]) \cdot \sin^2 \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) \cdot (D[\varepsilon_0] + D[\varepsilon_u]) +$$

$$+ (D[\text{Im } F_0(k_1, k_2, \dots, k_r)] + D[\text{Im } \Delta F_1(k_1, k_2, \dots, k_r)]) \cdot \cos^2 \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) \cdot (D[\varepsilon_0] + D[\varepsilon_u]) +$$

$$+ (D[E(\text{Re } F_0(k_1, k_2, \dots, k_r))] + D[E(\text{Re } \Delta F_1(k_1, k_2, \dots, k_r))]) \cdot \sin^2 \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right) +$$

$$+ (D[E(\text{Im } F_0(k_1, k_2, \dots, k_r))] + D[E(\text{Im } \Delta F_1(k_1, k_2, \dots, k_r))]) \cdot \cos^2 \left(2\pi \sum_{i=1}^r \frac{m_i k_i}{N_i} \right), \quad (12)$$

де

$$D[E(\text{Re } \Delta F_1(k_1, k_2, \dots, k_r))] = \sum_{s=0}^{2^r-2} \sum_{n_1=a_1}^{b_1-1} \sum_{n_2=a_2}^{b_2-1} \dots \sum_{n_r=a_r}^{b_r-1} D[\Delta x(n_1, n_2, \dots, n_r)] \times$$

$$\times \cos^2 \left(2\pi \sum_{i=1}^r \frac{n_i k_i}{N_i} \right) \cdot \left(t_1 D[\varepsilon_0] + t_2 D[\varepsilon_u] + \sum_{i=1}^r \prod_{j=i+1}^r (b_j - a_j) - 1 D[\varepsilon_0] + t_3 \cdot \sum_{l=\begin{cases} 1, s=0 \\ s, \text{інакше} \end{cases}}^{2^r-2} D[\varepsilon_0] \right),$$

$$D[E(\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r))] = \sum_{s=0}^{2^r-2} \sum_{n_1=a_1}^{b_1-1} \sum_{n_2=a_2}^{b_2-1} \dots \sum_{n_r=a_r}^{b_r-1} D[\Delta x(n_1, n_2, \dots, n_r)] \times \\ \times \sin^2 \left(2\pi \sum_{i=1}^r \frac{n_i k_i}{N_i} \right) \cdot \left(t_1 D[\varepsilon_0] + t_2 D[\varepsilon_u] + \sum_{i=1}^r \prod_{j=i+1}^r (b_j - a_j) D[\varepsilon_0] + t_3 \cdot \sum_{l=\begin{cases} 1, s=0 \\ s, \text{інакше} \end{cases}}^{2^r-2} D[\varepsilon_0] \right).$$

Грунтуючись на (11) та (12), отримуємо ДПФ, який має такий вигляд:
рекурентний вираз дисперсії похибок обчислення

$$D[E(F_1(k_1, k_2, \dots, k_r))] = D[E(\operatorname{Re} F_1(k_1, k_2, \dots, k_r))] + D[E(\operatorname{Im} F_1(k_1, k_2, \dots, k_r))] = \\ = (2D[\varepsilon_0] + D[\varepsilon_u]) \cdot D[F_1(k_1, k_2, \dots, k_r)] + D[E(\Delta F_1(k_1, k_2, \dots, k_r))] + D[E(F_0(k_1, k_2, \dots, k_r))], \quad (13)$$

де

$$D[E(\Delta F_1(k_1, k_2, \dots, k_r))] = D[\Delta x(n_1, n_2, \dots, n_r)] \cdot \left[\left(\prod_{i=1}^r N_i - \prod_{i=1}^r (N_i - m_i) \right) \times \right. \\ \left. \times \left(\left(t_1 + t_3 (2^{2^{r-1}} - 2^{r-1}) + \frac{1}{2} \right) D[\varepsilon_0] + t_2 D[\varepsilon_u] \right) + \frac{1}{2} D[\varepsilon_0] \sum_{s=0}^{2^r-2} \left(\prod_{i=1}^r (b_i - a_i) \right)^2 \right] = D[\Delta x(n_1, n_2, \dots, n_r)] \cdot K_1.$$

Дисперсію похибок обчислення ДПФ в з урахуванням того, що $E(F_0(k_1, k_2, \dots, k_r)) = 0$ для залежності від номера ітерації обчислення p $M_i = 0$, де $i = \overline{1, r}$:
визначаємо на підставі рекурентного виразу (13)

$$D[E(F_p(k_1, k_2, \dots, k_r))] = (2D[\varepsilon_0] + D[\varepsilon_u]) \sum_{i=1}^p D[F_i(k_1, k_2, \dots, k_r)] + \sum_{i=1}^p D[E(\Delta F_i(k_1, k_2, \dots, k_r))]. \quad (14)$$

Враховуючи, що для $l > N_i / m_i$, де $i = \overline{1, r}$,

$$D[F_l(k_1, k_2, \dots, k_r)] = D[x(n_1, n_2, \dots, n_r)] \prod_{i=1}^r N_i, \\ D[\Delta x(n_1, n_2, \dots, n_r)] = 2D[x(n_1, n_2, \dots, n_r)],$$

вираз (14) приймає такий вигляд:

$$D[E(F_p(k_1, k_2, \dots, k_r))] = p \left[(2D[\varepsilon_0] + D[\varepsilon_u]) \prod_{i=1}^r N_i + 2K_1 \right] \cdot D[x(n_1, n_2, \dots, n_r)]. \quad (15)$$

В основу аналізу точності рекурентних методів обчислення звичайного багатовимірного ДПХ покладений аналіз точності обчислення значень $H_1(k_1, k_2, \dots, k_r)$ та $H_1(-k_1, -k_2, \dots, -k_r)$ згідно з (3). Оскільки вирази обчислення значень $H_1(k_1, k_2, \dots, k_r)$ та $H_1(-k_1, -k_2, \dots, -k_r)$ ДПХ мають однакову структуру з виразами (5) та (6) обчислення дійсної та уявної частин ДПФ, то для аналізу точності обчислення ДПХ можуть бути використані вирази, подібні до виразів (7) – (12), в яких замість значень $\operatorname{Re} F_1(k_1, k_2, \dots, k_r)$ та

$\operatorname{Im} F_1(k_1, k_2, \dots, k_r)$ ДПФ використовуються відповідно значення $H_1(k_1, k_2, \dots, k_r)$ та $H_1(-k_1, -k_2, \dots, -k_r)$ ДПХ, замість значень $\operatorname{Re} F_0(k_1, k_2, \dots, k_r)$ та $\operatorname{Im} F_0(k_1, k_2, \dots, k_r)$ ДПФ – відповідно значення $H_0(k_1, k_2, \dots, k_r)$ та $H_0(-k_1, -k_2, \dots, -k_r)$ ДПХ, замість значень $\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r)$ та $\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r)$ – відповідно значення $\Delta H_1(k_1, k_2, \dots, k_r)$ та $\Delta H_1(-k_1, -k_2, \dots, -k_r)$, а замість значень

$\cos\left(2\pi\sum_{i=1}^r\frac{n_i k_i}{N_i}\right)$ та $\left(-\sin\left(2\pi\sum_{i=1}^r\frac{n_i k_i}{N_i}\right)\right)$ – $\cos\left(-2\pi\sum_{i=1}^r\frac{n_i k_i}{N_i}\right)$. В результаті рекурентний та ітераційний вирази суми дисперсій похибок відповідно значення $\cos\left(2\pi\sum_{i=1}^r\frac{n_i k_i}{N_i}\right)$ та обчислення значень ДПХ можна записати так:

$$D[E(H_1(k_1, k_2, \dots, k_r))] + D[E(H_1(-k_1, -k_2, \dots, -k_r))] = (D[H_1(k_1, k_2, \dots, k_r)] + D[H_1(-k_1, -k_2, \dots, -k_r)]) \times \\ \times (2D[\varepsilon_0] + D[\varepsilon_u]) + (D[E(\Delta H_1(k_1, k_2, \dots, k_r))] + D[E(\Delta H_1(-k_1, -k_2, \dots, -k_r))]) + \\ + (D[E(H_0(k_1, k_2, \dots, k_r))] + D[E(H_0(-k_1, -k_2, \dots, -k_r))]), \quad (16)$$

$$D[E(H_r(k_1, k_2, \dots, k_r))] + D[E(H_r(-k_1, -k_2, \dots, -k_r))] = \\ = (2D[\varepsilon_0] + D[\varepsilon_u]) \sum_{l=1}^r (D[H_l(k_1, k_2, \dots, k_r)] + D[H_l(-k_1, -k_2, \dots, -k_r)]) + \\ + \sum_{l=1}^r (D[E(\Delta H_l(k_1, k_2, \dots, k_r))] + D[E(\Delta H_l(-k_1, -k_2, \dots, -k_r))]). \quad (17)$$

Беручи до уваги, що для $l > N_i/m_i$, де $i = \overline{1, r}$,

$$D[H_l(k_1, k_2, \dots, k_r)] + D[H_l(-k_1, -k_2, \dots, -k_r)] = 2D[F_l(k_1, k_2, \dots, k_r)], \\ D[E(\Delta H_l(k_1, k_2, \dots, k_r))] + D[E(\Delta H_l(-k_1, -k_2, \dots, -k_r))] = 2D[E(\Delta F_l(k_1, k_2, \dots, k_r))],$$

сума дисперсій похибок обчислення значень ДПХ буде вдвічі більша за значення дисперсії похибки обчислення ДПФ за виразом (15), тобто середнє значення дисперсії похибки обчислення одного значення ДПХ співпадає з дисперсією похибки обчислення значення ДПФ.

Аналіз точності рекурентних методів обчислення модифікованого багатовимірного ДПФ здійснюється на підставі аналізу обчислення дійсної та уявної частин ДПФ згідно з (2), котрі для дійсної послідовності $x(n_1, n_2, \dots, n_r)$ є такі:

$$\operatorname{Re} F_1^u(k_1, k_2, \dots, k_r) = \operatorname{Re} F_0^u(k_1, k_2, \dots, k_r) + \operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r) \cdot \cos\left(2\pi\sum_{i=1}^r\frac{M_i k_i}{N_i}\right) + \\ + \operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r) \cdot \sin\left(2\pi\sum_{i=1}^r\frac{M_i k_i}{N_i}\right), \quad (18)$$

$$\operatorname{Im} F_1^u(k_1, k_2, \dots, k_r) = \operatorname{Im} F_0^u(k_1, k_2, \dots, k_r) + \operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r) \cdot \cos\left(2\pi\sum_{i=1}^r\frac{M_i k_i}{N_i}\right) - \\ - \operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r) \cdot \sin\left(2\pi\sum_{i=1}^r\frac{M_i k_i}{N_i}\right). \quad (19)$$

Рекурентні вирази знаходження дійсних та уявних частин ДПФ з врахуванням похибок вигляд:

$$Q(\operatorname{Re} F_1^u(k_1, k_2, \dots, k_r)) = \left\{ Q(\operatorname{Re} F_0^u(k_1, k_2, \dots, k_r)) + \left[Q(\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r)) \cdot \cos\left(2\pi\sum_{i=1}^r\frac{M_i k_i}{N_i}\right) \cdot (1 + \varepsilon_{u1}) + \right. \right. \\ \left. \left. + Q(\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r)) \cdot \sin\left(2\pi\sum_{i=1}^r\frac{M_i k_i}{N_i}\right) \cdot (1 + \varepsilon_{u2}) \right] \cdot (1 + \varepsilon_{o1}) \right\} \cdot (1 + \varepsilon_{o2}), \quad (20)$$

$$Q(\operatorname{Im} F_1^u(k_1, k_2, \dots, k_r)) = \left\{ Q(\operatorname{Im} F_0^u(k_1, k_2, \dots, k_r)) + \left[Q(\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r)) \cdot \cos\left(2\pi\sum_{i=1}^r\frac{M_i k_i}{N_i}\right) \cdot (1 + \varepsilon_{u3}) - \right. \right.$$

$$- Q(\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r)) \cdot \sin \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) \cdot (1 + \varepsilon_{u4}) \cdot (1 + \varepsilon_{o3}) \cdot (1 + \varepsilon_{o4}). \quad (21)$$

На підставі (20) та (21) можна отримати рекурентні вирази похибок обчислення дійсної та уявної частин ДПФ, які без врахування похибок обчислення вищих порядків є такі:

$$\begin{aligned} E(\operatorname{Re} F_1''(k_1, k_2, \dots, k_r)) &= \operatorname{Re} F_1''(k_1, k_2, \dots, k_r) \cdot \varepsilon_{o2} + \operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r) \cdot \cos \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) \cdot (\varepsilon_{u1} + \varepsilon_{o1}) + \\ &+ \operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r) \cdot \sin \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) \cdot (\varepsilon_{u2} + \varepsilon_{o1}) + E(\operatorname{Re} F_0''(k_1, k_2, \dots, k_r)) + \\ &+ E(\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r)) \cdot \cos \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) + E(\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r)) \cdot \sin \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right), \end{aligned} \quad (22)$$

$$\begin{aligned} E(\operatorname{Im} F_1''(k_1, k_2, \dots, k_r)) &= \operatorname{Im} F_1''(k_1, k_2, \dots, k_r) \cdot \varepsilon_{o4} + \operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r) \cdot \cos \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) \cdot (\varepsilon_{u3} + \varepsilon_{o3}) - \\ &- \operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r) \cdot \sin \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) \cdot (\varepsilon_{u4} + \varepsilon_{o3}) + E(\operatorname{Im} F_0''(k_1, k_2, \dots, k_r)) + \\ &+ E(\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r)) \cdot \cos \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) - E(\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r)) \cdot \sin \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right). \end{aligned} \quad (23)$$

З врахуванням (22) та (23) отримуємо рекурентні вирази дисперсій похибок обчислення дійсної та уявної частин ДПФ:

$$\begin{aligned} D[E(\operatorname{Re} F_1''(k_1, k_2, \dots, k_r))] &= D[\operatorname{Re} F_1''(k_1, k_2, \dots, k_r)] \cdot D[\varepsilon_o] + \left[D[\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r)] \cdot \cos^2 \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) + \right. \\ &+ D[\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r)] \cdot \sin^2 \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) \cdot (D[\varepsilon_u] + D[\varepsilon_o]) + D[E(\operatorname{Re} F_0''(k_1, k_2, \dots, k_r))] + \\ &+ D[E(\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r))] \cdot \cos^2 \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) + D[E(\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r))] \cdot \sin^2 \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right), \end{aligned} \quad (24)$$

$$\begin{aligned} D[E(\operatorname{Im} F_1''(k_1, k_2, \dots, k_r))] &= D[\operatorname{Im} F_1''(k_1, k_2, \dots, k_r)] \cdot D[\varepsilon_o] + \left[D[\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r)] \cdot \cos^2 \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) + \right. \\ &+ D[\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r)] \cdot \sin^2 \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) \cdot (D[\varepsilon_u] + D[\varepsilon_o]) + D[E(\operatorname{Im} F_0''(k_1, k_2, \dots, k_r))] + \\ &+ D[E(\operatorname{Im} \Delta F_1(k_1, k_2, \dots, k_r))] \cdot \cos^2 \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right) + D[E(\operatorname{Re} \Delta F_1(k_1, k_2, \dots, k_r))] \cdot \sin^2 \left(2\pi \sum_{i=1}^r \frac{M_i k_i}{N_i} \right). \end{aligned} \quad (25)$$

Базуючись на (24) та (25), рекурентний вираз дисперсії похибок обчислення ДПФ записується так:

$$\begin{aligned} D[E(F_1''(k_1, k_2, \dots, k_r))] &= D[E(\operatorname{Re} F_1''(k_1, k_2, \dots, k_r))] + D[E(\operatorname{Im} F_1''(k_1, k_2, \dots, k_r))] = \\ &= D[F_1''(k_1, k_2, \dots, k_r)] \cdot D[\varepsilon_o] + D[\Delta F_1(k_1, k_2, \dots, k_r)] \cdot (D[\varepsilon_u] + D[\varepsilon_o]) + \\ &+ D[E(\Delta F_1(k_1, k_2, \dots, k_r))] + D[E(F_0''(k_1, k_2, \dots, k_r))]. \end{aligned} \quad (26)$$

Ітераційний вираз дисперсії похибок p отримуємо на підставі рекурентного виразу обчислення ДПФ в залежності від номера ітерації (26) з урахуванням того, що

$E(F_0^u(k_1, k_2, \dots, k_r)) = 0$ для $M_i = 0$, де $i = \overline{1, r}$:

$$D[E(F_p^u(k_1, k_2, \dots, k_r))] = D[\varepsilon_0] \sum_{i=1}^p D[F_i^u(k_1, k_2, \dots, k_r)] + (D[\varepsilon_0] + D[\varepsilon_u]) \times \\ \times \sum_{i=1}^p D[\Delta F_i(k_1, k_2, \dots, k_r)] + \sum_{i=1}^p D[E(\Delta F_i(k_1, k_2, \dots, k_r))]. \quad (27)$$

Враховуючи вище визначені значення $D[F_i(k_1, k_2, \dots, k_r)]$ та $D[\Delta x(n_1, n_2, \dots, n_r)]$, а також те, що для $l > N_i/m_i$, де $i = \overline{1, r}$,

$$D[\Delta F_i(k_1, k_2, \dots, k_r)] = 2 \left(\prod_{i=1}^r N_i - \prod_{i=1}^r (N_i - m_i) \right) D[x(n_1, n_2, \dots, n_r)],$$

вираз (27) приймає такий вигляд:

$$D[E(F_p^u(k_1, k_2, \dots, k_r))] = p \left[D[\varepsilon_0] \prod_{i=1}^r N_i + 2 \left(\prod_{i=1}^r N_i - \prod_{i=1}^r (N_i - m_i) \right) \cdot (D[\varepsilon_0] + D[\varepsilon_u]) + 2K_1 \right] \times \\ \times D[x(n_1, n_2, \dots, n_r)]. \quad (28)$$

Для аналізу точності рекурентних методів обчислення модифікованого ДПХ за виразом (4) можуть бути використані вирази, подібні до виразів (20) – (25), з заміною дійсних та уявних частин ДПФ відповідними значеннями ДПХ, в результаті чого рекурентний та ітераційний вирази суми дисперсій похибок обчислення значень ДПХ можна записати так:

$$D[E(H_1^u(k_1, k_2, \dots, k_r))] + D[E(H_1^u(-k_1, -k_2, \dots, -k_r))] = \\ = (D[H_1^u(k_1, k_2, \dots, k_r)] + D[H_1^u(-k_1, -k_2, \dots, -k_r)]) \cdot D[\varepsilon_0] + \\ + (D[\Delta H_1(k_1, k_2, \dots, k_r)] + D[\Delta H_1(-k_1, -k_2, \dots, -k_r)]) \cdot (D[\varepsilon_u] + D[\varepsilon_0]) + \\ + D[E(\Delta H_1(k_1, k_2, \dots, k_r))] + D[E(\Delta H_1(-k_1, -k_2, \dots, -k_r))] + \\ + D[E(H_0^u(k_1, k_2, \dots, k_r))] + D[E(H_0^u(-k_1, -k_2, \dots, -k_r))], \quad (29)$$

$$D[E(H_p^u(k_1, k_2, \dots, k_r))] + D[E(H_p^u(-k_1, -k_2, \dots, -k_r))] = \\ = D[\varepsilon_0] \sum_{i=1}^p (D[H_i^u(k_1, k_2, \dots, k_r)] + D[H_i^u(-k_1, -k_2, \dots, -k_r)]) + \\ + (D[\varepsilon_u] + D[\varepsilon_0]) \sum_{i=1}^p (D[\Delta H_i(k_1, k_2, \dots, k_r)] + D[\Delta H_i(-k_1, -k_2, \dots, -k_r)]) + \\ + \sum_{i=1}^p (D[E(\Delta H_i(k_1, k_2, \dots, k_r))] + D[E(\Delta H_i(-k_1, -k_2, \dots, -k_r))]). \quad (30)$$

Беручи до уваги, що для $l > N_i/m_i$, де $i = \overline{1, r}$,

$$D[H_i^u(k_1, k_2, \dots, k_r)] + D[H_i^u(-k_1, -k_2, \dots, -k_r)] = 2D[F_i^u(k_1, k_2, \dots, k_r)], \\ D[\Delta H_i(k_1, k_2, \dots, k_r)] + D[\Delta H_i(-k_1, -k_2, \dots, -k_r)] = 2D[\Delta F_i(k_1, k_2, \dots, k_r)], \\ D[E(\Delta H_i(k_1, k_2, \dots, k_r))] + D[E(\Delta H_i(-k_1, -k_2, \dots, -k_r))] = 2D[E(\Delta F_i(k_1, k_2, \dots, k_r))],$$

сума дисперсій похибок обчислення значень ДПХ буде вдвічі більша за значення дисперсії похибки обчислення ДПФ за виразом (28), тобто середнє значення дисперсії похибки обчислення одного значення ДПХ співпадає з дисперсією похибки обчислення значення ДПФ.

Використовуючи (15) та (28), можна визначити СКЗ похибок обчислення звичайних та модифікованих багатовимірних ДПФ і ДПХ на стрибкових і ковзних інтервалах за всіма і одним виміром, а також їхні відношення до СКЗ

вхідного сигналу з урахуванням значень дисперсій відносних похибок операцій множення й додавання $D[\varepsilon_{\nu}]$ та $D[\varepsilon_o]$, котрі для $(b+1)$ -розрядних чисел мають такі значення [4]: $2^{-2b}/3$ – для випадку округлення прямого, оберненого й

доповняльного коду та утинання доповняльного коду; $4 \cdot 2^{-2b}/3$ – для випадку утинання прямого й оберненого коду.

Прийнявши $D[\varepsilon_{\nu}] = D[\varepsilon_o] = D[\varepsilon]$, вирази (15) та (28) можна записати так:

$$D[E(F_p(k_1, k_2, \dots, k_r))] = p \left[3 \prod_{i=1}^r N_i + 2K_2 \right] \cdot D[\varepsilon] \cdot D[x(n_1, n_2, \dots, n_r)], \quad (31)$$

$$D[E(F_p^{\nu}(k_1, k_2, \dots, k_r))] = p \left[\prod_{i=1}^r N_i + 4 \left(\prod_{i=1}^r N_i - \prod_{i=1}^r (N_i - m_i) \right) + 2K_2 \right] \cdot D[\varepsilon] \cdot D[x(n_1, n_2, \dots, n_r)], \quad (32)$$

де

$$K_2 = \left(\prod_{i=1}^r N_i - \prod_{i=1}^r (N_i - m_i) \right) \cdot \left(t_1 + t_3(2^{2r-1} - 2^{r-1}) + t_2 + \frac{1}{2} \right) + \frac{1}{2} \sum_{i=0}^{2^r-2} \left(\prod_{i=1}^r (b_i - a_i) \right)^2.$$

Для проведення порівняльного аналізу точності рекурентних методів обчислення багатовимірних ДПФ і ДПХ визначимо дисперсії похибок обчислення на підставі (31) та (32) для

$r > 2$ та $N_i \gg m_i$, прийнявши $N_i = N$ та $m_i = m$ для $i = \overline{1, r}$, в результаті чого отримуємо:

$$D[E(F_p(k_1, k_2, \dots, k_r))] = D[E(F_p^{\nu}(k_1, k_2, \dots, k_r))] \approx pkm^2 N^{2(r-1)} \cdot D[\varepsilon] \cdot D[x(n_1, n_2, \dots, n_r)], \quad (33)$$

де $k = r$ та $k = 1$ при обчисленні на інтервалах за всіма вимірами та одним виміром відповідно.

Враховуючи значення дисперсії вихідного сигналу, точність рекурентних методів обчислення багатовимірних ДПФ і ДПХ визначається через $pkm^2 N^{r-2} \cdot D[\varepsilon]$ і для $p = N$ та $k = m = 1$ складає $N^{r-1} \cdot D[\varepsilon]$, в той час як точність прямих методів обчислення багатовимірних ДПФ і ДПХ складає $N^r \cdot D[\varepsilon]/2$.

Таким чином, на підставі проведеного аналізу точності рекурентних методів обчислення ДПФ і ДПХ можна зробити такі висновки:

1. Точність рекурентних методів обчислення багатовимірних ДПФ збігається з точністю

рекурентних методів обчислення багатовимірних ДПХ.

2. Точність рекурентних методів обчислення звичайних багатовимірних ДПФ і ДПХ збігається з точністю рекурентних методів обчислення модифікованих багатовимірних ДПФ і ДПХ.

3. Точність рекурентних методів обчислення багатовимірних ДПФ і ДПХ на ковзних інтервалах в $N/2$ рази вища за точність прямих методів обчислення багатовимірних ДПФ і ДПХ.

4. Точність рекурентних методів обчислення багатовимірних ДПФ і ДПХ на ковзних інтервалах в m разів вища за точність рекурентних методів обчислення багатовимірних ДПФ і ДПХ на стрибкових інтервалах.

Список використаних джерел

1. Ярославский Л.П. Цифровая обработка сигналов в оптике и голографии: Введение в цифровую оптику. – М.: Радио и связь, 1987. – 296 с.
2. Цифровые анализаторы спектра / В.Н. Плотников, А.В. Белинский, В.А. Суханов, Ю.Н. Жигулевцев. – М.: Радио и связь, 1990. – 184 с.
3. Волинець В.І. Рекурентні методи

обчислення багатовимірних дискретних перетворень Фур'є та Хартлі // Вісник Вінницького політехнічного інституту. – 2003. – № 4. – С. 69-74.

4. Волинець В.І. Аналіз точності рекурентних методів обчислення дискретних перетворень Фур'є та Хартлі в арифметиці з плаваючою комою // Вісник Тернопільського державного технічного університету. – 2006. – Т. 11, № 4. – С. 151-160.

Надійшла до редколегії 05.03.08

УДК 519.863

Жук С.М., к.ф.-м.н., м.н.с.

Гарантоване оцінювання у лінійних дескрипторних різницевих рівняннях

У статті описано метод побудови гарантованих оцінок лінійних функцій, визначених на множині розв'язків лінійного різницевого рівняння, права частина якого і початкова умова є деякими, наперед невідомими елементами заданих множин. Спостереження за станом системи містять адитивний детермінований шум. Для дескрипторних рівнянь спеціальної структури запропоновано мінімаксий фільтр. Результат ілюструється чисельним прикладом.

Ключові слова: мінімаксий оцінки, фільтр Калмана, дескрипторні системи, невизначеність.

*E-mail: beetle@unicyb.kiev.ua

Статтю представив д.ф.-м.н., проф. Наконечний О.Г.

Introduction

Suppose that $\{x_k\}_0^{N+1}$ satisfies the linear descriptor difference equation

$$F_{k+1}x_{k+1} - C_k x_k = B_k f_k, k = \overline{0, N} \quad (1)$$

with the initial condition

$$F_0 x_0 = B_{-1} f_{-1}, \quad (2)$$

where F_k, C_k are $m \times n$ -matrixes, B_k is $m \times p$ -matrix. We'll be interested in building the minimax approximation of the linear function¹

$$\ell(\{x_k\}) := \sum_{k=0}^{N+1} (\ell_k, x_k)_n, \ell_k \in \mathbb{R}^n,$$

assuming that state x_k observations are given in the form of

$$y_k = H_k x_k + g_k, k = \overline{0, N}, \quad (3)$$

where $\{g_k\}_0^N$ is some deterministic noise, H_k is $q \times n$ matrix and $\{f_k\}_{-1}^N, \{g_k\}_0^N$ are some arbitrary elements of the ellipsoid

¹ $(\cdot, \cdot)_n$ denotes inner product in $\mathbb{R}^n$.

Serhiy M. Zhuk, Ph.D.

Guaranteed estimations for the linear difference descriptor systems

This paper describes an approach to the guaranteed estimation of linear functions, defined on the solutions domain of the linear descriptor difference equation (LDDE) assuming its right-hand part and initial condition to be arbitrary elements of the given set. The state observations are supposed to contain an unknown deterministic noise. The minimax recursive estimation is introduced for the linear descriptor system with a special structure.

Key Words: guaranteed estimation, observation, uncertainty, Kalman filtering, minimax, linear descriptor systems.

$$G := \{(\{f_k\}_{-1}^N, \{g_k\}_0^N) : (Q_{-1}f_{-1}, f_{-1})_p + \sum_{k=0}^N (Q_k f_k, f_k)_p + (R_k g_k, g_k)_q \leq 1\}, \quad (4)$$

where Q_k, R_k are some symmetric positively-defined matrixes with appropriate dimensions.

Now we introduce a notion of the minimax a-posteriori set. At first we denote by $\mathcal{N}$ a collection of pairs $(\{x_k\}_0^{N+1}, \{f_k\}_{-1}^N)$ satisfying (1)-(2). Then, we set

$$\mathcal{J}_r(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1}) := \sum_{k=0}^{N+1} (Q_{k-1} f_{k-1}, f_{k-1})_p + (R_k (y_k - H_k x_k), y_k - H_k x_k)_q, \quad (5)$$

where $y_{N+1} = 0, H_{N+1} = 0$. If $(\{f_k\}_{-1}^N, \{g_k\}_0^N) \in G$ and satisfies (1)-(3) for some $\{x_k\}_0^{N+1}$, then

$$(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1}) \in \mathcal{N}, \mathcal{J}_r(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1}) \leq 1$$

and vice-versa if $(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1})$ satisfies last inequality then

$$(\{f_k\}_{-1}^N, \{g_k := y_k - H_k x_k\}_0^N) \in G$$

It means that we can describe a set of all $\{x_k\}_0^{N+1}$ causing to appearance of given $\{y_k\}_0^N$ in (3) while $(\{f_k\}_0^N, \{g_k\}_0^N)$ run through some subset² of G . The collection

$$G_y \stackrel{\text{def}}{=} \{ \{x_k\}_0^{N+1} | (\{f_k\}_0^N, \{x_k\}_0^{N+1}) \in \mathcal{N}, \\ (\{f_k\}_0^N, \{y_k - H_k x_k\}_0^N) \in G \} \quad (6)$$

is called a-posteriori set. Because of uncertain $(\{f_k\}_0^N, \{g_k\}_0^N)$ we use minimax strategy for finding

optimal estimation $\widehat{\ell}(\{x_k\})$ from within L : for each $\{\tilde{x}_k\}_0^{N+1} \in G_y$ we calculate the greatest distance $\sigma(\{\tilde{x}_k\}_0^{N+1})$ between $\ell(\{\tilde{x}_k\})$ and L . Then we set $\widehat{\ell}(\{x_k\}) = \ell(\{\tilde{x}_k\})$, where $\{\tilde{x}_k\}_0^{N+1}$ has a minimal $\sigma(\{\tilde{x}_k\}_0^{N+1})$.

Definition 1. Linear function $\widehat{\ell}(\{x_k\})$ is called minimax a-posteriori estimation if

$$\inf_{\{\tilde{x}_k\} \in G_y} \sup_{\{x_k\} \in G_y} |\ell(\{x_k\}) - \widehat{\ell}(\{\tilde{x}_k\})| = \\ \sup_{\{x_k\} \in G_y} |\ell(\{x_k\}) - \widehat{\ell}(\{x_k\})|$$

The non-negative number

$$\hat{\sigma} \stackrel{\text{def}}{=} \sup_{\{x_k\} \in G_y} |\ell(\{x_k\}) - \widehat{\ell}(\{x_k\})|$$

is called minimax a-posteriori error.

Minimax a-posteriori estimation

Theorema 1. If $\ell(\{x_k\}) = \sum_{k=0}^{N+1} (\ell_k, x_k)_n$ and

$$\{\ell_k\}_0^{N+1} \in L := \{ \ell_k = F'_k z_k - C'_k z_{k+1} + H'_k u_k, \\ F'_{N+1} z_{N+1} = \ell_{N+1}, z_k \in \mathbb{R}^m, u_k \in \mathbb{R}^q \}$$

then

$$\widehat{\ell}(\{x_k\}) = \sum_{k=0}^{N+1} (\ell_k, \hat{x}_k)_n,$$

$$\hat{\sigma} = (1 - \sum_{k=0}^N (y_k, R_k (y_k - H_k \hat{x}_k)_q)^2)^{\frac{1}{2}} \times (\sum_{k=0}^{N+1} (\ell_k, p_k))^{\frac{1}{2}} \quad (8)$$

where $\{\hat{x}_k\}_0^{N+1}$ is a solution of

$$F'_k z_k - C'_k z_{k+1} = H'_k R_k (y_k - H_k \hat{x}_k), k = \overline{0, N} \\ F'_{k+1} \hat{x}_{k+1} - C'_k \hat{x}_k = B'_k Q_k^{-1} B'_k z_{k+1}, \\ F'_0 \hat{x}_0 = B'_{-1} Q_{-1}^{-1} B'_{-1} z_0, F'_{N+1} z_{N+1} = 0 \quad (10)$$

and $\{p_k\}_0^{N+1}$ is a solution of

$$F'_{k+1} p_{k+1} = C'_k p_k + B'_k Q_k^{-1} B'_k d_{k+1}, k = \overline{0, N}, \\ F'_k d_k = C'_k d_{k+1} + \ell_k - H'_k R_k H_k p_k, \\ F'_{N+1} d_{N+1} = \ell_{N+1}, F'_0 p_0 = B'_{-1} Q_{-1}^{-1} B'_{-1} d_0$$

Theorema 2. If $\text{rank } F'_k \equiv n$ and $B_k \equiv E$ then

for any $\ell \in \mathbb{R}^n$ minimax a-posteriori estimation $\widehat{\ell}(\{x_k\})$ of inner product $(\ell, x_N)_n$ can be represented

as $\widehat{\ell}(\{x_k\}) = (\ell, \hat{x}_{N,N})_n$ where

$$\hat{x}_{k|k} = P_{k|k} F'_k (Q_{k-1}^{-1} + C_{k-1} P_{k-1|k-1} C'_{k-1})^{-1} \times \\ \times C_{k-1} \hat{x}_{k-1|k-1} + P_{k|k} H'_k R_k y_k, \\ P_{k|k} = (F'_k (Q_{k-1}^{-1} + C_{k-1} P_{k-1|k-1} C'_{k-1})^{-1} F_k + \\ H'_k R_k H_k)^{-1}, P_{00} = (F'_0 Q F_0 + H'_0 R_0 H_0)^{-1}, \\ \hat{x}_{00} = P_{00} H'_0 R_0 y_0, \quad (11)$$

Example.

Let us set $m=2, n=3, l=4, N=200$,

$$f_{-1} = \frac{2}{149}, F_k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

$$C_k \equiv C = \begin{pmatrix} \frac{1}{40} & 0.5 & 0 \\ 0.1 & \frac{1}{4} & \frac{3}{10} \end{pmatrix}, H_k \equiv H = \begin{pmatrix} 0.6 & 0.96 & 2 \\ 0.001 & 2.3 & 0.6 \\ 1 & 0.1 & 1 \\ 0 & 0 & 0.23 \end{pmatrix}$$

and let's choose $\{f_k\}_0^N$ and $\{g_k\}_0^N$ from the unit

sphere. Simulated state $\{x_k\}_0^{N+1}$ and it's minimax a-posteriori estimation $\{\hat{x}_k\}_0^{N+1}$ are shown on the (7) figure 1.

²This subset consists of all pairs $(\{f_k\}_0^N, \{g_k\}_0^N) \in G$ satisfying (1)-(3) for some $\{x_k\}_0^{N+1}$.

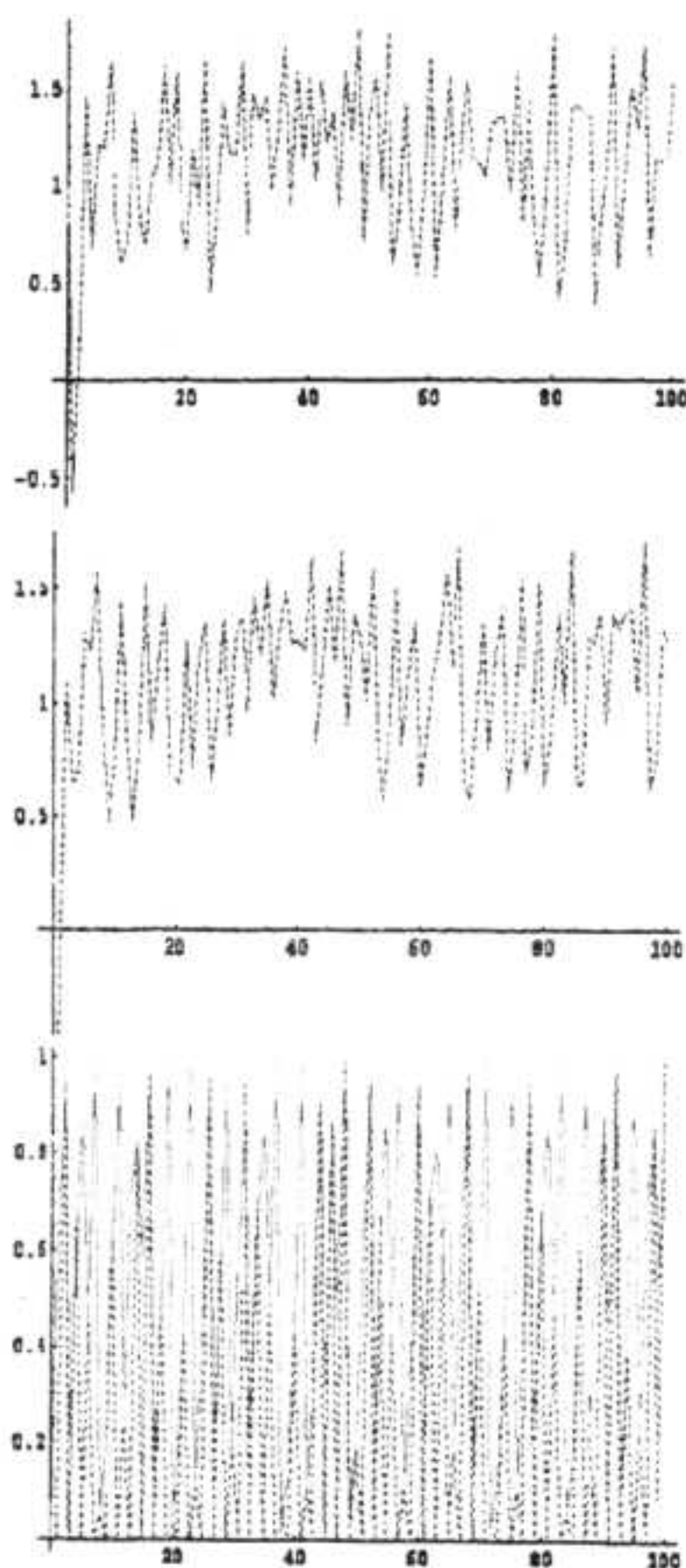


Figure 1: $\{\hat{x}_k\}_0^{N+1}$ (blue,dashed) and $\{x_k\}_0^{N+1}$.

Proofs.

Theorema 1 proof. The following lemma gives a criteria for the guaranteed estimation error finiteness.

Lemma 1.

$$\sup_{\{x_k\} \in G_y} |\ell(\{x_k\}) - \ell(\{\tilde{x}_k\})| < +\infty \Leftrightarrow \{\ell_k\} \in L$$

It's clear that $\sup_{d \in \mathbb{R}} |d - c| = D + |c|$ for any real D, c . This implies to

$$\begin{aligned} \sup_{\{x_k\} \in G_y} |\ell(\{x_k\}) - \ell(\{\tilde{x}_k\}_0^{N+1})| = \\ \frac{1}{2} [s(\{\ell_k\}_0^{N+1} | G_y) + s(-\{\ell_k\}_0^{N+1} | G_y)] + \\ |\ell(\{\tilde{x}_k\}_0^{N+1}) - \frac{1}{2} [s(\{\ell_k\}_0^{N+1} | G_y) - \\ s(-\{\ell_k\}_0^{N+1} | G_y)]| \end{aligned} \quad (12)$$

for $\{\ell_k\}_0^{N+1} \in L$, so

$$\begin{aligned} \widehat{\ell(\{x_k\})} = \frac{1}{2} [s(\{\ell_k\}_0^{N+1} | G_y) - s(-\{\ell_k\}_0^{N+1} | G_y)], \\ \hat{\sigma} = \frac{1}{2} [s(\{\ell_k\}_0^{N+1} | G_y) + s(-\{\ell_k\}_0^{N+1} | G_y)], \end{aligned} \quad (13)$$

$s(\cdot | G_y)$ denotes the support function of G_y .

Lemma 2. Vector $(\{\hat{f}_k\}, \{\hat{x}_k\}_0^{N+1})$, where $\hat{f}_k = B'_k z_{k+1}, k = \overline{-1, N}$ and $z_k, \hat{x}_k$ are some solutions of (10), gives a minimum to $\mathcal{J}_y$ on $\mathcal{N}$

$$\begin{aligned} \hat{J} \stackrel{\text{def}}{=} \inf_{(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1}) \in \mathcal{N}} \mathcal{J}_y = \\ \sum_{k=0}^N (R_k y_k, y_k - H_k \hat{x}_k)_q \\ \mathcal{J}_y(\{f_k - \hat{f}_k\}_{-1}^N, \{x_k - \hat{x}_k\}) = \\ \mathcal{J}_0(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1}) + \hat{J}, \end{aligned} \quad (14)$$

for any³ $(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1}) \in \mathcal{N}$. If we set

$$P(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1}) = \{x_k\}_0^{N+1},$$

$$\tilde{G} \stackrel{\text{def}}{=} \{(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1}) | \mathcal{J}_y(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1}) \leq 1\}$$

then $G_y = P(\tilde{G} \cap \mathcal{N})$ as it follows from G_y and G definitions. Formula (14) implies to

$$\tilde{G} \cap \mathcal{N} = (\hat{f}_k, \hat{x}_k) + G_0 \cap \mathcal{N},$$

where

$$G_0 \stackrel{\text{def}}{=} \{(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1}) : \mathcal{J}_0(\{f_k\}_{-1}^N, \{x_k\}_0^{N+1}) \leq 1 - \hat{J}\}$$

It's easy to see that $G_0 = -G_0$. Hence

$$s(\{\ell_k\}_0^{N+1} | G_y) = \sum_{k=0}^{N+1} (\ell_k, \hat{x}_k)_n + s(\ell_k | P(\tilde{G}_0 \cap \mathcal{N}))$$

for $\{\ell_k\}_0^{N+1} \in L$. Last formula with a regard to (13) implies to (7).

To prove (9) we need to find support function of the set $G_0 \cap \mathcal{N}$. This task [3, p.164,c.16.4.1] is

³ $\mathcal{J}_0 = \mathcal{J}_y$ for $\{y_k\}_0^N \equiv 0$.

equivalent to the minimization of G_0 -support function over affine set $P'(\{\ell_k\}) - \mathcal{N}^\perp$. It's easy to see that

$$\text{Arginf}_{P'(\{\ell_k\}) - \mathcal{N}^\perp} \subset \text{doms}(\cdot | G_0) \cap P'(\{\ell_k\}) - \mathcal{N}^\perp$$

Taking into account [3, p.136, T.13.5] we can show that

$$\text{doms}(\cdot | G_0) = \{H'_k R_k H_k x_k, k = \overline{0, N+1}\} (*)$$

so regarding to the structure of L it's not difficult to understand that the minimax error can be represented as (9).

Corollary 1. Suppose that all assumptions of the theorem 1 are fulfilled. To find a representation of

$\widehat{\ell(\{x_k\})}$ it is sufficient to find a solution $(\{\hat{f}_k\}, \{\hat{x}_k\}_0^{N+1})$ of the optimization task

$$\mathcal{J}_s(\{f_k\}_1^N, \{x_k\}_0^{N+1}) \rightarrow \inf_{\mathcal{N}}$$

Proof. We can represent the minimax a-posteriori estimation in terms of the a-posteriori set's support function (13). From the other hand each element of G_y could be treated as sum of $\{\hat{x}_k\}_0^{N+1}$ and some vector from balanced set $\mathcal{N} \cap G_0$.

Theorema 2 proof. Theorem 2 conditions implies L is equal to $(\mathbb{R}^n)^{N+1}$, so $(0, \dots, \ell) \in L$ for any $\ell \in \mathbb{R}^n$ and thus we can use corollary 1 from where and according to theorem (0) conditions we need to find a solution of

$$J_{N+1}(\{x_k\}) \stackrel{\text{def}}{=} \sum_{k=0}^{N+1} \|F_k x_k - C_{k-1} x_{k-1}\|_{Q_{k-1}}^2 + (15) \hat{f}_k = B'_k z_{k+1}, k = \overline{-1, N}.$$

$$\|y_k - H_k x_k\|_{R_k}^2, C_{-1} = 0, x_{-1} = 0$$

It is shown in [1] that we can obtain solution of (15) using recurrence process (11).

Lemma's 1 proof. Let $\{\ell_k\}_0^{N+1} \in L$. It's easy to see that

$$\sup_{\{x_k\} \in G_y} |\ell(\{x_k\}) - \ell(\{\tilde{x}_k\}_0^{N+1})| < +\infty \Leftrightarrow \sup_{\{x_k\} \in G_y} |\ell(\{x_k\})| < +\infty.$$

On the other hand if we set $B_{-1} = B, f_{-1} = f$ then after some simple calculations we obtain

$$\forall \{x_k\}_0^{N+1} \in G_y$$

$$\ell(\{x_k\}) = \sum_{k=0}^N (F'_k z_k - C'_k z_{k+1}, x_k)_n + (H_k x_k, u_k)_q + (z_{N+1}, F_{N+1} x_{N+1})_m = \sum_{k=0}^N (y_k - g_k, u_k)_q + \sum_{k=0}^{N+1} (B'_{k-1} z_k, f_{k-1})_p < +\infty,$$

because of (4). Now we'll assume that $\{\ell_k\}_0^{N+1} \notin L$.

It implies that $\ell(\{x_k\}) \neq 0$ for some

$$\{x_k\}_0^{N+1} : (\{F_k x_k, C_{k-1} x_{k-1}, H_k x_k\}) \equiv 0, \text{ so } \sup_{\{x_k\} \in G_y} |\ell(\{x_k\})| = +\infty.$$

Lemma's 2 proof.. Taking into account special structure of $\mathcal{J}_s$ and vector's projection theorem in Hilbert space [2] it's easy to prove that

$$\widehat{\mathcal{N}} \stackrel{\text{def}}{=} \text{Arginf}_{\mathcal{N}} \mathcal{J}_s \neq \emptyset$$

and

$$\sum_{k=0}^{N+1} (Q_{k-1} \hat{f}_{k-1}, f_{k-1})_p + (H'_k R_k (y_k - H_k \hat{x}_k), x_k)_n = 0,$$

$$\forall (\{f_k\}_1^N, \{x_k\}_0^{N+1}) \in \mathcal{N} \text{ if } (\{\hat{f}_k\}, \{\hat{x}_k\}_0^{N+1}) \in \widehat{\mathcal{N}}.$$

It implies

$$(Q\hat{f}, \{Q_k \hat{f}_k\}, \{H'_k R_k (y_k - H_k \hat{x}_k)\}) \in \mathcal{N}^\perp$$

$$\text{On the other hand } (\hat{f}, \{\hat{f}_k\}, \{\hat{x}_k\}_0^{N+1}) \in \mathcal{N}.$$

Now with a regard to structure of $\mathcal{N}$ we can show that exists $\{z_k\}$ which satisfies (10) and

References

- [1] Ishihara J.Y., Terra M.H., Campos J.C.T. Optimal recursive estimation for discrete-time descriptor systems // Int. J. of System Science."--- 2005."--- Vol. 36, N10."--- P. 1-22.
- [2] A.V.Balakrishnan Applied functional analysis // Springer-Verlag, 1976
- [3] R.Rockafellar Convex analysis // Princeton University Press, 1970
- [4] S.Zhuk Minimax estimations for linear descriptor difference equations // TWIM."--- 2005."--- N2."--- P. 10-20.

Received by editorial office at 7.02.2008

УДК 519.6

А.Л. Заворотний, к.ф.-м.н., В.С. Касьянюк,
к.ф.-м.н.

**Про один підхід до задачі відновлення
кривої за нечітко збуреними
координатами її точок**

В рамках теоретико-можливісного підходу розв'язана задача відновлення кривої при нечітких збуреннях в координатах її точок. Наведені результати обчислювального експерименту.

Ключові слова: теоретико-можливісний підхід, оптимізація за Парето, оцінювання.

Статтю представив професор, д.ф.-м.н. Белов Ю.А.

Задача відновлення кривої за скінченою множиною її точок не є суто теоретичною. Наявність додаткових умов, а саме: інформації про збурення в координатах відомих точок кривої, про природу цих збурень та ін., робить цю задачу важливою для практичного застосування. В такій постановці її можна використовувати, наприклад, при побудові контурів зон підповерхневих аномалій (зон залягання корисних копалин), при проведенні екологічного моніторингу (наприклад, для оконтурення зон розповсюдження забруднень), в геодезії, картографії, дослідженні земної поверхні з космосу, а також в багатьох інших областях. Важливим застосуванням цієї задачі є, зокрема, актуальна проблема визначення меж забруднення місцевості радіоактивними відходами.

Вимірювання характеристик, що визначають межі зони оконтурення, вимагають подальшого вдосконалення відомих методів і засобів як в напрямку модернізації та розширення спектру застосованої апаратури, так і розробки нових алгоритмів і програм обробки на ЕОМ результатів вимірювань з метою підвищення їх точності та продуктивності при проведенні робіт з локалізації аномальних зон.

При вимірюванні характеристик таких зон потрібно знаходити контури розповсюдження і об'єми тих речовин, які визначають ці зони, а також інші дані, що задаються скінченою множиною точок з похибками в їх координатах. На відміну від роботи [1], де такі похибки були випадковими величинами, в постановці, що розглядається в даній роботі, вважається, що похибки в координатах отриманої множини

Andriy L. Zavorotnyy, PhD, Veda S. Kasyanyuk,
PhD

**On one approach to curve recovery problem
by its fuzzy distorted coordinate**

Within the theory of possibilities approach the task of curve restoration by its fuzzy distorted coordinate set is solved. The computing experiment's results are submitted.

Key Words: the theory of possibilities approach, Pareto optimization, estimation

The problem of curve restoration by final set of its points is not academic. Presence of additional conditions, namely: the information about distortions in coordinates of known curve's points or the information about a distortion's nature etc., makes this problem important for practical application. In such statement it can be applied, for instance, to construction of subsurface anomalies arias' contours (zones of minerals bedding), to realization of ecological monitoring (for example, to contouring of pollution distribution zones), to geodesy, to cartography, to probing of the Earth surface from space, and also in many other areas. The important application of this problem is particularly urgent problem of delineation of radioactive pollution distribution zone.

Measurement of the characteristics which determine a zone delineation require of further improvement of both known methods and means in such directions as modernization and broaden out of a current technique's spectrum and development of new algorithms and PC programs for measurements results' processing with the purpose of increasing of its accuracy and productivity on localization of anomalies arias.

As a result of an anomalies arias' characteristics measuring it is expected to obtain contours of distribution and volumes of those substances, which form these zones, and also other data, which are performed as finite set of points with errors in their coordinates. In contrast to publication [1], where such errors were modeled as variates, within the article the errors in coordinates of the finite set of

мають нечітку природу.

Розглянемо задачу апроксимації контуру ABCDA (мал. 1). Для розв'язку цієї задачі, як і в [1], де дана задача розглядалась при умові випадкових похибок в даних, для побудови апроксиманти контуру послідовно апроксимуємо його ділянки AB, BC, CD та DA, використовуючи для цього поліноми степеня p .

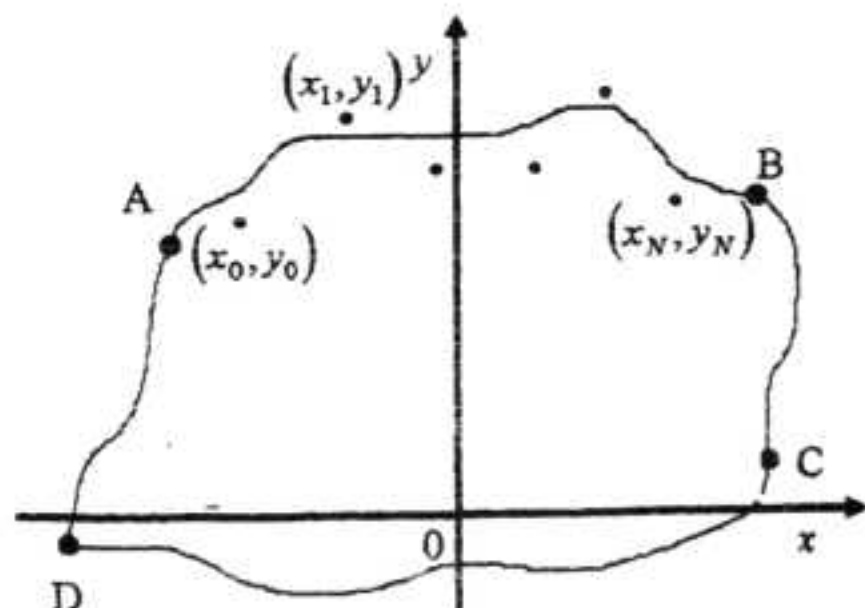


Рис. 1

Перетворення, що виконуються на кожній з ділянок, аналогічні з точністю до вибору ортогонального базису на площині, тому розглянемо одну з них, наприклад AB.

Припустимо, що на ділянці AB відома $N+1$ точка (x_i, y_i) , $i = \overline{0, N}$, координати яких спотворені нечіткими похибками.

Тоді проблема моделювання контуру за цими точками може розглядатись як задача оптимальної оцінки вектора параметрів $\vec{a}$ з лінійної системи

$$\vec{y} = G\vec{a} + \vec{v}, \quad (1)$$

де $\vec{a} = (a_0^*, \dots, a_p^*)^T$ – вектор невідомих коефіцієнтів полінома, $\vec{y} = (y_0^*, \dots, y_N^*)^T$ – вектор ординат відомих точок ділянки AB,

$$G = \begin{pmatrix} 1 & x_0 & \dots & x_0^p \\ 1 & x_1 & \dots & x_1^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & \dots & x_N^p \end{pmatrix}, \quad \vec{v} = (v_0^*, \dots, v_N^*)^T \text{ – вектор}$$

нечітких похибок, що спотворюють ординату відомих точок контуру. Зауважимо, що абсциса кожної з відомих точок x_i , $i = \overline{0, N}$ містить нечітку похибку: $x_i = x_i + \eta_i$.

Припустимо, що експертом заданий розподіл

measured points are interpreted in terms of fuzziness.

Let's consider an approximation problem of a contour ABCDA (fig. 1). To solve the problem, as well as in [1], where the given problem was considered in terms of probability, for delineation of the contour let's delineate consecutive its segments AB, BC, CD and DA, using for this purpose polynomial of a degree p

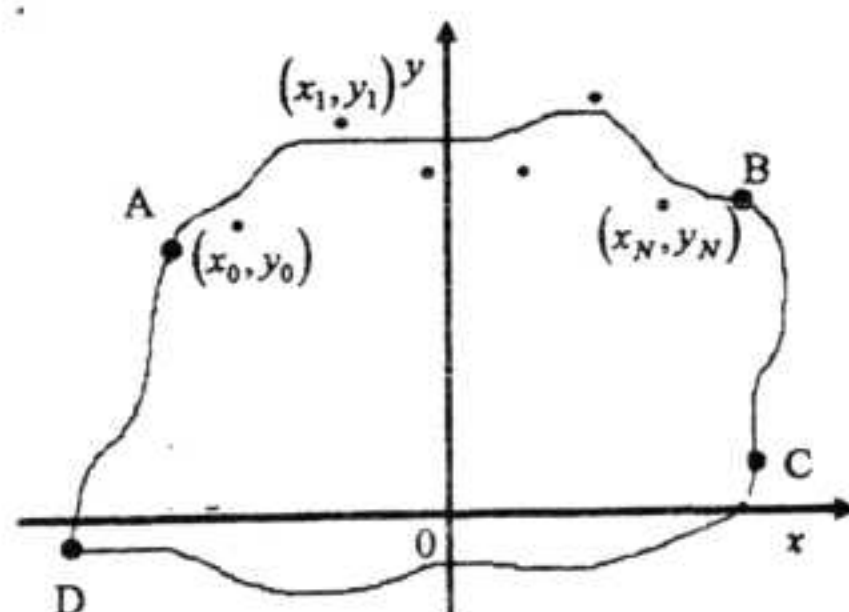


Fig. 1

The transformations, which are carried out on each of segments, are similar to within a choice of orthogonal basis on a plane, therefore we consider one of segments, for example AB.

Let's assume, that on a segment AB a known $N+1$ points (x_i, y_i) , $i = \overline{0, N}$. The points have coordinates, which is distorted by fuzzy errors.

Then the problem of contour modeling by these points can be considered as optimum estimation problem of a parameters vector $\vec{a}$ from linear system

$$\vec{y} = G\vec{a} + \vec{v}, \quad (1)$$

here $\vec{a} = (a_0^*, \dots, a_p^*)^T$ is a vector of unknown coefficients of the polynomial, $\vec{y} = (y_0^*, \dots, y_N^*)^T$ is a vector of ordinates of known points of a site AB,

$$G = \begin{pmatrix} 1 & x_0 & \dots & x_0^p \\ 1 & x_1 & \dots & x_1^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & \dots & x_N^p \end{pmatrix}, \quad \vec{v} = (v_0^*, \dots, v_N^*)^T \text{ is a vector of}$$

fuzzy errors, which distorted ordinates of contour's known points. Let's notice, that abscissa of each

нечітких елементів $\nu_i - \varphi^{\nu_i}(\cdot): R \rightarrow L$ та $\eta_i - \varphi^{\eta_i}(\cdot): R \rightarrow L$, $i = \overline{0, N}$, що реалізує принцип, природний для багатьох процесів вимірювання, – більший шум має меншу можливість. Тут R – простір дійсних чисел, L – шкала на сегменті $[0;1]$ з природнім порядком, який заданий нерівністю $\leq$ та двома правилами композиції: додаванням, що інтерпретується як $\max$, та множенням, що інтерпретується як $\min$ [2, 3]. Тоді можемо сказати, що найбільш можливі значення абсцис точок (x_i, y_i) , $i = \overline{0, N}$ є $x_i \in R$, $i = \overline{0, N}$, а можливість інших значень абсцис монотонно спадає з ростом норми відстані між ними та x_i , $i = \overline{0, N}$. Враховуючи визначення можливості події, введене в [2, 3], та умову монотонного незростання відносно норми аргументу, покладену на $\varphi^{\nu_i}(\cdot)$ та $\varphi^{\eta_i}(\cdot)$, отримуємо, що рівняння (1) можна переписати у вигляді

$$\bar{y} = X\bar{a} + \tilde{v}, \quad (2)$$

$$\text{де, } X = \begin{pmatrix} 1 & x_0 & x_0^p \\ 1 & x_1 & x_1^p \\ \vdots & \vdots & \vdots \\ 1 & x_N & x_N^p \end{pmatrix}, \quad \tilde{v} = \Lambda\bar{a} + \bar{v}, \quad \Lambda = G - X,$$

причому функція розподілу можливостей нечіткого елементу $\tilde{v} - \varphi^{\tilde{v}}(\cdot)$ також буде монотонно незростаючою відносно норми аргументу.

Будемо шукати розв'язок задачі (2) шляхом побудови перетворення B (в даному випадку матриця $(p+1) \times (N+1)$), яке б дозволило інтерпретувати вектор $B\bar{y}$ як найбільш точну версію $\bar{a}$. Для побудови перетворення B використаємо результати роботи [4] і поставимо задачу двокритеріальної оптимізації

$$\begin{cases} \|BX - E_{p+1}\| \rightarrow \min_B \\ \theta \sup_z \sup_u \min(\varphi^{\bar{y}}(z, \bar{a}), \theta l(\bar{a}, d(z))) \rightarrow \max_B \end{cases} \quad (3)$$

Іншими словами, (3) є задачею одночасної мінімізації норми операторної нев'язки та максимізації необхідності коректності оцінювання нечіткої величини, або ж інтегралу по необхідності від $l(\bar{a}, d(z))$, що записаний в

known point x_i , $i = \overline{0, N}$, contains a fuzzy error: $x_i = x_i + \eta_i$.

Let's assume, that some expert gave possibilities distribution functions of fuzzy elements $\nu_i - \varphi^{\nu_i}(\cdot): R \rightarrow L$ and $\eta_i - \varphi^{\eta_i}(\cdot): R \rightarrow L$, $i = \overline{0, N}$. For lots of measurement processes it's naturally to assume, that distribution functions realizes a principle - the greater noise has a smaller possibility. Here R is a space of real numbers, L - scale on a segment $[0;1]$ with the natural ordering relationship, which is given by an inequality $\leq$ and two rules of a composition: by addition, which is interpreted as $\max$, and multiplication, which is interpreted as $\min$ [2, 3]. Then we can say, that the most possible abscissas' values for points (x_i, y_i) , $i = \overline{0, N}$, are $x_i \in R$, $i = \overline{0, N}$, and the possibility of other abscissas' values monotonously falls down with growth of norm of distance between them and x_i , $i = \overline{0, N}$. Taking into account definition of event possibility submitted in [2, 3], and condition of monotonous nonincreasing concerning norm of argument imposed on $\varphi^{\nu_i}(\cdot)$ and $\varphi^{\eta_i}(\cdot)$, receive, that equation (1) is possible to rewrite as

$$\bar{y} = X\bar{a} + \tilde{v}, \quad (2)$$

$$\text{here } X = \begin{pmatrix} 1 & x_0 & x_0^p \\ 1 & x_1 & x_1^p \\ \vdots & \vdots & \vdots \\ 1 & x_N & x_N^p \end{pmatrix}, \quad \tilde{v} = \Lambda\bar{a} + \bar{v}, \quad \Lambda = G - X$$

and the possibilities distribution function $\varphi^{\tilde{v}}(\cdot)$ of fuzzy element $\tilde{v}$ also is be monotonously nonincreasing concerning norm of argument.

Let's search for the decision of the problem (2) by construction of transformation B (in this case it's a matrix $(p+1) \times (N+1)$), which would permit to interpret a vector $B\bar{y}$ as the exactest version $\bar{a}$. For construction of transformation B we use results of publication [4] and set up a bi-criterion optimization problem

$$\begin{cases} \|BX - E_{p+1}\| \rightarrow \min_B \\ \theta \sup_z \sup_u \min(\varphi^{\bar{y}}(z, \bar{a}), \theta l(\bar{a}, d(z))) \rightarrow \max_B \end{cases} \quad (3)$$

термінах теорії можливостей. В (3) E_{p+1} – одинична матриця розміру $(p+1) \times (p+1)$; $d(z)$ – стратегія оцінювання; $l(\bar{a}, d(z))$ – можливість відсутності похибки, що виникає при виборі $d(z)$ як значення $\bar{a}$ для кожного значення $\bar{a} \in R^{p+1}$ (нечітке відношення коректності); θ – інволюція, що визначена в [2, 3] – дуальний ізоморфізм, такий що $\theta(0)=1$, $\theta(1)=0$ і для довільних $c, b \in [0;1]$ відношення $c \leq b$, $\theta(c) \geq \theta(b)$ та $\theta(c) \lesssim \theta(b)$ еквівалентні ($\lesssim$ – відношення порядку, обернене до природного); $\varphi^{\bar{y}}(z, \bar{a})$ – сумісний розподіл нечітких векторів $\bar{y}$ та $\bar{a}$. Останнє, як легко бачити, дорівнює $\varphi^{\bar{v}}(z - X\bar{a})$. Оскільки $\bar{a}$ за умовою є довільним невідомим вектором простору R^{p+1} , його розподіл $\varphi^{\mu}(\bar{a})=1$. За нечітке відношення коректності $l(\bar{a}, Bz)$ приймемо розподіл $\varphi^{B\bar{v}}(B(z - X\bar{a}))$ нечіткої величини $B\bar{v}$, який можна виразити через $\varphi^{\bar{v}}(\cdot)$:

$$\varphi^{B\bar{v}}(B(z - X\bar{a})) = \max_w \left\{ \varphi^{\bar{v}}(w) \mid Bw = B(z - X\bar{a}) \right\}.$$

Очевидно, що в нашому випадку стратегія $d(z)$ це $Bz = \hat{\bar{a}}$.

Розв'язок задачі (3) методами, запропонованими в [4], приводить до оцінок вигляду $\hat{\bar{a}} = PX^-z$, де X^- – матриця, псевдообернена до X , P – будь-який ортогональний проектор з R^{p+1} в деякий простір R^k , $k = \overline{1, p+1}$. До речі, оцінка $\hat{\bar{a}} = X^-z$ ($P = E_{p+1}$, E_{p+1} – одинична матриця розмірності $(p+1) \times (p+1)$) максимально мінімізує норму операторної нев'язки.

Розглянемо кілька ілюстративних прикладів.

Нехай відомі спотворені похибками нечіткої природи координати чотирьох точок контуру, який потрібно відновити. Припустимо, що на думку експертів даний контур необхідно моделювати поліномом 5-го степеня. Зауважимо, що контур, за спотвореними координатами точок якого ведеться відновлення, був заданий поліномом 3-го степеня. Результати застосування розробленого методу представлені на мал. 2.

Результати застосування цього ж методу, але вже для випадку відновлення ділянки контуру,

In other words (3) is a problem of simultaneous minimization of operators misalignment's norm and maximization of necessity of fuzzy value assessment's correctness, or the same – integral by necessity from $l(\bar{a}, d(z))$ expressed in terms of the theory of possibilities. In (3) E_{p+1} is a unity matrix of the rank $(p+1) \times (p+1)$; $d(z)$ is an assessment strategy; $l(\bar{a}, d(z))$ is a possibility of error absence, which arises at a choice $d(z)$ as equivalent of $\bar{a}$ for each $\bar{a} \in R^{p+1}$ (fuzzy relation of a correctness); θ is an involution, which is defined in [2, 3] – dual isomorphism, such that $\theta(0)=1$, $\theta(1)=0$ and for the any $c, b \in [0;1]$ inequality relation $c \leq b$, $\theta(c) \geq \theta(b)$ and $\theta(c) \lesssim \theta(b)$ is equivalent (here $\lesssim$ is an ordering relationship, opposite to natural); $\varphi^{\bar{y}}(z, \bar{a})$ is a joint distribution of fuzzy vectors $\bar{y}$ and $\bar{a}$. Easy to see that the joint distribution is equaled to $\varphi^{\bar{v}}(z - X\bar{a})$. As long as by the data $\bar{a}$ is an undefined and unknown vector of R^{p+1} space, its distribution $\varphi^{\mu}(\bar{a})=1$. For the fuzzy relation of a correctness $l(\bar{a}, Bz)$ let's accept distribution $\varphi^{B\bar{v}}(B(z - X\bar{a}))$ of fuzzy value $B\bar{v}$, which could be expressed through $\varphi^{\bar{v}}(\cdot)$: $\varphi^{B\bar{v}}(B(z - X\bar{a})) = \max_w \left\{ \varphi^{\bar{v}}(w) \mid Bw = B(z - X\bar{a}) \right\}$. It is obvious that in our case the strategy $d(z)$ is $Bz = \hat{\bar{a}}$.

The decision of the problem (3) by methods offered in [4], leads to assessment $\hat{\bar{a}} = PX^-z$. Here X^- is a pseudoinverse matrix of X , P is any orthogonal projector from R^{p+1} to some R^k space, $k = \overline{1, p+1}$. By the way, the assessment $\hat{\bar{a}} = X^-z$ minimizes operators misalignment's norm as much as possible ($P = E_{p+1}$, E_{p+1} is a unity matrix of the rank $(p+1) \times (p+1)$).

Let's consider some illustrative examples.

Let for contour which needs to be restored is known four points' coordinates with errors of a fuzzy nature. Let's assume, that in the judgment of some expert the given contour is necessary to simulate by polynomial of 5-th degree. Notice, that the original contour is specified by polynomial of 3-rd degree. The results of the developed method application are submitted

заданої поліномом 1-го степеня, представлені на мал. 3. Припускалося, що експертам відома степінь полінома, яким потрібно моделювати ділянку контуру.

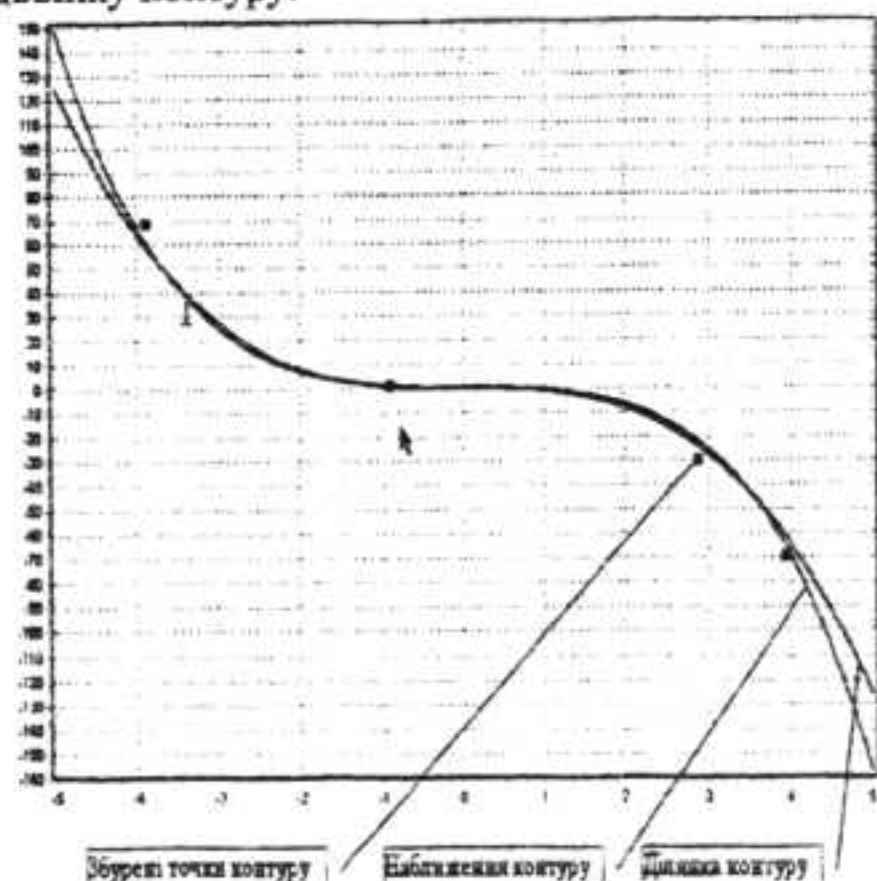


Рис. 2

В обох випадках ортогональний проектор P має той самий ранг, що й вектор $\hat{a}$.

Зауважимо, що нечіткі похибки в обох координатах відомих точок даний метод трансформує у нечіткі похибки тільки ординат. Ці нові похибки ординат відрізняються від початкових, але зберігають ту властивість, що розподіл можливостей їх значень є монотонно незростаючим відносно норми аргументу. При цьому вважається, що абсциси точок, за якими ведеться відновлення, відомі точно.

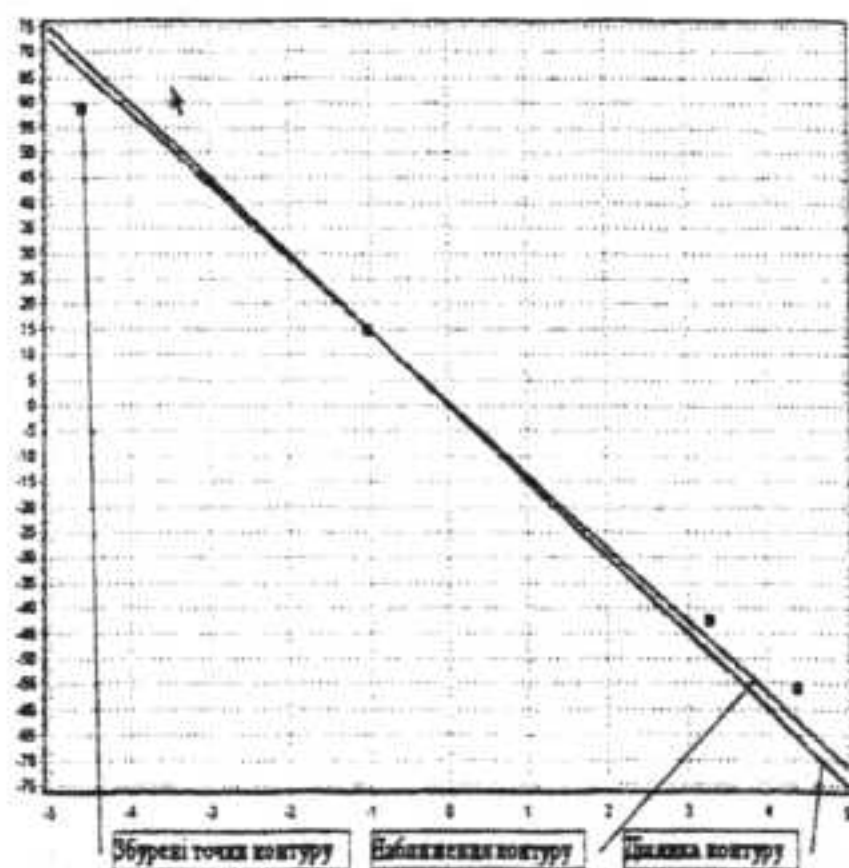


Рис. 3

on a fig. 2.

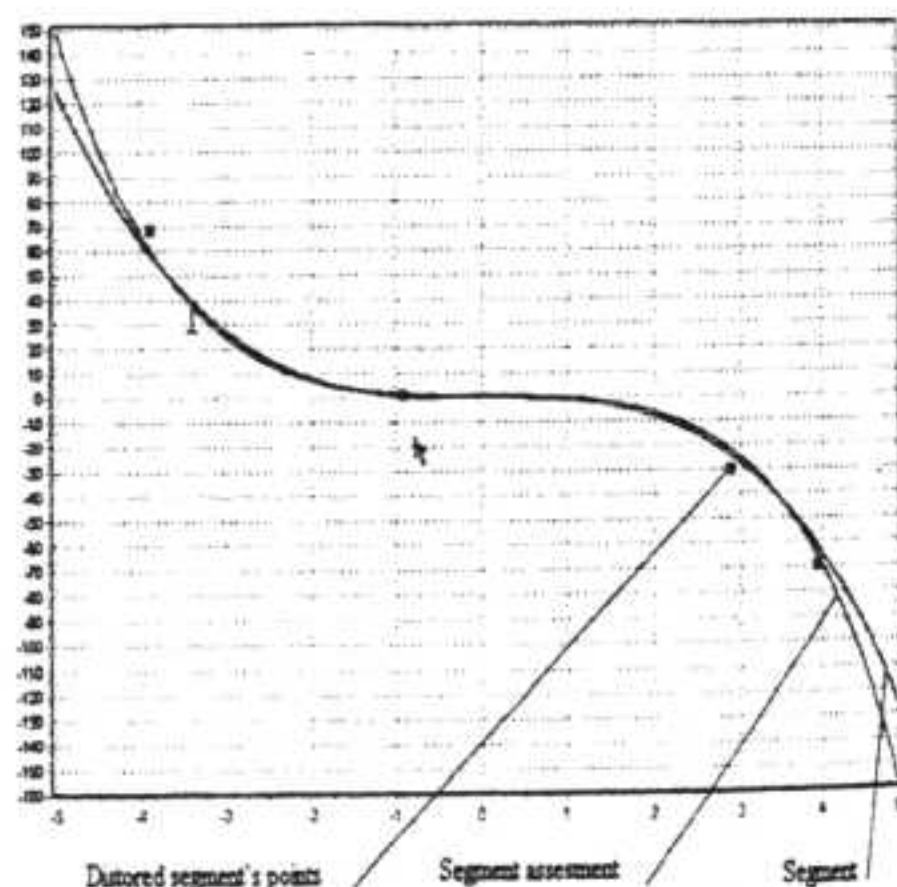


Fig. 2

The results of the same method application, however for the restoration of a contour's segment specified by polynomial of 1-st degree, are submitted on a fig. 3. It was supposed, that expert knew correct polynomial degree, with which it is necessary to simulate a contour's segment.

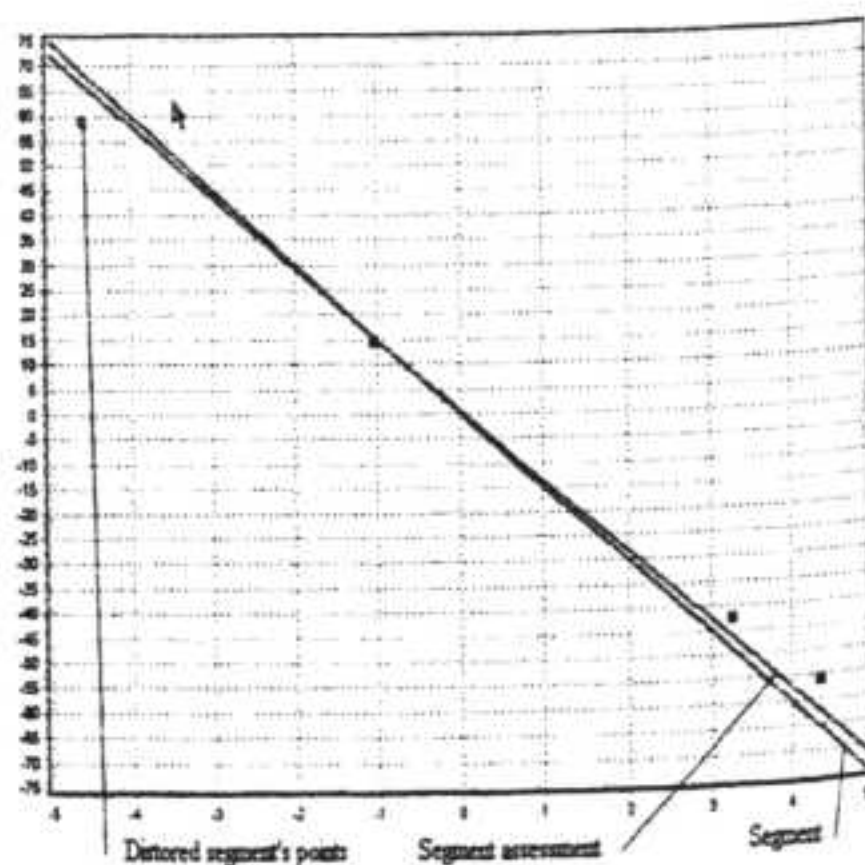


Fig. 3

In both cases the orthogonal projector is a unity matrix of the same rank as $\hat{a}$.

Let's notice, that the developed method transforms fuzzy errors in both coordinates of known

Завдяки тому, що постановка задачі враховує наявність похибок в даних та нечітку природу цих похибок, розроблений метод може бути успішно застосований в реальних умовах. Це, зокрема, демонструють представлені результати обчислювального експерименту. При цьому безсумнівною перевагою розробленого методу є невибагливість його до експертної інформації: необхідно відповісти лише на одне питання – чи буде зменшуватись можливість похибки із зростанням її норми. До того ж, як показано в [4], в рамках даного методу нескладно ще до побудови оцінок перевірити їх переконливість. Все це робить запропонований метод досить зручним інструментом в руках дослідника або інженера.

Список використаних джерел

1. Заворотний А.Л., Касьянюк В.С. *Парето-оптимальний підхід в задачі побудови контурів зон підповерхневих аномалій* // Київ, Вісник КНУ, серія фіз.-мат. науки, – 2001, – № 2, – С. 229-234.
2. Ю.П. Пытьев *Возможность. Элементы теории и применения* – М.: Эдиториал УРСС, 2000. – 192 с.
3. Ю.П. Пытьев *Неопределенные нечеткие модели и их применения* // Москва, Интеллектуальные системы, – 2003, – том 8, выпуск 1–4, – С. 147-310
4. Заворотний А.Л. *Теоретико-можливісний підхід до задачі редукції спотвореного значення лінійного оператора до значення іншого лінійного оператора* // Київ, Вісник КНУ, серія фіз.-мат. науки, – 2005, – № 3, – С. 268-274.

Надійшла до редакції 14.01.2008

points to fuzzy errors of ordinates. These new errors of ordinates differ from initial, whereas its save the behaviour, that the possibilities distribution of its values is monotonously nonincreasing concerning norm of argument. Thus it is considered, that abscissas of the points, on which the restoration is conducted, are known accurately.

Due to the fact that the statement of the problem takes into account presence of errors in the data and fuzzy nature of these errors the developed method can be successfully applied for real problems solving. It, in particular, is demonstrated by the submitted results of computing experiment. Thus doubtless advantage of the developed method is it's unpretentiousness about an expert information: it is necessary to answer only one a question - if the possibility of the errors is decrease while growing of its norm. In addition, as shown in [4], within the developed method it is simple to check up an assessment's persuasiveness even before it's calculation. All this makes the offered method easy-to-use tool for researcher or engineer.

References

1. Zavorotnyy A.L., Kasyanyuk V.S. *Pareto-optimum approach to modeling of undersurface anomaly area's contours problem* // Kyiv, Bulletin of University of Kyiv, Series: Physics & Mathematics, 2001, #2, 229-234 (in Ukrainian)
2. Pitiev Yu.P *The possibility. The elements of theory and applications*. URSS 2000. (in Russian), 192 p.
3. Pitiev Yu.P *Undefined fuzzy models and its applications* // Moscow, Intelligent systems, 2003, v. 8, iss. 1–4, 147-310 (in Russian)
4. Zavorotnyy A.L. *The theory of possibilities approach to the problem of reduction of the linear operator's distorted value to other linear operator's value* // Kyiv, Bulletin of University of Kyiv, Series: Physics & Mathematics, 2005, #3, 268-274 (in Ukrainian)

Has received at editorial board on 14.01.2008

УДК 681.3:004.056

Іванов І.Ю., аспірант

Про оптимізацію алгоритмів узгодження ключів в динамічних рівноправних групах

Розглядається питання ефективності алгоритмів групового узгодження ключів. Запропоновано спосіб обчислення вартості алгоритму для використання в конкретному середовищі. Наведено спосіб зменшення вартості виконання алгоритмів обміну ключами в високодинамічних групових середовищах.

Ключові слова: обмін ключами, динамічна група, вартість, оптимізація

*E-mail: Innokentiy@gmail.com

Статтю представив д.ф.м.н. проф. Анісімов А.В.

Останнім часом все більша увага приділяється розробці та аналізу алгоритмів забезпечення захищеного зв'язку в групах користувачів. Суттєвою складовою частиною будь-якого середовища, що має забезпечувати захищений обмін даними, є алгоритм узгодження спільного ключа між учасниками взаємодії. Протягом останніх років у літературі було запропоновано досить велику кількість алгоритмів узгодження ключів в групах користувачів, проте їхня ефективність практично не була досліджена у літературі. Тому задача аналізу ефективності існуючих алгоритмів є на сьогодні вкрай актуальною.

Під обчислювальним груповим середовищем (або обчислювальною групою) будемо розуміти множину обчислювальних пристроїв $M = \{m_i, i = 1, \dots, n\}$, що володіють спільним таємним ключем K , визначеним з використанням деякого алгоритму узгодження групового ключа, та можуть передавати між собою дані з метою розв'язання певної спільної задачі. Пристрої m_i називаються членами групи M , яка називається динамічною, якщо множина M не фіксована в часі, тобто до групи можуть приєднуватися нові члени та відокремлюватися існуючі. Група є рівноправною, якщо всі її члени мають однаковий вплив на функціонування групи. Рівноправність – суттєва вимога в багатьох

Ivanov I.Y.

On the optimization of key agreement algorithms in dynamic peer groups

The problem of effectiveness of group key agreement algorithms is considered. A method for calculating the total cost of key exchange algorithm for a particular environment is proposed. Means for decreasing the execution cost of key exchange algorithms in highly dynamic group environments is suggested.

Key words: key exchange, dynamic peer group, cost, optimization

In past years the growing interest in development and analysis of algorithms that provide secure group communications has been noticed. By the "secure group communication environment" term we understand the environment that provides such security features as confidentiality, authentication, access control etc. for the group of communicating peers. One of the core and most important parts of any secure group environment is group shared key agreement algorithm. In last several years there were proposed a big number of group key agreement algorithms, however their effectiveness almost has not been investigated. That's why the problem of analysis of effectiveness of the existing algorithms becomes quite actual today.

Computational group environment (or computational group) is a set of computational devices $M = \{m_i, i = 1, \dots, n\}$ that possess a shared secret key K , negotiated with the use of some group key agreement algorithm, who perform data exchange with the goal of solving some common problem. Devices m_i are referenced as members of group M . The group is called dynamic if M is not fixed in time, i.e. new members may join the group and existing members may leave it. The group is called peer group, if all members belonging to it have the same influence on the functioning of the group. Equality of rights is a reasonable requirement in many cases, as it allows minimizing the

випадках, оскільки вона дозволяє мінімізувати залежність групи від деякого її члена чи підмножини членів, а отже, зробити її більш стабільною та надійною.

Алгоритмом узгодження ключа називається алгоритм, що дозволяє створити та розповсюдити між учасниками взаємодії певну цифрову сутність K , що є відомою лише учасникам взаємодії (така цифрова сутність називається спільним ключем). Більшість сучасних алгоритмів узгодження ключа в групах користувачів базується на двосторонньому алгоритмі Діффі-Хеллмана (DH) [1]. Далі ми розглянемо три алгоритми – GDH [2], CKD [3] та BD [4], що також базуються на алгоритмі Діффі-Хеллмана, та проаналізуємо їх з точки зору ефективності. Було обрано саме ці алгоритми, оскільки вони якнайбільш підходять для використання у динамічних рівноправних групах користувачів та використовуються у декількох існуючих системах захищеної групової взаємодії. Далі ми наведемо лише описи алгоритмів; обґрунтування коректності та аналіз алгоритмів з точки зору забезпечуваного рівня безпеки можна знайти у відповідних джерелах ([2], [3], [4]).

Кожен алгоритм складається з двох під-алгоритмів: алгоритму злиття (M) та алгоритму розділення (P). Перший виконується при приєднанні до групи нових учасників, другий – при відокремленні від групи існуючих учасників взаємодії. У подальшому викладенні ми будемо користуватись наступними позначеннями: n – розмір існуючої групи, k – кількість учасників, що мають намір приєднатися, p – кількість учасників, що мають намір відокремитися, M_i – i -й учасник групи, r_i – експонента алгоритму DH, що відповідає i -му учаснику групи, L – множина учасників, що має намір залишити групу.

Алгоритм GDH (Group Diffie-Hellman).

GDH-M

1. M_n обирає нову експоненту $r_n' \in Z_q$, обчислює $g^{r_n' \cdot r_{n-1}'} \bmod p$ та надсилає отримане значення до M_{n+1} .
2. M_{n+j} обирає експоненту $r_{n+j} \in Z_q$, обчислює $g^{r_{n+j} \cdot r_{n-1}'} \bmod p$ та надсилає отримане значення M_{n+j+1} , $j \in [1, k-1]$.
3. При отриманні останнього значення, M_{n+k} надсилає його до групи широкомовним

dependency of the group on one or several group members and thus making the group more stable and robust.

Key agreement algorithm is an algorithm that allows generating and distributing some digital value K between group members in such way that it remains known only to the communicating parties (such digital value is called the shared key). Most of existing modern peer group key agreement algorithms are based on the classic two-party Diffie-Hellman (DH) key exchange algorithm [1]. Further in this paper we consider three algorithms, GDH [2], CKD [3] and BD [4], which are also based on Diffie-Hellman algorithm, and analyze them from the effectiveness point of view. We have chosen these three algorithms, as they can be effectively used in dynamic peer groups and are implemented in several existing secure group communication systems. We will only cite algorithm flow descriptions here; the proof of correctness and security analysis of the algorithms are available in the corresponding sources ([2], [3], [4]). All the mentioned algorithms are executed by the group members and do not rely on some trusted third party. GDH and BD algorithms are classified as contributory group key agreement algorithms (this feature makes them preferable to use in peer group environments); CKD is a centralized key agreement algorithm.

Each algorithm consists of two sub algorithms, merge algorithm (M) and partition algorithm (P). The first one is executed when new members are to join the existing group; the second one is run when some existing members are willing to leave the group. In the subsequent text we will use the following notation: n – the size of existing group, k – the number of members that wish to join the group, p – the number of members that wish to leave the group, M_i – i -th member of the group, r_i – DH exponent corresponding to the i -th member of the group, L – a set of members that wish to leave the group.

GDH Algorithm (Group Diffie-Hellman).

GDH-M

1. M_n generates a new exponent $r_n' \in Z_q$, computes $g^{r_n' \cdot r_{n-1}'} \bmod p$ and sends it to M_{n+1} .
2. M_{n+j} generates an exponent $r_{n+j} \in Z_q$, computes $g^{r_{n+j} \cdot r_{n-1}'} \bmod p$ and sends it to M_{n+j+1} , $j \in [1, k-1]$.
3. Having received the last accumulated value,

повідомленням.

4. При отриманні ширококомовного повідомлення кожен учасник $M_i, i \in [1, n+k-1]$ обчислює значення $g^{r_1 \dots r_{n-1} r_{n+k-1} / r_i} \bmod p$ та надсилає його до M_{n+k} .
5. При отриманні всіх значень-відповідей, M_{n+k} обирає нову експоненту r_{n+k} , обчислює множину значень $S = \{g^{r_1 \dots r_{n-1} r_{n+k} / r_i} \bmod p \mid \forall i \in [1, n+k-1]\}$ та надсилає їх до групи ширококомовним повідомленням.
6. При отриманні ширококомовного повідомлення, кожен учасник $M_i, i \in [1, n+k]$, обчислює ключ $K = (g^{r_1 \dots r_{n-1} r_{n+k} / r_i})^{r_i} \bmod p = g^{r_1 \dots r_{n-1} r_{n+k}} \bmod p$, що буде спільним для всіх учасників.

GDH-P

1. Контролер M_d обирає нову експоненту $r_d' \in Z_q$, обчислює множину значень $S = \{g^{r_d' / r_i} \bmod p \mid M_i \in L\}$ та надсилає їх до групи ширококомовним повідомленням.
2. При отриманні S кожен з учасників, що залишилися, обчислює ключ $K = (g^{r_d' / r_i})^{r_i} \bmod p = g^{r_d'} \bmod p$.

Алгоритм CKD (Centralized Key Distribution) CKD-M

1. Контролер M_1 обирає випадкове значення $r_1' \in Z_q$.
2. M_1 надсилає кандидатам на приєднання $M_{n+j}, j \in [1, k]$ значення $g^{r_1'} \bmod p$.
3. Кожен $M_{n+j}, j \in [1, k]$ обирає випадкове значення $r_{n+j}' \in Z_q$, обчислює значення $g^{r_{n+j}'}$ та надсилає його до M_1 .
4. M_1 обирає випадковий груповий ключ K , та для кожного $i \in [2, n+k]$ обчислює $K^{s^{r_i}}$ та надсилає всі обчислені значення ширококомовним повідомленням до групи.
5. При отриманні повідомлення кожен учасник групи може обчислити K .

CKD-P

1. Контролер M_1 обирає випадковий груповий ключ K , обчислює $K^{s^{r_i}}$ та надсилає всі обчислені значення

M_{n+k} broadcasts it to the group.

4. Having received a broadcast, each member $M_i, i \in [1, n+k-1]$ compute $g^{r_1 \dots r_{n-1} r_{n+k-1} / r_i} \bmod p$ and sends it to M_{n+k} .
5. Once all the answers are received, M_{n+k} generates new exponent r_{n+k} , computes a set of values $S = \{g^{r_1 \dots r_{n-1} r_{n+k} / r_i} \bmod p \mid \forall i \in [1, n+k-1]\}$ and broadcasts them to the group.
6. Having received a broadcast, each member $M_i, i \in [1, n+k]$ computes the key $K = (g^{r_1 \dots r_{n-1} r_{n+k} / r_i})^{r_i} \bmod p = g^{r_1 \dots r_{n-1} r_{n+k}} \bmod p$, which is shared between the parties.

GDH-P

1. The group controller M_d generates a new exponent $r_d' \in Z_q$, computes the set of values $S = \{g^{r_d' / r_i} \bmod p \mid M_i \in L\}$ and broadcasts them to the group.
2. Having received S , each remaining member computes the key: $K = (g^{r_d' / r_i})^{r_i} \bmod p = g^{r_d'} \bmod p$.

CKD Algorithm (Centralized Key Distribution) CKD-M

1. The group controller M_1 generates random value $r_1' \in Z_q$.
2. M_1 sends $g^{r_1'} \bmod p$ to members $M_{n+j}, j \in [1, k]$ willing to join the group.
3. Each $M_{n+j}, j \in [1, k]$ generates random value $r_{n+j}' \in Z_q$, computes $g^{r_{n+j}'}$ and sends it to M_1 .
4. M_1 generates random group key K , and for each $i \in [2, n+k]$ computes $K^{s^{r_i}} \bmod p, i \in [2, n+k]$, then it broadcasts all the calculated values to the group.
5. Once received the broadcast, every group member can compute K .

CKD-P

1. The group controller M_1 generates random group key K , computes $K^{s^{r_i}}$ and broadcasts all the values to the group.
2. Once received the broadcast, every group member can compute K .

широкомовним повідомленням до групи.

2. При отриманні повідомлення кожен учасник групи може обчислити K .

Алгоритм BD (Бурместера-Десмедта)

(виконується однаково для операцій приєднання та відокремлення від групи).

1. Кожен учасник M_i обирає випадкове значення $r_i' \in Z_q$, обчислює $Z_i = g^{r_i'} \bmod p$ та надсилає Z_i до групи широкомовним повідомленням.
2. Кожен учасник M_i при отриманні Z_{i-1} та Z_{i+1} обчислює значення

$$X_i = (Z_{i+1} / Z_{i-1})^{r_i} = g^{r_i+1^{r_i}-r_i^{r_i-1}} \bmod p$$
та надсилає його до групи широкомовним повідомленням.
3. Кожен учасник M_j при отриманні всіх X_i , $i \neq j$, обчислює

$$K = K_j = (Z_{j-1})^{r_j} X_j^{r_j-1} \dots X_{j+(n-2)} = g^{r_j+2^{r_j}+\dots+r_j^{r_j}} \bmod p$$

Суттєвими факторами, що впливають на ефективність алгоритмів, є кількість повідомлень, що передаються в ході виконання алгоритму (як простих, так і широкомовних), кількість операцій піднесення до степеня за модулем, та кількість операцій створення та перевірки цифрового підпису (останні необхідні у зв'язку з тим, що алгоритм Діффі-Хеллмана є вразливим до атак методом "людина посередині" [5], а тому підписуватись має кожне повідомлення, що передається в незашифрованому вигляді). Тому має сенс розглядати ефективність алгоритмів з точки зору саме цих показників. Проте, зазначимо, що порівнювати її на основі сумарної кількості всіх вищенаведених операцій є не зовсім коректним, оскільки деякі з операцій можуть виконуватись паралельно різними учасниками групи, зменшуючи, таким чином, загальний час виконання алгоритму. З метою усунення зазначеної некоректності введемо поняття раунду алгоритму:

Раундом є функціональна частина алгоритму узгодження групового ключа, що задовольняє наступним умовам:

- пересилка повідомлення мережею (простого або широкомовного) є раундом (відповідно, u-раунд та m-раунд),
- виконання операції піднесення до степеня за модулем є раундом (e-раунд),
- якщо декілька учасників можуть незалежно

BD Algorithm (Burmester-Desmedt)

(runs in the same way for merge and partition cases).

1. Each member M_i generates random value $r_i' \in Z_q$, computes $Z_i = g^{r_i'} \bmod p$ and broadcasts Z_i to the group.
2. Each member M_i , once Z_{i-1} and Z_{i+1} are received, computes

$$X_i = (Z_{i+1} / Z_{i-1})^{r_i} = g^{r_i+1^{r_i}-r_i^{r_i-1}} \bmod p$$
and broadcasts it to the group.
3. Having all X_i , $i \neq j$, each member M_j computes

$$K = K_j = (Z_{j-1})^{r_j} X_j^{r_j-1} \dots X_{j+(n-2)} = g^{r_j+2^{r_j}+\dots+r_j^{r_j}} \bmod p$$

It should be noted that the above description of the BD algorithms uses cyclic indices; i.e. the 0-th index corresponds to the n-th member, the (n+1)-th index corresponds to the 1st member, the (n+2)-th index corresponds to the 2nd member and so on.

The effectiveness of the algorithms depends on the following factors: the number of messages transferred over the network during protocol execution (both unicast and multicast), the number of modular exponentiation operations, and the number of signature generation and verification operations (the need for latter is caused by the vulnerability of Diffie-Hellman algorithm to a "man-in-the-middle" attack [5], that's why each message transferred over the network in unencrypted way must be signed and validated by the recipient(s)). That's why it makes sense to consider these values when discussing the effectiveness of the algorithms. However, we stress that it is not correct enough to compare the exact number of the corresponding operations, as some of them can be executed in parallel by different group members, thus decreasing the overall time of algorithm execution. In order to exclude the mentioned inaccuracy, let's introduce the concept of algorithm round:

Round is a functional part of the group key agreement algorithm that satisfies the following conditions:

- Sending message over the network (either unicast or multicast) is a round (u-round and m-round, accordingly),
- Modular exponentiation execution is a round (e-round),
- If several members can execute several modular exponentiation operations independently in

виконати операції піднесення до степеня за модулем одночасно, то всі ці операції відносяться до одного e-раунду,

- виконання операції підписування або перевірки підпису є раундом (відповідно, s-раунд та v-раунд),
- якщо декілька учасників можуть незалежно виконати операції підписування або перевірки підпису, то такі операції відносяться до одного s- або v- раунду.

У таблиці наведені кількості відповідних раундів у алгоритмах.

Таблица 1

Алгоритм	n_u	n_m	n_e	n_s	n_v
GDH-M	$n+2k-1$	2	$n+2k+1$	$k+3$	$n+2k+1$
GDH-P	0	1	$n-p$	1	1
CKD-M	k	2	$n+2k$	3	$k+2$
CKD-P	0	1	$n-p$	1	1
BD-M	0	$2n+2k$	3	2	$n+k+1$
BD-P	0	$2n-2p$	3	2	$n-p+1$

Table 1

Очевидно, що загальний час виконання алгоритму є сумою часів виконання раундів, з яких він складається. Узагальнюючи, сумарна вартість виконання алгоритму може бути записана наступним чином:

$$T = \tau_u n_u + \tau_m n_m + \tau_e n_e + \tau_s n_s + \tau_v n_v,$$

де $\tau_u, \tau_m, \tau_e, \tau_s, \tau_v$ – вартості виконання одного, відповідно, u-, m-, e-, s- та v-раунду, а n_u, n_m, n_e, n_s та n_v – кількості відповідних раундів у ході виконання алгоритму. Зазначимо, що термін “вартість” тут застосовується у найбільш широкому сенсі та залежить від особливостей конкретного середовища. Частковим і найбільш природним видом вартості може виступати час виконання раунду, проте вартість може також включати такі фактори, як рівень навантаження на мережу та кількість використовуваної оперативної пам’яті.

Наведена формула дозволяє порівнювати вартості алгоритмів узгодження спільного ключа, враховуючи особливості певного середовища. Використаємо її для порівняння ефективності розглянутих алгоритмів, визначивши вартість як загальний час їх виконання (у мс). Скористаємось наступними характеристиками середовища, отриманими експериментально (час виконання криптографічних операцій замірювався на

parallel, all these operations belong to the same e-round.

- Execution of signature generation or signature verification operation is a round (s-round and v-round, accordingly),
- If several members can execute several signing or verification operations independently in parallel, all these operations belong to the same s- or v-round.

The following table shows the number of rounds for the discussed algorithms:

It is obvious that the overall algorithm execution time is a sum of execution times of all the rounds forming the algorithm. In general, the total algorithm execution cost can be evaluated in the following way:

$$T = \tau_u n_u + \tau_m n_m + \tau_e n_e + \tau_s n_s + \tau_v n_v,$$

where $\tau_u, \tau_m, \tau_e, \tau_s, \tau_v$ – the costs of execution of single u-, m-, e-, s- and v-round, n_u, n_m, n_e, n_s and n_v – the number of corresponding rounds during algorithm execution. We note that the ‘cost’ concept is used here in the widest sense and depends on the exact characteristics of the particular environment. A particular (and the most natural) interpretation of the cost is round execution time, however, the cost might also include other factors, such as network load, CPU load and amount of memory used.

The above formula allows comparing the costs of DH-based group key agreement algorithms according to the specifics of particular environment. Let’s use it to compare the effectiveness of considered algorithms, defining the cost as the total time of their execution (in milliseconds). We will use the following environment characteristics obtained in an experimental way (the time of execution of cryptographic operations was measured on AMD Athlon 64x2 2.21GHz, 1GB RAM machine, the time of execution of network operations was calculated as arithmetic mean of

комп'ютері AMD Athlon 64x2 2.21GHz, 1GB RAM, час виконання мережевих операцій обчислено як середнє арифметичне значення часу передачі повідомлення мережею Інтернет на 5 географічно віддалених вузлів). В якості алгоритму цифрового підпису використано алгоритм RSA [6] з довжиною ключа, рівною довжині модуля схеми Діффі-Хеллмана.

Таблиця 2

Розмір ключа	$\tau_{\text{д}}$, мс	$\tau_{\text{м}}$, мс	$\tau_{\text{в}}$, мс	$\tau_{\text{р}}$, мс	$\tau_{\text{н}}$, мс
1024	136	440	65	16	0
2048			515	125	8
4096			4015	960	16

Table 2

За наведених значень ми отримаємо наступні формули для обчислення часу виконання алгоритмів за певних значень n та k :

Таблиця 3

	1024 біт	2048 біт	4096 біт
GDH-M	$201n + 418k + 857$	$659n + 1443k + 1642$	$4167n + 9294k + 7655$
GDH-P	$65n - 65p + 456$	$515n - 515k + 573$	$4015n - 4015k + 1416$
CKD-M	$65n + 266k + 928$	$515n + 1174k + 1271$	$4015n + 8182k + 3792$
CKD-P	$65n - 65p + 456$	$515n - 515k + 573$	$4015n - 4015k + 1416$
BD-M	$880n + 880k + 227$	$888n + 888k + 1803$	$896n + 896k + 13981$
BD-P	$880n - 880p + 227$	$888n - 888p + 1803$	$896n - 896p + 13981$

Table 3

Наступні графіки ілюструють залежність часу виконання кожного з алгоритмів приєднання (за наведених умов середовища) від кількості учасників, що приєднуються, при фіксованому n (для модулів довжини 1024 та 4096 біт відповідно).

time of message transmission over Internet to 5 hosts physically located in different geographic locations). We used RSA [6] as signature generation and verification algorithm with private key length equal to the length of Diffie-Hellman scheme module.

Using the above values, we get the following formulas for calculating the time of algorithm execution for the particular values of n and k :

The following charts illustrate the dependence of merge algorithms execution times on the number of members being added to the group for a fixed n (for modules lengths of 1024 bits and 4096 bits accordingly).

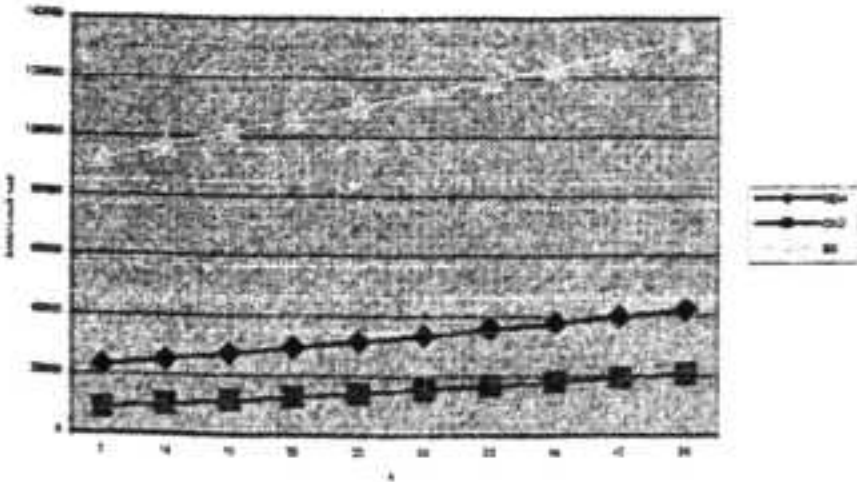


Рис. 1

З таблиці 1 та з графіків (рис. 1) можна побачити, що використання алгоритмів CKD та GDH є більш виправданим в середовищах, що забезпечують швидке виконання операцій

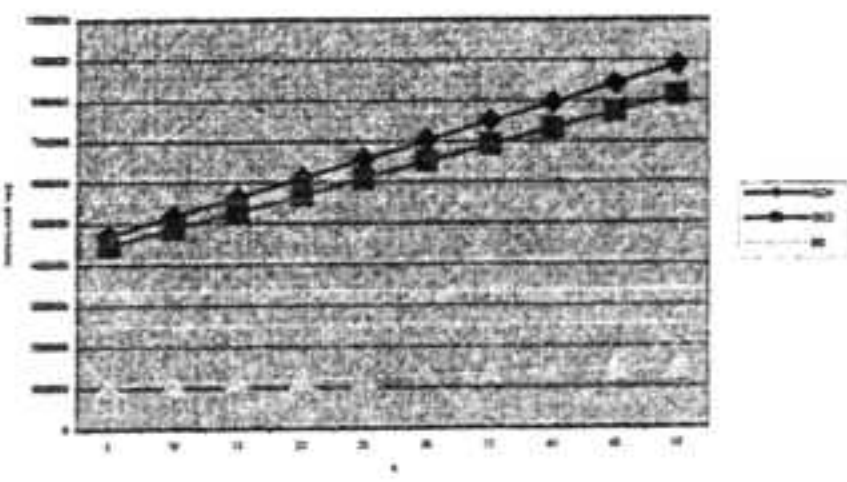


Рис. 1

Table 1 and the charts (pic. 1) show that the use of CKD and GDH algorithms is more suitable for environments that provide fast execution of modular exponentiation operations and digital signature

піднесення до степеня за модулем та цифрового підпису. Виконання алгоритму BD вимагає великої кількості m -раундів, а тому цей алгоритм є неефективним у середовищах, де передача широкомовного повідомлення виконується повільно. В той же час, алгоритм BD забезпечує константу кількість e -раундів (3), що робить привабливим його використання в середовищах, де операції піднесення до степеня за модулем не можуть бути виконані швидко. Алгоритми GDH та CKD показують майже однакову ефективність за даних умов. Враховуючи факт, що CKD в повній мірі не відповідає вимозі щодо максимальної рівноправності учасників групової взаємодії через наявність контролера, на якого покладаються спеціальні функції, GDH виглядає більш привабливим до використання серед цих двох алгоритмів.

З наведених оцінок також можна побачити, що у випадку високої динамічності групи та великої кількості учасників взаємодії, алгоритми узгодження ключа створюють суттєве навантаження як на обчислювальні ресурси учасників, так і на мережу. Це суттєво обмежує сферу використання алгоритмів потужними обчислювальними машинами, тому задача збільшення їх ефективності є досить актуальною на сьогодні. В якості одного з можливих способів вирішення цієї задачі в умовах високодинамічних групових середовищ, для яких вимога функціонування в реальному часі не є суттєвою, ми пропонуємо використовувати підхід, що базується на "затриманій" зміні складу групи. Його суть полягає в об'єднанні певної кількості послідовних індивідуальних операцій приєднання або відокремлення від групи в одну еквівалентну операцію приєднання або відокремлення множини учасників. Учасник, що має намір приєднатися до групи, спочатку додається до тимчасової групи кандидатів на приєднання. При настанні певного моменту часу ініціюється алгоритм узгодження ключів і учасники тимчасової групи, таким чином, стають дійсними учасниками групової взаємодії. Аналогічним чином виконується від'єднання підмножини дійсних учасників від групи.

Порівняємо вартості класичних алгоритмів та алгоритмів з "затриманою" зміною складу. Різниця Δ між вартостями k послідовних індивідуальних операцій приєднання та однієї операції з приєднання k нових учасників буде

generation and verification. From the other side, execution of Burmester-Desmedt algorithm requires a big number of m -rounds, and that's why this algorithm is not effective in environments that do not provide support for fast transmission of broadcast messages. Besides, BD algorithm requires the constant number of e -rounds (3) that makes it a preferable choice in environments that do not guarantee execution of fast modular exponentiation operations. GDH and CKD algorithms expose similar performance within the chosen environment. Taking into account the fact that CKD does not provide the complete equality of rights of the participants of group communication due to the presence of the controller member who performs special functions (key generation and distribution), GDH is more preferable choice between these two algorithms.

The above estimations show that in the case of high group dynamics level and big number of members taking part in communication, group key agreement algorithms expose significant computational resources and network load. This fact constrains the area of application of these algorithms to powerful computational machines, and that's why the problem of increasing their efficiency is quite actual today. One of the possible solutions for this problem in case of highly dynamic group environments which do not set up a claim for real-time functioning is to use the approach based on "delayed" group membership change. Its main idea is to replace a number of consecutive individual join or leave operations with a single equivalent merge or partition operation of a set of members. The member that is willing to join the group is first added to the temporary intermediate group of candidates for joining the group. At some moment in time, key agreement algorithm is invoked and the members of intermediate group thus become the actual members of the group communication. Partition operation is performed in similar way.

Let's compare the costs of the classic algorithms and algorithms with "delayed" membership change. The difference Δ between the costs of k consecutive individual merge operations and a single operation of adding k new members defines the positive effect of using the "delayed" membership change algorithm instead of classic algorithms. The corresponding differences are shown in the following table.

визначати позитивний ефект від використання алгоритмів “з затримкою” замість класичних алгоритмів. Різниці, відповідні алгоритмам, що розглядаються, наведені у наступній таблиці.

Таблиця 4

Table 4

	Δn_u	Δn_m	Δn_e	Δn_r	Δn_v
GDH-M	$\frac{1}{2}(k^2 + (2n-3)k - 2(n-1))$	$2k-2$	$\frac{1}{2}(k^2 + (2n+1)k - 2(n+1))$	$3k-3$	$\frac{1}{2}(k^2 + (2n+1)k - 2(n+1))$
GDH-P	0	$p-1$	$\frac{1}{2}(-p^2 + (2n+1)p - 2n)$	$p-1$	$p-1$
CKD-M	0	$2k-2$	$\frac{1}{2}(k^2 + (2n-1)k - 2n)$	$3k-3$	$2k-2$
CKD-P	0	$k-1$	$\frac{1}{2}(-p^2 + (2n+1)p - 2n)$	$p-1$	$p-1$
BD-M	0	$k^2 + (2n-1)k - 2n$	$3k-3$	$2k-2$	$\frac{1}{2}(k^2 + (2n+1)k - 2(n+1))$
BD-P	0	$-p^2 + (2n+1)p - 2n$	$3p-3$	$2p-2$	$\frac{1}{2}(-p^2 + (2n+3)p - 2(n+1))$

Залежність Δ від k при зафіксованому n ($n=50$) наведена на графіку:

Dependence of Δ on k with a fixed n ($n=50$) is shown on the following chart:

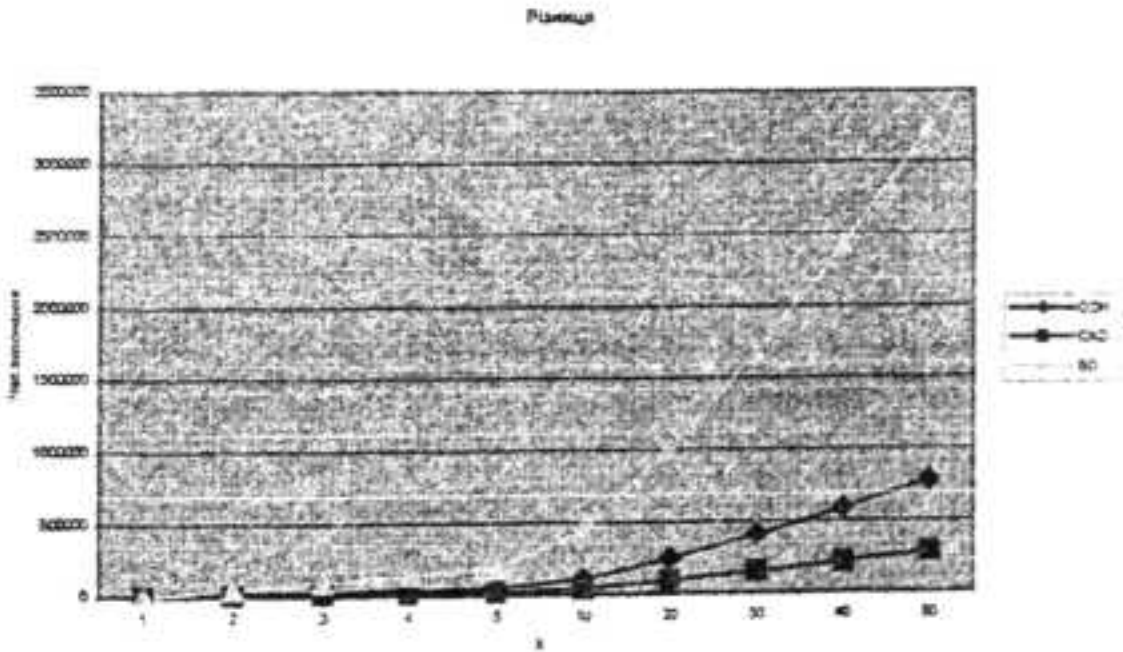


Рис. 2

Pic. 2

Можна побачити, що виграш від використання модифікацій алгоритмів з “затриманим” приєднанням швидко зростає зі зростанням k . Затримане приєднання є ефективним навіть при приєднанні двох нових учасників до групи (так, при використанні алгоритму CKD в умовах тестового середовища, при розмірі модуля 1024 біт, “затримане” приєднання двох нових учасників відбудеться на 4 секунди швидше, ніж їх окреме приєднання. Водночас, приєднання десяти нових членів за цим же алгоритмом відбудеться з вииграшем

The chart shows that the profit of using “delayed” membership change algorithms modification is greatly increased with the growth of k . The “delayed” join is effective even if two new members are joining the group (using CKD algorithm in the considered environment with DH module size of 1024 bits, the “delayed” merge of two members runs in 4 seconds faster than two consecutive individual join operations. At the same time, the profit of “delayed” join of ten new members is about 40 seconds).

The size of time delay between two consecutive

близько 40 секунд).

Розмір часової затримки між послідовними виконаннями алгоритму узгодження ключа має обиратися, виходячи з особливостей конкретного середовища і вимог до нього. В якості часткових випадків можна навести наступні:

- Фіксований розмір затримки. Протокол зміни складу групи виконується через рівні фіксовані проміжки часу.
- Гранична затримка. Протокол зміни складу виконується при досягненні проміжною групою певного розміру (або певної частки від розміру основної групи).
- Плаваюча затримка, що визначається на основі багатьох факторів, як то кількість поточних учасників в групі, рівень динамічності групи, поточні завдання групи і т.д.

Список використаних джерел

1. *W. Diffie, M. Hellman*. New Directions in Cryptography // IEEE Trans. Inform. Theory IT-22. – 1976. – P.644-654.
2. *M. Steiner, G. Tsudik, M. Waidner*. Key Agreement in Dynamic Peer Groups // IEEE Transactions on Parallel and Distributed Systems. – 2000. – 11, №8. – P.769-780.
3. *Y. Amir, Y. Kim, C. Nita-Rotaru, J. Schultz, J. Stanton, G. Tsudik*. Secure Group Communication using Robust Contributory Key Agreement // IEEE Transactions on Parallel and Distributed Systems. – 2004. – 15, №5. – P.468-480.
4. *M. Burmester, Y. Desmedt*. A Secure and Efficient Conference Key Distribution System // Advances in Cryptology. – EUROCRYPT'94. – 1995. – P.275-286.
5. *H.J. Highland*. A Clipper Primer // Computer Fraud and Security Bulletin. – 1994. – 1994, №5. – P.13-18.
6. *R. Rivest, A. Shamir, L. Adleman*. A Method for Obtaining Digital Signatures and Public-Key Cryptosystems // Communications of the ACM. – 1978. – 21, №2. – P.120-126.

invocations of key agreement algorithm should be taken basing on specifics of the particular environment and its requirements. In particular, the following approaches can be used:

- Fixed delay. The membership change algorithm is initiated in every fixed time interval.
- Threshold-based delay. The membership change algorithm is initiated when the size of intermediate group reaches some threshold value (that could be also expressed in terms of defined part of the general group).
- Flexible delay, calculated on the basis of many factors, such as current group size, the level of group dynamics, the actual problems being solved by the group etc.

References

1. *W. Diffie, M. Hellman*. New Directions in Cryptography // IEEE Trans. Inform. Theory IT-22. – 1976. – P.644-654.
2. *M. Steiner, G. Tsudik, M. Waidner*. Key Agreement in Dynamic Peer Groups // IEEE Transactions on Parallel and Distributed Systems. – 2000. – 11, №8. – P.769-780.
3. *Y. Amir, Y. Kim, C. Nita-Rotaru, J. Schultz, J. Stanton, G. Tsudik*. Secure Group Communication using Robust Contributory Key Agreement // IEEE Transactions on Parallel and Distributed Systems. – 2004. – 15, №5. – P.468-480.
4. *M. Burmester, Y. Desmedt*. A Secure and Efficient Conference Key Distribution System // Advances in Cryptology – EUROCRYPT'94. – 1995. – P.275-286.
5. *H.J. Highland*. A Clipper Primer // Computer Fraud and Security Bulletin. – 1994. – 1994, №5. – P.13-18.
6. *R. Rivest, A. Shamir, L. Adleman*. A Method for Obtaining Digital Signatures and Public-Key Cryptosystems // Communications of the ACM. – 1978. – 21, №2. – P.120-126.

Надійшла до редколегії 14.03.2008

УДК 510.65:510.67

Лялецький О.О., аспірант

Про типи збіжності послідовностей та топологічні властивості частково упорядкованих множин

В роботі проведено вивчення трьох понять типів збіжності послідовностей та їх взаємодії з топологічними властивостями частково упорядкованих множин. Це може дати можливість ввести поняття неперервності функції засобами, відмінними від традиційних, та побудувати новий клас моделей лямбда-числення.

Ключові слова: частково упорядкована множина, збіжність послідовності, топологія, неперервність.

mail: foraal@mail.ru

Статтю представив д.ф.-м.н., проф. Гаращенко Ф.Г.

Топологія є основним відомим підходом для завдання моделей лямбда-числення. Перша така модель, що була знайдена Д.Скоттом у 1969 році в категорії алгебраїчних решіток і неперервних функцій, була успішно використана для того, щоб показати, що всі нерозв'язні лямбда-терми можуть бути несуперечливо ототожнені. Після роботи Д.Скотта, велике число моделей лямбда-числення, що базуються на синтаксично вільних функціях, були введені в різних категоріях доменів і семантично класифіковані відповідно до тих функцій, що їх представляють. Зауважимо, що в 1972 році В.Скотт запропонував топологічну модель у термінах категорії, об'єктами якої є повні решітки, а морфізмами є функції, які неперервні за Скоттом. Надалі, ця топологічна семантика була узагальнена на один спеціальний клас декартово замкнених категорій, обумовлених за допомогою суцільно теоретико-категорних методів.

Однак усі ці моделі є структурно еквівалентно складними в тому розумінні, що денотаційні семантики є еквівалентно неповними, тобто вони не відбивають усі можливі операційні моделі лямбда-числення. Сказане, зокрема, вказує на те, що, незважаючи на великий обсяг роботи, пророблений вченими на даний час в області побудови і дослідженні семантики лямбда-числення, багато фундаментальних питань побудови моделей

O. O. Lyaletsky, a PhD student

On types of sequences convergence and topological properties of partially-ordered sets

Three notions of types of sequences convergence and their interconnection with topological properties of partially-ordered sets are investigated in the paper. This may give a possibility for defining the notion of continuity of functions by means distinguished from the traditional ones and for constructing a new class of models for the lambda-calculus.

Key Words: partially-ordered set, convergence of sequence, topology, continuity.

лямбда-числення, як суцільно екзистенціальних, так і інтенсіальних, усе ще залишаються відкритими. Таким чином, при побудові інтенсіальних моделей одне з ключових місць займає саме поняття неперервності, що, однак, може вводитися багатьма способами. Тому природно припустити, що нееквівалентним чином змінивши топологічне поняття неперервності, можна побудувати новий клас моделей лямбда-числення, що, можливо, не є підкласом топологічних чи навіть неперервних лямбда-моделей (звичайно, за умови, що змінене поняття неперервності визначає, зокрема, нетопологізовні класи функцій).

Автор даної роботи пропонує вводити поняття неперервні функції як функції, що зберігають границю послідовності. Оскільки поняття границі послідовності можна ввести різними способами над частково упорядкованими множинами, то воно дає можливість визначити нові поняття неперервності функції. Отже, в першу чергу встає потреба в вивченні типів збіжності послідовностей та їх взаємодії з топологічними властивостями частково упорядкованих множин. І тільки після цього можна буде перейти к вивченню неперервностей функції, які базуються на них.

Нижче досліджуються три різних поняття збіжності послідовності. Усі ці поняття збіжності мають ту загальну рису, що вони усі опираються

на поняття границі монотонної (тобто зростаючої чи спадної) послідовності; у випадку зростаючої послідовності ця границя за визначенням дорівнює супремуму множини всіх її елементів, а у випадку спадної - інфімуму.

Послідовності та типи їх збіжності

Введемо поняття зв'язаності елементів частково упорядкованої множини (чум надалі).

Припустимо, що в нас мається чум $\langle A, \leq \rangle$, і нехай $A = \{a_i: i < \alpha\}$ є повним упорядкуванням A за ординалами. Для кожної підмножини B множини A і кожного ординату β , що не є строго меншим α , ми таким чином вводимо множину B^β :

(I) $B^0 = B$; (II) $B^\beta = \{x \in A \mid \exists \text{ ординат } \gamma < \beta, \text{ такий, що } \exists y \in B^\gamma \text{ і або } x \leq y, \text{ або } y \leq x\}$.

Елементи x і y чум $\langle A, \leq \rangle$ називаються зв'язаними тоді і тільки тоді, коли $y \in \{x\}^\alpha$ для деякого α .

Легко бачити, що введене відношення зв'язаності є відношенням еквівалентності. Його класи еквівалентності називаються класами зв'язаності. Отже, кожне чум є об'єднання диз'юнктивного сімейства його класів зв'язаності.

Чум, що має єдиний клас зв'язаності, називається зв'язаним.

Частково упорядкована множина $\langle A, \leq \rangle$ називається (ко-)повною тоді і тільки тоді, коли кожна (ко-)спрямована непорожня множина має $\sup$ ($\inf$) (тобто супремум (інфімум)).

Чум, що є одночасно повною і ко-повною, називається дуальне повною (д-повною для стислості) частково упорядкованою множиною (д-п-чум).

Нехай $\langle A, \leq \rangle$ є чум, і B_1 і B_2 - його підмножини. Тоді говориться, що B_1 мажорирує B_2 (відносно $\leq$) тоді і тільки тоді, коли для кожного елемента b_1 з B_1 існує елемент b_2 з B_2 , такий, що має місце $b_1 \leq b_2$.

Далі буде використовуватися наступні пропозиції.

Лема 1. Для кожної спрямованої підмножини деякого частково упорядкованої множини $\langle A, \leq \rangle$ з того, що існує супремум її підмножини B , випливає, що існує підмножина C , мажоруюча B , що є максимальним ланцюгом у B і рівним $\sup B$.

Наслідок. Довільна частково упорядкована множина містить супремуми усіх своїх непорожніх спрямованих підмножин тоді і тільки тоді, коли воно містить супремуми всіх своїх непорожніх ланцюгів.

При розгляді своїх моделей Д. Скотт побудував деякі спеціальні топології на повних

решітках. Пізніше, ці топології були поширені на випадок повних частково упорядкованих множин. Нижче буде розглядатися ще більш загальний випадок.

Для кожен повної чум $\langle A, \leq \rangle$ розглянемо сімейство підмножин O множини A , щодо якого виконуються такі умови:

1. Ко-спадковість (що означає, що якщо елемент x належить A , те всі елементи з $A, \leq$ більші, ніж x , також належать A).

2. Для кожного спрямованої підмножини $B \subseteq A$ з того, що $\sup B \in O$, випливає, що перетинання $B \cap O$ є непорожньою множиною.

Легко переконатися в тому, що множини з будь-якого сімейства, що задовольняє перерахованим вище умовам для чум $\langle A, \leq \rangle$, задовольняють всім аксіомам відкритих множин. Тому ми можемо ввести топологію $S_\leq$. Такі топології звичайно називаються топологіями Скотта.

Головною властивістю топологій Скотта є те, що умова функціональної неперервності відносно топологій Скотта може бути виражене в термінах збереження супремума множин, для яких він існує.

Теорема (Д. Скотт). Нехай $\langle A_1, \leq_1 \rangle$ і $\langle A_2, \leq_2 \rangle$ є повними частково упорядкованими множинами. Тоді функція $f: A_1 \rightarrow A_2$ є неперервною відносно топологій $S_{\leq_1}$ і $S_{\leq_2}$ тоді і тільки тоді, коли для кожної непорожньої $S_{\leq_1}$ -спрямованої підмножини B_1 множини A_1 має місце рівність $f(\sup B_1) = \sup f(B_1)$, де другий супремум береться відносно $S_{\leq_2}$.

Тепер, використовуючи відношення часткового порядку, ми введемо наше поняття неперервної функції. Для цього, у першу чергу, ми визначимо поняття збіжності послідовності, а потім визначимо неперервність функції як властивість, що "зберігає" усі послідовності, що сходяться. Для того, щоб визначити поняття збіжності, ми узагальнюємо наступне стандартне поняття математичного аналізу, що оперує тільки з послідовностями дійсних чисел, індексованих натуральними числами:

1. Послідовність збігається до точки b (у звичайній топології) тоді і тільки тоді, коли всі її монотонні (зростаючі чи спадні) підпослідовності сходяться до b .

2. Кожна (нескінченна) підпослідовність послідовності є конфінальною.

Границя будь-якої зростаючої (спадної) послідовності, у свою чергу, визначається як супремум (інфімум) множини її елементів.

Таке визначення границі послідовності має наслідком те, що послідовність збігається до

точки b тоді і тільки тоді, коли всі її монотонні конфінальні підпослідовності сходяться до b (і множина таких під послідовностей є непорожньою).

Для цього, звичайно, ми повинні уточнити поняття досліджуваної на виводимість послідовності, яке пропонується використовувати переформулювання звичайних (узагальнених) послідовностей, розглянутих у рамках загальної топології, головних образом, у теорії збіжності Морлі-Сміта.

У цьому зв'язку, під послідовністю S розуміється трійка $\langle A, \langle I, D \rangle, s \rangle$, у якій A є множиною (звичайно званим носієм послідовності S , $\langle I, D \rangle$ є спрямованою чум (називаною множиною індексів послідовності S), і $s: I \rightarrow A$ є функція.

У випадку, коли множина індексів є фіксованою, послідовність S зазвичай ідентифікується з функцією s . Тому скрізь нижче ми, для зручності, S будемо ідентифікувати з $\langle s: I \rightarrow A, D \rangle$, а елементи множини $s(I)$ будуть називатися елементами S .

Потужність послідовності дорівнює, за визначенням, потужності множини її індексів.

Надалі буде передбачатися, що потужність кожної послідовності позитивна, тобто що вона містить принаймні один елемент.

Послідовності, індексовані початковим сегментом натуральних чисел, називаються простими й позначаються, як правило, $a_0, a_1, \dots, a_n, \dots$, де $a_0, a_1, \dots, a_n$ є елементи послідовності.

Нехай $S_1 = \langle s_1: I_1 \rightarrow A, D_1 \rangle$ і $\langle s_2: I_2 \rightarrow A, D_2 \rangle$ є послідовності над тим самим носієм A . S_1 називається підпослідовністю послідовності S_2 тоді і тільки тоді, існує функція $\pi: S_1 \rightarrow S_2$, що є ін'єктивною, монотонною (відносно D_1 і D_2) функцією, яка зберігає елементи (тобто для кожного індексу i_1 з I_1 має рівність $s_1(i_1) = s_2(i_1\pi)$). Якщо існує функція $\pi: S_1 \rightarrow S_2$, що задовольняє цим умовам і така, що для кожного індексу i_1 з I_1 і індексу i_2 з I_2 , що задовольняє $i_2 = S2(i_1\pi)$, тоді S_1 зветься конфінальною підпослідовністю послідовності S_2 .

Тепер ми готові ввести визначення границі послідовності. Для нього дуже важливим є поняття часткового порядку над носієм, що повинне задовольняти деяким додатковим умовам для супремума і інфімуму. Ця умова вимагає, щоб носій був дуальне повною чум (д-п-чум).

Поняття зростаючої (спадної) послідовності вводиться стандартним образом. При цьому, послідовність називається монотонною тоді і тільки тоді, коли вона є зростаючою чи спадною.

Легко перевіряється, що всі елементи кожної зростаючої (спадної) послідовності утворюють (ко-)спрямовану множину.

Нехай $S = \langle s: I \rightarrow A, D \rangle$ є послідовністю над д-п-чум $\langle A, D \rangle$. Якщо S є зростаюча послідовність, то вважаємо, що за визначенням її границя ($\lim$) дорівнює супремуму множини $s(I)$ його елементів (він повинний існувати). Якщо S є спадною послідовністю, то її границя ($\lim$) дорівнює інфімуму множини $s(I)$ його елементів.

Остаточно маємо наступне: якщо S є довільна послідовність (включаючи випадок, коли S є монотонна послідовність), те говорить, що S слабо збігається до границі ($\lim$) b тоді і тільки тоді, коли множина монотонних конфінальних підпослідовностей послідовності S не є порожньою, і кожна така підпослідовність збігається до b . У цьому випадку b називається слабою границею (weak limit) послідовності S , і ця ситуація позначається $\lim_w S = b$.

Послідовність S називається слабо збіжною (відносно $\leq$) тоді і тільки тоді, коли вона слабо збігається до деякої слабкої границі; у протилежному випадку S називається слабо розбіжною.

Легко перевіряється коректність цього визначення, тобто що для зростаючої послідовності, що супремум множини елементів будь-якої зростаючої послідовності збігається з границею (супремумом множини елементів) кожної її конфінальної і монотонної підпослідовності, і що для спадної послідовності має місце аналогічна умова.

Слабка збіжність послідовностей над д-п-чум має наступні основні властивості:

- для кожного д-п-чум $\langle A, \leq \rangle$ будь-яка послідовність над нею збігається не більш, ніж до однієї границі;

- для кожного індексу i послідовності $S = \langle s: I \rightarrow A, D \rangle$ і $\langle s \upharpoonright J, D \cap (J)^2 \rangle$, де J позначає множину $\{j \in I \mid \langle i, j \rangle \in D\}$ і $s \upharpoonright J$ є звуження функції s на множину $s \upharpoonright J$, або одночасно слабо сходяться до той же самою границі, або одночасно розходяться;

- якщо множина індексів послідовності S має найбільший елемент i , то вона слабо збігається до точки $s(i)$ (зокрема, кожна скінчена позитивна послідовність слабо збігається до елемента, що відповідає найбільшому індексу послідовності);

- якщо послідовність S_1 (над д-п-чум) слабо збігається до b , то будь-яка її підпослідовність S_2 є слабо збіжною тоді і тільки тоді, коли множина всіх монотонних підпослідовностей послідовності S_2 є непорожньою, причому, S_2 також слабо збігається до b .

Помітимо, що завдяки лемі 1, ми можемо звести дослідження слабкої збіжності послідовностей до дослідження слабкої збіжності тільки послідовностей з лінійно упорядкованими множинами індексів.

Дійсно, довільна послідовність S над д-п-чум $\langle A, \leq \rangle$ слабо збігається до b тоді і тільки тоді, коли множина всіх монотонних конфінальних підпослідовностей з лінійно упорядкованою множиною індексів є непорожньою, і усі вони сходяться до b .

Сказане вище і визначення слабкої границі послідовності мають наслідком наступні критерії, що підмножина A_I д-п-чум $\langle A, \leq \rangle$ є повною, ко-повною, чи д-повною:

- A_I є повною тоді і тільки тоді, коли A_I містить границі всіх зростаючих послідовностей над ним;

- A_I є повною тоді і тільки тоді, коли A_I містить границі всіх спадних послідовностей над ним;

- A_I є дуальне повною тоді і тільки тоді, коли A_I містить границі всіх послідовностей, що сходяться, над ним, що еквівалентно тому, що A_I містить усі границі всіх монотонних послідовностей з лінійно упорядкованими множинами індексів.

У випадках, коли частково упорядкована множина $\langle A, \leq \rangle$ не є лінійною, існують послідовності S , що містять елементи, що не беруть участь у визначенні слабкої границі, тобто послідовності, що не містять елементи, які не належать деякої конфінальної підпослідовності послідовності S . Такі елементи будуть називатися нейтральними.

Оскільки звичайний порядок на дійсних числах є лінійним, така ситуація не виникає в стандартному аналізі, і вона потребує окремого вивчення, що приводить до двох інших понять збіжності послідовностей.

Говориться, що послідовність $S = \langle s: I \rightarrow A, D \rangle$ над д-п-чум $\langle A, \leq \rangle$ середньо збігається до точки $b \in A$ тоді і тільки тоді, коли S слабо збігається до b і існує такий її індекс i , що кожен індекс j , який задовольняє $\langle i, j \rangle \in D$, не є нейтральним. У цьому випадку ми пишемо $\lim_a S = b$.

Якщо послідовність S слабо збігається до точки b і взагалі не містить нейтральних елементів, то S називається сильно збігається (відносно $\leq$) до b , і цей факт записується як $\lim_s S = b$.

Очевидно, що сильна збіжність послідовності до деякої точки b імплікує її збіжність у

середньому до тієї ж точки b , а середня збіжність імплікує слабку збіжність до b .

Транком T_S послідовності S над д-п-чум $\langle A, \leq \rangle$ є її підпослідовність, одержувана елімінацією в S усіх її нейтральних елементів.

Легко, бачити, що послідовність слабо збігається до точки b тоді і тільки тоді, коли її транк сильно збігається до b .

Тепер ми хочемо дати деяку топологічну характеристику трьом введеним вище поняттям збіжності. З цією метою по кожному д-п-чум $\langle A, \leq \rangle$ ми будемо так званий оператор замикання $\sigma_{lim-w}: P(A) \rightarrow P(A)$ у такий спосіб: для кожного B з A вираження $\sigma_{lim-w}(B)$ позначає множину всіх елементів b з A , таких, що існує послідовність S над B , що задовольняє рівності $\lim-w(S) = b$.

Оператори замикання σ_{lim-a} і σ_{lim-s} для двох інших типів збіжності визначаються аналогічно.

Якщо один із введених операторів задовольняє аксіомам оператора топологічного замикання, то в нас виникає задача порівняння $\lim-w$, $\lim-a$, і $\lim-s$ зі збіжністю в топології, що відповідає розглянутому оператору замикання. Вивчення цієї задачі проводиться нижче.

Насамперед помітимо, що для кожного д-п-чум $\langle A, \leq \rangle$, оператори $\lim-w$, $\lim-a$, і $\lim-s$ збігаються з оператором, що відображає кожен підмножину B з A у множину $\{b \mid \exists C (C \subseteq B \wedge (b = \sup C \vee b = \inf C))\}$. Тому далі, ми цей оператор позначимо σ_{lim} (чи $\sigma_{lim-\leq}$ у випадку, коли з контексту не ясно, як задане відношення $\leq$).

Очевидно, що для кожного д-п-чум $\langle A, \leq \rangle$ має місце рівність $\sigma_{lim-\leq}(\emptyset) = \emptyset$, і що він є зростаючим (тобто для кожної підмножини B з A має місце включення $B \subseteq \sigma_{lim-\leq}(B)$). Більш того, легко можна довести, що оператор $\sigma_{lim-\leq}$ зберігає скінченні об'єднання підмножин з A .

Отже, залишається тільки перевірити властивість ідемпотентності відносно $\sigma_{lim-\leq}$. Але ця властивість не обов'язкова має місце. Для того, щоб переконатися в цьому, давайте побудуємо наступний клас чум.

Чум $\langle A, \leq \rangle$ називається *FT-чум* тоді і тільки тоді, коли його носій може бути представлений у виді $A = B \cup (\cup \{N_b \mid b \in B\})$, де:

- (1) B є непорожній ланцюг, називана боул у випадку, коли в точності одне з властивостей $\sup B \notin B$ і $\inf B \notin B$ має місце, причому, при виконанні однієї з цих умов елемент (рівний $\sup B$ чи $\inf B$) називається коренем;

- (2) для кожного b з B , N_b є непорожній ланцюг називана нидл і така, що або $\sup N_b = b$, ($\sup N_b \notin$

N_b і $\inf B \in B$), або $\inf N_b = b$, ($\inf B \notin N_b$ і $\sup N_b \in N_b$);

(3) якщо $\inf B$ є корінь і нидл N_b є спрямованим униз, то $\inf B$ є непорівнянним з кожним елементом з N_b (чи має місце дуальна властивість).

Помітимо, що корінь будь-якого FT -чум не належить цьому чум. Також звертаємо увагу на те, що в загальному випадку елементи різних нидл можуть бути як порівнянними, так і не порівнянними.

Топологічні властивості частково упорядкованих множин

Припустимо, що мається д-п-чум $\langle A', \leq \rangle$, що включає в себе FT -чум $\langle A, \leq \cap A'^2 \rangle$ і нехай $r \in$ його корінь. Тоді $\sigma_{lim}(\cup \{N_b \mid b \in B\}) = A \setminus \{r\}$. У той же час $\sigma_{lim}(\sigma_{lim}(\cup \{N_b \mid b \in B\})) = A$, що зазначає, що оператор σ_{lim} не є ідемпотентним. Наступна пропозиція затверджує, що FT -чум і їхні корені дають "найпростіші" приклади д-п-чум у випадку, коли відповідні оператори замикання не є ідемпотентними.

Твердження 1. Оператор σ_{lim} є ідемпотентним для д-п-чум $\langle A, \leq \rangle$ тоді і тільки тоді, коли $\langle A, \leq \rangle$ не містить FT -подчум. Більш точно, якщо D є така підмножина множини A , що має місце $\sigma_{lim}(D) \neq \sigma_{lim}(\sigma_{lim}(D))$, то для кожного елемента r з $\sigma_{lim}(\sigma_{lim}(D))$ існує FT -підчум, що міститься в $\sigma_{lim}(D)$, коренем якого є r .

Доведення. Якщо $\langle A, \leq \rangle$ являється ізоморфним деякому FT -чум, то оператор σ_{lim} не є ідемпотентним. Отже, якщо $\langle A, \leq \rangle$ містить ізоморфну копію одного з цих д-п-чум, то σ_{lim} також не є ідемпотентним.

Тепер припустимо, що для д-п-чум $\langle A, \leq \rangle$ оператор σ_{lim} не є ідемпотентним. Тоді існує така підмножина D множини A і такий елемент r , що $r \in \sigma_{lim}(\sigma_{lim}(D))$. Але тому що має місце $r \notin \sigma_{lim}(D)$, то існує така спрямована підмножина B множини $\sigma_{lim}(D)$, що або $r = \sup B$, або $r = \inf B$.

Для визначеності розглянемо перший випадок.

Відповідно з лемою 1, без утрати загальності, ми можемо припустити, B є ланцюг і що $(\inf B) \in B$. Більш того, ми можемо припустити, що B не містить жодного елемента з D (якщо B не задовольняє цій умові, то ми можемо замінити його підланцюгом $B \setminus D$, супремум якого також дорівнює r). Для кожного елемента b з B існує спрямована (чи ко-спрямована) множина N_b , така, що $\sup N_b = b$ ($\inf N_b = b$ відповідно). Виходить,

множина B , розглянута як боул, і ланцюги вигляду N_b , розглянуті як нидл, утворюють FT -чум.

Для д-п-чум $\langle A, \leq \rangle$ позначимо σ_{lim}^* найменшу ідемпотентну функцію (відносно поточкового порядку на множині функцій), що є більшою, ніж σ_{lim} , і яка є шуканою. Виходить, для закінчення доказу нам треба показати існування такої функції.

Нехай $\langle A, \leq \rangle$ є спрямованою чум, і B є її підмножиною. Позначимо B° найменшу множину, що містить B і таку, що має місце рівність $\sigma_{lim}(B) = B$. Оскільки сімейство надмножин множини B , що задовольняють цій рівності, не є порожнім (воно містить A) і є замкнутим щодо довільних перетинань, множина B° існує для кожної підмножини B множини A . Ця властивість доводить існування функції σ_{lim}^* для кожного д-п-чум $\langle A, \leq \rangle$. Q.E.D.

Тепер ми можемо побудувати B° таким же самим чином, що використовується при покроковій побудові алгебраїчного замикання підмножини алгебр.

Розглянемо зростаючий ланцюг множини $B_0 \subseteq B_1 \subseteq \dots \subseteq B_\alpha \subseteq \dots$, індексовану ординалами і задану в такий спосіб:

$$B_0 = \sigma_{lim}(B);$$

$$B_\alpha = \sigma_{lim}(\cup \{B_\beta \mid \beta < \alpha\}).$$

Цей ланцюг стабілізується, починаючи з деякого ординала α_B (тобто α_B є найменший ординал, що задовольняє рівності $B_{\alpha_B} = B_{\alpha_B+1}$), і, отже, $B_{\alpha_B} = B$, має місце для кожного ординала γ , більшого, ніж α_B). Помітимо, що на відмінність від операторів алгебраїчного замикання, α_B не обов'язково менше чи дорівнює ω (першого нескінченного ординала).

Твердження 2. Для кожної підмножини B д-п-чум $\langle A, \leq \rangle$ має місце рівність $B^\circ = B_{\alpha_B}$.

Доведення. Очевидно. Q.E.D.

Крім множин виду B° , ми також будемо розглядати такі множини вигляду $B^>$ і $B^<$, що є, відповідно, найменше повними і ко-повними надмножинами множини B .

На підставі останнього твердження і способу побудови множини B° , ми можемо вказати аналогічний спосіб побудови множин $B^>$ і $B^<$.

Твердження 1 констатує, що при конструюванні "регулярного" розширення $B_{\alpha-1}$ з B_α ми може обмежитися тільки додаванням коренів FT -чум, що є підмножинами множини B_α (замість визначення $B_{\alpha-1}$ як $\sigma_{lim}(B_\alpha)$). Q.E.D.

Твердження 3. Для кожного позитивного ординала α множина B_α складається тільки з

елементів, що є коренями FT -чум, що входять у $\cup \{B_\beta \mid \beta < \alpha\}$.

Доведення (схема). Індукція по ординалам, у якій база індукції й індукційний перехід доводяться тим же самим чином, що й у випадку твердження 1. Q.E.D.

Наслідок. Для кожної підмножини B д-п-чум $\langle A, \leq \rangle$ множина B° є найменшою надмножиною множини B , що містить корені всіх його FT -підчум.

Легко бачити, що для кожної підмножини B д-п-чум $\langle A, \leq \rangle$ множина B° є σ_{lim} -замкненою (тобто B° є нерухомою точкою оператора σ_{lim}). Вже установлені властивості σ_{lim} імплікують, що сімейство $\{B^\circ \mid B \subseteq A\}$ задовольняє всім аксіомам замкнутих підмножин топологічного простору. В зв'язку з цим, відповідна топологія буде позначатися $T_\leq$.

Дослідимо структуру топології $T_\leq$.

Опукла підмножина C д-п-чум $\langle A, \leq \rangle$ називається досить опуклою тоді і тільки тоді, коли вона замкнена щодо операцій узяття супремума і інфімуму спрямованих і ко-спрямованих підмножин множини C , відповідно.

Твердження 4. Топологія $T_\leq$ є більш сильною ніж найменша топологія $S_\leq$ або $S_{\leq 1}$ (≤ 1 - це теж саме, що $\leq^1$), що включає топології Скотта $S_{\leq i}$ і $S_{\leq 1}$ для кожного д-п-чум $\langle A, \leq \rangle$. Довільна $T_\leq$ замкнена підмножина B множини $A \in (S_\leq \leq \text{або } S_{\leq 1})$ -замкненою тоді і тільки тоді, коли вона може бути представлена як скінченне об'єднання опуклих підмножин множини A .

Доведення. В силу принципу дуальності для доказу першої частини досить показати, що топологія Скотта $S_\leq$ є більш слабкою, ніж $T_\leq$.

Дійсно, довільна підмножина B з $A \in S_\leq$ замкненою тоді і тільки тоді, коли вона є спадковою і замкненою відносно супремумів спрямованих підмножин з B . Але спадковість, зокрема, імплікує, що B є замкненою відносно інфімумів довільних підмножин з A . Отже, $S_\leq$ або $S_{\leq 1}$ є більш слабкою топологією, чим $T_\leq$.

Доведемо другу частину. Для цього спочатку помітимо, що досить опуклі підмножини множини A утворюють замкнену базу топології $S_\leq$ або $S_{\leq 1}$. (Дійсно, замкнені підмножини кожної з топологій $S_\leq$ і $S_{\leq 1}$ задовольняють цим властивостям). Таким чином, кожна підмножина, що може бути представлена як скінченне об'єднання опуклих підмножин, що зберігають супремуми та інфімуми спрямованих і ко-спрямованих підмножин, є $S_\leq$ або $S_{\leq 1}$ -замкненою, і навпаки. Q.E.D.

Наслідок. Існують д-п-чум $\langle A, \leq \rangle$, такі, що топології $(S_\leq \text{ або } S_{\leq 1})$ і $T_\leq$ не співпадають.

Доведення. Потрібно розглянути довільну д-п-чум, що містить нескінченне диз'юнктивне сімейство F сильно опуклих підмножин, таких, що їх об'єднання $\cup F$ не дорівнює всьому д-п-чум, і об'єднання кожної пари різних множин не є сильно опуклим. Потім слід застосувати твердження 4. Q.E.D.

Нехай $S = \langle s: I \rightarrow A_1, D \rangle$ є послідовність, і $f: A_1 \rightarrow A_2$ є функція. Тоді $f(S)$ позначає послідовність $\langle s \circ f: I \rightarrow A_2, D \rangle$ (вираження $s \circ f$ позначає композицію функцій s і f), що називається образом S відносно f , а S - прообразом $f(S)$. Очевидно, що образ і прообраз зберігають властивість "бути конфінальною підпослідовністю конфінальної послідовності".

Резюме

В роботі було проведено вивчення трьох понять типів збіжності послідовностей та їх взаємодії з топологічними властивостями частково упорядкованих множин. Це дає можливість ввести поняття неперервності функції засобами, відмінними від прийнятих в поточний час, що ведуть до побудови нових топологічних моделей. Автор планує опублікувати продовження досліджень в цьому напрямку окремою статтею.

Список використаних джерел

1. Scott D. S. Data types as lattices // SIAM J. of Comp. – 1976. – 5(3). – P.522-587.
2. Александров П.С. Введение в теорию множеств и общую топологию. – 2-е изд. стер. – М.: УРСС, 2004. – 367 с.
3. Гретцер Г. Общая теория решеток. – М.: Мир, 1982 – 454 с.
4. Барендрегт Х. Лямбда-исчисление. Его синтаксис и семантика. – М.: Мир, 1985. – 606 с.
5. Энгелькинг Р. Общая топология. – М.: Мир, 1986. – 752 с.
6. Lyaletsky A. On Types of Continuity Defined by Partially Ordered Sets with Usual Topological Continuity // BSL. – 2005. – 11 (2). P.284-285.

Надійшла до редколегії
11.01.2008

УДК 519.21

А.А. Муzychuk

Оцінка вартості та моменту впровадження інвестиційних проектів в умовах ризику

У статті розглянуто задачу оптимального вибору інвестиційного проекту з множини допустимих в умовах ризику. За припущенням, виконання проекту можна відкласти до деякого моменту в майбутньому. Запропоновано підхід до обчислення прибутку за умови затримання інвестиції та його складових. Досліджено поведінку величини прибутку інвестиційних проектів з однаковою вартістю, але різною швидкістю повернення грошових потоків, в залежності від стану зовнішнього середовища.

Ключові слова: ризик, грошові потоки

E-mail: andriy@pancha.lviv.ua

1. Вступ. Постановка задачі

Проблема оптимального вибору інвестиційних проектів в умовах невизначеності є предметом дослідження багатьох авторів, зокрема в працях [1-4]. Джерелом ризику для особи, що приймає рішення (ОПР), є нестабільне зовнішнє середовище. Його економічні, політичні, демографічні та інші фактори впливають на результат інвестування – прибуток інвестора. Щодо врахування наявної інформації про поведінку зовнішнього середовища в майбутньому можна виділити три типи задач прийняття рішень: в умовах ігнорування невизначеності, в умовах неповної інформації та в умовах ризику. У задачі першого типу розподіл ймовірностей станів зовнішнього середовища невідомий. В такому випадку, використовуючи коефіцієнт відношення ОПР до ризику, можна побудувати суб'єктивний ймовірнісний розподіл, який є ваговим вектором оператора впорядкованого зваженого середнього [2].

У цій статті розглянемо задачу оптимального вибору в умовах ризику інвестиційного проекту з множини допустимих. Зовнішні фактори, що впливають на прибуток інвестора, будемо називати станами зовнішнього середовища, множину яких позначимо як J . Через j_t позначимо стан, в якому перебуває зовнішнє середовище на момент часу t . Вважаємо, що інвестиційний проект має такі характеристики: момент впровадження τ , вартість I , тривалість

A.A. Muzychuk

Investment value and launching time evaluation under risk

In the paper an optimal selection of investment project from available set under risk is discussed. It is assumed that launching of the project possible to postpone to future moment of time. An approach to calculating profit from the investment delay and its parts is proposed. The behavior of profit value of the projects with equal price but different speed of cash flow return depended on the environment state is explored.

Key Words: risk, cash flow

T та грошові потоки $(f_{\tau+1}, \dots, f_{\tau+T})$ утворені в дискретні моменти часу $\{\tau+1, \dots, \tau+T\}$.

Щодо даних задачі зробимо такі припущення:

- 1) множина J є скінченною;
- 2) визначена функція $r: J \rightarrow [0, 1]$;
- 3) $\forall t: 0 \leq t \leq \tau$ відомі ймовірності $P(j_t \in J)$;
- 4) $\tau \geq 0$.

Перше припущення дає змогу врахувати найважливіші фактори, які впливають на процес прийняття рішень. Економічний зміст функції $r(\cdot)$ полягає в тому, що кожному стану зовнішнього середовища $j \in J$ ставиться у відповідність значення $r(j)$, яке є відсотковою ставкою дисконтування. Будемо вважати, що стан зовнішнього середовища $i \in J$ є кращий за стан $j \in J$, якщо $r(i) < r(j)$.

Припущення 3) є важливим, оскільки інвестиційне рішення приймають в умовах ризику, то є відомими ймовірності $P(j_t \in J)$ перебування зовнішнього середовища на момент часу t в стані $j_t \in J$. Позначимо $P(j_t = i \in J) = p_i(t)$.

В припущенні 4) стверджується, що ОПР може відкласти впровадження проекту до деякого моменту τ в майбутньому. Доцільність введення такого припущення розглянемо на наступному прикладі.

Приклад 1. Розглянемо проект А вартістю $I = 100$ грн., який утворює грошові потоки величиною 39.096 грн. на рік протягом трьох років. Припустимо, що ОПР може впроваджувати ПІ в моменти $\tau = 0$ або $\tau = 1$. Тобто інвестор має можливість відтермінувати проект на 1 рік.

Припустимо, що на момент $t = 0$ зовнішнє середовище перебуває в стані j_0 і $r(j_0) = 8\%$. Нехай наступного року очікується збільшення або зменшення значення відсоткової ставки дисконтування на 2% з ймовірністю 0.5. Тобто на момент часу $t = 1$ зовнішнє середовище буде перебувати в стані $j_1 = k$ з відсотковою ставкою $r(k) = 10\%$ і з ймовірністю $p_k(t) = 0.5$, або в стані $j_1 = l$ з відсотковою ставкою $r(l) = 6\%$ і з ймовірністю $p_l(t) = 0.5$. Вважаємо, що інвестор є нейтральний до ризику.

Знайдемо чистий дисконтований прибуток інвестора утворений проектом А при умові негайного інвестування $\tau = 0$

$$\frac{39.096}{(1,08)} + \frac{39.096}{(1,08)^2} + \frac{39.096}{(1,08)^3} - 100 = 0,754.$$

У випадку відтермінування вкладення коштів в проект до моменту $\tau = 1$ прибуток ОПР буде визначатись як

$$\frac{1}{1,08} \left[\frac{1}{2} \left(\frac{39.096}{(1,06)} + \frac{39.096}{(1,06)^2} + \frac{39.096}{(1,06)^3} - 100 \right) + \frac{1}{2} \left(\frac{39.096}{(1,1)} + \frac{39.096}{(1,1)^2} + \frac{39.096}{(1,1)^3} - 100 \right) \right] = 0.801.$$

Чистий приведений прибуток 0.801 проекту А, впровадження якого затримано до моменту $\tau = 1$, переважає прибуток 0.754 утворений внаслідок негайного інвестування на момент $\tau = 0$. Таким чином, можна зробити висновок, що інвестування в проект А краще відкласти до моменту $\tau = 1$ в майбутньому.

Отже, в роботі будемо розглядати задачу, коли момент впровадження τ інвестиційного проекту може не співпадати з моментом прийняття рішень $t = 0$.

2. Визначення вартості проекту з можливістю його відтермінування

Чистий прибуток ОПР, утворений проектом відтермінованим до моменту τ в майбутньому позначимо як $V(j_\tau)$. Тоді

$$V(j_\tau) = \sum_{t=\tau}^T B(j_\tau, t) \cdot f_t - I, \quad (1)$$

де j_τ - стан зовнішнього середовища на момент часу $\tau \geq 0$, $j_\tau \in J$, а $B(j_\tau, t) = (1 + r(j_\tau))^{-t}$ - дисконтуючий множник.

Нехай інвестування в проект відкладають до деякого моменту в майбутньому, причому проект впроваджують незалежно від стану зовнішнього середовища. Для оцінки такої дії будемо враховувати прибуток від затримання фінансового потоку, утвореного інвестицією. Оскільки можливий такий випадок, що стан економіки буде несприятливий для проекту, то для уникнення збитків, будемо допускати зміну інвестиційного рішення. При такому підході прибуток $\delta(j_0)$ при затриманні інвестування, за умови перебування зовнішнього середовища на момент $t = 0$ у стані $j_0 \in J$, буде складатися з прибутку, утвореного затриманими грошовими потоками $\delta^{CFD}(j_0)$, і прибутку від зміни рішення $\delta^{DM}(j_0)$:

$$\delta(j_0) = \delta^{CFD}(j_0) + \delta^{DM}(j_0). \quad (2)$$

Якщо стан зовнішнього середовища $j_\tau \in J$ є таким, що інвестуючи ОПР несе збитки, то рішення про впровадження такого проекту в момент часу τ відхиляють, вважаючи $V(j_\tau) = 0$. Для решти станів j_τ буде виконуватись $V(j_\tau) > 0$. Тоді за умови перебування зовнішнього середовища в момент часу $t = 0$ у стані $j_0 \in J$ майбутні виплати від затримання інвестування до моменту часу τ можна подати у вигляді

$$\begin{aligned} \delta(j_0) &= B(j_0, \tau) \cdot E(\max\{0, V(j_\tau)\}) = \\ &= B(j_0, \tau) \cdot \sum_{i \in J} \max\{0, V(i)\} p_i(\tau), \end{aligned} \quad (3)$$

де $V(i)$ обчислюють за формулою (1).

Якщо рішення про інвестування в проект у деякий момент $\tau > 0$ в майбутньому приймають не зважаючи на стан економіки, прибуток від затримання грошових потоків інвестиції можна визначити як

$$\begin{aligned} \delta^{CFD}(j_0) &= B(j_0, \tau) \cdot E(V(j_\tau)) = \\ &= B(j_0, \tau) \cdot \sum_{i \in J} V(i) \cdot p_i(\tau). \end{aligned} \quad (4)$$

Тоді на основі співвідношень (2), (3) і (4) можна отримати прибуток внаслідок зміни інвестиційного рішення

$$\begin{aligned}\delta^{DM}(j_0) &= \delta(j_0) - \delta^{CFD}(j_0) = \\ &= B(j_0, \tau) \cdot \sum_{i \in J} \max\{-V(i), 0\} p_i(\tau). \quad (5)\end{aligned}$$

Отже, прибуток ОПР залежить від вибору моменту впровадження проекту. Тоді, при умові негайного інвестування ($\tau = 0$), величина прибутку становить $V(j_0)$, а у випадку $\tau > 0$ - $\delta(j_0)$. Оскільки ОПР зацікавлена в отриманні максимального прибутку, то за умов виконання припущення 4) прибуток інвестора $F(j_0)$ можна визначити як

$$F(j_0) = \max\{\delta(j_0), V(j_0)\}. \quad (6)$$

3. Порівняння двох проектів з однаковою вартістю і різною швидкістю повернення грошових потоків

В умовах прикладу 1 проаналізуємо проект В, який теж має вартість 100 грн., але утворює грошовий потік величиною 56.5 грн. на рік протягом двох років. Знайдемо чистий дисконтований прибуток ОПР утворений проектом В при умові негайного інвестування $\tau = 0$

$$\frac{56.5}{(1.08)} + \frac{56.5}{(1.08)^2} - 100 = 0.754.$$

У випадку відтермінування вкладення коштів в проект до моменту $\tau = 1$ прибуток ОПР буде визначатись як

$$\begin{aligned}&\frac{1}{1.08} \left[\frac{1}{2} \left(\frac{56.5}{(1.06)} + \frac{56.5}{(1.06)^2} - 100 \right) + \right. \\ &\left. + \frac{1}{2} \left(\frac{56.5}{(1.1)} + \frac{56.5}{(1.1)^2} - 100 \right) \right] = 0.761.\end{aligned}$$

Оскільки виконується нерівність $0.761 > 0.754$, то можна зробити висновок, що впровадження В краще відкласти на рік. Зазначимо, що проект А є привабливішим для інвестування ніж В, оскільки $0.801 > 0.761$.

Обидва проекти А і В мають однакову прибутковість за умов негайного інвестування і різну при умові відтермінування впровадження. Така різниця утворилась внаслідок різної швидкості повернення грошових потоків, які в свою чергу по різному реагують на можливі зміни стану зовнішнього середовища в майбутні моменти часу.

Для дослідження більш загального випадку розглянемо два проекти П1 і П2, які мають

однакову вартість, моменти впровадження, тривалість і чистий дисконтований прибуток за умови негайного інвестування:

$$I_1 = I_2 = I, \quad V_1(j_0) = V_2(j_0).$$

Нехай проекти в дискретні моменти часу $\{\tau + 1, \dots, \tau + T\}$ утворюють грошові потоки $(f_{1\tau+1}, \dots, f_{1\tau+T})$ та $(f_{2\tau+1}, \dots, f_{2\tau+T})$, причому другий проект утворює більш ранні грошові потоки, ніж перший.

Через $X_1(t)$ та $X_2(t)$ позначимо сумарний грошовий потік на момент часу t утворений проектами П1 та П2 відповідно. Тоді

$$X_1(t) = \sum_{k=1}^t f_{1k}, \quad X_2(t) = \sum_{k=1}^t f_{2k}. \quad (7)$$

Отже,

$$\begin{aligned}V_1(j_0) &= \sum_{t=1}^T B(j_0, t) \cdot \Delta X_1(t) - I = \\ &= \sum_{t=1}^T B(j_0, t) \cdot \Delta X_2(t) - I = V_2(j_0), \quad (8)\end{aligned}$$

де $\Delta X_i(t) = X_i(t) - X_i(t-1) = f_{it}$, $i = 1, 2$, $t > 0$.

Дослідимо, який із проектів має більшу вартість з врахуванням можливості відкласти інвестування до деякого моменту $\tau > 0$ у майбутньому. Розглянемо різницю потоків $g(t) = X_2(t) - X_1(t)$. Оскільки проект П2 утворює більш ранні грошові потоки, ніж проект П1, то існує таке t^* ($0 < t^* < T$), що $g(t)$ є зростаючою на проміжку $[0, t^*]$ і спадною на $[t^*, T]$. Будемо вважати, що проект П2 має більшу негайність або більшу швидкість повернення, якщо він утворює більш ранні грошові потоки.

Якщо обидва проекти відтерміновані до деякого моменту $\tau > 0$ в майбутньому, то через $\delta_1(j_0)$ та $\delta_2(j_0)$ позначимо прибуток від затримання 1-ї та 2-ї інвестиції відповідно, а через $F_1(j_0)$ та $F_2(j_0)$ - прибутковості проектів, які визначені за допомогою (6).

Для подальшого дослідження важливим є припущення про рух відсоткових ставок в майбутні моменти часу. Нехай розглядається відтермінування впровадження проекту до деякого моменту $\tau > 0$ в майбутньому. Якщо середня ставка дисконтування $E(r(j_\tau))$ на цей момент є більшою ніж відсоткова ставка дисконтування $r(j_0)$ на момент прийняття рішень $\tau = 0$, то оптимальним є негайне інвестування, і розглядати виплати від затримання

інвестування непотрібно. Отже, необхідною умовою для дослідження відкладення впровадження інвестиційного проекту до деякого моменту в майбутньому є виконання умови

$$E(r(j_\tau)) \leq r(j_0), \quad (9)$$

що еквівалентно

$$\sum_{i \in J} r(i) p_i(\tau) \leq r(j_0).$$

Лема 1. Нехай $V_1(j_0) = V_2(j_0)$ і функція $g(t) = X_2(t) - X_1(t)$ є зростаюча по t на проміжку $[0, t^*]$ і спадна на проміжку $[t^*, T]$ для деякого t^* ($0 < t^* < T$).

Тоді справедливі наступні твердження:

1. якщо $r(j_\tau) \leq r_0$, то

$$V_2(j_\tau) \geq V_2(j_0) \text{ і } V_1(j_\tau) \geq V_2(j_\tau),$$

2. якщо $r(j_\tau) \geq r_0$, то

$$V_2(j_\tau) \leq V_2(j_0) \text{ і } V_1(j_\tau) \leq V_2(j_\tau).$$

Доведення леми 1 є громіздким, тому в статті його не подаємо.

Лему 1 можна інтерпретувати з допомогою графіка залежності вартості відтермінованого проекту $V(j_\tau)$ від стану зовнішнього середовища $j_\tau \in J$ або відповідної ставки дисконтування $r(j_\tau)$.

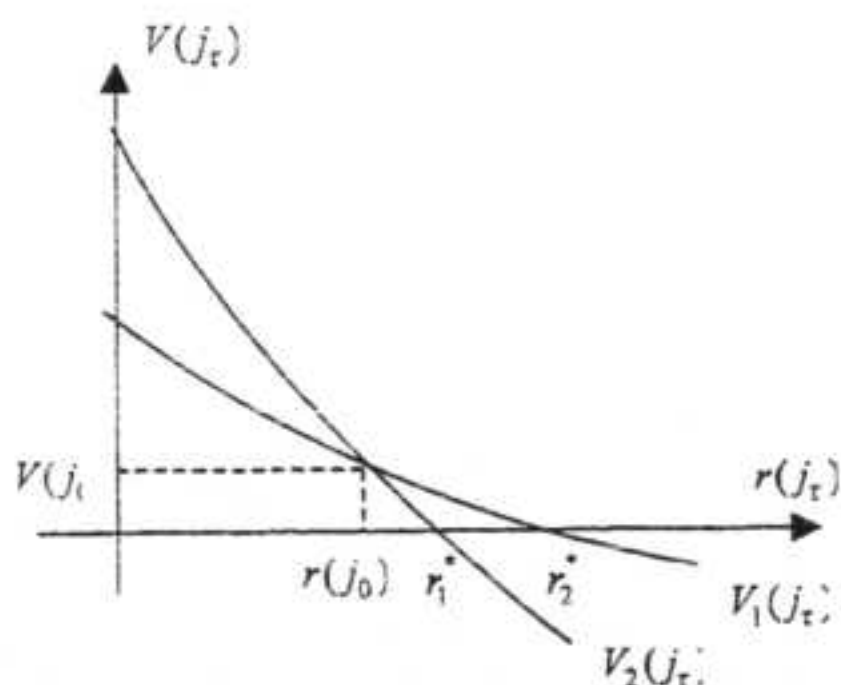


Рис.1. Залежність прибутковості відтермінованих проектів і станів зовнішнього середовища

За умовою леми $V_1(j_0) = V_2(j_0)$. Якщо виконання проекту відбувається в момент часу $\tau > 0$, то обидва проекти будуть мати однакові виплати у випадку $r(j_\tau) = r(j_0)$. Якщо зовнішні чиники в майбутньому покращаться, то відсоткова ставка дисконтування зменшиться. Тоді проект з більшою негайністю фінансового потоку буде мати більшу величину виплат, оскільки він є більш чутливим до змін середовища (пункт 2 леми 1). Якщо відсоткова ставка зростає до деякого значення r_1^* , то обидва проекти будуть мати невід'ємні виплати, причому більші виплати має проект з меншою швидкістю повернення, оскільки на нього менше впливають відсоткові ставки. У випадку, якщо $r \in [r_1^*, r_2^*]$, то додатні виплати має тільки проект з меншою швидкістю повернення. Подальші зростання відсоткової ставки призведуть до отримання від'ємних виплат в обох проектах, але в такому випадку рішення про інвестування відхиляється і виплати рівні 0. Оскільки будь-який з цих випадків є можливий, то не відразу очевидно, який з проектів має більшу вартість при затриманні впровадження інвестиційного проекту.

Лема 2. Припустимо, що виконуються умови леми 1 і для $\tau > 0$ справедлива нерівність $E(r(j_\tau)) \leq r(j_0)$. Тоді

$$\delta_1(j_0) \geq \delta_2(j_0).$$

Доведення леми 2 є громіздким, тому в роботі не подається (для доведення потрібно врахувати (2)).

Наслідок 1. Якщо в умові леми 2 нерівність $E(r(j_\tau)) \leq r(j_0)$ не виконується, тоді

$$\delta_1^{DM}(j_0) \geq \delta_2^{DM}(j_0).$$

З наслідку 1 випливає, що вартість зміни рішення є спадною функцією від швидкості повернення грошових потоків. Тобто ймовірність відхилення інвестиції з більшою швидкістю повернення грошових потоків є більшою, ніж ймовірність відхилення інвестиції з меншою негайністю, оскільки, як видно з рис.1, рішення про відхилення проекту з більшою негайністю приймається у випадку зростання відсоткової ставки до величини r_1^* , а для проекту з меншою негайністю таке рішення приймається при зростанні ставки до значення $r_2^* > r_1^*$. Більше того, коли приймається рішення про відхилення проекту, то величина збереженого капіталу

зменшується із збільшенням негайності грошових потоків.

Якщо величина грошових потоків фіксована, то очікуваний прибуток на момент $t=0$ від інвестування в момент τ залежить тільки від ймовірнісного розподілу j_τ .

Наслідок 2. Нехай виконується нерівність $E(r(j_\tau)) \leq r(j_0)$. Тоді для фіксованого грошового потоку $(f_{\tau+1}, \dots, f_{\tau+T})$ утвореного проектом вартістю I виконується нерівність

$$V(j_\tau) > V(j_0).$$

Доведення. Оскільки $B(j, t)$ є спадною по t , то з нерівності Єнсена випливає $E(B(j_\tau, t)) > B(j_0, t)$.

З наслідку 2 випливає, що невизначеність відсоткової ставки збільшує очікуване теперішнє значення фінансового потоку утвореного інвестиційним проектом в момент часу $\tau > 0$. Більше того, така різниця є більшою для проектів з меншою швидкістю повернення, оскільки більша частина фінансового потоку сконцентрована у майбутні моменти часу, тобто є більш чутливою до змін станів зовнішнього середовища.

Теорема Нехай виконуються умови леми 2. Тоді справедлива нерівність

$$F_1(j_0) \geq F_2(j_0).$$

Доведення випливає з леми 2 і з формули (6).

Таким чином, прибутковість проекту є спадною функцією від швидкості повернення грошових потоків. Зауважимо, що не завжди просто оцінити, який із проектів має більшу швидкість повернення грошових потоків. Особливо це стосується грошових потоків однакової довжини. Якщо виконуються припущення 3), то можна використати коефіцієнт швидкості повернення $\nu(\cdot)$ грошового потоку $(f_{\tau+1}, \dots, f_{\tau+T})$ при умові, що зовнішнє середовище на момент часу τ перебуває в стані j_τ , запропонованого в [5]

$$\nu(f_{\tau+1}, \dots, f_{\tau+T} | j_\tau) = \sum_{t=1}^T f_{\tau+t} \cdot \nu_t,$$

де

$$\nu_t = 1 - \left(\sum_{k=1}^t B(j_\tau, k) \right) \cdot \left(\sum_{s=1}^T (T-s+1) \cdot B(j_\tau, s) \right)^{-1},$$

$$B(j_\tau, t) = (1 + r(j_\tau))^{-t}.$$

4. Вибір оптимального інвестиційного проекту та моменту його впровадження

Нехай для ОПР на момент часу $t=0$ є досяжними m інвестиційних проектів вартістю I_i , які утворюють відповідні грошові потоки $(f_{i,\tau+1}, \dots, f_{i,\tau+T})$ ($i = \overline{1, m}$) в дискретні моменти часу $\{\tau+1, \dots, \tau+T\}$. Припустимо, що виконується припущення 4), тобто інвестор має можливість відкласти впровадження кожного з проектів до деякого моменту τ в майбутньому. Прибутковість проекту $i = \overline{1, m}$ за умови негайного інвестування позначимо як $V_i(j_0)$, а за умови затримання впровадження проекту - $\delta_i(j_0)$.

Тоді оптимальний проект i^* і момент впровадження τ^* визначаються наступним чином

$$(i^*, \tau^*) = \arg \max_i \max_{\tau} \{V_i(j_0), \delta_i(j_0)\},$$

де $k = \overline{1, m}$, $\tau > 0$, $V_i(j_0)$ і $\delta_i(j_0)$ обчислюються за формулами (1) і (3) відповідно.

Отже, розглянуто задачу вибору інвестиційного проекту з можливістю його відтермінування в умовах ризику. Досліджено як змінюється прибутковість проектів в залежності від швидкості повернення грошових потоків і від впливу зовнішнього середовища. Знайдено умови, при яких можна розглядати можливість відтермінування. Показано, що між прибутковістю проекту і швидкістю повернення грошових потоків в умовах ризику є обернено пропорційна залежність.

Список використаних джерел

1. *Ingersoll, J.E. and S.A.Ross* Waiting to invest: Investment and uncertainty // *Journal of Business*. – 1992. – № 65. – P.1-29.
2. *Ronald R.Yager* Including decision attitude in probabilistic decision making // *International Journal of Approximate Reasoning*. – 1999. – №21. – P.1-21.
3. *Кігель В.Р.* Методи і моделі прийняття рішень у ринковій економіці: Монографія. – К.: Цул, 2003. – 202с.
4. *Матвійчук А.В.* Аналіз і управління економічним ризиком. – К.: Цул, 2005. – 224 с.
5. *Музичук А., Масна О.* Критерій порівняння банком інвестиційних проектів з однаковим NPV // *Банківська система України в умовах глобалізації фінансових ринків: міжнар. наук. конф. Черкаси, 18-20 жовт. 2006 р.* – Черкаси, 2006. – С. 97-100.

Надійшла до редколегії 29.01.08

УДК 681.3.062

Поляков С.А.*

Огляд стандартів мови SQL

Описана історія створення мови SQL, розглянуті існуючі стандарти мови в тому числі останній SQL 2003. Наведена класифікація операторів SQL згідно з останнім стандартом. Розглянуті перспективні напрямки подальшого розвитку мови.

Ключові слова: SQL, реляційні бази даних, реляційна алгебра, СУБД, РСУБД, SEQUEL

*E-mail: polyakov@unicyb.kiev.ua

Статтю представив: д.ф.м.н., с.н.с. Буй Д.Б.

Мова SQL базується на реляційній моделі даних, розробленій Коддом (1923-2003 рр.) у 1969 р. У той час Кодд був науковим співробітником в компанії IBM. Він шукав нові шляхи для обробки баз даних, які містять велику кількість даних. Як математик за фахом він використовував у своїй роботі результати з теорії множин та математичної логіки першого порядку. Розроблена їм нова модель даних була оприлюднена у червні 1970 у статті "A Relational Model of Data for Large Shared Databases" [1]. Назва моделі даних базується на терміні "реляція", який використовується у теорії множин. Слід сказати, що сьогодні поширена ще одна версія походження назви реляційної моделі даних. А саме: назва моделі походить від того факту, що таблиці в реляційній базі даних можуть мати зв'язок (relationship) між собою. Але ця версія є помилковою. Дійсно, виразна сила та гнучкість моделі пояснюються, значною мірою, наявністю зв'язків між таблицями та можливістю моделювання цих зв'язків операціями реляційної алгебри.

Незабаром після оприлюднення роботи Кодда університети та науково-дослідні лабораторії почали працювати над розробкою мови, яка могла би використовуватись для роботи з базами даних реляційного типу. Ці роботи привели до появи декількох мов запитів до реляційних баз даних на початку та всередині 1970-х рр., а пізніше і до появи мови SQL, розробленої в лабораторії IBM "Santa Teresa Research Laboratory" в San Jose, California.

Polyakov S.A.*

The survey of the SQL language standards.

The paper introduces to the history of the SQL language and describes its standards including SQL 2003. It represents classification of the SQL clauses according to the last standard – SQL 2003. The paper considers the perspective ways of the SQL growth as well.

Keywords: SQL, relational data base, relational algebra, DBMS, RDBMS, SEQUEL

В цій лабораторії на початку 1970-х рр. почав виконуватись дуже важливий для подальшого розвитку реляційних баз даних проект під назвою "Syster/R". Його метою було показати життєздатність концепції реляційної моделі даних та напрацювати досвід в проектуванні та реалізації систем управління реляційними базами даних (СУБД). В результаті в 1974-1975 рр. був розроблений успішний мінімальний прототип СУБД. Разом з розробкою прототипу науковці працювали над мовою запитів для реляційних баз даних. Результати роботи, яка виконувалась в рамках цього проекту, були важливими для подальшого комерційного розвитку реляційних баз даних. В 1974 р. Дональд Чамберлін (Dr. Donald Chamberlin) та його колеги розробили мову Structured English Query Language (SEQUEL) [2]. Ця мова дозволяла писати запити, в основі яких є реляційна алгебра, до бази даних у вигляді, дуже близькому до природних англійських речень. Чамберлін та його колеги розробили першу реалізацію цієї мови під назвою SEQUEL-XRM.

Успіх цієї мови привів до продовження досліджень та подальшого розвитку мови. Науковці провели повну ревізію мови та в 1976-1977 рр. розробили нову версію, яку назвали SEQUEL/2. Пізніше вони вимушені були перейменувати мову в SQL (Structured Query Language) або SQL Query Language внаслідок юридичних причин. Як з'ясувалось, хтось інший вже використовував акронім SEQUEL. До сьогодні багато людей все ще вимовляють SQL як *sequel*, хоча офіційно назва вимовляється як

es-cue-el. В SQL було впроваджено декілька нових засобів, таких як багатотабличні запити та можливість одночасного доступу до бази даних багатьох користувачів.

Незабаром після появи SQL IBM почала новий, більш амбіційний проект, метою якого було розробити прототип СУРБД, який був би спроможний показати можливість використання реляційних систем на практиці. Новий прототип назвали System R. В його основу була покладена велика підмножина мови SQL. Потім, коли розробка першої черги була завершена, діючий прототип інсталивали на комп'ютерах компанії та на комп'ютерах деяких клієнтів для проведення тестування та оцінки придатності системи для використання в реальному житті. Протягом цього часу система та мова SQL набули чисельних модифікацій, базуючись на набутому досвіді експлуатації та відгуках користувачів. IBM закрила проект в 1979 р. Був зроблений висновок про життєздатність реляційної моделі та наявності у неї комерційного потенціалу. Але на той час техніка не дозволяла розробити систему, придатну для створення реальних баз даних. Як згадує Джон Віскас (John Viescas), який був свідком демонстрації системи в кінці 1970-х рр.: "Вона мала багато разючих факторів, але на комп'ютерах, наявних на той час, навіть простий запит виконувався протягом декількох хвилин. Вона очевидно мала гарний потенціал, але безумовно вимагала кращого обладнання та програмного забезпечення для використання цього продукту в комерційній сфері" [3, с. 55]. Одним з найбільш важливих наслідків цього проекту була поява мови SQL.

Робота, яка проводилась в лабораторії IBM, викликала великий інтерес в світі. Позитивні якості нової реляційної моделі даних широко обговорювались на семінарах та в наукових журналах. Наприкінці 70-х рр. стало відомо, що IBM зацікавлена реляційною технологією і розробляє систему, яка базується на реляційній моделі даних і мові SQL. Це породило багато розмов, коли IBM закінчить розробку і продемонструє готовий продукт. Деякі компанії почали розробляти власні системи, намагаючись першими вийти на новий ринок. В 1977 р. групою інженерів було засновано компанію Relational Software, Inc. в Menlo Park, California для створення нової системи на базі SQL. Вони розробили її в 1979 р. і назвали Oracle. Це була перша комерційна реляційна СУБД. Вона випередила продукт IBM на два роки. До того ж

Oracle працював на мінікомп'ютерах VAX компанії Digital, які були дешевше за меінфрейми компанії IBM. Relation Software була перейменована в Oracle Corporation. На сьогодні вона є одним з провідних постачальників реляційних систем.

У той же самий час Майкл Стонебракер (Michael Stonebraker), Євгеній Вонг (Eugene Wong) та інші співробітники лабораторії Каліфорнійського університету в Берклі також проводили дослідження реляційної моделі даних. Вони розробили прототип реляційної СУБД, який назвали INGRES. Система використовувала свій язык запитів до бази даних, який називався Query Language (QUEL). В порівнянні з SQL він був більш структурованим, але менш подібним до англійських речень. Пізніше INGRES був переведений на використання SQL, коли стало зрозуміло, що SQL становиться фактичним стандартом для мов програмування баз даних. Науковці залишили Берклі в 1980 р. та заснували компанію під назвою Relational Technology, Inc. В 1981 р. вони анонсували комерційну версію INGRES. Ця компанія зараз є частиною Computer Associates, Inc. В 2004 р. СУБД стала вільно розповсюджуватись на умовах open-source ліцензії.

IBM анонсувала свій продукт тільки в 1981 р. та почала його розповсюджувати в 1982 р. під назвою SQL/Data System (SQL/DS). В 1983 р. компанія анонсувала новий продукт під назвою Database 2 (DB2) [4]. Він почав розповсюджуватись в 1985 р. і на сьогодні входить в трійку ведучих комерційних СУБД.

Така велика зацікавленість мовою SQL та майже одночасна поява кількох діалектів мови в продуктах різних компаній неминуче ставили питання про розробку стандарту. Але не було консенсусу стосовно того, хто буде розробляти стандарт та який діалект SQL буде взято за його основу. Тому кожний розробник продовжував розвивати свій діалект мови. Така ситуація приводила до несумісності продуктів і була дуже незручною для користувачів, які вимагали підвищити сумісність різних діалектів мови. Незабаром з'явився неофіційний стандарт мови. Це була маленька підмножина сьогоденного стандарту, яка включала тільки ті елементи, які були подібні у тодішніх діалектах мови. Але він надавав базові критерії, які дозволяли оцінювати чисельні реляційні системи, що з'являлись на ринку.

В 1982 р. American National Standards Institute (ANSI) доручив технічному комітету по базам

даних X3H2, який входить в состав X3, розробити пропозиції стосовно офіційного стандарту мови програмування реляційних баз даних. X3 це уповноважений ANSI форум (ANSI-accredited forum) розробників у сфері інформаційних технологій, зараз перейменованій в International Committee for Information Technology Standards, або INCITS (промовляється як "insights"). До складу X3H2 входять ведучі експерти у сфері баз даних та представники головних компаній-розробників SQL-продуктів. Спочатку комітет почав розробляти стандарт, базуючись на QUEL, мові програмування СУБД INGRES. Але зростаючий вплив SQL та зусилля IBM по розвитку саме цієї мови привели до того, що в основу стандарту була покладена мова SQL.

Запропонований комітетом стандарт значною мірою був оснований на діалекті мови в СУБД DB2 компанії IBM. Комітет працював над декількома версіями стандарту протягом наступних двох років і навіть в деякій мірі покращив SQL. Але після цього запропонований стандарт став несумісним з існуючими діалектами мови. Тому комітет вирішив, що запропоновані покращення не переважають втрату сумісництва, і повернувся до початкового варіанту стандарту. ANSI ратифікував стандарт в 1986 р. під назвою "ANSI X3.135-1986 Database Language SQL"; цей стандарт більш відомий як SQL/86.

Стандарт SQL/86 визначав тільки найменший загальний знаменник, якому повинна задовольняти реалізація мови для того, щоб претендувати на сумісність зі стандартом. По суті він надав офіційний статус елементам мови, які були подібні в різних діалектах SQL та в різних реалізаціях реляційних баз даних. Стандарт заложив основи, спираючись на які був можливий подальший розвиток мови.

International Organization for Standardization (ISO) в свою чергу прийняв свій стандарт 1987 р., який повністю співпадав зі стандартом SQL/86. Він називався "ISO 9075-1987 Database Language SQL."

Але, незважаючи на появу стандарту, сама мова була ще далека від завершення. Невдовзі SQL/86 підпав під широку критику з боку фахівців. Серед недоліків стандарту називалися надмірність мовних засобів, коли один і той же запит можна було записати декількома способами, недостатня підтримка деяких реляційних операцій, відсутність засобів визначення цілісності посилань на дані (referential integrity). Хоча комітет X3H2 знав про ці недоліки, але було прийняте рішення все ж таки оприлюднити стандарт. Комітет виходив з того, що

недосконалий стандарт в усякому разі краще чим відсутність будь-якого стандарту. У відповідь на критику стосовно відсутності засобів визначення цілісності посилань ISO випустив вдосконалену версію стандарту в середині 1989 р. під назвою "ISO 9075:1989 Database Language SQL with Integrity Enhancements". Трохи пізніше комітет ANSI випустив свою версію під назвою "X3.135-1989 Database Language SQL with Integrity Enhancements". Ця версія стандарту відома під назвою SQL/89. В опублікованому стандарті була відсутня частина, яка б специфікувала вбудований SQL. Тому ANSI розробила наступний стандарт, який називався "ANSI X3.168-1989 Database Language Embedded SQL" і був опублікований в тому ж 1989 р. Цікаво зазначити, що ISO вирішила не публікувати відповідний стандарт через відсутність достатньої зацікавленості з боку міжнародної спільноти. Ситуація змінилась тільки в 1992 р., коли була опублікована наступна версія стандарту.

SQL/86 та SQL/89 не були завершеною специфікацією мови. Їм бракувало деяких важливих частин, без яких використання SQL в комерційних системах було неможливе. Наприклад, була відсутня специфікація операторів модифікації схем таблиць та вилучення таблиць. Аналогічна ситуація була з операторами надання повноважень на допустимі операції з таблицями. Було специфіковано оператор надання повноважень, але були відсутні оператори зміни повноважень та їх вилучення. Але без таких операторів не могла існувати жодна комерційна СУБД. Все це вимушувало розробників баз даних впроваджувати свої версії відсутніх в стандарті операторів, які відрізнялись один від одного.

На той час коли SQL/89 було опубліковано, в комітетах велась робота по підготовці наступного релізу мови. Планувалося, що в наступному релізі буде завершено специфікацію SQL та будуть додані нові мовні конструкції, які вже використовувалися багатьма провідними розробниками реляційних баз даних. Головною метою нової роботи була специфікація "повноцінної" мови, а не найменшого загального знаменника, як це було раніше. Тому до стандарту були включені конструкції, які ще не набули широкого поширення, а деякі навіть не були реалізовані. Комітети ANSI та ISO опублікували свої стандарти в 1992 р. під назвами "X3.135-1992 Database Language SQL" та "ISO/IEC 9075:1992 Database Language SQL" відповідно. Цей стандарт відомий як SQL/92. Він

був не тільки значно більший ніж SQL/89, але і значно ширший за змістом. Він включав операції модифікації структури бази даних, функції для маніпулювання рядками символів, функції маніпулювання датами та часом, також були розширені засоби надання повноважень користувачам. Це був значний крок уперед порівняно з попередніми стандартами.

Для того щоб полегшити підтримку нового стандарту розробниками баз даних, комітет визначив три рівня стандарту – початковий рівень (Entry SQL), проміжний рівень (Intermediate SQL) та повний рівень (Full SQL).

Специфікація початкового рівня дуже близька до SQL/89, в ній були виправлені деякі помилки попереднього стандарту. Крім того вона включає засоби, які полегшують перехід на новий стандарт з SQL/89. Загалом задовольнити вимогам першого рівня було досить просто за рахунок його близькості до попереднього стандарту.

Специфікація проміжного рівня включає більшість рис нового стандарту. До цього рівня були включені елементи мови, які забезпечують кращу підтримку концепцій реляційної моделі; були перевизначені незрозумілі або надлишкові синтаксичні конструкції. Також були включені мовні конструкції, які вже були представлені в продуктах головних розробників реляційних баз даних. Крім того враховувалось, щоб нові конструкції можна було відносно легко реалізувати.

Специфікація повного рівня співпадала з усім SQL/92. Вона включала мовні конструкції, які мали важливий характер для користувачів, але були досить складні в реалізації для більшості постачальників баз даних. Повна сумісність з цим рівнем не була вимогою стандарту.

Постачальники баз даних почали втілювати специфікації SQL/92 в свої продукти. Але кожна компанія розширяла мову своїми доповненнями, які виходили за межі стандарту. Ці доповнення називалися *розширеннями*. Наприклад, постачальник міг додати до базових типів даних свої типи, такі як геометричні типи. Таки розширення збільшували функціональність окремої реалізації мови, були корисними для користувачів та дозволяли постачальникам дистанціювати свою версію від версій інших компаній. Наслідком стала поява багатьох діалектів мови та складності для розробників прикладних програм, коли вимагалась підтримка кількох версій СУБД.

В 1997 р. організація ANSI X3 змінила назву на National Committee for Information Technology Standards (NCITS). Відповідно технічний комітет, який займався стандартизацією SQL, був названий ANSI NCITS-H2. Складність стандарту зростала швидкими темпами. Робота була завершена в 1999 р. Це була третя значна ревізія стандарту, відома під назвою SQL3 або SQL: 1999 [5]. Вона складалася з п'ятьох частин [6, 7, 8, 9, 10].

ISO/IEC 9075-1: Framework (SQL/Framework). Ця частина містила визначення базових понять, таких як таблиця, рядок, схема, каталог, вбудовані типи даних, NULL значення, та ін., опис та стислу характеристику усіх інших частин стандарту, систему умовних позначень та погоджень, використаних подалі в стандарті.

ISO/IEC 9075-2: Foundation (SQL/Foundation). Ключова частина стандарту, яка містить специфікацію ядра SQL. Вона задає мови маніпулювання даними (DML – Data Manipulation Language) та визначення даних (DDL – Data Definition Language). В документі специфікуються символічні та числові типи даних, булевський тип даних, типи даних для роботи з датами, часом та інтервалами часу. Також вводяться типи рядків таблиць, масивів, доменів та типи, визначені користувачем (UDT – User Defined Types). Визначаються посилкові (reference) типи. Надається специфікація функціональних залежностей та ключів кандидатів (candidate keys); визначаються тригери.

Множина всіх операторів SQL розбита на декілька класів.

- Оператори роботи зі схемами. Використовуються для створення, модифікації та видалення схем які є аналогами типів даних в звичайних мовах програмування та об'єктів схем, тобто структур даних відповідаючи цим схемам. До таких схем та об'єктів відносяться таблиці, представлення (view), домени, типи визначені користувачем, обмеження на цілісність даних, тригери, та ін.

- Оператори роботи з даними. Використовуються для виконання запитів та додавання, видалення і зміни рядків в таблицях.

- Оператори завдання транзакцій. Як слідує із назви, ініціюють та завершують транзакції.

- Оператори керування процесом виконання програми. До них відноситься оператор RETURN, який задає значення, що повертається функцією.

- Оператори встановлення з'єднання. До них відносяться оператори, які дозволяють ініціювати та завершити з'єднання з базою даних та дозволяють SQL-клієнтам перемикає підключення від одної бази даних до іншої.

- Оператори керування сесією SQL-клієнта. Вони дозволяють задавати значення параметрів сесії.

- Оператори для діагностики. Надають різноманітну діагностичну інформацію.

ISO/IEC 9075-3: Call Level Interface (SQL/CLI). В цій частині розглядається, яким чином оператори SQL викликаються з програмам, написаних на інших мовах програмування. Розглянуті в цій частині функції дозволяють виділяти та звільняти ресурси, встановлювати та розривати зв'язок між SQL-клієнтом та SQL-сервером. Описуються функції для виконання SQL-операторів, включаючи підготовку операторів для наступного виконання, одержання діагностичної інформації та інформації стосовно властивостей реалізації версії SQL, що використовується.

ISO/IEC 9075-4: Persistent Stored Modules (SQL/PSM). В документі описуються додаткові оператори SQL, які дозволяють писати запити, використовуючи виключно оператори SQL/PSM та не вдаючись до засобів інших мов програмування. Таки як декларація змінних оператором DECLARE, присвоювання значень змінним оператором SET, оператори розгалуження CASE та IF, оператори циклювання LOOP, WHILE, REPEAT, оператор циклювання по таблиці FOR, та ін. Тобто мова стає обчислювальне повною (computationally complete).

ISO/IEC 9075-5: Host Language Bindings (SQL/Bindings). Документ специфікує три можливих метода зв'язування операторів SQL, які вбудовані в прикладні програми, з SQL-сервером. До таких методів відноситься динамічний SQL, вбудований SQL та пряме звернення до SQL-операторів (direct invocation of SQL).

Динамічний SQL є методом включення операторів SQL в одну з численних мов програмування, який дозволяє створювати та виконувати SQL-оператори в процесі виконання програми, а не на фазі компіляції.

Вбудований SQL є методом включення операторів SQL в мову програмування, при якому оператори SQL компілюються разом з іншими операторами мови. В процесі компіляції вони замінюються викликами функцій, які виконують оператори SQL.

Пряме звернення до операторів SQL дозволяє виконувати оператори SQL напряду.

В 2003 р. був прийнятий новий стандарт, який називається SQL: 2003 [11]. В ньому були зроблені незначні зміни у вже існуючих специфікаціях, а також додані нові поняття та операції. З'явилися нові можливості пов'язані з XML. Були додані віконні функції (window functions), генератори послідовностей (sequence generator) та два нових типи даних для колонок – обчислювальні колонки, та ідентифікуючі колонки (auto-generated values, identity-columns). Був доданий новий оператор MERGE. Оператор CREATE TABLE був розширений двома новими формами, такими як CREATE TABLE AS та CREATE TABLE LIKE. Було вилучено два типа даних – BIT та BIT VARYING. Новий стандарт складався з наступних частин [12, 13, 14, 15, 16, 17, 18, 19, 20].

Part 1: Framework (SQL/Framework). Ця частина була поновлена порівняно з попереднім стандартом.

Part 2: Foundation (SQL/Foundation). Базується на стандарті 1992 р. Поновлена в 1999 р. та в 2003 р. Класифікація операторів SQL була розширена специфікацією операторів маніпулювання збереженими процедурами та оператора CALL. В цю частину був включений розділ SQL/Bindings, який раніше виступав окремим документом.

Part 3: Call-Level Interface (SQL/CLI). Ця частина була сформована в 1995 р. та поновлена в 1999 р., 2003 р. Вона відповідає специфікації інтерфейсу ODBC компанії Microsoft.

Part 4: Persistent Stored Modules (SQL/PSM). Порівняно з попередньою версією стандарту специфікації збережених процедур (stored procedures) та оператора CALL було перенесено в частину Foundation. Специфікації, що залишилися, були поновлені.

Part 9: Management of External Data (SQL/MED). Новий документ визначає, яким чином SQL-оператори можуть звертатися до зовнішніх даних за допомогою таких типів даних як зовнішні таблиці та спеціальні контейнерні типи даних (datalink data types).

Part 10: Object Language Bindings (SQL/OLB). Надає специфікації, яким чином SQL вбудовується в мову Java.

Part 11: Information and Definition Schema (SQL/Schemata). Цей документ раніше був розділом частини Foundation; він визначає інформаційну схему та схему визначень (Information and Definition schema), за допомогою

яких можна отримати інформацію про наявні в базі даних об'єкти, такі як таблиці, колонки і т.д.

Part 13: Routines and Types Using the Java™ Programming Language (SQL/JRT). Новий документ визначає, яким чином мова Java може бути використана всередині СУБД.

Part 14: XML-Related Specifications (SQL/XML). Мабуть найбільш важливе розширення стандарту; визначає, як мова XML може використовуватись в SQL. Ця частина розроблялась з врахуванням стандарту W3C мови XQuery V1.1.

Станом на початок 2008 р. це останній завершений стандарт мови SQL. Зараз повним ходом іде робота над новою версією стандарту. Планується суттєво переробити документ SQL/XML, де описується яким чином XML-дані можуть бути імпортовані та записані в реляційну базу даних, засоби маніпуляції такими даними,

представлення реляційних даних у XML-форматі. Також специфікується, яким чином XQuery – мова запитів до XML-даних – може бути інтегрована в SQL.

Міжнародний комітет також працює над SQL/MM – мультимедійним стандартом SQL, який буде складатися з п'яти частин: Framework, Full Text, Spatial, Still Image та Data Mining.

В кінці 2007 р. Microsoft представила новий проект LINQ (Language Integrated Query) [21], який є SQL-подібною мовою запитів, інтегрованою в стандартні мови програмування і функціонує як частина .Net Framework 3.5. Проект дозволяє ліквідувати розрив між мовами, орієнтованими на роботу з множинами, такими як SQL, та звичайними мовами програмування, орієнтованими на роботу з окремими даними та послідовностями даних.

Список використаних джерел

1. *E. F. Codd* A relational model of Data for Large Shared Data Banks.// Comm. ACM.- vol. 13, no. 6 (June 1970).- p. 377 – 387.
2. *Donald D. Chamberlin, Raymond F. Boyce*.- SEQUEL: a Structured English Query Language.// SIGMOD Workshop, Vol 1. – 1974. – P. 249-264
3. *John L. Viescas, Michael J. Hernandez*. SQL Queries for Mere Mortals: a hands-on guide to data manipulation in SQL. - Pearson Education, Inc.- Second edition, 2008
4. *К.Деїт* Руководство по реляционной СУБД DB2.- Москва, Финансы и Статистика, 1988
5. *Jim Melton and Alan Simon*. SQL:1999 Understanding Relational Language Components. - San Francisco: Morgan Kaufmann Publishers, 2002
6. ANSI/ISO/IEC International Standard (IS) Database Language SQL—Part 1: SQL/Framework
7. ANSI/ISO/IEC International Standard (IS) Database Language SQL—Part 2: Foundation (SQL/Foundation)
8. ANSI/ISO/IEC International Standard (IS) Database Language SQL—Part 3: Call Level Interface (SQL/CLI).
9. ANSI/ISO/IEC International Standard (IS) Database Language SQL—Part 4: Persistent Stored Modules (SQL/PSM)
10. ANSI/ISO/IEC International Standard (IS) Database Language SQL—Part 5: Host Language Bindings (SQL/Bindings)
11. *Andrew Eisenberg, Jim Melton, Krishna Kulkarni, Jan-Eike Michels, Fred Zemke*. SQL:2003// SIGMOD Record, Vol. 33, No. 1.- March 2004
12. ISO/IEC 9075, Framework
13. ISO/IEC 9075-2, Foundation
14. ISO/IEC 9075-3, Call-level interface
15. ISO/IEC 9075-4, Persistent Stored Modules
16. ISO/IEC 9075-9, Management of External Data
17. ISO/IEC 9075-10, Object Language Bindings
18. ISO/IEC 9075-11, Information and Definition Schemas
19. ISO/IEC 9075-13, Routines and Types Using the Java Programming Language
20. ISO/IEC 9075-14, XML-Related Specifications
21. *Paolo Pialorsi, Marco Russo*. Introducing Microsoft LINQ. Microsoft Press, 2007

Надійшла до редколегії 09.06.2008

УДК 004.7
Романюк С.І.

Архітектура мережних засобів дистанційного навчання

В статті описуються архітектурні рішення інтеграції локальної комп'ютерної мережі гуртожитку в загальноуніверситетську мережу з метою створення умов для дистанційного навчання студентів університету.

Ключові слова: дистанційне навчання, локальні мережі, корпоративні мережі, мережна архітектура, канали зв'язку.

*E-mail: ats@univ.kiev.ua

Статтю представив д.т.н., проф. Гаращенко Ф.Г.

Вступ.

Комп'ютерні мережі, локальні, корпоративні і глобальні, є важливим фактором підвищення якості навчального процесу. Застосування мереж та мультимедійних периферійних пристроїв разом зі стрімким зростанням ємності запам'ятовуючих пристроїв й швидкості процесорів та каналів зв'язку привело до появи таких понять як електронне навчання та дистанційна освіта. З розвитком телекомунікаційного обладнання та розширенням масштабів комп'ютерних мереж електронна освіта виходить за межі комп'ютерних класів, покриває спочатку території університетів, а згодом країн і континентів. Розвиток мереж і засобів доступу до них вивів спілкування з освітніми ресурсами за межі аудиторій. Першими проявами цієї тенденції ще у 90-ті роки минулого століття стала поява комп'ютерних кафе та клубів, де учні та студенти не лише гралися, а й блукали найвіддаленішими куточками Всесвітнього павутиння.

Київський університет – унікальний науковий і освітній центр, який об'єднує в собі властивості високоінтелектуального і індустріально розвиненого організму. Існування такого унікального центру, з одного боку, значно полегшує проектування комп'ютерної мережі, яка охоплює і об'єднує різні її частини за рахунок наявності значних інтелектуальних і виробничих ресурсів. З іншого боку, побудова всеохоплюючої і всезнаючої обчислювальної мережі пов'язано із значними кількісними і якісними затратами. Вихід із ситуації, яка склалася може бути знайдений тільки по шляху створення інтегрованої навчально-науково-

Romanyuk S.I.

Network architecture for distance learning

The article describes architecture solutions for hostel local network integration to the global university network directed to the creation distance learning conditions for university students.

Key Words: distance learning, local network, corporative network, network architecture, communication channels.

виробничої обчислювальної мережі, яка використовує для інформаційного транспорту відповідні високошвидкісні комп'ютерні мережі. Розбудова комп'ютерної мережі Київського національного університету імені Тараса Шевченка розпочалася ще у 80-х роках [1] минулого століття. Першочерговими завданнями при цьому були:

- об'єднання персональних комп'ютерів в навчальних лабораторіях в локальну мережу;
- забезпечення загального доступу до навчальних та системних ресурсів на сервері мережі з будь-якого робочого місця в навчальних приміщеннях факультету;
- можливість одержання навчальної та наукової інформації зі світової мережі Інтернет;
- надання Інтернет послуг на робочих місцях.

Створення і розвиток університетської комп'ютерної мережі надає факультетам, інститутам і кафедрам доступ до університетської і світової системи інформаційних комунікаційних послуг. Викладачі, наукові співробітники, аспіранти і студенти отримують широкий доступ до розподілених інформаційних систем, баз даних, обчислювальних ресурсів, як в нашій країні, так і за кордоном. В комплексі робіт по створенню університетської комп'ютерної мережі і її інформаційному наповненню бере участь не тільки інформаційно-обчислювальний центр університету, центр телекомунікацій, але і інформаційно-обчислювальні сектори факультетів і інститутів, підрозділи ректорату. До робіт

також залучались великі фірми, які спеціалізуються на комп'ютерних технологіях і засобах комунікацій.

В даний час в університеті встановлено більше 5000 комп'ютерів (в основному персональних), із них понад 90% підключено до університетської мережі. 13 начальних корпусів університету з'єднані між собою швидкісними каналами зв'язку, які зв'язують встановлені в кожному корпусі сервери, роутери, інше мережне обладнання.

Для вирішення питань створення, підтримки і модернізації інфраструктури комп'ютерної мережі в 2005 році був створений центр телекомунікації університету. Задачами централізованого управління мережею, підключенням і реєстрацією користувачів, питаннями захисту інформації та іншими питаннями, пов'язаними з адмініструванням мережі, формуванням загальноуніверситетських баз даних і організацією університетських інформаційних серверів займається створений у складі інформаційно обчислювального центру університету сектор мережних технологій.

Головні завдання мережі дистанційного навчання. Затверджена Вченою Радою університету стратегія розвитку навчальних комп'ютерних інформаційних ресурсів передбачає формування єдиної відкритої освітньо-наукової мережі, в рамках якої передбачається інтегрувати зусилля всіх навчальних підрозділів університету.

Інформаційне наповнення мережі включає в себе:

- Інформацію довідкового і інформаційного характеру для студентів і абітурієнтів;
- Методичні і загальноосвітні матеріали;
- Електронні навчальні і наукові видання, матеріали конференцій;
- Науково-технічну, наукову і економічну інформацію у вигляді персональних сайтів спеціалістів університету.

При розгляді перспектив розвитку систем комп'ютерної обробки даних в університеті робиться акцент на наступних важливих напрямках:

- повне оснащення всіх навчальних підрозділів університету сучасною комп'ютерною технікою із врахуванням необхідних функціональних можливостей з інтеграцією в університетську мережу;
- створення інтегрованої загально-університетської обчислювальної системи на базі сучасних архітектурних рішень з

підключенням всіх інформаційних і обчислювальних ресурсів;

- подальше розширення і розвиток каналів зв'язку і комп'ютерних мереж.

Базовими компонентами такої проблемно-орієнтованої інформаційної мережі служать:

- освітньо-наукова комп'ютерна мережа у вигляді сегмента Internet;
- базові сервіси Internet: WWW, FTP, E-mail і інші;
- об'єднання різних типів інформаційних серверів (WWW, FTP, SQL і т.і.), як в складі окремих цілісних інформаційних вузлів, так і розпорошених по різних інформаційних вузлах корпоративної мережі;
- ідеологія інформаційних сховищ і архітектура клієнт-сервер;
- розвинута система електронного бібліотечного обслуговування, яка включає: мережу сховищ електронних видань з розвинутою системою доступу, як з окремих робочих місць, так і через електронні читальні зали;
- технологічна і інструментальна база підготовки і випуску електронної інформації (електронних видань);
- організація і супровід тематичних мережних телеконференцій та форумів;
- розвинена система пошукових, навігаційних та інших проблемно-тематичних засобів в WWW-просторах.

Основними інфраструктурними рішеннями при створенні університетської освітньо-наукової проблемно-орієнтованої мережі є:

- центральний проблемно-орієнтований інформаційний вузол університету;
- корпоративні інформаційні проблемно-орієнтовані сервери;
- видавництва електронних видань -- Web-лабораторії;
- проблемно-орієнтовані "електронні" читальні зали - комплекси робочих місць з відкритим доступом до мережі Internet;
- проблемно-орієнтовані комп'ютерні класи у вигляді локальних мережних сегментів з виходом в Internet.

Навчальна мережа факультету. Розбудова комп'ютерної мережі університету розпочалася у 1993 році зі створення мережі південного кампусу, у якому зосереджені всі фізико-математичні факультети. Спочатку, мережа об'єднала 3 факультети: фізичний, радіофізичний

та кібернетики, але вже через рік до неї приєдналися і факультети центрального кампусу розташовані у головному та гуманітарному навчальних корпусах.

Для організації дистанційного навчання, забезпечення доступу студентів до електронних баз даних бібліотек і інших установ у другій половині 90-х років на факультетах починають створюватися підключені до мережі комп'ютерні класи.

Здешевлення обчислювальної техніки та інформаційного трафіку дозволило перенести Інтернет-послуги в інші місця концентрації молоді. Комп'ютери ставали дійсно персональними, що надало студентам можливість придбання їх для власних потреб вдома чи в гуртожитку.

Ще з кінця 90-х років комп'ютерне господарство у гуртожитках почало проходити ті ж етапи розвитку, що і в навчальних корпусах: спочатку лінійні з'єднання коаксіальним кабелем (навіть найкоротшим шляхом - по зовнішнім стінам), потім встановлення активного мережного обладнання, самостійне підключення до Інтернет через телефонні лінії і, нарешті, повномасштабне підключення до загальноуніверситетської мережі. Враховуючи обмеженість ресурсу місць в комп'ютерних

класах навчальних корпусів, керівництвом університету було прийнято рішення про організацію комп'ютерних класів у гуртожитках для доступу студентів до мережі університету та Інтернет.

Інші важливі фактори, такі як зменшення аудиторного навантаження, масове працевлаштування студентів та їх вічна звичка працювати по ночах, також сприяли тому, що місцем найбільшого попиту Інтернет доступу у 2000-х роках стають студентські гуртожитки.

Розглянемо поетапний розвиток архітектури комп'ютерної мережі навчального підрозділу на прикладі одного з найбільш комп'ютеризованих факультетів – факультету кібернетики.

Детальний опис комп'ютерної мережі корпусу факультету кібернетики дано у роботі [2], тому у даній статті зупинимось на специфіці створення мережі гуртожитку і її підключення до загальноуніверситетських ресурсів. Мережа корпусу реалізовувалась засобами локальної інтеграції, оскільки найдовші сегменти були віддалені від серверів, у тому числі і загальноуніверситетських, лише на 200-300 метрів. Для підключення гуртожитку треба було крім локальних засобів організувати віддалений доступ на відстань до 4,3 км (рис.1).

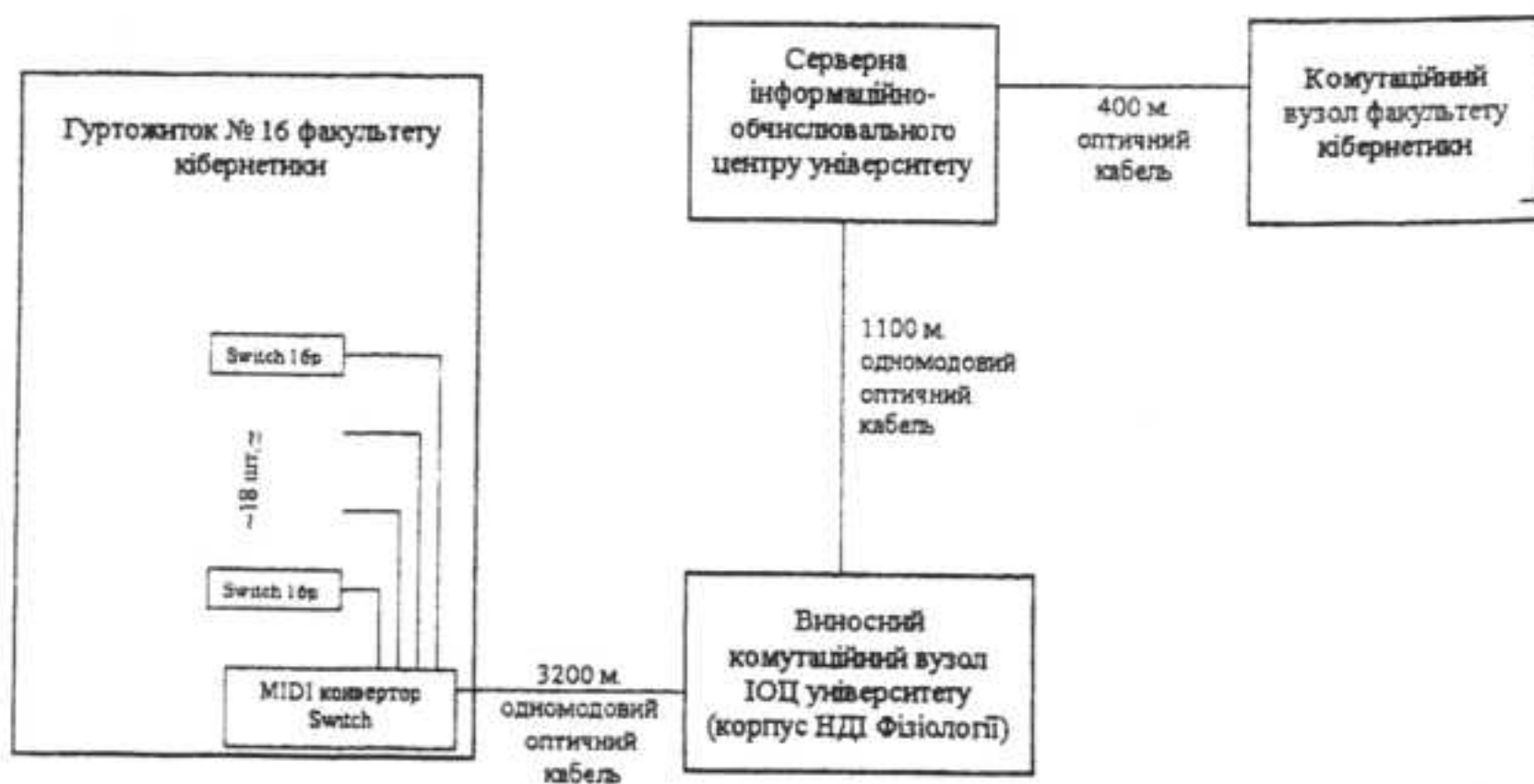


Рис.1. Структурна схема підключення факультету кібернетики та гуртожитку факультету кібернетики до загальноуніверситетської комп'ютерної мережі

Локальна комп'ютерна мережа гуртожитку факультету (рис.2) була побудована під час проведення капітального ремонту гуртожитку у

2004 році. Гуртожиток має два крила по 9 поверхів. На кожному поверсі встановлено ХАБ на 16 портів до якого підключено 12 кімнат. Вся

мережа гуртожитку виконана кабелем UTP 5 категорії. Хаби всіх поверхів підключені до центрального комутатора, який встановлено на 1 поверсі.

Всього у гуртожитку підключено до комп'ютерної мережі 216 кімнат та адміністрація гуртожитку. До 2007 року мережу гуртожитку було підключено до загальноуніверситетської мережі по безпосередньому зв'язку з використанням однієї пари проводів у існуючих мідних кабелях типу ТПП, які прокладено з гуртожитку до кросу відомчої АТС університету (ВАТС) та з кросу ВАТС до серверної інформаційно-обчислювального центру

університету(ІОЦ). Підключення здійснювалось за допомогою двох модемів 2 Мбіт, що дало змогу організувати канал передачі даних між ІОЦ університету та гуртожитком із швидкістю до 2 Мбіт/с.

Адміністрування мережі гуртожитку здійснюється централізовано черговим адміністратором ІОЦ університету. По мірі підключення комп'ютерів у проживаючих в гуртожитку канал на гуртожиток із швидкістю 2 Мбіт/с перестав задовольняти всі потреби користувачів у передачі даних і виникла потреба у збільшенні каналу на гуртожиток.

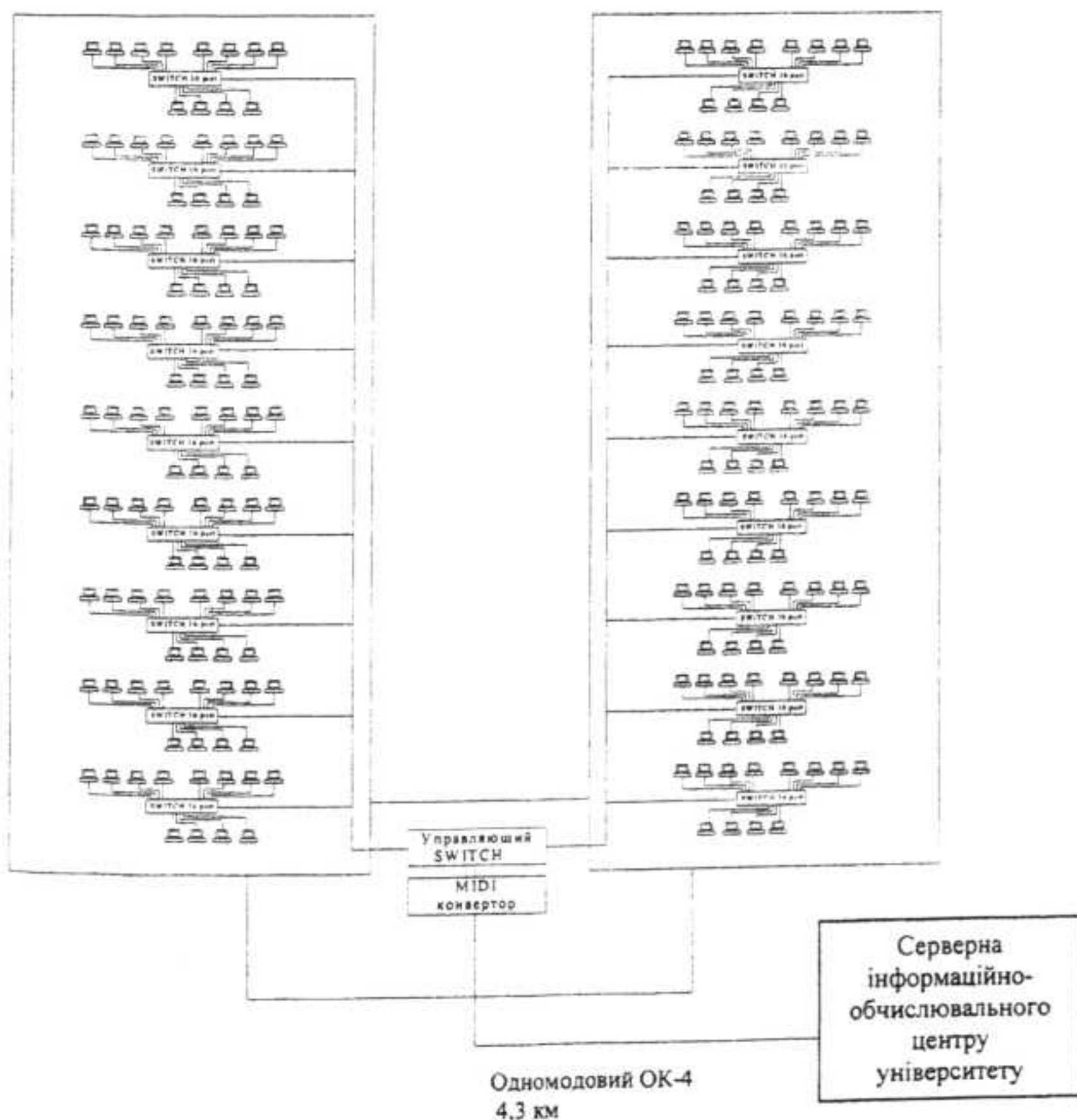


Рис.2. Структурна схема мережі гуртожитку факультету кібернетики

Враховуючи велику кількість користувачів мережі університету та для забезпечення належного рівня передачі даних у 2005 році керівництвом університету було прийнято рішення про необхідність заміни існуючих каналів передачі даних які мають швидкість передачі 2 Мбіт/с між підрозділами та ІОЦ університету на високошвидкісні канали із швидкістю до 100 Мбіт/с. Для переключення всіх підрозділів на високошвидкісні канали у тому числі гуртожитку факультету було виконано проект де передбачалось прокладення сучасних оптоволоконних ліній з використанням одномодового волоконно-оптичного кабелю та встановлення у підрозділах конв'єкторів і комутаторів з пропускною швидкістю 100 Мбіт/с. Вартість переоснащення значна тому роботи виконувались поетапно і були закінчені у 2007 році. У гуртожитку факультету встановлено оптичний конв'єктор на комутатор із швидкістю передачі даних 100 Мбіт/с, проведено необхідне налагодження мережі і на даний час потреби студентства по забезпеченню високошвидкісними каналами передачі даних задоволені.

Висновки. Робота з кінцевими користувачами ведеться по трьом напрямкам. Проводиться відбір стандартних апаратних і програмних засобів для підключення до Internet по комутованим телефонним лініям з метою забезпечення можливості роботи на різних платформах, з різними операційними системами і надання повного сервісу мегамережі Internet (E-mail, FTP, WWW и др). При цьому користувач може мати можливість виступати не тільки, як користувач, але і як генератор інформаційних послуг.

Підготовка методичних матеріалів і інструкцій для кінцевого користувача є задачею, яку найбільш ефективно може вирішити вища школа. Викладачі і методисти університету готують посібники, проводять обмін досвідом їх використання. Наявна комп'ютерна мережа використовується для розповсюдження методичних матеріалів.

Досвід втілення і експлуатації комп'ютерної техніки показав, як позитивні, так і негативні властивості апаратних і програмних рішень.

Позитивними моментами є:

- ефективність по відношенню показників продуктивності / вартість;
- використання доступних і зрозумілих користувачу інтерфейсів;
- гнучкість, простота і дешевизна супроводження і нарощування системи;

- достатність можливостей для вирішення більшості завдань.

Негативними моментами є:

- недостатня кількість робочих станцій;
- недостатність ресурсів для організації ефективного виконання різних задач в режимах з великою кількістю користувачів та поставлених завдань;
- відсутність апаратних засобів і програмного забезпечення для роботи в операційних системах UNIX, VM/SP, а також для організації ефективної роботи з університетською інтегрованою базою даних;
- недостатність обчислювальних ресурсів для вирішення складних обчислювальних задач моделювання і графічної візуалізації результатів.

На даний момент в області комунікаційних систем Київським університетом отримано ряд практичних результатів, значення яких полягає у створенні концептуальних, технічних і організаційних засад для втілення мережних інформаційних технологій в підрозділах університету. Змістом запланованих робіт є розвиток наявної системи в цілях перетворення її в повномасштабну корпоративну мережу, яка забезпечує інформаційні потреби науки і освіти. При цьому як пріоритетні вибрані наступні напрямки:

- створення інформаційного наповнення мережі;
- побудова сучасної високошвидкісної базової мережі університету;
- науково-методична підтримка втілення мережних інформаційних технологій.

Список використаних джерел

1. *Tretyak Oleg V., Shevchenko Volodymyr P.* Kiev University Network – Goals and Principles. Proceedings of AFCEA Europe Seminar “The Information Technology Contribution to the Building of a Safe Regional Environment” Section 2. “Information Technologies – Technical Assessment”. - Kiev, 1998. - P.28-31.
2. *Романюк С.І., Цикунов В.С., Шевченко В.П.* Програмні та апаратні засоби побудови комп'ютерної мережі факультету // Вісник Київського університету. Серія. фіз.-мат. науки. 2007 - № 3 - с.216 - 222.

Надійшла до редколегії 29.04.2007

УДК 531.37; 538.31

Рубльов¹ Б.В., д. ф.-м. н., проф.,
Чурюмова Т.К., інж., Шарманська В.Б., студ.

Площа взаємного розташування трьох фігур на площині

В статті подається геометрична інтерпретація розподілу трьох випадкових величин як розташування трьох фігур на площині. Задача їх можливого взаємного розміщення розв'язана повністю.

Ключові слова: перетин та об'єднання множин, площа області, опукла фігура.

¹E-mail: b_v@mail.univ.kiev.ua, tata@univ.kiev.ua, vsharmanska@unicyb.kiev.ua

Статтю представив чл.-кор. НАНУ, д. ф.-м. н., проф. Ляшко С.І.

При вивченні задач статистичної обробки експериментів виникають геометричні задачі. Так, вважаючи, що нам відомий розподіл трьох випадкових величин із своїми функціями розподілу, виникають проблеми вивчення їх можливого взаємного розташування. Саме розв'язанню однієї з таких проблем присвячена ця стаття.

Нам відома ймовірність попадання досліджуваної точки в області A_1 , A_2 та A_3 , яка пропорційна площі відповідних областей. При цьому треба оцінити максимальне значення такого виразу:

$$\min\{p(A_1A_2) - p(A_1A_2A_3), p(A_1A_3) - p(A_1A_2A_3), p(A_3A_2) - p(A_1A_2A_3)\}$$

у випадку довільних чи опуклих множин. В статті ці усі поняття переведені на мову геометричних понять.

Постановка проблеми

Основні позначення. Нехай є три множини на площині – M_1 , M_2 , M_3 . Позначимо таким чином множини, що утворюються внаслідок взаємного перетину цих множин:

$$F_i = M_i \setminus (M_j \cup M_k),$$

$$F_{ij} = (M_i \cap M_j) \setminus (M_i \cap M_j \cap M_k),$$

$$F_{ijk} = M_i \cap M_j \cap M_k, \text{ де } i, j, k \in \{1, 2, 3\} \text{ та } i, j, k - \text{різні.}$$

Відповідно до цього будемо позначати через S_α площу області F_α , де α – відповідний індекс.

Позначимо через S – площу $M_1 \cup M_2 \cup M_3$ та розглянемо величину:

$$f = \frac{\min\{S_{12}, S_{13}, S_{23}\}}{S} \quad (1)$$

B.V. Rublyov, D.Sci.(Phys.-Math.), professor,
T.K. Churiumova, eng., V.B. Sharmanska, student

Area of mutual location of three figures on the plane

In this paper geometrical interpretation the distribution of three random variables as location of the three figures on the plane is described. Problem of the possible mutual location of them is completely solved.

Key words: intersection and union of the sets, area of the region, convex figure.

When studying the problems of statistical processing of experiments appear geometrical problems. Considering that distribution of three random variates with their distribution functions are known, there are problems in studying of their possible mutual location. This article is devoted exactly to decision one of such problems.

We know a probability of hitting the explored point into regions A_1 , A_2 and A_3 , which is directly proportional to corresponding regions area. It is necessary to estimate the maximum value of such quantity:

$$\min\{p(A_1A_2) - p(A_1A_2A_3), p(A_1A_3) - p(A_1A_2A_3), p(A_3A_2) - p(A_1A_2A_3)\}$$

in case of arbitrary or convex sets. In this article all of these concepts are translated into geometry concepts.

Problem statement

Basic designations. Let M_1 , M_2 , M_3 be three sets on the plane. Let's designate the sets, which are formed as a result of mutual intersection of these sets, through:

$$F_i = M_i \setminus (M_j \cup M_k),$$

$$F_{ij} = (M_i \cap M_j) \setminus (M_i \cap M_j \cap M_k),$$

$$F_{ijk} = M_i \cap M_j \cap M_k, \text{ here } i, j, k \in \{1, 2, 3\} \text{ and } i, j, k \text{ are different.}$$

According to this designations we denote the area of region F_α as S_α , here α is a appropriate index. Let's designate the area of $M_1 \cup M_2 \cup M_3$ through S and consider the quantity:

$$f = \frac{\min\{S_{12}, S_{13}, S_{23}\}}{S} \quad (1)$$

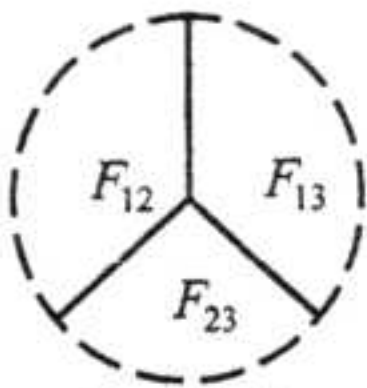


Рис.1 Fig.1

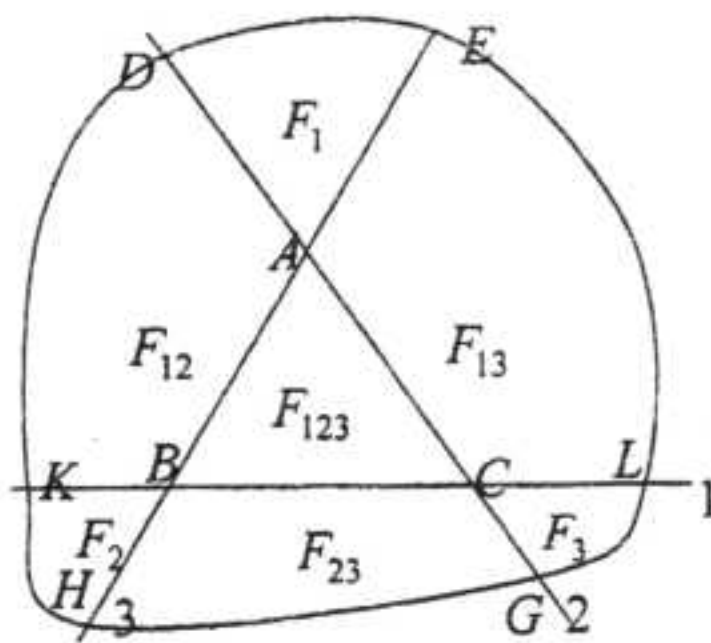


Рис.2 Fig. 2

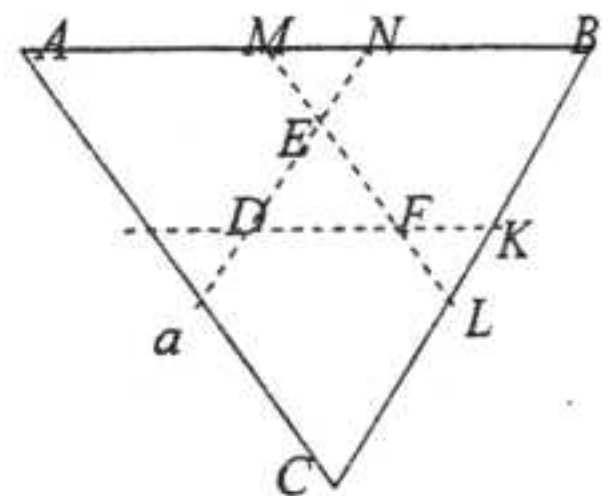


Рис.3 Fig.3

Треба з'ясувати, за яких умов ця величина буде досягати максимуму та чому дорівнює цей максимум.

Загальну проблему доречно розглянути за певних додаткових обмежень, які роблять її різноманітною.

Випадок 1. Довільні фігури M_1, M_2, M_3 .

Випадок 2. Опуклі фігури M_1, M_2, M_3 .

Випадок 3. Опуклі фігури M_1, M_2, M_3 , які в об'єднанні утворюють задану множину M .

а) M – одиничний квадрат, тобто квадрат зі стороною 1.

б) M – одиничне коло, тобто коло радіусу 1.

Випадок 1

Оскільки множини S_{12}, S_{13}, S_{23} є диз'юнктними, тобто вони попарно не перетинаються, тому $S = S_{12} + S_{13} + S_{23}$ і

$$f = \frac{\min\{S_{12}, S_{13}, S_{23}\}}{S} \leq \frac{S_{12} + S_{13} + S_{23}}{3S} = \frac{1}{3}.$$

Ця оцінка справджується для будь-якого випадку.

Таким чином, якщо вдасться побудувати приклад розташування множин M_1, M_2, M_3 , для якого $f = \frac{1}{3}$, то це буде шуканий максимум.

Розглянемо множини (Рис.1) $M_1 = F_{12} \cup F_{13}$, $M_2 = F_{12} \cup F_{23}$, $M_3 = F_{13} \cup F_{23}$, де кожна F_a – це третина круга.

Таким чином, для випадку 1 задача розв'язана. Якщо до цього пункту застосувати підпункти а) і б) із випадку 3, то задача не стає складнішою. Для довільної фігури достатньо розбити її на три частини однакової площі, кожні дві з яких зв'язні, і ми одержимо потрібний приклад.

Випадок 2

Цей випадок ми будемо розглядати для опуклих багатокутників. Після чого спробуємо

We need to understand the condition for maximum of this variable and to determine it.

It is appropriate to consider the general problem in terms of certain additional restrictions, which make it diverse.

Case 1. Random figures M_1, M_2, M_3 .

Case 2. Convex figures M_1, M_2, M_3 .

Case 3. Convex figures M_1, M_2, M_3 , which in union form the given set M .

а. M – unit square, the length of the side of a square is 1.

б. M – unit circle (with radius 1).

Case 1

The sets S_{12}, S_{13}, S_{23} are disjunctive, it means they don't intersect in pairs, so that's why $S = S_{12} + S_{13} + S_{23}$ and

$$f = \frac{\min\{S_{12}, S_{13}, S_{23}\}}{S} \leq \frac{S_{12} + S_{13} + S_{23}}{3S} = \frac{1}{3}.$$

This estimation is executed in any case.

Thus, if it is possible to construct an example of location of sets M_1, M_2, M_3 , for which $f = \frac{1}{3}$, then it will be a required maximum.

Let's consider the sets (Fig. 1) $M_1 = F_{12} \cup F_{13}$, $M_2 = F_{12} \cup F_{23}$, $M_3 = F_{13} \cup F_{23}$, where each of F_a is a third of circle.

Thus, for case 1 problem is solved. The problem doesn't become more difficult, if we try to apply subitems a. and b. from case 3 to this case. In case of random figure is enough to divide it in three parts of equal area, each two of parts are coherent, and we become a necessary example.

Case 2

Let's consider in case 2 convex polygons. After that we will try to generalise it in case of the arbitrary

узагальнити його на випадок довільних опуклих множин. Спочатку розглянемо загальні властивості множин M_1, M_2, M_3 , за яких цей максимум може досягатись, або навпаки.

Властивість 1. Розглянемо вплив множин F_i , $i = \overline{1,3}$ на значення функції f з умови (1). Зрозуміло, що відповідні значення S_i , $i = \overline{1,3}$ входять лише в знаменник цього виразу.

Таким чином, якщо нам вдасться зменшити бодай одне із цих значень, то значення функції f збільшиться.

Властивість 2. Множина F_{123} так само повинна бути якомога меншою, оскільки величина S_{123} також входить лише у знаменник виразу (1). Легко бачити, приклади „оптимальних” розташувань фігур у випадку 1 досягалися за умови $S_1 = S_2 = S_3 = S_{123} = 0$.

Властивість 3. Розглянемо множину F_{123} . Оскільки вона є перетином трьох опуклих багатокутників, то вона так само є опуклим багатокутником.

З трьох одержаних властивостей маємо, що опукла фігура F_{123} повинна бути трикутником. Кожна сторона цього трикутника належить відповідній фігурі. З опуклості цих фігур, вони усі повністю розташовані по один бік від прямої, що містить сторону трикутника F_{123} .

Розв'язання випадку 2

Побудуємо множину F_{123} , яка повинна бути трикутником. Без обмежень загальності можемо зробити цю множину правильним трикутником. Позначимо його ABC . З опуклості фігур вони будуть розташовані по один бік від відповідних сторін $\triangle ABC$ (Рис.2).

Розглянемо криволінійний трикутник ADE , що відповідає F_1 . З доведеного вище зрозуміло, що F_i повинні бути якомога меншими, а F_j – більшими. Якщо вважати, що точки перетину з прямими, що містять сторони $\triangle ABC$ нам відомі, то це досягається тим, що ми з'єднаємо точки D і E , L і G , H і K відповідно відрізками. Це зменшить F_i . Якщо ці відрізки зробити прямими, вони в перетині утворять максимально можливі опуклі множини F_j . Таким чином, шуканий екстремум досягається на такому розташуванні фігур.

Оскільки у нас в знаменнику досліджуваної

convex sets. At first let's consider the general properties of sets M_1, M_2, M_3 , which this maximum can be reached by, or on the contrary.

Property 1. Let's consider influence of the sets F_i , $i = \overline{1,3}$ on value of the function f from condition (1). It is clear, that the appropriate values S_i , $i = \overline{1,3}$ respectively influence only on the denominator of this expression.

Thus, if we manage to reduce at least one of these values, then value of the function f will increase.

Property 2. The set F_{123} similarly must be as small as possible, because value of S_{123} is also included only in the denominator of expression (1).

It is easy to see, that examples of „optimal” locations of figures in case 1 were reached by condition $S_1 = S_2 = S_3 = S_{123} = 0$.

Property 3. Let's consider the set F_{123} . As it is an intersect of three convex polygons, it is similarly a convex polygon.

From these three received properties we have, that convex figure F_{123} should be a triangle. Each side of this triangle belongs to appropriate figure. From convexity of these figures follows that all of them are completely located on the one side of the line that contains the side of triangle F_{123} .

The solution of case 2

Let's construct the set F_{123} , which should be a triangle. Without generally restrictions we can make this set a regular triangle. Let's designate it ABC . From convexity of the figures follows that they will be located on the one side of the appropriate side of triangle $\triangle ABC$ (Fig. 2).

Let's consider curvilinear triangle ADE , which corresponds to F_1 . As proved above, it is clear, that F_i should be as small as possible, and F_j – as big as possible. Suppose that the intersect points with lines, which contain the sides of triangle $\triangle ABC$, are known, then we reach this by joining the point D with E , L with G , H with K accordingly with segments. F_i decrease. If we remake these segments to lines, then by intersect they form the maximally possible convex sets F_j . Thus, the required extremum is reached by this location of figures.

In the denominator of explored function f we

функції f знаходиться площа усієї області об'єднання трьох множин, то легко одержати наступну властивість. Якщо S_{13} – найбільша, а S_{23} – найменша з S_{ij} , і при цьому $S_{13} > S_{23}$, то коливанням відповідної прямої (в даному разі це пряма, що проходить через LG) можна досягти збільшення S_{23} , зменшення S_{13} із збереженням загальної площі S . Але ж тоді відбудеться збільшення функції f .

Таким чином, ми маємо, що при екстремальному розташування фігур повинна виконуватись така властивість.

Теорема 1. Для досягнення максимуму функції f необхідно, щоб виконувались рівності $S_{12} = S_{13} = S_{23}$.

Тепер залишається остаточно дослідити вказану конструкцію. Афійними перетвореннями можна зробити в основі досліджень правильні трикутники (Рис.3). Позначимо сторону великого трикутника $AB = a$, а відрізок $BN = x$. Будемо вважати, що a – фіксоване, а максимум шукатимемо по змінній x . З умов побудови відповідних множин зрозуміло, що $\frac{1}{3}a \leq x \leq \frac{1}{2}a$. Якщо позначити $EF = b$, то з подібності трикутників легко одержати, що $b = 3x - a$.

$$S_{BNEFK} = 2 \cdot \frac{\sqrt{3}}{4} x^2 - \frac{\sqrt{3}}{4} b^2 = \frac{\sqrt{3}}{4} (2x^2 - (3x - a)^2) \\ = \frac{\sqrt{3}}{4} (-7x^2 + 6ax - a^2) \rightarrow \max \quad \text{при} \quad x = \frac{2}{7}a.$$

Відповідне шукане значення максимуму функції

$$\text{складає: } f = \frac{S_{BNEFK}}{S_{ABC}} = \frac{\frac{\sqrt{3}}{4} \cdot \frac{2}{7}a^2}{\frac{\sqrt{3}}{4}a^2} = \frac{2}{7}.$$

Висновки

В загальному випадку задача повністю розв'язана. Тепер треба продовжити вивчення випадку 3, коли множини, що розглядаються, утворюють задану множину. Після повного розв'язання цієї проблеми для випадку трьох множин, виникне потреба її узагальнення на випадок більшої кількості множин, але ця задача спочатку вимагає точної постановки.

Список використаних джерел

1. Кендалл М. Дж, Стюарт А. "Теория распределений". - М., "Наука", 1966.
2. I. Couso, S. Montes, P. Gil, Central tendency measures for random sets, in: Proceedings of EUROFUSE-SIC'99, Budapest, Hungary, 1999.

have area of the union of three sets, that's why it is easy to receive the following property. If S_{13} is the biggest one among S_{ij} , and S_{23} - the smallest one, and therewith $S_{13} > S_{23}$, then through the fluctuation of appropriate line (in this case on it is the line LG) we can make S_{23} bigger and S_{13} smaller with the same total square S . But in this case value of the function f increases in size.

Thus we have, that this property must be executed at extreme location of the figures.

Theorem 1. A necessary condition to reach maximum of the function f is the feasibility of equalities $S_{12} = S_{13} = S_{23}$.

Now it is necessary to research the stated construction. By affine transformations we can make regular triangles in research basis (Fig. 3).

Let's designate the side of the biggest triangle $AB = a$, and segment $BN = x$. Let's assume: a is fixed and maximum is found at variable x . From the conditions of construction the appropriate sets is clear, that $\frac{1}{3}a \leq x \leq \frac{1}{2}a$.

Let $EF = b$, then from the similar triangles we have, that $b = 3x - a$.

$$S_{BNEFK} = 2 \cdot \frac{\sqrt{3}}{4} x^2 - \frac{\sqrt{3}}{4} b^2 = \frac{\sqrt{3}}{4} (2x^2 - (3x - a)^2) \\ = \frac{\sqrt{3}}{4} (-7x^2 + 6ax - a^2) \rightarrow \max \quad \text{at } x = \frac{2}{7}a.$$

Appropriate required maximum of the function is:

$$f = \frac{S_{BNEFK}}{S_{ABC}} = \frac{\frac{\sqrt{3}}{4} \cdot \frac{2}{7}a^2}{\frac{\sqrt{3}}{4}a^2} = \frac{2}{7}.$$

Conclusions

In general case the problem is fully solved. Now it is needed to prolong the study of case 3, when the appropriate sets form the given set. After the complete solution of this problem for the case of three sets, it would be necessary to generalize it in case of more number of sets, but this problem demands at first the exact statement.

References

1. Кендалл М. Дж, Стюарт А. "Теория распределений". - М., "Наука", 1966.
2. I. Couso, S. Montes, P. Gil, Central tendency measures for random sets, in: Proceedings of EUROFUSE-SIC'99, Budapest, Hungary, 1999.

Надійшла до редколегії 07.02.08

УДК 681.3.062

Сільвейструк Л.М.*

**Розширена модель „сутність-зв'язок”:
уточнення типів сутностей суперклас і
підклас та типу зв'язку суперклас/підклас**

Розглядаються і уточнюються такі
поняття розширеної моделі „сутність-
зв'язок”: тип сутності суперклас, тип
сутності підклас, тип зв'язку суперклас/підклас,
тип зв'язку *isa*, концепції
уточнення/узагальнення і категоризації.

Ключові слова: тип сутності суперклас,
тип сутності підклас, тип зв'язку
суперклас/підклас, одиничне успадкування,
множинне успадкування, вибіркоче
успадкування.

*E-mail: slm-klm@rambler.ru

Статтю представив: д.ф.-м. н., с.н.с. Буй Д.Б.

Починаючи з 80-х років минулого сторіччя,
стали швидко розповсюджуватися нові типи
додатків баз даних: мультимедійні додатки,
інструменти автоматизованої підготовки
виробництва, інструменти автоматизованого
проектування. Для того, щоб створювати моделі
даних для цих баз даних, стало недостатньо
понять моделі „сутність-зв'язок”, через це
з'являється розширена модель „сутність-
зв'язок” (*Enhanced Entity-Relationship model*) або
EER-модель.

Розширена модель „сутність-зв'язок”
включає всі концепції моделі „сутність-зв'язок”
плюс власні концепції. Як і у випадку моделі
„сутність-зв'язок” не існує єдиного
загальноприйнятого стандарту для розширеної
моделі „сутність-зв'язок”, але є набір загальних
конструкцій, які лежать в основі більшості
варіантів моделі. Опису цих загальних
конструкцій, а саме типу сутності суперклас,
типу сутності підклас, типу зв'язку
суперклас/підклас, типу зв'язку *isa*, концепцій
уточнення/узагальнення і категоризації згідно [1,
гл. 2; 2, част. III, гл. 14; 3, част. III, гл. 12; 4, част.
II, гл. 3; 5, част. IV] та формалізації моделі
присвячена дана робота.

Питання, пов'язані з формалізацією
розширеної моделі „сутність-зв'язок”, можна
знайти у роботах Елмасрі [5] та Чена [6-8]. В даних
роботах уточнення здійснюються за допомогою
теорії множин, зокрема, теорії решіток.

Уточнення типів сутностей підклас Чен
здійснює за допомогою предикатів. Особливу

Silveystruk L.N.*

**Enhanced model „entity-relationship”
refinement of superclass and subclass
entities types, superclass/subclass
relationship type**

Such notions of enhanced model „entity-
relationship” – superclass entity type, subclass
entity type, superclass/subclass relationship type,
isa relationship type, specialization/generalization
and categorization conceptions – are considered
and refined.

Key words: superclass entity type, subclass
entity type, superclass/subclass relationship type,
single inheritance, multiple inheritance, selective
inheritance.

увагу він приділяє розгляду типу сутності
суперклас та відповідних типів сутностей підклас
при обов'язковому обмеженні участі; на множині
типів сутностей, що входять до ієрархії типів
деякого типу сутності, задає відношення
часткового порядку і розглядає відповідні
решітки. У роботі [8] запропоновано алгоритм
ефективної організації даних та їх пошук за
допомогою булевих решіток.

В свою чергу Елмасрі також уточнює типи
сутностей підклас за допомогою предикатів, але
він передбачає існування типів сутностей
підклас, які не задаються предикатами.

Продовжуючи розпочату у роботі [9]
формалізацію моделі, далі в термінах теорії
відношень уточнюються обмеження, що
накладаються на тип сутності суперклас та його
типи сутностей підклас. Розглядається задача
задання відношення для отримання типів
сутностей підклас відповідного типу сутності
суперклас. Також за допомогою графів
прослідковується процес успадкування
(одиничного та множинного) типів сутностей у
межах ієрархії типів окремого типу сутності та
моделі загалом.

Слід зазначити, що поняття моделі не мають
загальноприйнятої та чіткої інтерпретації; більш
того, існує суттєвий термінологічний різнобій;
тому наведемо у наступній таблиці варіанти, що
зустрічаються в доступній російськомовній
літературі [1-4].

Такий термінологічний різнобій зумовлений,
зокрема, неточностями перекладу відповідних

джерел. Якщо притримуватися термінології роботи [9], то доцільно суперклас називати типом сутності суперклас, підклас – типом сутності підклас, зв’язок суперклас/підклас – типом

зв’язку суперклас/підклас, зв’язок isa – типом зв’язку isa.

Таблиця 1

Ключові поняття розширеної моделі „сутність-зв’язок” та їхні назви

Поняття	1 варіант	2 варіант	3 варіант	4 варіант
Суперклас	Супертип	Базовий клас	Суперклас	Надтип
Підклас	Підтип	Підклас	Підклас	Підтип
Зв’язок суперклас/підклас	–	–	Зв’язок суперклас/підклас	–
Зв’язок isa	Зв’язок isa	Зв’язок isa	–	Зв’язок типу „Є” („Есть”)
Джерело	2, част. III, гл. 14	1, гл. 2	3, част. III, гл. 12	4, част. II, гл. 3

Типи сутностей суперклас та підклас, типи зв’язку суперклас/підклас та isa

Поняття типів сутностей суперклас та підклас було добавлено у розширену модель „сутність-зв’язок” після публікації роботи Дж. Сміта та Д. Сміта [10]. Причиною стали випадки, коли тип сутності містить визначені сутності, які мають спеціальні властивості, що не мають інші сутності даного типу. Тому корисно розділяти такий тип сутності на декілька спеціальних типів сутностей, кожний з яких має назву – *тип сутностей підклас (subclass entity type)*, а вихідний тип сутності називають *типом сутності суперклас (superclass entity type)*.

Тип сутності суперклас – тип сутності, який включає одну або декілька різних допоміжних сукупностей його сутностей, що повинні бути представлені у моделі даних.

Тип сутності підклас – допоміжна сукупність сутностей деякого типу сутності (типу сутності суперклас), яка повинна бути представлена у моделі даних.

Уточнення одного і того ж типу сутності може проводитися на основі різних відмінних особливостей, тобто один тип сутності суперклас може мати декілька множин типів сутностей підклас, які відображають різні способи групування сутностей типу сутності суперклас.

Кожна сутність типу сутності підклас є сутністю відповідного типу сутності суперклас, але сутність типу сутності суперклас не обов’язково є сутністю деякого типу сутності підклас. Зауважимо, що тип сутності підклас має містити щонайменше одну сутність, інакше не має сенсу створювати даний тип сутності підклас.

Нижчеподана особливість типів сутностей досить важлива при розгляді та уточненні понять розширеної моделі „сутність-зв’язок”, але в

літературі на ній не акцентують увагу, хоча використовують: один і той же тип сутності не може бути одночасно типом сутності суперклас та типом сутності підклас даного типу сутності суперклас (змістовно, рефлексивності тут нема).

Типи сутностей суперкласи та підкласи будемо інтерпретувати як множини, а сутності вказаних типів – як елементи даних множин. Нехай множина E відповідає типу сутності суперклас, а непорожні множини $E_1, E_2, \dots, E_n$ – відповідним типам сутностей підклас даного типу сутності суперклас, тоді $\bigcup_{i=1}^n E_i \subseteq E$.

Далі розглядається тип сутності суперклас та відповідні типи сутностей підклас, які відповідають деякому способу групування сутностей суперкласу. При створенні типу сутності суперклас та відповідних йому типів сутностей підклас на ці типи сутностей накладаються обмеження. Розглянемо наступні обмеження:

- *обмеження участі (participation constraint)*;
 - *обмеження неперетину (disjoint constraint)*.
- Обмеження участі* визначає, чи кожна сутність типу сутності суперклас відноситься до деякого типу сутності підклас.

Обмеження участі може бути *обов’язковим (mandatory)* або *необов’язковим (optional)*. При обов’язковій участі кожна сутність типу сутності суперклас повинна бути сутністю деякого типу сутності підклас. При необов’язковій участі деяка сутність типу сутності суперклас може не бути сутністю жодного типу сутності підклас.

Обмеження неперетину описує зв’язок між сутностями типів сутностей підклас (які пов’язані з одним типом сутностей суперклас) і вказує, чи може сутність типу сутності суперклас належати тільки до одного або до декількох типів

сутностей підклас (іншими словами, чи можуть різні типи сутностей підклас перетинатися).

Обмеження неперетину може бути *неперетинаючим* (*disjoint*) або *перетинаючим* (*nondisjoint*). Якщо обмеження неперетину є неперетинаючим (або іншими словами типи сутностей підклас є неперетинаючими), то кожна сутність типу сутності суперклас може бути сутністю не більше одного типу сутності підклас. Якщо обмеження неперетину є перетинаючим (або іншими словами типи сутностей підклас є перетинаючими), то сутність типу сутності суперклас може бути сутністю декількох типів сутностей підклас.

Зрозуміло, що дане обмеження можна застосовувати, якщо тип сутності суперклас має більше одного типу сутності підклас.

Так як вказані два обмеження є логічно незалежними характеристиками утворення типів сутностей суперклас та підклас, то при їх сумісному використанні виділяють чотири наступні різні види обмежень:

- обов'язкове та неперетинаюче;
- необов'язкове та неперетинаюче;
- обов'язкове та перетинаюче;
- необов'язкове та перетинаюче.

При обов'язковому та неперетинаючому обмеженні будь-яка сутність типу сутності суперклас повинна належати одному і тільки одному типу сутності підклас даного типу сутності суперклас.

При необов'язковому та неперетинаючому обмеженні кожна сутність типу сутності суперклас або належить єдиному типу сутності підклас даного типу сутності суперклас, або не належить жодному типу сутності підклас даного типу сутності суперклас.

При обов'язковому та перетинаючому обмеженні будь-яка сутність типу сутності суперклас повинна належати деякому типу сутності підклас (необов'язково одному) даного типу сутності суперклас.

При необов'язковому та перетинаючому обмеженні кожна сутність типу сутності суперклас або належить деякому типу сутності підклас (необов'язково одному) даного типу сутності суперклас, або не належить жодному типу сутності підклас даного типу сутності суперклас.

Розглядаючи обов'язкове обмеження участі, зустрічаємо поняття покриття множини. Очевидно, що обов'язкове та перетинаюче обмеження можна уточнити за допомогою

покриття підмножинами $E_1, E_2, \dots, E_n$ множини

E , тобто $E = \bigcup_{i=1}^n E_i$ і $E_i \neq \emptyset$, де $i = 1, 2, \dots, n$.

Підмножини $E_1, E_2, \dots, E_n$ утворюють розбиття множини E , якщо розглядати обов'язкове та неперетинаюче обмеження, тобто

$E = \bigcup_{i=1}^n E_i$, $E_i \neq \emptyset$ і $E_i \cap E_j = \emptyset$, де $i \neq j$ та $i, j = 1, 2, \dots, n$.

Якщо розглядати необов'язкове обмеження участі, то необов'язкове та перетинаюче обмеження можна уточнити за допомогою покриття підмножинами $E_1, E_2, \dots, E_n$ множини

E або її власної підмножини, тобто $\bigcup_{i=1}^n E_i \subseteq E$ і $E_i \neq \emptyset$, де $i = 1, 2, \dots, n$.

Нарешті, необов'язкове та неперетинаюче обмеження можна уточнити за допомогою розбиття підмножинами $E_1, E_2, \dots, E_n$ множини

E або її власної підмножини, тобто $\bigcup_{i=1}^n E_i \subseteq E$, $E_i \neq \emptyset$ і $E_i \cap E_j = \emptyset$, де $i \neq j$ та $i, j = 1, 2, \dots, n$.

У роботі [9] обмеження кардинальності, які накладалися на бінарні типи зв'язків, уточнювалися за допомогою функціональності та проєкції відповідних відношень. Тому по аналогії дамо уточнення обмеження неперетину та обмеження участі.

Розглянемо множину $E' = E \cup \{E_1, E_2, \dots, E_n\}$ і задамо на ній бінарне відношення $<$:
$$e < E_i \stackrel{\text{def}}{\Leftrightarrow} e \in E_i, \text{ тобто множина } E_i, i = 1, 2, \dots, n$$

(тип сутності підклас) містить елемент e множини E (типу сутності суперклас).

Обмеження неперетину можна уточнити за допомогою функціональності відношення $<$.

Якщо відношення $<$ функціональне, то відповідне обмеження неперетину є неперетинаючим (типи сутностей підклас даного типу сутності суперклас є неперетинаючими).

Якщо відношення $<$ не функціональне, то відповідне обмеження неперетину є перетинаючим (типи сутностей підклас даного типу сутності суперклас є перетинаючими).

Обмеження участі можна уточнити за допомогою проєкції відношення $<$. Якщо $\pi_1^2(<) = E$, то обмеження участі обов'язкове; у

випадку, коли $\pi_1^2(\prec) \subseteq E$, обмеження участі необов'язкове.

У табл. 2 показані різні види обмежень, що накладаються на тип сутності суперклас і його типи сутностей підклас, та їх уточнення.

Таблиця 2

Уточнення обмежень, що накладаються на тип сутності суперклас та його типи сутностей підклас

	$\prec$ функціональне	$\prec$ не функціональне
$\pi_1^2(\prec) = E$	Обов'язкове та неперетинаюче обмеження (типи сутностей підклас утворюють розбиття типу сутності суперклас)	Обов'язкове та перетинаюче обмеження (типи сутностей підклас утворюють покриття типу сутності суперклас)
$\pi_1^2(\prec) \subseteq E$	Необов'язкове та неперетинаюче обмеження (типи сутностей підклас утворюють розбиття типу сутності суперклас або його власної підмножини сутностей)	Необов'язкове та перетинаюче обмеження (типи сутностей підклас утворюють покриття типу сутності суперклас або його власної підмножини сутностей)

Обмеження, що накладаються на тип сутності суперклас та його типи сутностей підклас, застосовуються для кожного способу групування сутностей типу сутності суперклас. Очевидно, що не має сенсу застосовувати дані обмеження до типів сутностей підклас різних логічно-незалежних способів групування сутностей типу сутності суперклас, так як тоді, взагалі кажучи, не отримується корисна інформація.

Розглядаючи тип сутності суперклас та його типи сутностей підклас, цікаво розглянути методи, за допомогою яких можна їх задати. У літературі (дивись, наприклад, [5, част. IV, с. 103]) запропоновано метод *визначаючого предиката* (*defining predicate*) та *визначаючого атрибута* (*defining attribute*). Тобто для кожного фіксованого способу групування сутностей типа сутностей суперклас виділяють деякі визначаючі атрибути, а для кожного типа сутностей підклас деякий визначаючий предикат, заданий на цих визначаючих атрибутах (точніше на значеннях атрибутів). Для того, щоб визначити приналежність сутності типа сутності суперклас відповідному типу сутності підклас, необхідно проаналізувати значення предиката, пов'язаного з типом сутності підклас: якщо значення істинне, то приналежність має місце, якщо хибне, то не має.

Виникає задача: чи можливо задати деяке відношення на множині, що інтерпретує тип сутності суперклас, і за ним отримати множини, що інтерпретують типи сутностей підклас.

Дана задача має розв'язок для частинного випадку – для типу сутності суперклас і його типів сутностей підклас з обов'язковим та неперетинаючим видом обмежень, яке на них накладається. Як зазначалось раніше, даний вид обмежень можна уточнити за допомогою

розбиття. Спираючись на класичні теоретико-множинні результати, відношення еквівалентності дозволяє однозначно розбивати множину, на якій задане відношення, на неперетинаючі класи, які містять еквівалентні між собою елементи (дивись, наприклад, [11]). Ці класи еквівалентності і інтерпретують типи сутностей підклас.

Потрібно відмітити, що при побудові типів сутностей підклас деякого типу сутності суперклас проєктант в моделі явно вказує кількість відповідних типів сутностей підклас. В той час як відношення еквівалентності, в загальному випадку, не дозволяє задавати кількість необхідних класів еквівалентності (точніше кажучи, оцінити потужність фактор-множини).

Перейдемо від частинного випадку до більш загального, тобто знімемо умову транзитивності та узагальнимо поняття еквівалентності до поняття толерантності [11]. Добре відомо, що між відношеннями толерантності і покриттями множини існує тісний зв'язок (нагадаємо, що за допомогою покриття вище уточнювався тип сутності суперклас та його типи сутностей підклас з обов'язковим та перетинаючим видом обмежень). Однак на цьому шляху виникає проблема при побудові класів толерантності (елементів покриття): якщо по покриттю відношення толерантності будується однозначно, то обернена задача побудови покриття по відношенню толерантності має, взагалі кажучи, декілька розв'язків [11]. Таким чином, при заданні покриття відношенням толерантності потрібні додаткові засоби.

Тип зв'язку суперклас/підклас (*superclass/subclass relationship type*) – це тип

зв'язку між типом сутності суперклас та його типами сутностей підклас.

Деякі автори даний тип зв'язку називають *типом зв'язку клас/підклас* (*class/subclass relationship type*) [5, част. IV, с. 87].

Даний тип зв'язку є $(n+1)$ -арним типом зв'язку (тобто це тип зв'язку між одним типом сутності суперклас та його n типами сутностей підклас), де $n \geq 1$.

Уточнення типу зв'язку суперклас/підклас між типами сутностей $E, E_1, E_2, \dots, E_n$ проводиться в два етапи: визначення серед типів єдиного суперкласа (нехай для визначеності E , решта типів є підкласами) та співставлення типам $E, E_1, E_2, \dots, E_n$ множин $E, E_1, E_2, \dots, E_n$ відповідно, причому множини $E_1, E_2, \dots, E_n$ формують покриття множини E або її власної підмножини.

Тип зв'язку *isa* (*isa relationship type*) – це тип зв'язку суперклас/підклас між типом сутності суперклас та його (одним з багатьох, взагалі кажучи) типом сутності підклас.

Даний тип зв'язку є бінарним типом зв'язку „один до одного” (1:1) і його уточнення можна здійснювати за допомогою відношень.

Згідно роботи [9] бінарний тип зв'язку уточнюється за допомогою бінарного відношення, причому бінарний тип зв'язку буде виду „один до одного”, коли відповідне відношення та обернене до нього відношення функціональні. Тип зв'язку *isa* можна уточнити за допомогою тотожного відношення, в даному випадку вигляду $\Delta_{E_1} = \{ \langle e, e \rangle | e \in E_1 \}$ (діагональ на множині E_1). Тривіально, що тотожне відношення та обернене до нього (яке співпадає з вихідним відношенням) є функціональними, тобто тип зв'язку *isa* – це дійсно бінарний тип зв'язку „один до одного”.

Зазначимо, що в доступній російськомовній літературі не має чітко окреслених даних типів зв'язків [1-4], причому у праці [3] використовують термін „тип зв'язку суперклас/підклас” в розумінні типу зв'язку *isa*.

Тип зв'язку суперклас/підклас деякого типу сутності суперклас можна зобразити за допомогою ациклічного орієнтованого графа (точніше дерева одиничної глибини), де множина вершин – множина типів сутностей (типу сутності суперклас та його типів сутностей підклас), множина дуг – множина пар вершин (початком дуги буде вершина, що відповідає

типу сутності підклас, кінцем дуги – вершина, яка відповідає типу сутності суперклас).

На рис. 1 зображено граф, який ілюструє тип сутності суперклас E та його типи сутностей підклас $E_1, E_2, \dots, E_n$ і відповідний тип зв'язку суперклас/підклас.

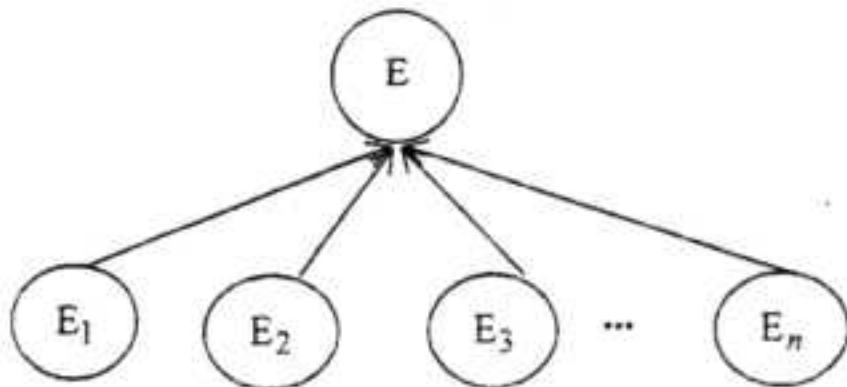


Рис. 1. Граф типу зв'язку суперклас/підклас типу сутності суперклас E

Так як типи сутностей підклас є типами сутностей, то вони можуть мати свої типи сутностей підклас. Очевидно, що поняття типу сутності суперклас та типу сутності підклас відносні. Так, один тип сутності в одному типі зв'язку суперклас/підклас може бути типом сутності підклас, а в іншому – типом сутності суперклас.

Тип сутності суперклас разом із своїми типами сутностей підклас, типами сутностей підклас своїх типів сутностей підклас і т.п. утворює так звану *ієрархію типів* (*type hierarchy*) даного типу сутності. Поняття ієрархії типів у літературі зустрічається під різними іменами, наприклад, ієрархія уточнення, ієрархія узагальнення, ієрархія IS-A [3, част. III, гл. 12].

Для того, щоб зобразити ієрархію типів даного типу сутності, необхідно об'єднати графи типів зв'язків суперклас/підклас типів сутностей суперклас (дерева одиничної глибини), які входять до даної ієрархії типів. В результаті, якщо коректно розбити тип сутності на відповідні типи сутностей підклас та суперклас, то граф ієрархії типів буде ациклічним орієнтованим графом.

На рис. 2 зображено граф ієрархії типів типу сутності E . Вершини графа відповідають типам сутностей, а дуги – типам зв'язку *isa*. Коренева вершина відповідає типу сутності, для якої будується відповідна ієрархія типів.

Сутність типу сутності підклас є сутністю відповідного типу сутності суперклас, тобто вона представляє той же об'єкт реального світу, що і тип сутності суперклас; тому тип сутності підклас успадковує атрибути та типи зв'язків

типу сутності суперклас, а також може мати власні атрибути та типи зв'язків (іноді їх називають *специфічні* або *локальні* атрибути та типи зв'язків).

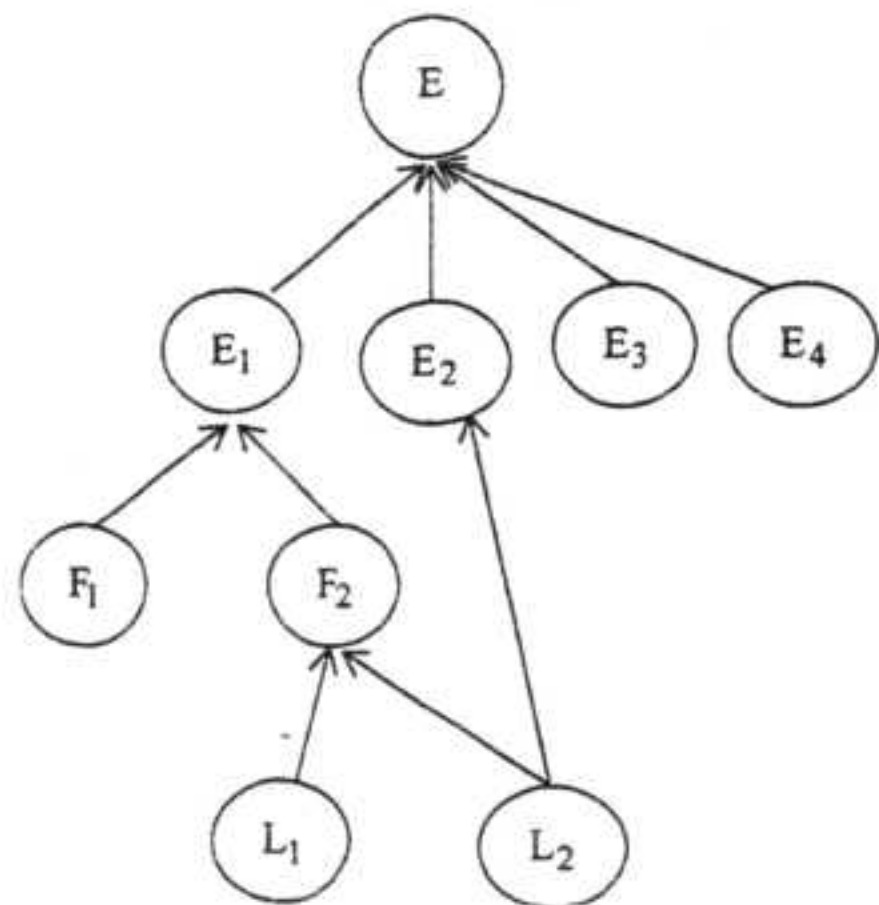


Рис. 2. Граф ієрархії типів типу сутності E

Розглянемо наступні форми успадкування (*inheritance*): *одиничне успадкування* (*single inheritance*) та *множинне успадкування* (*multiple inheritance*).

Одиничне успадкування означає, що тип сутності підклас має рівно один тип сутності

суперклас і успадковує атрибути та типи зв'язків саме даного типу сутності суперклас.

Розглянемо множинне успадкування. Тип сутності, який є типом сутності підклас декількох типів сутностей суперклас, називається *спільним типом сутності підклас* (*shared entity type subclass*) даних типів сутностей суперклас. Це означає, що його елемент повинен бути елементом всіх типів сутностей суперклас, для яких він є типом сутності підклас. В даному випадку до спільного типу сутності підклас застосовують множинне успадкування, при якому атрибути та типи зв'язків типів сутностей суперклас успадковуються даним спільним типом сутності підклас.

Для того, щоб прослідкувати множинне та одиничне успадкування типів сутностей, які входять до ієрархії типів деякого типу сутності, необхідно побудувати відповідний граф ієрархії типів (рис. 2).

Аналізуючи орієнтований граф ієрархії типів даного типу сутності, можна виділити такі його особливості:

- якщо вершині інцидентна одна вихідна дуга, то до відповідного типу сутності підклас застосовують одиничне успадкування;
- якщо вершині інцидентно більше однієї вихідної дуги, то до відповідного типу сутності підклас застосовують множинне успадкування.

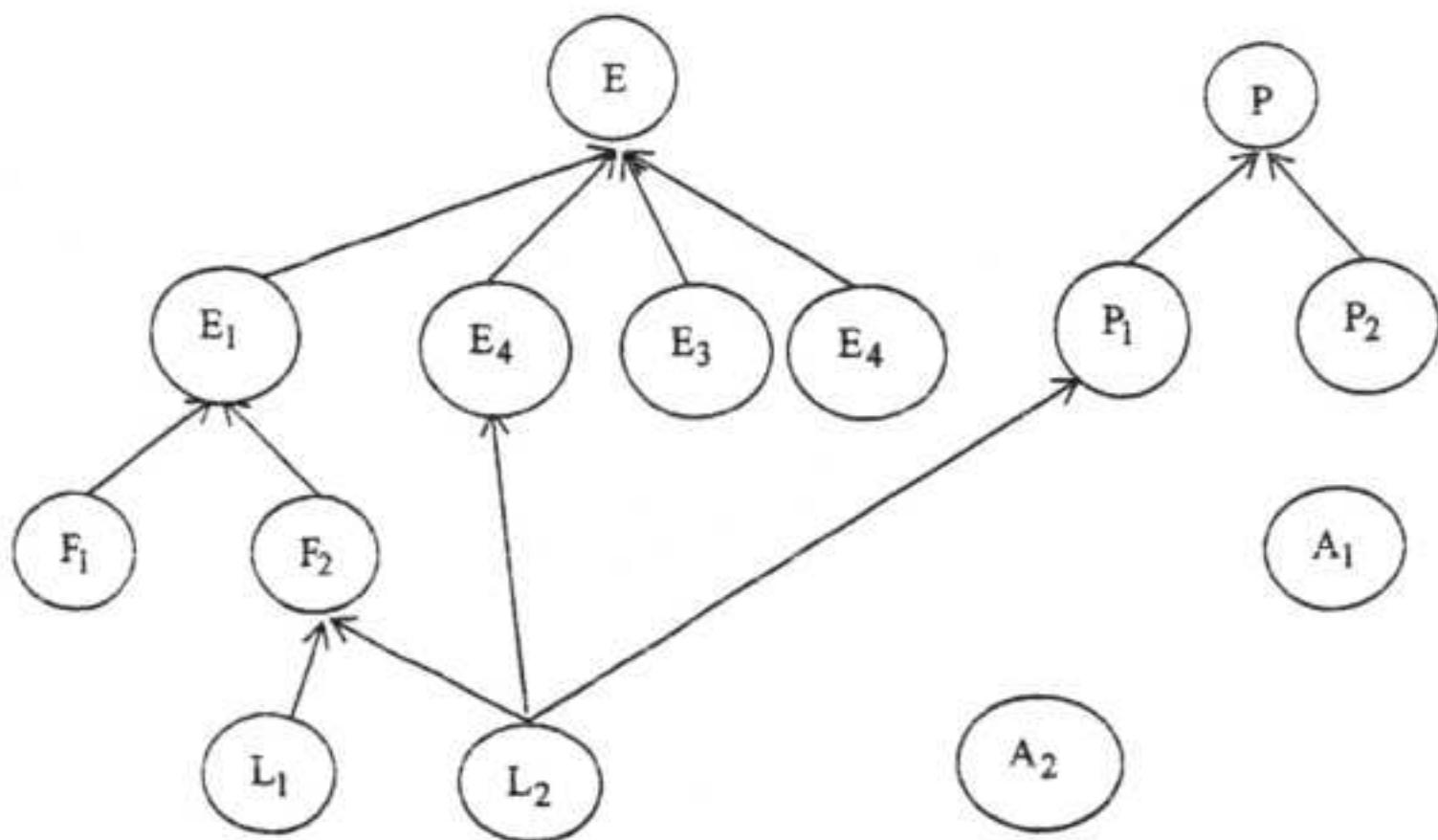


Рис. 3. Граф ієрархії типів моделі

Для того, щоб прослідкувати множинне та одиничне успадкування у межах всієї моделі,

необхідно побудувати ієрархію типів моделі, тобто модель, яка складається з всіх типів

сутностей моделі і типів зв'язків суперклас/підклас, які пов'язують відповідні типи сутностей суперклас та типи сутностей підклас. Ієрархія типів моделі отримується шляхом об'єднання графів типів зв'язку суперклас/підклас та ізольованих вершин, що відповідають типам сутностей, які не є учасниками типів зв'язку суперклас/підклас даної моделі (рис. 3).

Аналізуючи даний орієнтований граф ієрархії типів моделі, можна прослідкувати одиничне та множинне успадкування у межах моделі, аналогічно як це було зроблено для ієрархії типів деякого типу сутності. Зауважимо: до типів сутностей, яким відповідають ізольовані вершини (тобто степінь вершини дорівнює 0) та тупикові вершини (тобто їм інцидентні тільки вхідні дуги), не застосовують жодного виду успадкувань.

Концепція уточнення/узагальнення

У процесі виділення типу сутності суперклас та типів сутностей підклас використовують два підходи – уточнення та узагальнення.

Уточнення (specialization) – це процес підкреслення відмінностей між сутностями типів сутностей шляхом виділення їх особливих характеристик.

Узагальнення (generalization) – це процес прибирання відмінностей між сутностями типа сутності шляхом виділення їх спільних характеристик.

Уточнення представляє собою низхідний підхід для визначення типу сутності суперклас і пов'язаними з ним типів сутностей підклас. Множина типів сутностей підклас визначається на основі деяких відмінних характеристик окремих сутностей типу сутності суперклас.

Узагальнення представляє собою висхідний підхід, який дозволяє створювати узагальнений тип сутності суперклас на основі різних вихідних типів сутностей підклас. Процес узагальнення можна розглядати як протилежний (дуальний) до процесу уточнення, тому весь цей підхід у моделюванні відомий під загальною назвою *уточнення/узагальнення*.

Відповідно до вищеприписаного методу визначаючого предиката та визначаючого атрибута під час процесу уточнення та узагальнення виділяють *тип сутності підклас, визначений предикатом (predicate-defined subclass entity type)*, деякі автори називають його *типом сутності підклас, визначеним умовою (condition-defined subclass entity type)*.

В моделі існують також *типи сутностей підклас, визначені користувачами (user-defined subclass entity type)* [5, част. IV, с. 92]. В цьому випадку, сутності типу сутності суперклас розподіляються у відповідні типи сутностей підклас, визначені користувачем, індивідуально, тобто на розсуд проектувальника моделі.

Концепція категоризації

Розглянемо специфічну форму успадкування – *вибіркове успадкування (selective inheritance)*. Як раніше зазначалось, у моделі можуть бути *спільні типи сутностей підклас*. Інколи необхідно змодельовувати ситуацію, коли сутності спільного типу сутності підклас успадковують атрибути та типи зв'язків того єдиного типу сутності суперклас, яким вони породжені. При цьому даний тип сутності підклас називається *типом сутності категорія (category entity type)* або *підкласом об'єднуючого типу (subclass a union type)*, а дане моделювання – *категоризацією*.

Тип сутності категорія має мінімум два типи сутності суперклас.

Тип зв'язку між типами сутностей суперклас та типом сутності категорія також називають *типом зв'язку суперклас/підклас*.

Даний тип зв'язку є $(n+1)$ -арним типом зв'язку (тобто це тип зв'язку між одним типом сутності категорія та його n типами сутностей суперклас), причому $n \geq 2$.

Типи сутностей суперкласи та категорії будемо інтерпретувати як множини, а їхні сутності – як елементи даних множин. Припустимо, що множина E відповідає типу сутності категорія, а множини $E_1, E_2, \dots, E_n$, $n > 1$, – типам сутностей суперклас даного типу

сутності категорія, тоді, $E \subseteq \bigcup_{i=1}^n E_i$.

До типу сутності категорія, застосовують вибіркове успадкування, тобто сутність даного типу сутності категорія успадковує атрибути та типи зв'язків єдиного типу сутності суперклас, до якого дана сутність належить.

Тип сутності категорія може мати *власні* атрибути та типи зв'язків.

Відмінності між спільним типом сутності підклас та типом сутності категорія:

- спільний тип сутності підклас можна уточнити за допомогою перетину його типів сутностей суперклас. Тип сутності категорія можна уточнити за допомогою об'єднання його типів сутностей суперклас;

- спільний тип сутності підклас успадковує атрибути та типи зв'язків всіх своїх типів сутностей суперклас, тобто дані атрибути (типи зв'язків) отримуються об'єднанням атрибутів (типів зв'язків) всіх відповідних типів сутностей суперклас. Тип сутності категорія успадковує атрибути та типи зв'язків одного з своїх типів сутностей суперклас для конкретної сутності;
- власні атрибути та типи зв'язків спільного типу сутності підклас використовуються для підкреслення особливостей визначених сутностей типів сутностей суперклас (які належать відповідному перетину). Власні атрибути та типи зв'язків типу сутності категорія використовуються, щоб показати спільні атрибути та типи зв'язків всіх сутностей типів сутностей суперклас.

Аналізуючи вищеподані відмінності, можна сказати, що поняття спільного типу сутності підклас та типу сутності категорія є дуальними (двоїстими).

Кожна сутність типу сутності категорія є сутністю єдиного типу сутності суперклас, тому на тип сутності категорія та відповідні типи сутностей суперклас накладається лише обмеження участі.

Обмеження участі може бути обов'язкове або необов'язкове. При обов'язковій участі всі сутності типів сутностей суперклас повинні бути сутностями відповідного типу сутності категорія,

тобто в попередніх позначеннях $E = \bigcup_{i=1}^n E_i$. При

необов'язковій участі (деяка) сутність (деякого) типу сутності суперклас може не бути сутністю відповідного типу сутності категорія, тобто

$$E \subseteq \bigcup_{i=1}^n E_i.$$

При обов'язковій участі сутностей типів сутностей суперклас можна використати і уточнення/узагальнення для представлення сутностей (нагадаймо, що у випадку уточнення/узагальнення вимагається виконання

включення $E \supseteq \bigcup_{i=1}^n E_i$, зокрема рівності). Вибір

методу представлення залежить від розробника.

Типи сутностей суперклас та підклас (категорії) доцільно вводити в модель „сутність-зв'язок”:

- по-перше, вони дозволяють не описувати декілька раз аналогічні сутності, завдяки чому економиться час проектувальника, а діаграми сутностей та зв'язків стають більш зручними для сприйняття;
- по-друге, вони дозволяють ввести у модель більший об'єм семантичної інформації у формі, прийнятної для користувачів моделі.

Список використаних джерел

1. Гарсиа-Молина Г., Ульман Дж., Уидом Дж. Системы баз данных. Полный курс.: Пер. с англ. – Москва: „Вильямс”, 2004. – 1088 с.
2. Дейт Дж. Введение в системы баз данных, 8-е издание.: Пер. с англ. – Москва: „Вильямс”, 2005. – 1328 с.
3. Коннолли Т., Бегг К., Страчан А. Базы данных: проектирование, реализация и сопровождение. Теория и практика, 3-е изд.: Пер. с англ. – Москва: „Вильямс”, 2003. – 1440 с.
4. Крёмке Д. Теория и практика построения баз данных. 8-е изд. – Санкт-Петербург: „Питер”, 2003. – 800 с.
5. Elmasri R., Navathe S. Fundamentals of database systems, 4-th edition. – Pearson: Addison-Wesley, 2004. – 1030 p.
6. Chen P. P., Li M. The lattice structure of entity sets // Proceedings of the 5th International Conference on Entity-Relationship Approach. – 1987. – P. 217-229.
7. Chen P. P., Yang A. The application of Boolean lattice to data manipulation on entity types and their subtypes. Technical report, Computer Science Department, Louisiana State University. – 1988. – 14 p.
8. Chen P. P., Yang A. Efficient data retrieval and manipulation using Boolean entity lattice // Data & Knowledge Engineering. – 1996. – Vol. 20. – P. 211-226.
9. Буй Д. Б., Сільвейструк Л. М. Модель „сутність-зв'язок”: формалізація сутностей та зв'язків // Вісник Київського університету. Сер.: фіз.-мат. науки. – 2006. – Вип. 3. – С. 143-152.
10. Smith, J., and Smith, D. Database abstractions: Aggregation and generalization // ACM Transactions on Database Systems. – 1977. – P. 105-133.
11. Шрейдер Ю.А. Равенство, сходство, порядок. – Москва: Наука, 1971. – 255 с.

Надійшла до редакції 10.01.2008

УДК 519.63

Філатов М.С. *

**Автоматизована система адаптивного
тріангулювання із згущенням багатозв'язних
та складених областей**

Розроблено алгоритм та програмний комплекс адаптивного тріангулювання із згущенням багатозв'язних і складених областей з криволінійними межами.

Ключові слова: адаптивне тріангулювання, тріангуляція, багатозв'язна область, складена область, метод скінченних елементів, згущення.

* E-mail: mfilatov@gmail.com

1. Вступ. В загальному випадку проблема оптимального розбиття області є досить складною. На кожен елемент розбиття можуть накладатись досить жорсткі обмеження. Крім того, область тріангуляції може мати невеликі відносно розмірів самої області вирізи, щілини, містити в собі інші підобласті. Це значно ускладнює алгоритм тріангуляції, оскільки потребує додаткових досліджень, природа яких в загальному випадку пов'язана із методами теорії розпізнання образів.

Сучасні прикладні задачі вимагають високої точності обчислень, розбиття області досліджень повинно містити велику кількість елементів [1-3]. Можливе досить часте повторне розбиття області, обумовлене її динамічними змінами. Враховуючи всі ці обставини, приходимо до висновку, що тільки автоматизований програмний комплекс в змозі ефективно вирішувати поставлені задачі, оскільки числові розрахунки з використанням ручного розбиття області були б занадто громіздкими і неефективними.

На даний час розроблено значну кількість програмних комплексів, які орієнтовані на розв'язання задач по дослідженню різних процесів в областях складної конфігурації. Під час їх створення багато уваги приділяється саме задачі алгоритмічної та програмної реалізації генерації сітки в багатозв'язних областях складної геометричної форми та областях, що складаються з окремих підобластей.

2. Постановка задачі. Розглянемо задачу про тріангуляцію багатозв'язної області або області, складеної з деяких підобластей (складеної

M.S. Filatov *

**The automated system of an adaptive
triangulation with thickening of multicoherent
and complex areas**

Designed and implemented algorithm and program complex for adaptive triangulation of complex and multiply connected areas with curvilinear boundaries.

Key Words: adaptive triangulation, triangulation, multicoherent area, complex area, finite elements method, thickening.

області) [2-4]. Дана задача широко використовується для апроксимації різних поверхонь, скінченно-елементного покриття різних областей при розв'язуванні крайових задач методом скінченних елементів та багатьох інших прикладних задачах.

Нехай задано довільну багатозв'язну область Ω , обмежену криволінійними межами $\Gamma_i, i = \overline{1, n}$, які задані рівняннями в явному, неявному або параметричному вигляді чи достатньо густими таблицями точок, або область Ω , що складається з k підобластей $\Omega_i, i = \overline{1, k}$ (рис.1, 2) [5]. Більше того, задана область може бути багатозв'язно-складена (багатозв'язна область, складена із деяких підобластей). Потрібно покрити дану область Ω трикутниками, що не перетинаються та максимально наближені до рівносторонніх, площа кожного з яких не перевищує певного значення, що визначається функцією густини покриття

$$S: \Omega \rightarrow \mathbb{R} \quad (1)$$

Дана функція задає максимально допустимі розміри трикутників для кожної з точок області.

Крім того, трикутники, що покривають задану область, адаптовані до межі області із згущенням їх в окремих підобластях розглядуваної області або біля окремих її меж [6].

Потрібно провести тріангуляцію області Ω . При цьому на саму область повинно накладатись якомога менше обмежень стосовно опуклості, вирізів, підобластей.

Дану задачу будемо називати задачею адаптивного тріангулювання із згущенням заданої області.

Статтю представив проф., д.т.н. Гарашенко Ф.Г.

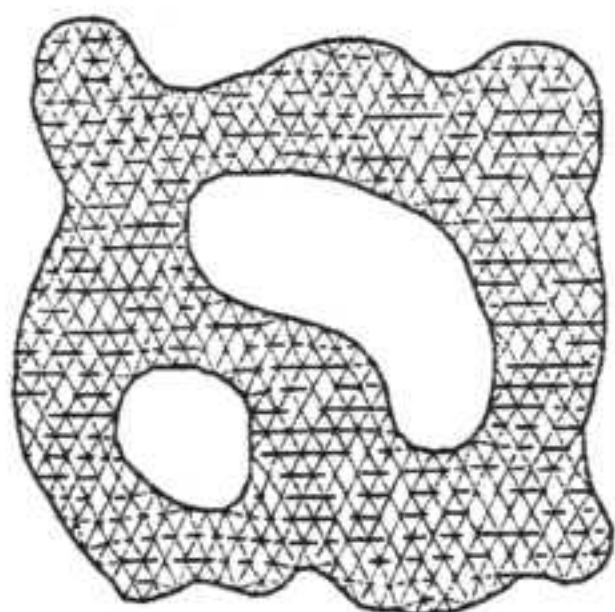


Рис.1. Триангуляція багатозв'язної області.

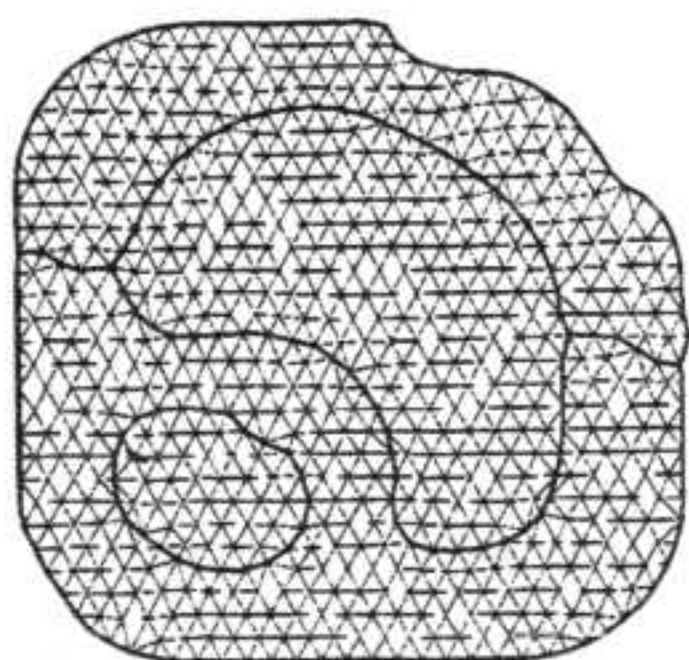


Рис.2. Триангуляція складеної області.

Аналогічна задача триангулювання заданої області з використанням рівномірного покриття розглянута в роботі [4].

Введемо критерій, що дозволяє оцінити якість триангулювання, тобто який показує, наскільки трикутники, що покривають область, наближені до рівносторонніх

$$R = \left(\frac{\pi}{3} - \alpha\right)^2 + \left(\frac{\pi}{3} - \beta\right)^2 + \left(\frac{\pi}{3} - \gamma\right)^2, \quad (2)$$

де α, β, γ – кути трикутника, a, b, c – відповідно його сторони (рис. 3).

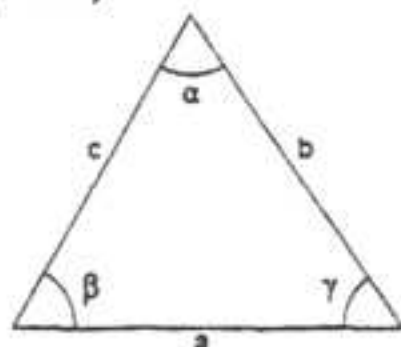


Рис. 3. Трикутний елемент.

Знаючи довжини сторін трикутника, легко можна обчислити його кути за допомогою наступних формул

$$\alpha = \arccos\left(\frac{c^2 + b^2 - a^2}{2bc}\right)$$

$$\beta = \arccos\left(\frac{a^2 + c^2 - b^2}{2ac}\right)$$

$$\gamma = \arccos\left(\frac{a^2 + b^2 - c^2}{2ab}\right)$$

Для того, щоб перевірити, чи площа трикутника не перевищує максимально допустимої, використаємо простий критерій

$$S_T \leq \min_{\Omega_T} S, \quad (3)$$

де S_T – площа трикутного елемента T ; Ω_T – частина області Ω , що належить трикутному елементу T ; S – функція густини покриття (1).

Поставимо у відповідність області Ω , що складається з окремих підобластей, функцію такого вигляду:

$$f(x, y) = \begin{cases} 0, & \text{якщо } M(x, y) \notin \Omega, \\ N, & \text{якщо } M(x, y) \in \Omega_N, \end{cases}$$

де Ω_N , $N = \overline{1, m}$ – підобласті області Ω ,

$$\text{причому } \Omega = \bigcup_{N=1}^m \Omega_N, \quad \bigcap_{N=1}^m \Omega_N = \emptyset, \quad M(x, y) -$$

точка підобласті Ω_N .

Перевагою такого задання функції в даній постановці задачі є те, що ми можемо триангулювати довільні об'єкти. Задавати об'єкти можна за допомогою зображень, отриманих під час аеро- чи фотозйомки, а також задаючи межі підобластей за допомогою аналітичних функцій.

У випадку багатозв'язної області Ω , заданої обмежувачими кривими $\Gamma_i, i = \overline{1, n}$ функція $f(x, y)$ набуде вигляду

$$f(x, y) = \begin{cases} 1, & \text{якщо } \Gamma_1(x, y) \leq 0, \\ 2, & \text{якщо } \Gamma_2(x, y) \leq 0, \\ \dots & \\ 0, & \text{інакше.} \end{cases}$$

Розглянемо приклад співставлення функції $f(x, y)$ для випадку задання поверхні біметалічної деталі, зображеної на рис. 4.

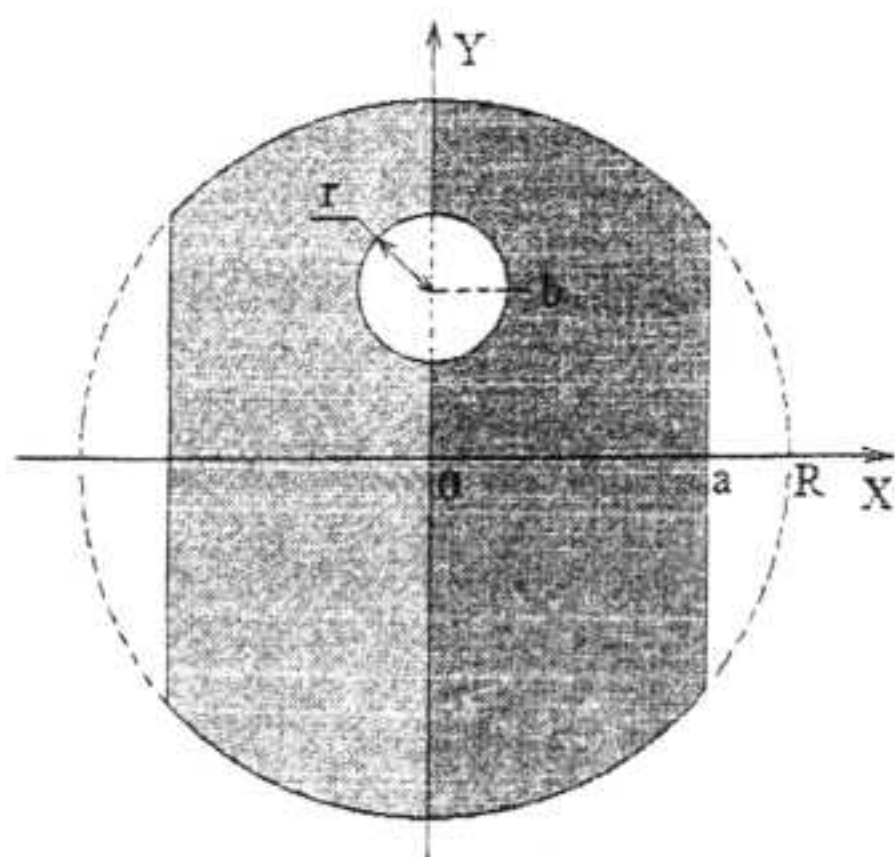


Рис.4. Двохзв'язна область, складена із 2-х підобластей.

$$f(x, y) = \begin{cases} 1, & \text{якщо } (x < 0) \cap (x \geq -a) \cap (x^2 + y^2 \leq R^2) \cap \\ & \cap (x^2 + (y - b)^2 \geq r^2), \\ 2, & \text{якщо } (x \geq 0) \cap (x \leq a) \cap (x^2 + y^2 \leq R^2) \cap \\ & \cap (x^2 + (y - b)^2 \geq r^2), \\ 0, & \text{інакше.} \end{cases}$$

В більшості випадків, програмні додатки, що виконують триангуляцію, - це комплекси для дослідження процесів у певних областях. Їх основне призначення - розв'язання задач математичної фізики, а триангуляція є лише складовою частиною їх математичних моделей. Це не дозволяє використовувати результати триангуляції, отримані за допомогою інших програмних комплексів, порівнювати методи, оснований на триангуляції, використати дещо іншу методику МСЕ. Взагалі, найчастіше результатом роботи таких комплексів є числові дані певної величини, що характеризують той чи інший процес, який перебігає в даній області.

В зв'язку з цим, постає проблема збереження результатів триангуляції для подальшого їх використання.

Питання триангуляції і комплекс суміжних із нею проблем, як вже відмічалось раніше, мають дуже багато способів вирішення. Вибір алгоритму триангуляції не завжди строго регламентує методи чисельних обрахунків, типи використовуваних даних, їх структуру. Проектування структури даних, наявних інтерфейсів, повністю визначає подальший процес розробки програмного додатку. Тому проектуванню варто приділити особливу увагу. У

зв'язку з цим, проведемо повний аналіз описаного нижче алгоритму.

3. Розробка і аналіз алгоритму триангуляції. Суть пропонованого алгоритму триангуляції полягає в наступному. Задана область Ω покривається регулярною трикутною сіткою із деякими, наперед визначеними розмірами трикутників (рис. 5). Регулярна трикутна сітка являє собою сітку, побудовану з правильних трикутників. Таке початкове покриття є досить грубим, тому потребує додаткового аналізу, в результаті якого виявляються всі елементи, які несуть в собі дефекти сітки (наприклад, коли розміри початково вибраних трикутників значно перевищують розміри вирізів або підобластей даної області).

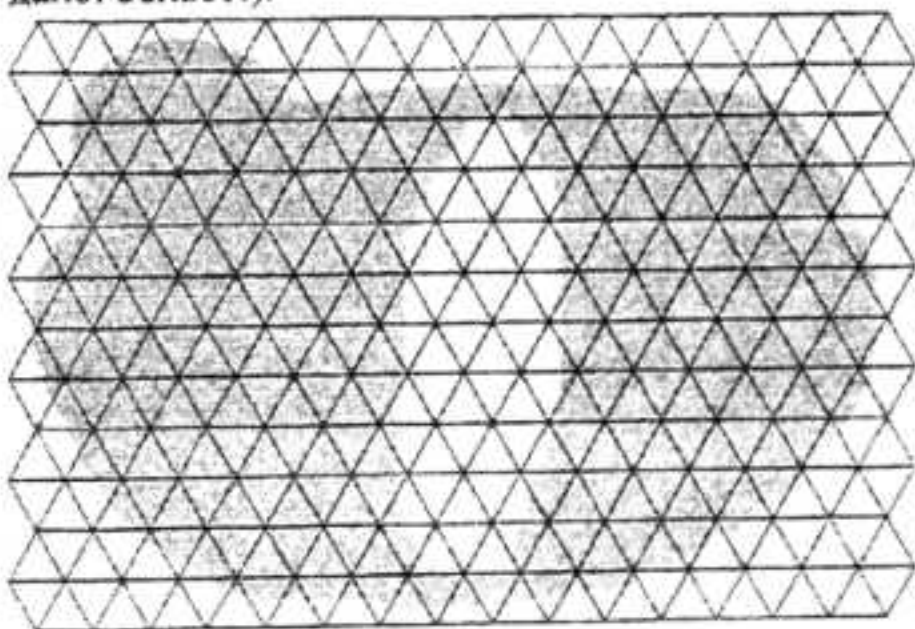


Рис.5. Регулярна трикутна сітка, накладена на задану область.

Завершальна обробка сітки передбачає згладжування і одночасно вирівнювання елементів по межі області. Розглянемо кожен з цих етапів детально.

Здійснюємо прорахунок вузлів регулярної сітки згідно деякого алгоритму. Програмний прорахунок вузлів сітки не викликає особливих труднощів. Після накладання такої сітки на певну область необхідно відкинути так звані "зайві" елементи - трикутники, які повністю виходять за межі заданої області. Критерій відбору елементів можна вибрати такий: для покриття області залишаються тільки ті трикутники, у яких хоча б одна з вершин належить розглядуваній області.

Зокрема, для двозв'язної кільцевої області (області з вирізом) початкове накладання має вигляд, зображений на рис. 5. Після відкидання "зайвих" трикутників, отримуємо адаптовану регулярну сітку (сітку без згущення) до заданої області (рис. 6).

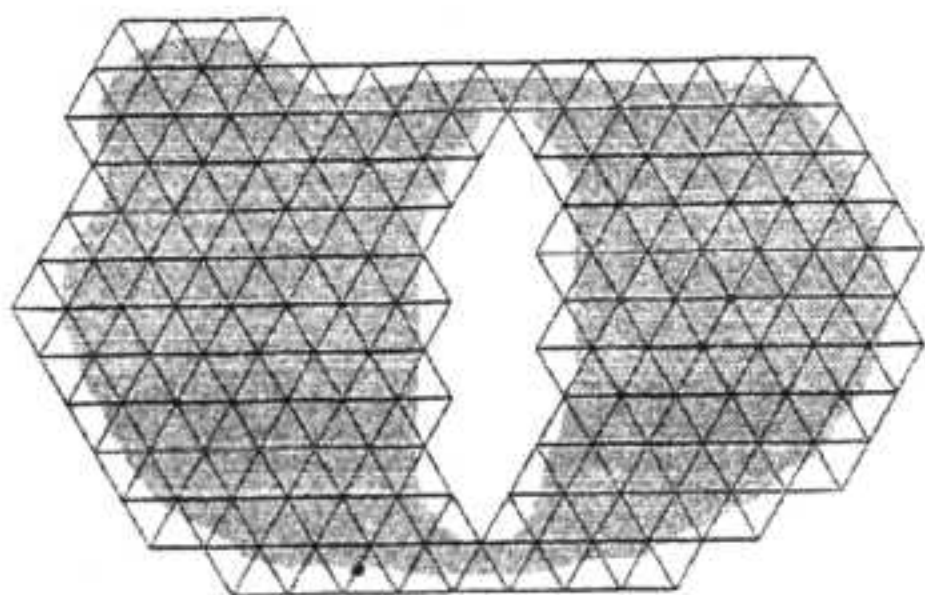


Рис.6. Адаптація регулярної сітки до заданої області.

Дану процедуру не можна назвати завершеним розбиттям області, оскільки отримана таким чином сітка потребує корекції. Можливі випадки, коли деякі трикутники будуть містити в собі дефекти, тобто сітка з такими елементами не буде придатна для подальшого використання в МСЕ. Це може бути через ряд причин.

Можливий випадок, коли область містить в собі відносно малі вирізи чи підобласті (порівняно з трикутними елементами початкового покриття). Тоді деякі трикутники, розміри яких більші за вирізи чи підобласті, будуть накладатися на них або містити їх в собі.

Одним із способів вирішення цієї проблеми є згущення сітки в місцях дефектів. Згущення являє собою поділ вже існуючих трикутників на трикутники менших розмірів. Програмна реалізація цього завдання не є такою тривіальною, як може здатися на перший погляд, оскільки поділ трикутного елемента впливає на всі суміжні елементи (рис. 7), що особливо відчутно при одночасному поділі багатьох трикутників (рис. 8).

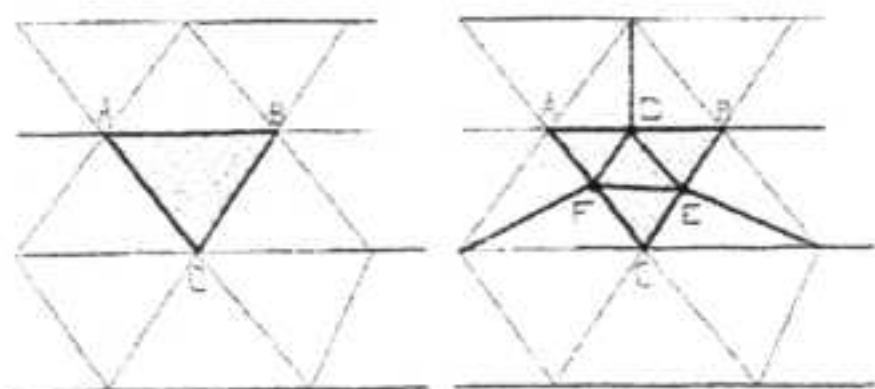


Рис.7. Найпростіший випадок розбиття трикутника.

При подібненні трикутника ABC (рис. 7) на сторонах AB, BC, CA утворюються нові вершини – D, E, F, координати яких обчислюються за формулами

$$\bar{D} = \frac{\bar{A} + \bar{B}}{2}, \quad \bar{E} = \frac{\bar{B} + \bar{C}}{2}, \quad \bar{F} = \frac{\bar{C} + \bar{A}}{2},$$

де $\bar{A}, \bar{B}, \bar{C}, \bar{D}, \bar{E}, \bar{F}$ – координати відповідних їм вузлів сітки.

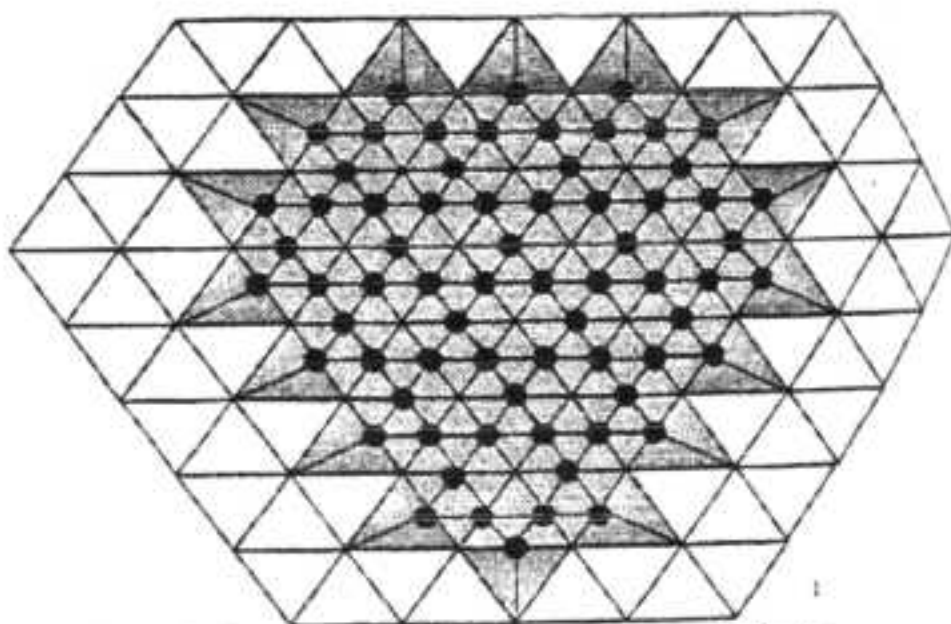


Рис.8. Розбиття частини трикутної сітки.

Проведемо згущення за наступним алгоритмом:

1. Сформуємо множину X всіх трикутних елементів, що не задовільняють критерій (3) або містять вирізи чи підобласті.
2. Сформуємо множину Y всіх трикутних елементів, що не увійшли в множину X , але є суміжними з елементами множини X .
3. Розбиваємо всі трикутні елементи множини X на чотири нові трикутні елементи (так як зображено на рис. 7 для трикутника ABC).
4. Розбиваємо трикутні елементи з множини Y на два елементи поділом сторони, що суміжна з елементом множини X (так як зображено на рис. 7 для суміжних до A трикутників).
5. Наступним важливим етапом в алгоритмі тріангуляції є згладжування сітки та її адаптація до межі області, суть якої полягає в ітераційному усередненні координат вершин із сусідніми.

Нові координати вузлів сітки обраховуємо за формулою [5]

$$X_i = \alpha X_i + (1 - \alpha) \bar{X}_i, \quad (4)$$

де $\bar{X}_i$ – точка, координати якої обчислюються як середнє арифметичне координат всіх суміжних вузлів до вузла i , $\alpha \in [0, 1]$. Причому для вузлів, що належать області – $\alpha = 0$, а для тих, що не належать області – α близьке до нуля, але вузол має залишатися за межами області. Ітеративно проводимо обчислення до досягнення заданої точності.

На етапі покращення сітки розглядаємо всі пари суміжних трикутників і перевіряємо, чи поворот

спільного ребра дозволить отримати покращену сітку, яка більш відповідає заданому критерію (2), тобто $R_{\Delta ABD} + R_{\Delta BCD} > R_{\Delta ABC} + R_{\Delta ACD}$.

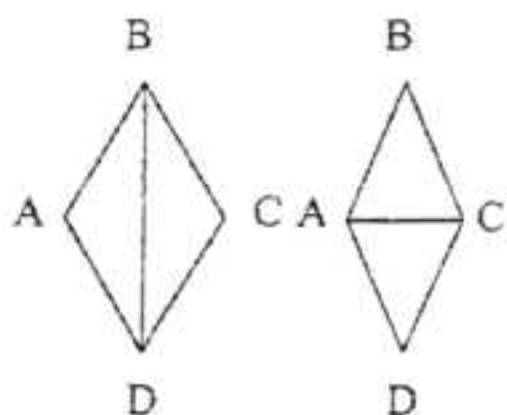


Рис.10. Приклад пари суміжних трикутників до і після повороту спільного ребра.

З рис.10. видно, що після повороту спільного ребра BD трикутників ABD і BCD отримаємо трикутники ABC і ADC, які значно краще наближені до рівносторонніх і значно краще задовільняють заданий критерій (3), ніж попередні.

Крім того, передбачено, що результуюча сітка повинна задовільняти критерій Делоне [4], тому на завершення алгоритму проводиться перевірка виконання цього критерію.

Якість триангуляції оцінюється за мінімальним кутом для всіх трикутників.

4. Програмна реалізація алгоритму. Автоматизована система адаптивного триангулювання із згущенням багатозв'язних та складених областей розроблена в сучасному середовищі програмування Visual Studio 2005 з використанням об'єктно-орієнтованого підходу.

У програмному додатку реалізовано:

- діалоговий режим задання конфігурації області;
- задання фізичних властивостей матеріалів підобластей;
- передбачена можливість ручного контролю параметрів триангуляції;
- підтримка повністю автоматичного режиму триангуляції;
- експорт згенерованої сітки та фізичних параметрів у файл або обчислювальний модуль;
- експорт зображення області та згенерованої сітки в графічний файл;
- експертний режим візуального контролю якості згенерованої сітки.

Значна увага приділена сервісному оформленню програми у вигляді систем меню, панелі керування, наглядному візуальному представленню підобластей та згенерованої сітки.

4.1. Об'єктна модель. Детальний аналіз алгоритму показав, що для його реалізації

необхідно вдало підібрати структури даних. Вони повинні бути достатньо гнучкими для виконання всіх попередньо описаних кроків, без написання громіздких методів.

Виходячи із принципів об'єктно-орієнтованого програмування, організуємо роботу алгоритму за допомогою класів та об'єктів.

Основні класи та інтерфейси розробленої автоматизованої системи наступні:

IArea – Інтерфейс для задання областей;

PhotoArea – Клас для роботи з областями заданими фотографічно;

FunctionalArea – Клас для роботи з областями заданими функціями;

Mesh – клас для представлення сітки скінченних елементів;

Edge – клас, що представляє ребро скінченного елемента;

Point – точка, вузол;

FiniteElement – скінченний елемент;

Material – клас для опису матеріалів;

Boundary – клас для представлення меж підобластей;

ICriteria – інтерфейс для критеріїв;

DeloneCriteria – клас критеріїв Делоне;

AreaCriteria – клас критеріїв площі;

MinimumAngleCriteria – клас критеріїв мінімального кута;

MeshRefinement – клас для покращення (оптимізації) сітки;

MeshDensity – клас для опису очікуваної густини розбиття;

MeshBlur – клас для згладжування сітки;

XmlMeshExporter – клас для експорту сітки в іншу програму, файл, тощо.

5. Результати числових експериментів.

На рис. 11-14 візуально представлені основні кроки алгоритму триангуляції заданої області.

Завантаження області триангуляції і початкове накладання регулярної сітки зображене на рис.11.

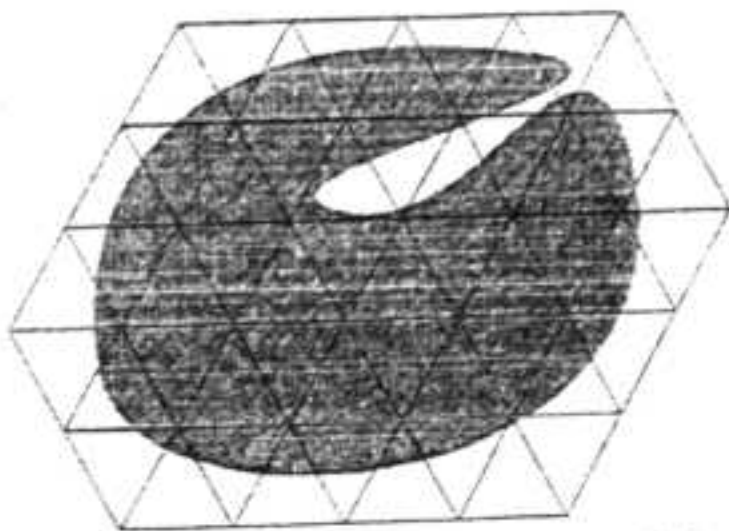


Рис.11. Накладання регулярної трикутної сітки на задану область.

Після видалення дефектів сітки і згущення в місцях дефектів, отримуємо:

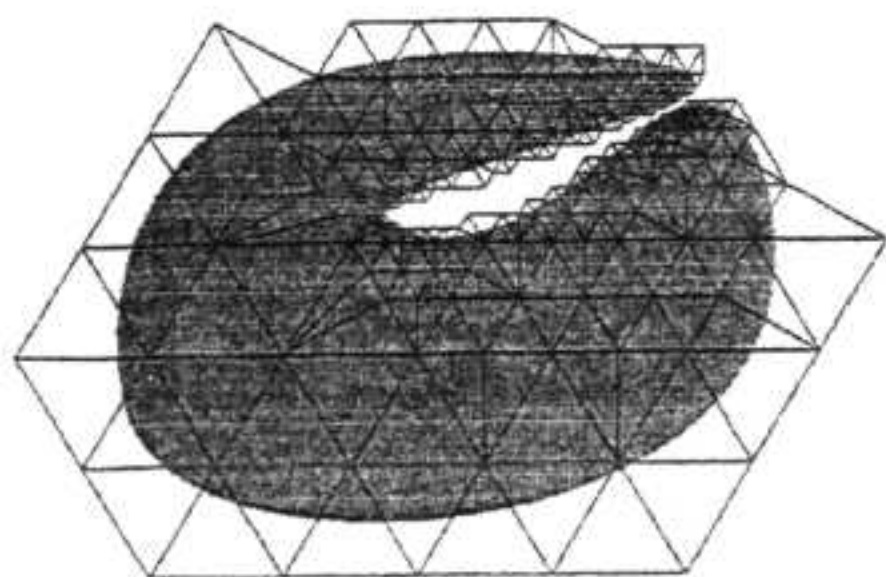


Рис. 12. Згущення регулярної сітки.

Після завершального згладжування сітка скінченних елементів набуває відповідного вигляду, зображеного на рис.13.

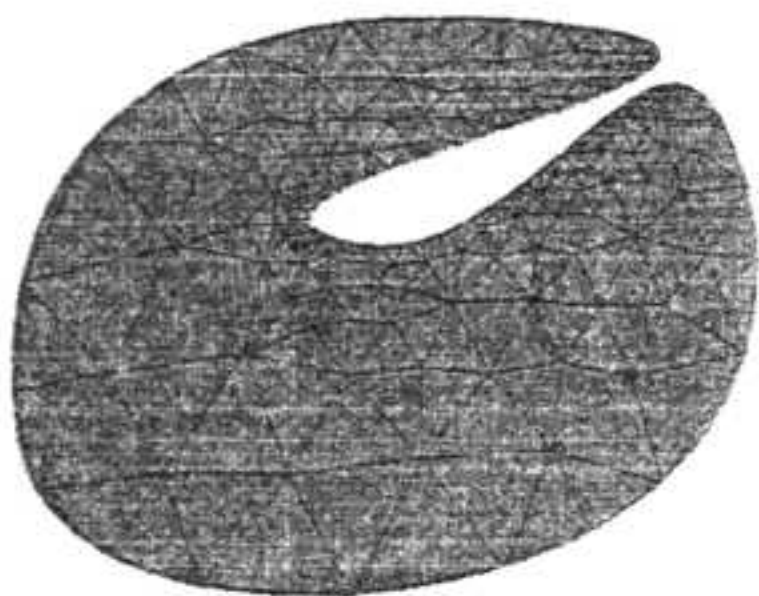


Рис. 13. Результати згладжування та адаптації трикутної сітки до межі заданої області.

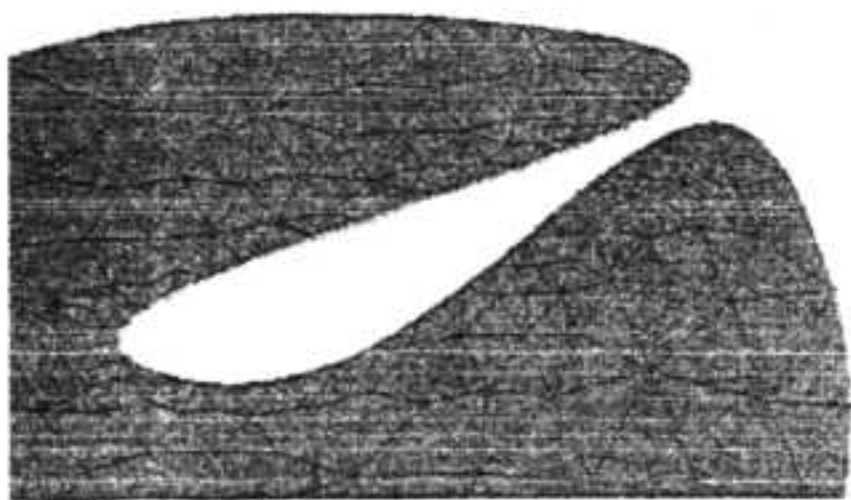


Рис. 14. Згущення дрібних скінченних елементів (в збільшеному вигляді).

На рис.14 у збільшеному вигляді показано фрагмент розглядуваної області, що покритий сіткою трикутних елементів, адаптованої до межі області із згущенням.

6. Висновки. Таким чином, в даній роботі розроблено алгоритм та програмний комплекс адаптивного триангулювання із згущенням багатозв'язних, складених та багатозв'язно-складених областей з криволінійними межами.

Даний алгоритм генерації сітки заданої області може бути успішно використаний при розв'язанні багатьох задач математичної фізики МСЕ та для розв'язання багатьох інших прикладних задач [1, 7].

Автор вдячний проф. Власюку А.П. за постановку задачі та цінні поради при написанні роботи.

Список використаних джерел

1. Сергиенко И.В., Скопецкий В.В., Дейнека В.С. Математическое моделирование и исследование процессов в неоднородных средах. - Киев: Наук. думка, 1991. - 432 с.
2. Зенкевич О., Морган К., Конечные элементы и аппроксимация. - М.: Мир, 1986. - 318 с.
3. Стренг Г., Фикс Дж. Теория метода конечных элементов - М.: Мир, 1989. - 512 с.
4. Скворцов А.В. Триангуляция Делоне и ее применение.- Томск; изд-во Томского университета, 2002. - 128 с.
5. Власюк А.П., Філатов М.С. Автоматизована система скінченно-елементного покриття багатозв'язних та складених областей//Вісник Національного університету водного господарства та природокористування. Вип. 4(36), ч.2. - Рівне, 2006. - С.68-73.
6. Філатов М.С. Автоматизована система адаптивного триангулювання багатозв'язних та складених областей // Тез. Міжнар. школи-семінару «Прогнозування та прийняття рішень в умовах невизначеності». Тези доп.-Новий світ, вересень 2007. - С.129-130.
7. Сегерлинд Л.Д. Применение метода конечных элементов. - М.: Мир, 1979. - 393 с.

Надійшла до редколегії 31.01.08.

УДК 627.324.2/3:532.546

Філатова І.А.*

Математичне моделювання напружено-деформованого стану багат шарового ґрунтового масиву при наявності рівня ґрунтових вод

Проведено математичне моделювання напружено-деформованого стану багат шарового ґрунтового масиву при наявності в ньому рівня ґрунтових вод. В результаті розв'язку відповідної крайової задачі отримано переміщення, напруження, деформації та величину осідання ґрунтового масиву. Проведено числові експерименти для конкретних даних.

Ключові слова: математичне моделювання, ґрунтовий масив, напружено-деформований стан, вільна поверхня, переміщення, деформація, напруження, осідання.

* E-mail: ifilatova@gmail.com

1. Вступ. У зв'язку з розвитком промислового та цивільного будівництва зростає вплив техногенних факторів на навколишнє середовище. При прогнозуванні довговічності служіння різних об'єктів, визначення їх стійкості та надійності важливу інформацію надають знання напружено-деформованого стану (НДС) основ споруд, на яких вони побудовані.

Важливий вплив на стійкість споруд має положення рівня ґрунтових вод (вільної поверхні). Зміна положення рівня ґрунтових вод може призвести до зміни НДС ґрунтових основ об'єктів і врешті до негативних явищ, які відбуваються в основах цих споруд.

Так, наприклад, фільтраційні деформації ґрунту внаслідок інтенсивної відкачки із свердловин, призводили до утворення тріщин земної поверхні, які досягали значних розмірів.

Пониження вільної поверхні ґрунтових вод виникає внаслідок зменшення інфільтрації опадів у ґрунт, забудови і асфальтування території, дренажу товщі ґрунту, каналізаційними колекторами, різноманітними підземними комунікаціями, лініями метрополітену тощо.

Inna A. Filatova.*

Mathematical modelling of tensely deformed condition of a multispherical soil file at presence of a level of subsoil waters

Mathematical modelling tensely deformed condition of a multispherical soil file is lead at presence of a level of subsoil waters. As a result of the decision of a respective regional problem have received moving, a pressure, deformations and size of subsidence of a soil file. It is lead numerical experiments for concrete data.

Key Words: mathematical modelling, soil file, tensely deformed condition, free surface, moving, deformation, pressure, subsidence.

Рух вільної поверхні ґрунтових вод веде до виникнення різноманітних об'ємних сил і зміни напружено-деформованого стану ґрунтового масиву [1-3]. Про важливість даної проблеми свідчить статистика аварій споруд, пов'язаних із порушенням несучої здатності їх ґрунтових основ. Найчастіше вони відбувалися при зміні напружено-деформованого стану ґрунтових масивів внаслідок дії техногенних факторів. При прогнозуванні довговічності експлуатації різних об'єктів, визначення їх стійкості та надійності необхідно знати напружено-деформований стан основ споруд. Тому важливої уваги заслуговує задача визначення напружено-деформованого стану ґрунтових масивів з урахуванням зміни рівня ґрунтових вод [4].

2. Постановка задачі. Нехай дано масив ґрунту, що розміщений на скельній основі і складається з n горизонтальних шарів, в одному з яких присутня вільна поверхня - рівень ґрунтових вод (РГВ). Шари ґрунту вважаються пружно-деформованими в межах лінійної теорії пружності [1-3] з різними пружними сталими (коефіцієнтами Ламе) $\lambda_i, \mu_i, i = \overline{1, n}$, на які діє лише сила тяжіння, а у випадку змоченого ґрунту враховується ще й архімедова сила.

Статтю представив проф. д-р.н. Гарашенко О.Г.

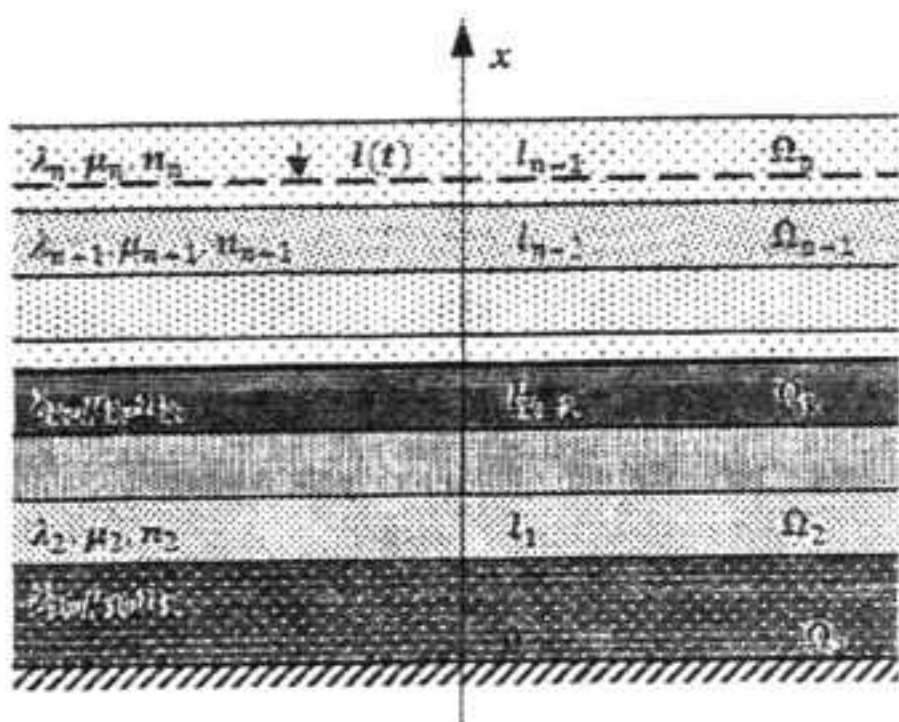


Рис.1 Багатошаровий масив ґрунту при наявності в ньому рівня ґрунтових вод.

Кожен з шарів ґрунту характеризується пружними сталими $\lambda_i, \mu_i, i = \overline{1, n}$, пористостями $n_i, i = \overline{1, n}$ та питомими вагами ґрунтів $\gamma_i, i = \overline{1, n}$.

Нехай вільна поверхня РГВ знаходиться на рівні l_k в k -му шарі ґрунту. Тоді шари ґрунту, розміщені вище l_k , знаходяться в природному стані, а нижче l_k – в зваженому стані (при повній вологості). Вільна поверхня РГВ вважається нерухомою.

Потрібно розрахувати НДС в шарах ґрунту, які знаходяться в природному стані та в шарах ґрунту нижче РГВ, які знаходяться в зваженому стані, при умові, що відсутнє переміщення нижньої та верхньої меж ґрунту або лише нижньої межі ґрунту (скельної основи).

3. Математична модель задачі. На основі вищесказаного, НДС в шарах ґрунту в природному стані та тих, що знаходяться нижче РГВ, у стаціонарному одновимірному випадку описується рівнянням рівноваги лінійної теорії пружності [1,3,4]

$$(\lambda_i + 2\mu_i) \frac{d^2 u_i}{dx^2} = X_i, i = \overline{1, n}, \quad (1)$$

де $x \in \bigcup_{i=1}^n (l_{i-1}, l_i), l_0 = 0, l_n = l$.

Тут $u_i, i = \overline{1, n}$ – переміщення вздовж осі ОХ в шарах, що знаходяться відповідно в

зваженому $\left(x \in \bigcup_{i=1}^k (l_{i-1}, l_i) \right)$ та в природному

$\left(x \in \bigcup_{i=k+1}^n (l_{i-1}, l_i) \right)$ стані;

$$X_i = \begin{cases} \gamma_{ж,i}, i = \overline{1, k}, \\ \gamma_{пр,i}, i = \overline{k+1, n}, \end{cases} \quad (2)$$

Тут $\gamma_{ж,i}, \gamma_{пр,i}$ – питомі ваги ґрунтів в зваженому та природному стані, причому

$$\gamma_{ж,i} = \gamma_{нас,i} - \gamma_p, i = \overline{1, k}, \quad (3)$$

де $\gamma_{нас,i}, i = \overline{1, k}$ – питомі ваги ґрунтів в насиченому стані, які обчислюються за формулою

$$\gamma_{нас,i} = \gamma_{г,i} + n_i \cdot \gamma_p, i = \overline{1, k}, \quad (4)$$

де $\gamma_{г,i}, i = \overline{1, k}$ – питомі ваги ґрунтів в сухому стані, γ_p – питома вага рідини, $\gamma_{пр,i}$ – питомі ваги ґрунтів в природному стані, причому

$$\gamma_{пр,i} = \rho_{пр,i} \cdot g, \quad (5)$$

де $\rho_{пр,i}$ – густини ґрунтів у природному стані, g – прискорення вільного падіння.

Підставляючи значення (4) в (3), отримаємо

$$\gamma_{ж,i} = \gamma_{г,i} - (1 - n_i) \gamma_p, i = \overline{1, k}. \quad (6)$$

Крайові умови для переміщень мають вигляд

$$u_i(0) = 0, \quad (7)$$

$$u_n(l) = 0 \text{ або } \frac{du_n(l)}{dx} = 0. \quad (8)$$

Умови спряження записуються так:

$$u_i(l_i) = u_{i+1}(l_i), i = \overline{1, n-1}, \quad (9)$$

$$(\lambda_i + 2\mu_i) \frac{du_i(l_i)}{dx} = (\lambda_{i+1} + 2\mu_{i+1}) \frac{du_{i+1}(l_i)}{dx} \quad (10)$$

$i = \overline{1, n-1}.$

Отже, НДС поставленої вище задачі описується математичною моделлю (1)-(10).

У випадку переміщень верхньої межі ґрунту врахуємо, що

$$l \leq l_0, l(t) - l_0 = u_n(l(t)), l(0) = l_0.$$

Перейдемо до безрозмірних величин $\bar{x}, \bar{l}, \bar{l}_i, \bar{u}_i, i = \overline{1, n}$, згідно формул

$$\bar{x} = \frac{x}{l_0}, \bar{u}_i = \frac{u_i}{l_0}, i = \overline{1, n}, \bar{l} = \frac{l}{l_0}, \bar{l}_i = \frac{l_i}{l_0}, i = \overline{1, n-1},$$

де $\bar{l} < 1$ для будь-якого $t > 0$.

Тоді у безрозмірних змінних математична модель НДС ґрунту матиме вигляд (надалі рисочки над безрозмірними змінними для простоти опущені)

$$\frac{d^2 u_i}{dx^2} = a_i, \quad i = \overline{1, n}, \quad (11)$$

$$u_1(0) = 0, \quad (12)$$

$$u_i(l_i) = u_{i+1}(l_i), \quad i = \overline{1, n-1}, \quad (13)$$

$$(\lambda_i + 2\mu_i) \frac{du_i(l_i)}{dx} = (\lambda_{i+1} + 2\mu_{i+1}) \frac{du_{i+1}(l_i)}{dx}, \quad (14)$$

$$i = \overline{1, n-1},$$

$$u_n(1) = 0 \text{ або } \frac{du_n(l)}{dx} = 0, \quad (15)$$

$$\text{де } a_i = \frac{X_i}{\lambda_i + 2\mu_i}, \quad i = \overline{1, n}.$$

4. Розв'язок задачі. Загальний розв'язок крайової задачі (11)-(15) має вигляд

$$u_i(x) = \frac{a_i x^2}{2} + c_{2i-1}x + c_{2i}, \quad i = \overline{1, n}, \quad (16)$$

де $c_i, i = \overline{1, 2n}$ - невідомі сталі.

Для визначення невідомих сталих $c_i, i = \overline{1, 2n}$ використаємо крайові умови (12), (15) та умови спряження (13), (14).

Так наприклад, у випадку задання на верхній межі ґрунту граничної умови 1-го роду, маємо наступні співвідношення для визначення невідомих коефіцієнтів в (16):

$$\text{із умови (12) випливає, що} \quad c_2 = 0; \quad (17)$$

враховуючи, що $u_n(1) = 0$, маємо

$$c_{2n-1} + c_{2n} = -\frac{a_n}{2}; \quad (18)$$

на основі умов спряження (13), (14),

отримаємо

$$l_i(c_{2i-1} - c_{2i+1}) + c_{2i} - c_{2i-2} = 0, 5l_i^2(a_{i+1} - a_i), \quad (19)$$

$$i = \overline{1, n-1},$$

$$\begin{aligned} (\lambda_i + 2\mu_i)(a_i l_i + c_{2i-1}) = \\ (\lambda_{i+1} + 2\mu_{i+1})(a_{i+1} l_i + c_{2i+1}), \quad i = \overline{1, n-1}. \end{aligned} \quad (20)$$

Як відомо [1-3], величина деформації обчислюється за формулою

$$\varepsilon = \frac{du}{dx}$$

і згідно (16) набуде вигляду

$$\varepsilon_i(x) = a_i x + c_{2i-1}, \quad i = \overline{1, n}. \quad (21)$$

Напруження знаходяться згідно формули

$$\sigma = (\lambda + 2\mu) \frac{du}{dx}$$

і в даному випадку матимуть вигляд

$$\sigma_i(x) = (\lambda_i + 2\mu_i)(a_i x + c_{2i-1}), \quad i = \overline{1, n}. \quad (22)$$

Таким чином, задача НДС в шарі ґрунту при наявності в ньому РГВ у випадку задання на верхній межі ґрунту граничної умови 1-го роду розв'язана повністю і дається формулами (16)-(22).

Розглянемо випадок задання на верхній межі ґрунту граничної умови 2-го роду. Загальний розв'язок задачі має вигляд (16), причому невідомі константи обчислюються за формулами

$$c_2 = 0, \quad (23)$$

$$c_{2n-1} = -a_n l, \quad (24)$$

$$c_{2i-1} = \frac{\lambda_{i+1} + 2\mu_{i+1}}{\lambda_i + 2\mu_i} (a_{i+1} l_i + c_{2i+1}) - a_i l_i, \quad (25)$$

$$i = \overline{n-1, 1},$$

$$c_{2i+2} = \frac{(a_i - a_{i+1})l_i^2}{2} + (c_{2i-1} - c_{2i+1})l_i + c_{2i}, \quad (26)$$

$$i = \overline{1, n-1}.$$

Величини деформацій і напружень в шарах ґрунту обчислюються за формулами (21), (22) з врахуванням (24)-(26).

Зазначимо, що у зв'язку із тим, що відбувається переміщення верхньої межі ґрунту, то спостерігається її осадка, яка обчислюється, як відомо, згідно формули

$$l(t) - l_n = \int_0^{l(t)} \varepsilon(x(t)) dx = u_n(l(t)). \quad (27)$$

Використовуючи формулу (16), маємо

$$u_n(l(t)) = \frac{a_n l^2(t)}{2} + c_{2n-1} l(t) + c_{2n}. \quad (28)$$

Обрахувавши значення коефіцієнтів c_{2n-1}, c_{2n} за допомогою формул (23)-(26) та підставивши отримані результати у (28), (27), після деяких перетворень, отримаємо

$$l(t) = \frac{-b_1 + \sqrt{b_1^2 - 2a_n b_2}}{a_n}, \quad (29)$$

де b_1 і b_2 - деякі відомі сталі.

Як приклад, розглянемо випадок двох шарів ґрунту при наявності РГВ в одному з них (наприклад, нижньому). В даному випадку $n=3$ та руху верхньої межі розв'язок задачі прийме вигляд

$$\begin{aligned} u_1(x) &= \frac{a_1 x^2}{2} + c_1 x + c_2, \quad 0 < x < l_1, \\ u_2(x) &= \frac{a_2 x^2}{2} + c_3 x + c_4, \quad l_1 < x < l_2, \\ u_3(x) &= \frac{a_3 x^2}{2} + c_5 x + c_6, \quad l_2 < x < l, \end{aligned} \quad (30)$$

де $c_2 = 0$, $c_5 = -a_3 l$,

$$c_{2i-1} = \frac{\lambda_{i+1} + 2\mu_{i+1}}{\lambda_i + 2\mu_i} (a_{i+1} l_i + c_{2i+1}) - a_i l_i, \quad (31)$$

$i = \overline{2, 1}$

$$c_{2i+2} = \frac{(a_i - a_{i+1}) l_i^2}{2} + (c_{2i-1} - c_{2i+1}) l_i + c_{2i}, \quad (32)$$

$i = \overline{1, 2}$.

На основі (28) при $n=3$ знаходимо величину осідання ґрунту

$$u_3(l(t)) = \frac{a_3 l^2(t)}{2} + c_5 l(t) + c_6. \quad (33)$$

$$\text{де } l(t) = (-b_1 + \sqrt{b_1^2 - 2a_n b_1}) / a_n, \quad (34)$$

$$b_1 = \frac{\lambda_3 + 2\mu_3}{\lambda_1 + 2\mu_1} a_3 l_1 + \frac{\lambda_3 + 2\mu_3}{\lambda_2 + 2\mu_2} (a_3 l_2 - a_3 l_1) - a_3 l_2 + 1, \quad (35)$$

$$\begin{aligned} b_2 &= \frac{a_1 + a_2}{2} l_1^2 + \frac{a_2 + a_3}{2} l_2^2 - a_2 l_1 l_2 + \\ &+ \frac{\lambda_2 + 2\mu_2}{\lambda_1 + 2\mu_1} (a_2 l_1 l_2 - a_2 l_1^2) - l_0 - \\ &- \frac{\lambda_3 + 2\mu_3}{\lambda_1 + 2\mu_1} a_3 l_1 l_2 + \frac{\lambda_3 + 2\mu_3}{\lambda_2 + 2\mu_2} (a_3 l_1 l_2 - a_3 l_2^2). \end{aligned} \quad (36)$$

У випадку фільтрації з вільної поверхні крім (1)-(10) маємо

$$V_i = -k_i(x) \frac{dh_i}{dx}, \quad \frac{dV_i}{dx} = 0, \quad i = \overline{1, k}, \quad (37)$$

$$h_1(0) = H_1, \quad (38)$$

$$h_i(l_i) = h_{i+1}(l_i), \quad i = \overline{1, k-1}, \quad (39)$$

$$k_i(l_i) \frac{dh_i(l_i)}{dx} = k_i(l_i) \frac{dh_{i+1}(l_i)}{dx}, \quad i = \overline{1, k-1}, \quad (40)$$

$$h_n(l) = H_2. \quad (41)$$

Тут $V_i(x)$, $h_i(x)$, $k_i(x)$ - відповідно швидкість фільтрації, п'єзометричний напір та коефіцієнт фільтрації в i -му шарі, H_1 , H_2 - напори відповідно на нижній та верхній межах ґрунту.

Розв'язок задачі НДС при наявності фільтрації з вільної поверхні має вигляд у випадку граничних умов 2-го роду на верхній межі

$$u_i(x) = \frac{a_i x^2}{2} + a'_i \int_{l_{i-1}}^x \left(\int_{l_{i-1}}^z \frac{ds}{k(s)} \right) dz + c_{2i-1} x + c_{2i},$$

$i = \overline{1, k}$,

$$u_i(x) = \frac{a_i x^2}{2} + c_{2i-1} x + c_{2i}, \quad i = \overline{k+1, n}, \quad (42)$$

де

$$c_2 = 0, \quad (43)$$

$$c_{2n-1} = -a_n, \quad (44)$$

$$\begin{aligned} c_{2i-1} &= \frac{\lambda_{i+1} + 2\mu_{i+1}}{\lambda_i + 2\mu_i} (a_{i+1} l_i + c_{2i+1}) - \\ &- a_i l_i - a'_i \int_{l_{i-1}}^{l_i} \frac{dz}{k(z)}, \quad i = \overline{1, n-1}, \end{aligned} \quad (45)$$

$$\begin{aligned} c_{2i+2} &= \frac{a_i - a_{i+1}}{2} l_i^2 + a'_i \int_{l_{i-1}}^{l_i} \left(\int_{l_{i-1}}^z \frac{ds}{k(s)} \right) dz + \\ &+ l_i (c_{2i-1} - c_{2i+1}) + c_{2i}, \quad i = \overline{1, n-1}, \end{aligned} \quad (46)$$

$$a_i = \frac{\gamma_{\text{жл}} - \gamma_p}{\lambda_i + 2\mu_i}, \quad i = \overline{1, k}, \quad (47)$$

$$a'_i = \frac{\gamma_p (H_2 - H_1)}{(\lambda_i + 2\mu_i) \int_0^{l_i} \frac{ds}{k(s)}}, \quad i = \overline{k+1, n}. \quad (48)$$

5. Результати числових розрахунків.

Нижче наведені результати числових розрахунків для розглянутого випадку ($n=3$) при наступних вхідних даних: $\lambda_1 = 7500$ кг/м²; $\lambda_2 = 13500$ кг/м²; $\lambda_3 = 17000$ кг/м²; $\mu_1 = 5000$ кг/м²; $\mu_2 = 9000$ кг/м²; $\mu_3 = 11500$ кг/м²;

$\gamma_{ж.1} = 11,5 \text{ кг/м}^3$; $\gamma_{ж.2} = 10,5 \text{ кг/м}^3$; $\gamma_{пр.} = 17,0 \text{ кг/м}^3$; $l_1 = 3 \text{ м}$; $l_2 = 7 \text{ м}$; $l_3 = l = 10 \text{ м}$.

На рис.2, 4 зображені відповідно переміщення, а на рис.3,5 - напруження у випадку задання на верхній межі ґрунту граничної умови 1-го та 2-го роду в зважених та природному станах.

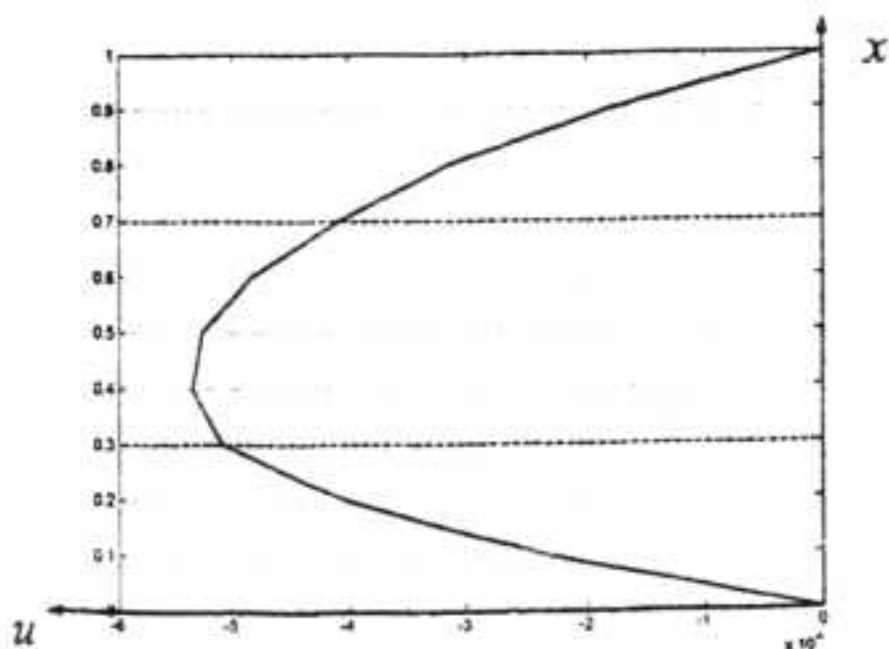


Рис.2. Графік переміщень по глибині ґрунту (I).

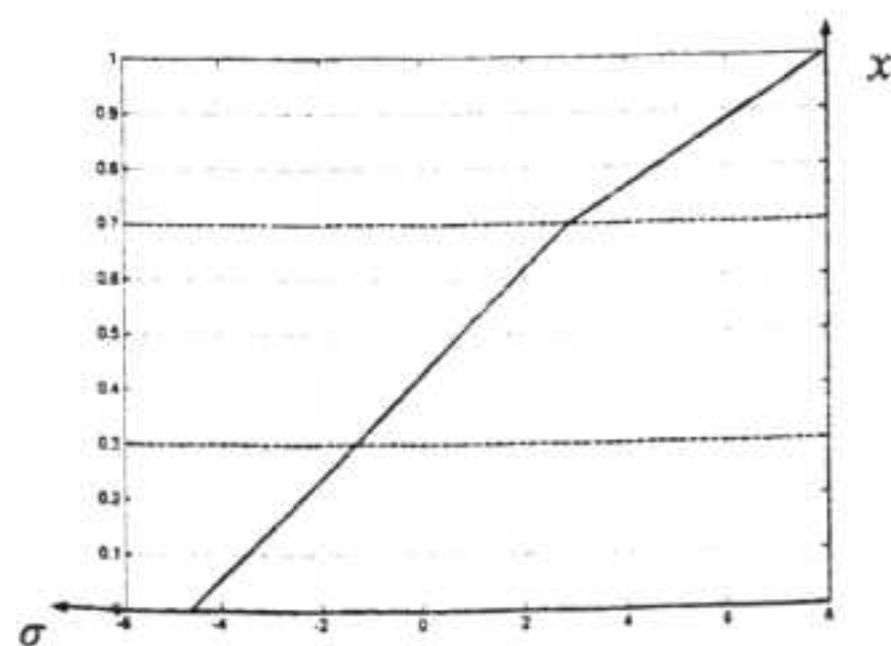


Рис. 3. Графік напружень по глибині ґрунту (I).

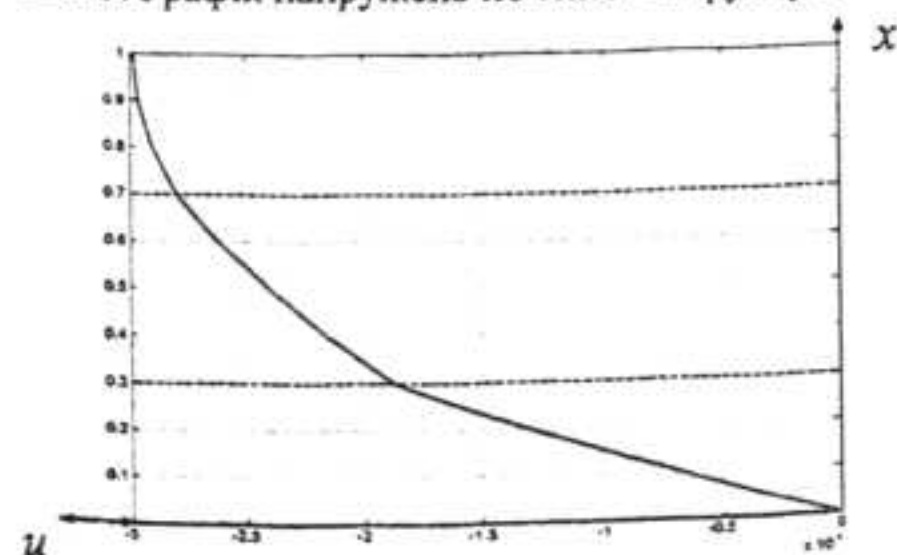


Рис.4. Графік переміщень по глибині ґрунту(II).

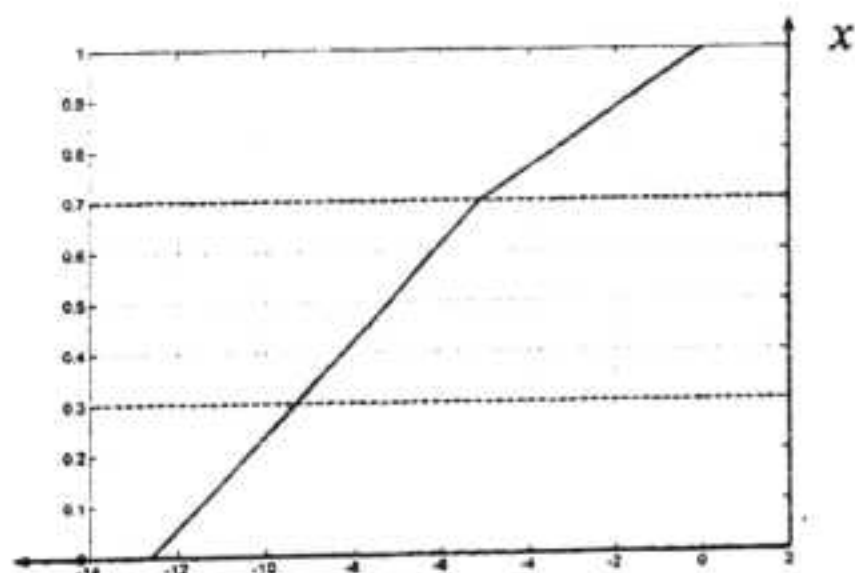


Рис.5. Графік напружень по глибині ґрунту (II).

6. Висновки. Таким чином, в роботі проведено математичне моделювання НДС багат шарового масиву ґрунту при наявності в ньому РГВ. Отримано аналітичні розв'язки відповідних крайових задач при відсутності переміщень нижньої межі та наявності (відсутності) переміщення верхньої межі ґрунту, на основі яких проведено числові розрахунки.

Отримані розв'язки дають можливість виконати оцінку напружено-деформованого стану ґрунтового масиву при наявності в ньому рівня ґрунтових вод. Вони можуть бути використані для оцінки стану ґрунтових основ гідротехнічних та енергетичних споруд.

Список використаних джерел

1. Сергиенко И.В., Скопешкий В.В., Дейнека В.С. Математическое моделирование и исследование процессов в неоднородных средах.-Киев: Наукова думка, 1991. – 432 с.
2. Флорин В.А. Основы механики грунтов. Т.2. Госстройиздат. 1961.-543 с.
3. Дейнека В.С., Сергиенко И.В., Скопешкий В.В. Модели и методы решения задач с условиями сопряжения. К.: Наукова думка, 1998.-615 с.
4. Кузло М.Т., Філатова І.А. Про деякі математичні моделі НДС ґрунтових масивів у процесі руху вільної поверхні ґрунтових вод //В зб.: вісник Національного університету водного господарства та природокористування, Рівне 2005. – Вип. 2(30). – С. 282-287.
5. Філатова І.А., Кузло М.Т. Математичне моделювання НДС ґрунтових масивів при наявності рівня ґрунтових вод //В зб.: вісник Національного університету водного господарства та природокористування, Рівне 2006. – Вип. 4(36), Ч.2 – С. 203-208.

Надійшла до редколегії 31.01.08

РАДІОФІЗИКА

УДК 519.9

Алексеев О.С.¹ студент, Мусатенко Ю.С.² канд.
ф.-м. наук

Порівняння мір подібності кривих для задачі розпізнавання підписів

Дану роботу присвячено розробці алгоритму міри подібності кривих для задачі розпізнавання підписів. За основу алгоритму було використано алгоритм динамічної корекції часу. В роботі запропоновано порівнювати не лише часові залежності координат пера, а й додатково враховувати залежність швидкостей пера від часу та локального кута нахилу від часу. Проведено порівняння алгоритму з відомими з літератури подібними алгоритмами та продемонстровано перевагу запропонованого. Досягнуто 98%-ну надійність розпізнавання в задачі розпізнавання підпису.

Ключові слова: метод найменших квадратів, алгоритм динамічної корекції часу, лінійна регресія, розпізнавання підпису.

¹ E-mail: alexeyev@univ.kiev.ua

² E-mail: ysmu@univ.kiev.ua

Вступ

Проблема порівняння двох кривих зустрічається при вирішенні багатьох задач: розпізнавання голосу, порівняння структури білків, порівняння форм об'єктів, аналіз тенденцій на фондових ринках та інших. Під порівнянням кривих розуміють визначення міри подібності між двома кривими. Із визначенням міри подібності між двома кривими пов'язують два головних питання - що є мірою подібності та наскільки точно та швидко ми можемо її розрахувати. Досить часто зустрічаються питання щодо виділення кривої з наявного набору, яка більш за всіх схожа на задану криву, та чи можна сформувати групи кривих із заданим граничним значенням подібності. Зазвичай криві розглядаються як послідовність лінійних відрізків на площині.

На сьогоднішній день найчастіше зустрічаються наступні міри подібності між кривими: відстань Хаусдорфа [3], відстань між кривими, що змінюються [4], та відстань Фречета [5]. В літературі найбільше уваги приділено відстані Фречета. Дана відстань розраховується як деяка мінімізація максимального значення

Alexeyev A.S.¹ student, Musatenko Y.S.² Ph.D

Comparison of curve similarity measures for signature verification applications

The paper is devoted to development of algorithm for curve similarity measure intended for signature verification applications. Algorithm is built as an improvement of dynamic time warping approach. It is suggested to compare not only pen coordinates but account pen velocity and curve local angle temporal dependencies. Algorithm performance is compared to known similar algorithms. Superior results of the suggested method are demonstrated in task of signature recognition. The achieved signature recognition accuracy is 98%.

Keywords: least squares fitting, dynamic time warping, linear regression, signature verification.

відстані між точками двох кривих і тому чутлива до малих змін вхідних даних. Цього можна уникнути, якщо перейти від визначення максимального значення до визначення середнього значення або суми. Найвідомішою такою мірою є відстань DTW (Dynamic Time Warping). Ідея даного алгоритму полягає у визначенні положення схожих ділянок двох кривих, та подальшого розрахунку суми відстані між точками схожих ділянок.

Значення суми відстаней між точками є неочевидною мірою. Щоб прийняти рішення, чи схожі дві криві чи ні, потрібно вводити значення граничної суми відстані. У [7] було запропоновано відійти від розгляду відстаней, до визначення розширеного коефіцієнта змішаної кореляції (РКЗК) ER^2 , що визначає наявність лінійного зв'язку між кривими, використовуючи регресійний аналіз. Причому для визначення послідовності пар точок, за якими будується залежність, використовувати DTW.

З точки зору вирішення задачі порівняння двох підписів алгоритм [7] може використовувати ще декілька джерел інформації.

Їх використання задля покращення надійності порівняння кривих і пропонується в даній роботі.

А саме, не враховується інформація щодо кутових характеристик підпису через використання евклідової відстані між точками двох кривих як міри подібності кривих у відповідних точках під час пошуку розрахунку DTW. Також не враховується динаміка виконання підпису.

Навіть якщо інформація щодо динаміки пера була врахована під час розрахунку DTW, то вона не враховується при розрахунку РКЗК. Ми пропонуємо розраховувати РКЗК не тільки для залежностей координат пера від часу, а й для залежностей швидкостей від часу.

Значення РКЗК, розраховане для різних величин одночасно, не враховує різницю масштабу величин, що порівнюються. Наприклад, якщо розраховувати значення РКЗК за залежностями координат пера та швидкостей пера від часу, зміни швидкості пера на фоні координати у кінцевому результаті не відчувуються. Нами пропонується замість розрахунку РКЗК розраховувати окремо коефіцієнти змішаної кореляції (КЗК) для кожного типу даних, а потім розраховувати середньгеометричне отриманих значень. КЗК мають однакову нормалізацію незалежно від того з якого типу і масштабу даних вони пораховані і тому не потребують додаткового зважування чи масштабування перед обрахунком середнього.

Робота складається з 4 розділів. У вступі окреслено мету даної роботи. В розділі 1 подаються короткі теоретичні відомості щодо задачі визначення подібності двох кривих. Розділ 2 описує алгоритм розрахунку міри подібності кривих. Розділ 3 містить в собі експериментальну частину та результати розпізнавання. Розділ 4 містить висновки роботи.

1. Методи визначення параметрів кривих

В цьому розділі коротко розглядаються відомі методи, використовувані в задачах вимірювання подібності кривих, на яких ґрунтується пропонований алгоритм. А саме: міри подібності на основі лінійної регресії та пошуку відповідності точок кривих на основі динамічної корекції часу.

1.1 Лінійна регресія та міри подібності кривих на її основі

Маємо дві довільні послідовності $X = (x_1, x_2, \dots, x_n)$ та $Y = (y_1, y_2, \dots, y_n)$. Лінійна регресія статистично визначає розподіл точок

$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ в X - Y просторі. Якщо послідовності точок X та Y мають чіткий лінійний зв'язок $Y \approx \beta_0 + \beta_1 X$, то ми можемо сподіватися на те, що точки будуть знаходитись вздовж так званої лінії регресії.

Для визначення регресійного коефіцієнта між X та Y спочатку розглянемо наступну модель зв'язку між X та Y :

$$Y = \beta_0 + \beta_1 X + u, \text{ де } u - \text{помилка.} \quad (1)$$

Зазначимо, що $u = (u_1, u_2, \dots, u_n)$. Далі знайдемо коефіцієнти β_0, β_1 за принципом мінімуму середньоквадратичної похибки, тобто мінімуму значення наступного виразу:

$$\sum_{i=1}^n u_i^2 = \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i))^2 \quad (2)$$

З геометричної точки зору знаходимо лінію регресії, що визначається коефіцієнтами β_0, β_1 , як лінію в просторі X - Y , що лежить найближче до усіх точок. Нехай

$$Q(\beta_0, \beta_1) = \sum_{i=1}^n u_i^2 \quad (3)$$

Для знаходження коефіцієнтів необхідно знайти екстремум функції Q , тобто $\frac{\partial Q}{\partial \beta_0} = 0, \frac{\partial Q}{\partial \beta_1} = 0$. Рішення даної задачі можна представити у наступному вигляді:

$$\beta_0 = \bar{Y} - \beta_1 \bar{X} \quad (4)$$

$$\beta_1 = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{X})^2}, \text{ де} \quad (5)$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i, \bar{Y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (6)$$

Визначивши коефіцієнти β_0, β_1 , можна визначити лінію регресії. Але залишається ще одне питання: наскільки добре лінія апроксимує вихідну послідовність. Іншими словами, наскільки близько дана лінія знаходиться до вихідних точок. Як міру помилки апроксимування використовують значення наступного виразу:

$$R^2 = 1 - \frac{\sum_{i=1}^n u_i^2}{\sum_{i=1}^n (y_i - \bar{Y})^2}. \quad (7)$$

Враховуючи (4, 5):

$$R^2 = \frac{\left(\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y}) \right)^2}{\sum_{i=1}^n (x_i - \bar{X})^2 \sum_{i=1}^n (y_i - \bar{Y})^2}, \quad (8)$$

R^2 називають коефіцієнтом змішаної кореляції (КЗК). R^2 має наступні властивості:

– $R^2(X, X) = 1$;

– з виразу (8) видно, що $R^2(X, Y) = R^2(Y, X)$;

– $R^2 \in [0, 1]$. Чим ближче значення R^2 до 1, тим більше точок прямує до лінії регресії, тим сильніша лінійна залежність між двома послідовностями.

Спираючись на властивості R^2 , R^2 можна визначити як надійність лінійного зв'язку. Також міру R^2 можна вважати мірою подібності. З (рис. 1) можна побачити те, що чим більше схожі дві послідовності, тим більше значення R^2 . Так для послідовностей X_1 та X_2 $R^2 = 0,91$. А для послідовностей X_3 та X_4 $R^2 = 0,31$.

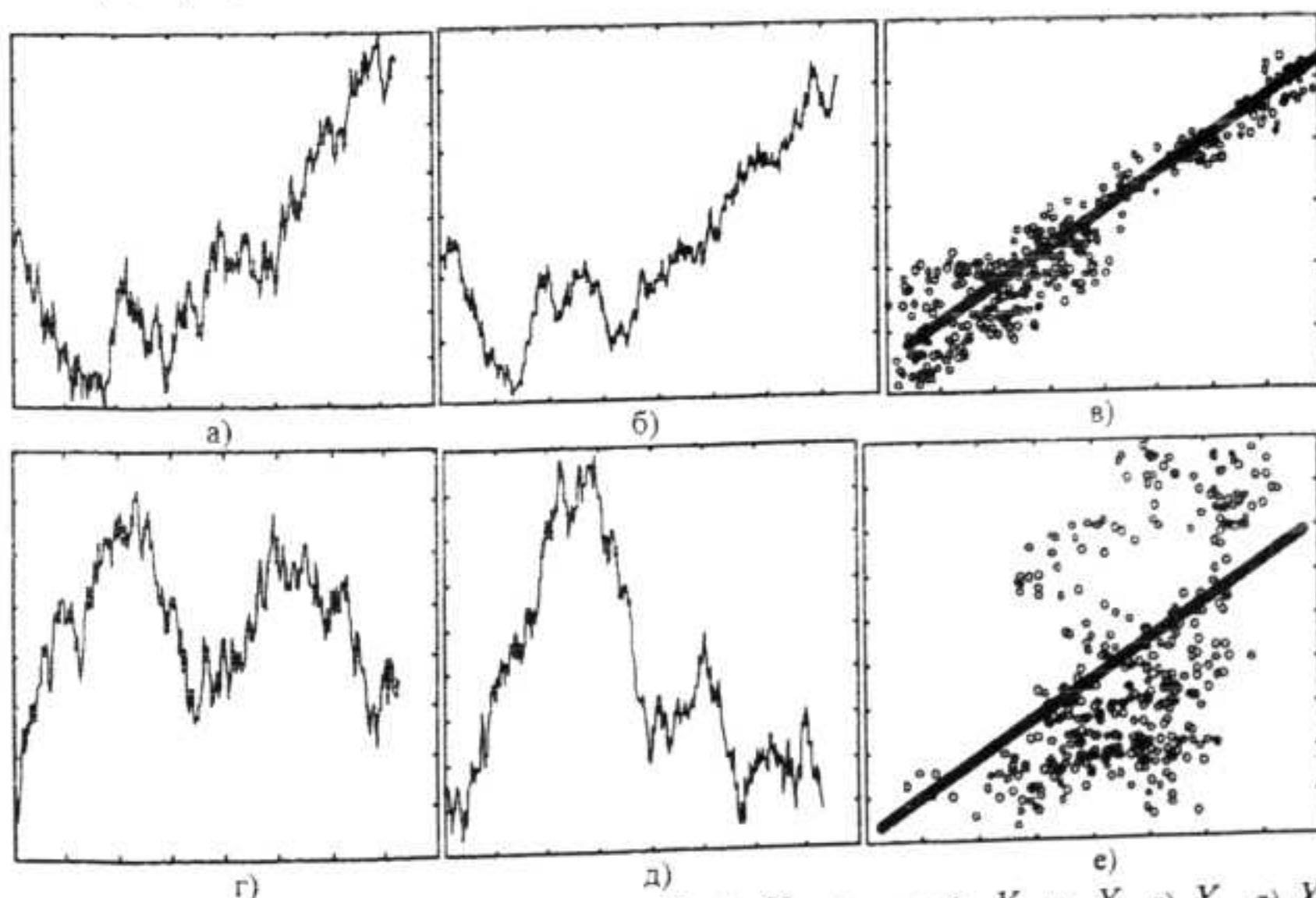


Рис. 1 Дві пари послідовностей, та їх розподіл: а) X_1 , б) X_2 , в) розподіл X_2 на X_1 , г) X_3 , д) X_4 , е) розподіл X_4 на X_3

Також потрібно зазначити те, що значення R^2 є більш наочною мірою схожості, ніж деяка середня відстань чи помилка.

Звичайно, просту лінійну регресію можна застосовувати лише для одновимірних послідовностей. Але ж існує велика кількість багатовимірних послідовностей. Наприклад, для задачі розпізнавання підписів послідовність є багатовимірною – дві просторові координати, тиск, інше. Для визначення подібності таких

послідовностей у [7] була визначена наступна міра:

$$ER^2 = \frac{\left[\sum_{j=1}^M \left(\sum_{i=1}^n (x_{ij} - \bar{X}_j)(y_{ij} - \bar{Y}_j) \right) \right]^2}{\sum_{j=1}^M \sum_{i=1}^n (x_{ij} - \bar{X}_j)^2 \sum_{j=1}^M \sum_{i=1}^n (y_{ij} - \bar{Y}_j)^2}. \quad (9)$$

ER^2 називають розширеним коефіцієнтом змішаної кореляції (РКЗК). У [7] значення даного коефіцієнту і визначає міру подібності двох

кривих. Через певні недоліки РКЗК в цій системі міра подібності двох підписів визначається як середньгеометричне значень КЗК для залежностей координат пера та швидкостей пера від часу.

1.2. Алгоритм динамічної корекції часу (DTW)

Міра подібності кривих на основі РКЗК розраховується послідовним порівнянням точок

двох послідовностей. Тому така міра безпосередньо не може бути застосована для кривих, що містять часові зсуви відповідних елементів, або для кривих, що відрізняються наявністю окремих елементів взагалі. При розрахунку даної міри потрібно визначати які саме ділянки необхідно порівнювати. Для пошуку схожих ділянок може бути застосовано алгоритм динамічної корекції часу - Dynamic time warping (DTW) [1].

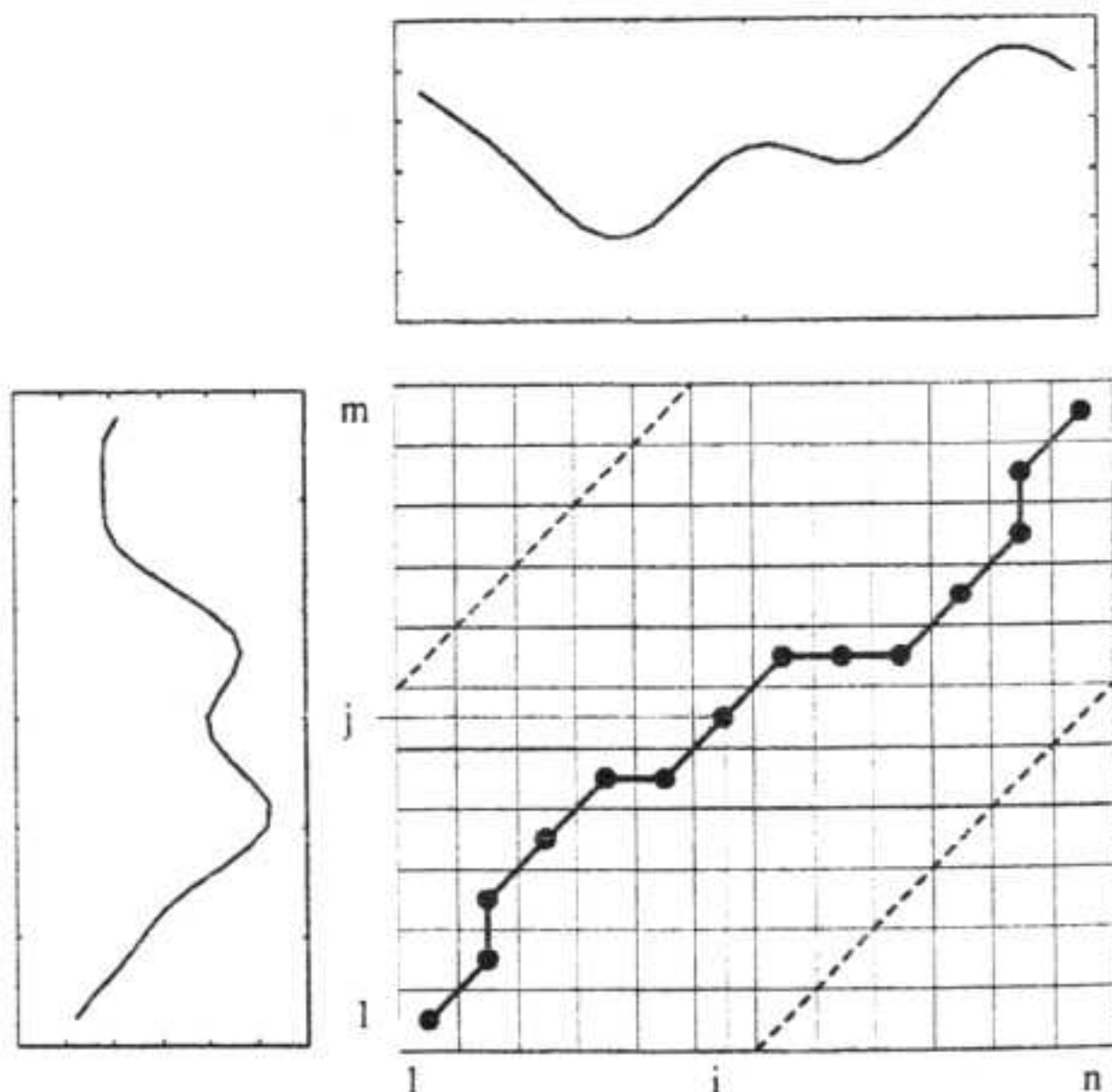


Рис.2 Матриця подібності окремих елементів двох кривих

Нехай X, Y – дві послідовності довжиною n та m відповідно:

$$\begin{aligned} X &= x_1, x_2, \dots, x_n \\ Y &= y_1, y_2, \dots, y_m \end{aligned} \quad (10)$$

Побудуємо матрицю розмірністю $[n, m]$, де елемент матриці (i, j) – це значення відмінності кривих X та Y у відповідних точках i та j . На рис. 2 зображений приклад такої матриці.

Шлях W – це послідовність елементів матриці, що визначає шлях від точки $(1, 1)$ до точки (n, m) , що має найменшу загальну відмінність окремих ділянок послідовностей X та Y . Іншими словами – це найкоротший шлях від точки $(1, 1)$ до точки (n, m) , що визначається найменшою помилкою.

$$DTW(X, Y) = \min \left(\sum_{k=1}^K w_k \right), \text{ де}$$

$$\max(n, m) \leq K < m + n - 1. \quad (11)$$

Алгоритм побудови шляху W може бути реалізований за допомогою динамічного програмування, щоб уникнути проблем, пов'язаних з рекурсивним розрахунком. Розрахунок починається з точки $(1, 1)$, (рис. 2). Для кожної наступної точки (i, j) розраховується загальна помилка $\gamma(i, j)$ шляху від точки $(1, 1)$ до даної точки у вигляді:

$$\gamma(i, j) = d(x_i, y_j) + \min \{ \gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1) \}. \quad (12)$$

Визначивши шлях W , тепер можна між собою (рис. 3).
визначити ті ділянки, які потрібно порівнювати

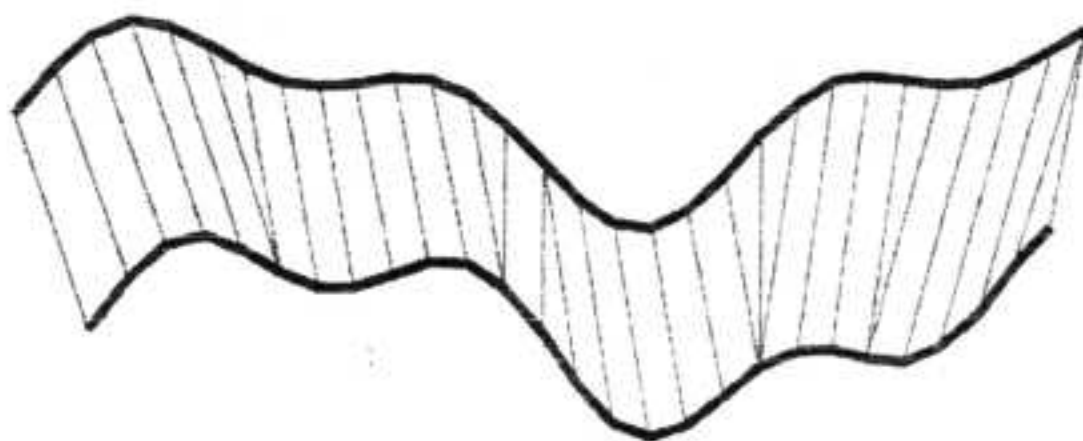


Рис. 3 Визначення схожих ділянок двох кривих для порівняння

2. Алгоритм вимірювання подібності кривих

Роботу запропонованого алгоритму можна умовно розділити на два етапи. Перший етап – це підготовка підписів до розпізнавання. складається з наступних кроків:

- нормування;
- згладжування;
- нормування за кількістю точок у послідовності.

Другий етап – безпосередньо визначення значення подібності між кривими двох підписів, складається з таких кроків:

- пошук схожих ділянок за DTW;
- розрахунок значення коефіцієнта лінійного зв'язку.

2.1 Попередня обробка вхідних даних

Нормування підпису – це приведення лінійних розмірів підписів до одного значення. Координати кожної точки перераховуються наступним чином:

$$Y = \frac{Y - \bar{Y}}{\sigma(Y)}, \quad X = \frac{X - \bar{X}}{\sigma(X)}. \quad (13)$$

Наступний крок – згладжування. Для згладжування кривої використовується апроксимація поліномом 3-го ступеню методом найменших квадратів. Звичайно використання поліному 3-го ступеню для згладжування кривих набагато більших порядків недостатньо. Тому було вирішено розбити всю криву на ділянки та кожну ділянку згладжувати окремо. За допомогою методу найменших квадратів можна побудувати криву, що найкращим чином описує сукупність точок. Зазвичай, помилка апроксимування збільшується від центру ділянки до її кінців. Тому при відновленні цілої кривої з окремих ділянок на стику двох сусідніх ділянок

з'являються розриви. Для зменшення величини стрибків згладжування окремої ділянки вихідної кривої відбувається із врахуванням точок сусідніх ділянок.

Останнім кроком підготовки підпису до розпізнавання є нормування за кількістю точок. Після згладжування підпис являє собою залежність X координати від часу та залежність Y координати від часу. Звідси можна розрахувати значення координат необхідної кількості точок, що буде визначати криву підпису.

2.2 Розрахунок значення подібності кривих

Остаточне значення подібності двох кривих підписів розраховується за допомогою регресійного аналізу, через визначення КЗК. Значення подібності – це число від 0 до 1. 0 відповідає випадку зовсім різних підписів, а 1 відповідає випадку абсолютно схожих підписів. Як зазначалось вище, такі значення міри відмінності є набагато прозорішими і зрозумілішими, ніж деяке значення середньої відстані чи помилки.

Зрозуміло, що застосування регресійного аналізу на пряму до кривих, що відрізняються лише часовим зсувом окремих ділянок, не має ніякого сенсу. Спочатку потрібно визначити відповідність точок двох кривих через визначення положення схожих ділянок. Для визначення схожих ділянок двох підписів та формування пар відповідних точок кривих використовувався алгоритм DTW.

Перед застосуванням алгоритму DTW потрібно згенерувати матрицю відмінностей двох підписів у відповідних точках, тобто потрібно визначити міру подібності окремих ділянок. У [7] матрицю відмінностей будують лише за значеннями евклідової відстані між точками кривих. Зрозуміло, що для задачі розпізнавання підпису цього недостатньо, оскільки дуже

важливо враховувати динаміку виконання підпису, а також нахил кривої. У роботі пропонується розраховувати міру подібності окремих ділянок у наступному вигляді (рис. 4):

$$d(i, j) = \frac{1}{K-1} \sum_{k=1}^{K-1} (a_k + b_k), \text{ де} \quad (14)$$

a – різниця довжин,

b – відстань, що відповідає куту повороту,

K – кількість точок в одній ділянці для порівняння.

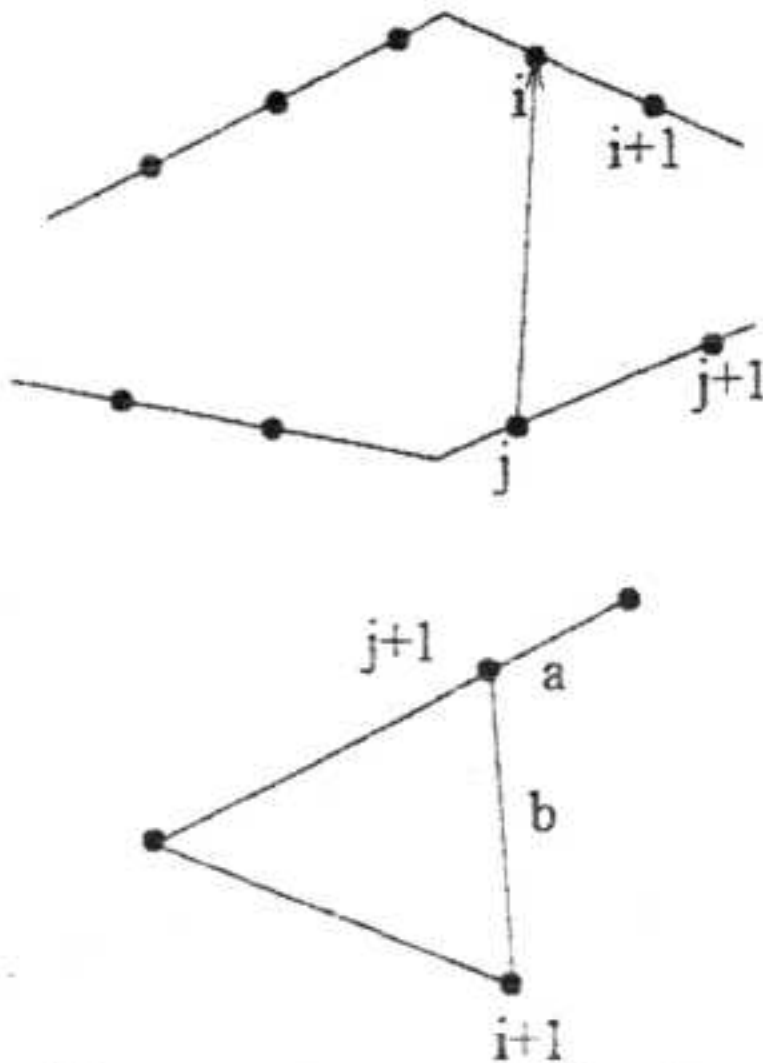


Рис. 4 Визначення значення подібності окремих ділянок підпису

Запропонована міра (14) враховує:

- різницю швидкостей у відповідних точках двох кривих;
- різницю локальних нахилів двох кривих.

На рис.5 зображено матрицю відмінностей двох підписів, та оптимальний шлях, знайдений за алгоритмом DTW. Після цього застосовуємо регресійний аналіз для оцінки подібності двох кривих. Значення КЗК розраховується для чотирьох залежностей – дві просторові координати та швидкості їх зміни. Для визначення коефіцієнта змішаної кореляції одразу для декількох залежностей у [7] було запропоновано використовувати коефіцієнт РКЗК. Але даний коефіцієнт, застосований до різних величин одночасно, не враховує різницю масштабу величин, що порівнюються. Тому визначалось значення КЗК для кожної залежності, а потім розраховувалось

середньогометричне отриманих значень. Це вирішує проблему масштабу оскільки КЗК є однаково нормалізованими.

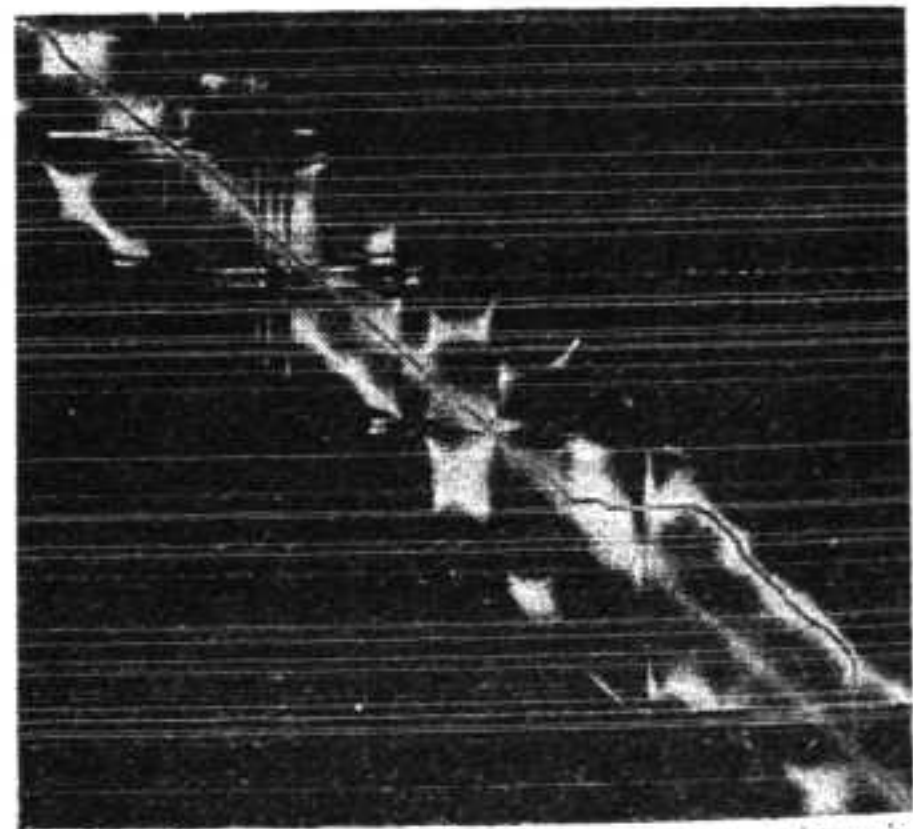


Рис.5 Матриця порівняння двох підписів, оптимальний шлях

На рис. 6 показано зв'язок між ділянками двох кривих, що вважаються подібними, для яких буде розраховано КЗК.

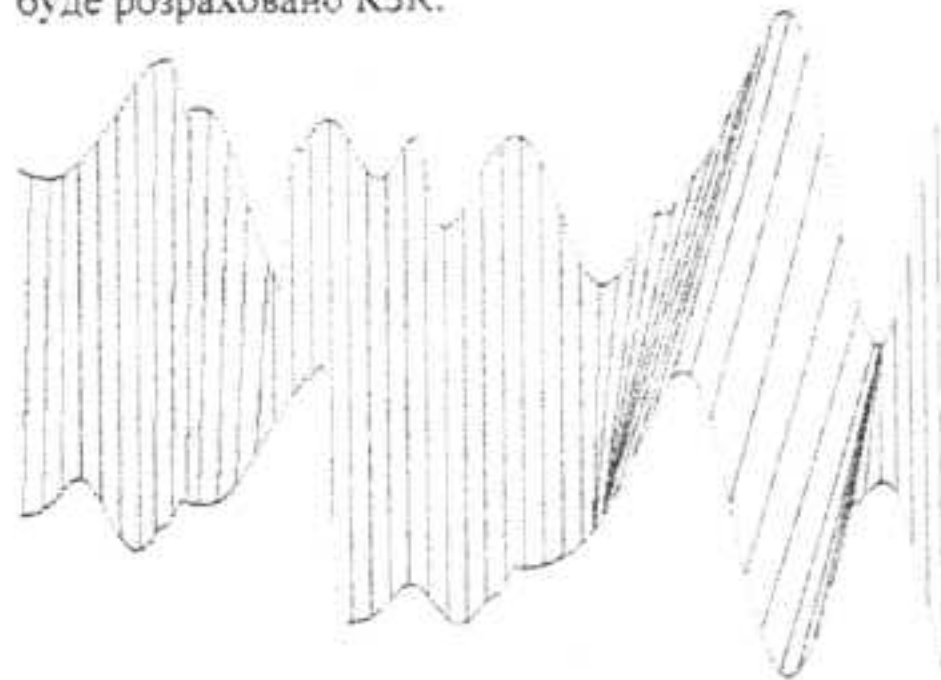


Рис. 6 Вибір схожих ділянок двох підписів

3. Експериментальні результати

Оцінку міри подібності двох кривих була проведена для задачі розпізнавання підписів людини. Для задачі розпізнавання визначають два типи помилок. Перший тип – помилково детектовані підписи (хибна тривога), другий – помилково відхилені підписи (пропуск цілі).

Використовуючи вище згаданий алгоритм проведено розпізнавання 2228 підписів 64 людей. Результати роботи даного алгоритму ми порівнювали з результатами роботи вихідного алгоритму [7] та з результатами роботи системи розпізнавання, що базується на виділенні ознак [2, 6]. Результати зображено на рис. 7. 3

наведеного рисунку видно, що алгоритм, надійність - 98%, порівняно з вихідним запропонований у даній роботі, має більшу алгоритмом, надійність якого складає 97%.

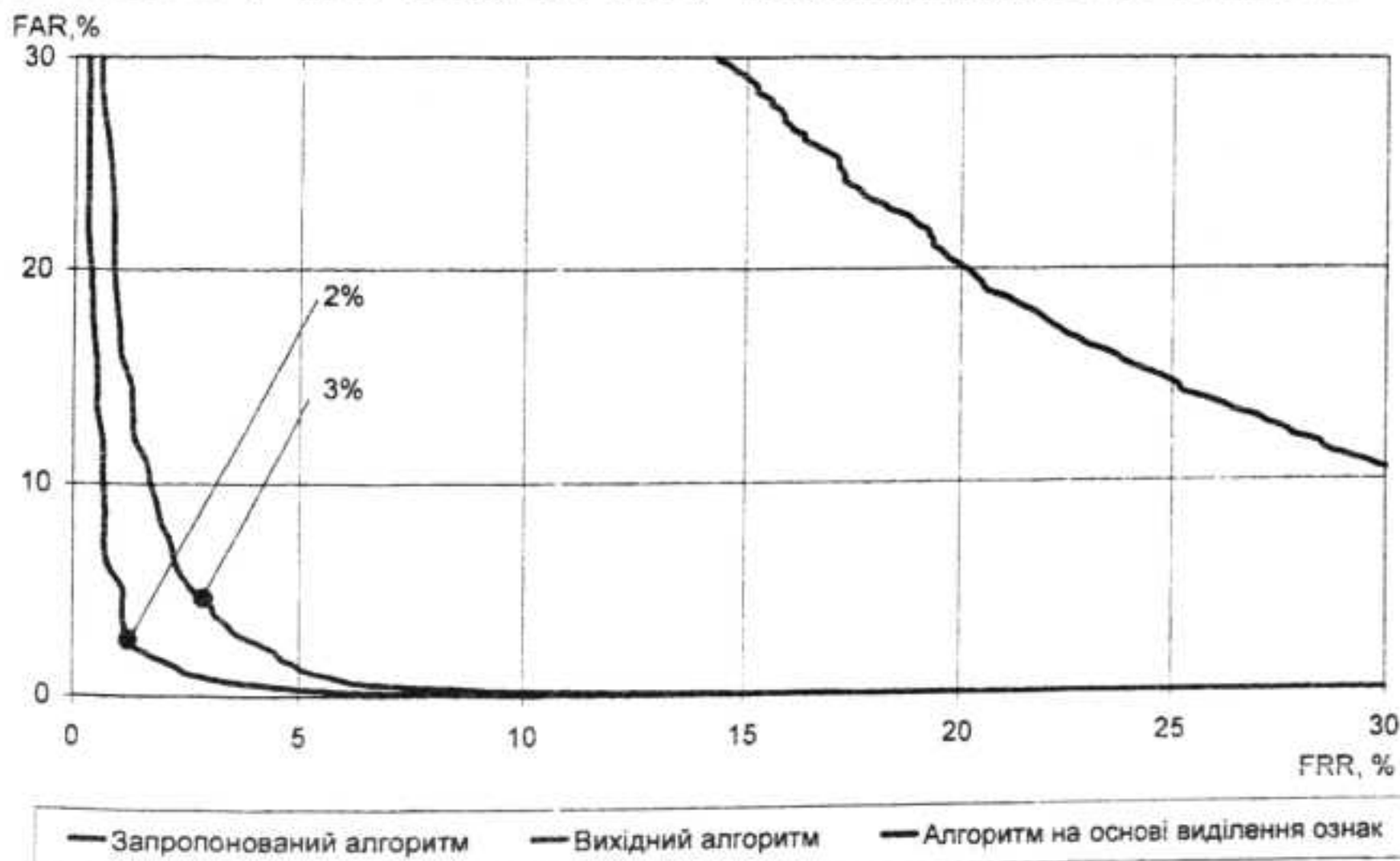


Рис. 7 Кількість помилкових детектувань (FAR) від кількості помилкових відхилень (FRR)

Висновки

Запропоновано алгоритм визначення міри подібності двох кривих для задачі розпізнавання підписів.

Запропоновано алгоритм розпізнавання підпису на основі алгоритму динамічної корекції часу з урахуванням залежності швидкостей пера від часу та залежності локального кута нахилу від часу. Враховуючі цих залежностей, вдалося довести надійність розпізнавання до рівня 98%, зменшивши при цьому ймовірність помилки на 30%, порівняно з подібними алгоритмами, відомими з літератури.

Список використаних джерел

1. Alon Efrat, Quanfu Fan. Curve Matching, Time Warping, and Light Fields: New Algorithms for Computing Similarity between Curves // Journal of mathematical imaging and vision. – 2007. – 3. – P. 203-216.
2. Юрій С. Мусатенко¹, Олександр С. Алексєєв, Віталій М. Глущенко, Людмила А. Мусатенко. Порівняння якості ознак для задач розпізнавання підпису // Вісник Київського Університету. – Сер.: Фізико-математичні науки. – 2006. – Вип. 3. – С. 366-375.
3. H. Alt, B. Behrends, and J. Blömer. Approximate matching of polygonal shapes // Ann. Math. Artif. Intell. – 1995. – 13. – P. 251-266.
4. E. M. Arkin, L. P. Chew, D. P. Huttenlocher, K. Kedem, and Joseph S. B. Mitchell. An efficiently computable metric for comparing polygonal shapes // IEEE Trans. Pattern Anal. Mach. Intell. – 1991. – 13. – P. 209-216.
5. K. Buchin, M. Buchin and C. Wenk. Computing the Frechet Distance Between Simple Polygons in Polynomial Time // Proc. 22th ACM Symp. on Computational Geometry (SoCG). – 2006. – P. 80-88.
6. Gluschenko V.M., Musatenko Yu.S. Human signature recognition: sign quality comparison // Proc. of the 4th Int. Young Scientists' Conference on Applied Physics. Kyiv (Ukraine). – 2004. – P. 238-239.
7. Hansheng Lei, Srinivas Palla, Venu Govindaraju. ER2: an Intuitive Similarity Measure for On-line Signature Verification // Proc. of the Ninth International Workshop on Frontiers in Handwriting Recognition (IWFHR'04). – 2004. – 0. – P. 191 – 195.

Надійшла до редколегії 1.04.08

УДК 533.951

Голдаєвич Є.Л., к.ф.-м.н., Мартиш* Є.В.,
к.ф.-м.н., доцент

Перехід у нерівноважний стан розряду з ковзаючою дугою

Досліджено вплив електричного поля на розряд з ковзаючою дугою і показано, що воно може відігравати головну роль у досягненні нерівноважного стану такої плазми. Запропоновано, що більша напруженість поля спричиняє іонізаційну нестійкість в плазмі, що відбивається у складній часово-просторовій структурі такого розряду, включаючи швидкий перехід між рівноважним та нерівноважним станами. Останній має низьку поступальну температуру і порівняно високу електронну. Це є можливою базою для більшої керованості нановиробництва у таких розрядах.

Ключові слова: дуговий розряд, іонізаційна нестійкість, нерівноважність, нановиробництво.

*E-mail: emart@univ.kiev.ua

Карбонові нанотрубки привертають велику увагу дослідників та фахівців з інженерних технологій з часів свого відкриття [1]. Їхні фізичні та електронні властивості, які поєднуються з їх хімічною інертністю, роблять ці структури потенційно дуже корисними для застосувань у якості польових емітерів, електронів, наноелектродів, фільтруючих середовищ та супергідрофобних поверхонь. На сьогодні можна вважати доведеним (теоретично і експериментально), що технології нановиробництва за допомогою плазми мають незаперечні переваги перед такими, що використовують нейтральний газ або інші засоби [2]. Серед них: осадження з реагуючої пари (CVD), епітаксія молекулярним пучком (MBE), нанесення кластерним пучком або «мокра» хімія, електронний пучок, мікролітографія, плазмені технології і таке інше.

Серед найбільш уживаних сьогодні плазмових методів є дуговий розряд, лазерне випаровування та осадження з реагуючої пари, підсилене плазмою (PECVD). Відзначимо, що вперше такі структури були синтезовані на дуговій установці, подібній до такої, на якій Кречмер вперше продукував фулерени C_{60} .

Goldayevich E.L., Ph.D., Martysh* E.V.,
Ph.D., associate professor

A non-equilibrium state achieving for discharge with gliding arc

The effect of the electric field in plasma discharge with gliding arc is investigated and shown to be important in achieving non-equilibrium state. It is proposed that a higher field causes ionization instability within the plasma, which results in complex spatial-temporal structure of discharge, including the fast transition between equilibrium and non-equilibrium states. The last is characterized by low translational and relatively high electron temperature. That is a possible base for more controlled nanofabrication.

Key Words: arc discharge, ionization instability, nonequilibrium, nanofabrication.

Зауважимо, що плазма, яка використовується у вказаних вище технологіях, зазвичай поділяється на рівноважну і нерівноважну, частіше за єдиною ознакою – температури компонент (тобто плазма є рівноважна, якщо всі температури майже однакові і навпаки, якщо існує помітна різниця). Кожна плазма має свої переваги і недоліки у технологічних застосуваннях.

Термічна (рівноважна) плазма зустрічається у розрядах при високому (типово – від кількох Тор до атмосферного) тиску. В термічній плазмі це значить, що T_e (електронна температура) $\sim T_i$ (іонна температура) $\sim T_n$ (температура нейтралів) ~ 1 eV. В свою чергу, це значить, що оточуючий нейтральний газ може бути гарячим достатньо, щоб розплавити або навіть випарувати твердотільні матеріали. В такій плазмі ступінь іонізації буде суттєвим, в багатьох випадках досягаючи повної іонізації (навіть з двічі чи тричі іонізованими компонентами) у центральних високотемпературних областях, але набагато нижчою біля границі або додаткових областях.

З іншого боку, в низькотемпературній термічно нерівноважній плазмі, електронна температура принаймні на порядок величини вища від температури іонів та нейтралів.

Наприклад, у ВЧ чи розряді постійного струму низького тиску, іони та нейтралі перебувають при кімнатній температурі, тобто $T_i \sim T_n \sim 0,026$ eV, у той час, як електронна температура сягає 1–2 eV. В цьому випадку можна очікувати зовсім іншу взаємодію плазми з твердим тілом, тому що нейтральний газ є холодним і „лагідним” до твердого стану. У такій плазмі ступінь іонізації є порівняно низьким, у межах 10^{-5} – 10^{-2} .

Особливою перевагою нерівноважної плазми є можливість спрямованого руху заряджених компонент під впливом електричного поля. Це дає можливість анізотропного щавлення поверхонь та керованого нанесення будівельних одиниць з наноансамблів. В термічній плазмі велика частота зіткнень означає, що спрямований рух є набагато менш поширеним, за виключенням дуже великих наноансамблів, таких як наночастинки та нанотрубки, що розглядались у [3]. В інших аспектах нерівноважна плазма має перевагу при використанні в ряді процесів, які включають температурно чутливі матеріали, такі як полімери, пластики та ультра-тонкі нанопрошарки.

Головною перевагою термічної плазми є доступність високих температур (високих у порівнянні з тими, що доступні у системах, базованих на горінні – а не з термоядерною плазмою), з температурами більш ніж 20 000 K для дугових розрядів та 10 000 K для індуктивних ВЧ та плазових факелів постійного струму. Високі температури дозволяють використовувати тверді, рідкі, газові суспензійні або розчинні попередники, з можливістю повної дисоціації хімічних реактантів в атоми. Той факт, що термічні плазми перебувають при тиску біля атмосферного, визначає можливість отримання компонент з великими густинами. Більш того, плазмова температура практично не залежить від хімічних реакцій (хоча у випадку великого навантаження – нагрівання та випаровування може суттєво охолодити плазму) та існує дуже широкий вибір реактивних або інертних середовищ. Навпаки, технології, базовані на горінні, залежать від хімічних реакцій, що нагрівають газ, тому атмосфера обов'язково містить гази, що горять та продукти згоряння. Висока температура термічної плазми означає також, що тут можливо отримати дуже круті температурні градієнти, наприклад при розширенні в зону низького тиску, при введенні холодної поверхні або при змішуванні холодного газу з плазовим потоком. Нуклеація у газовій фазі при високих швидкостях охолодження стає

чітко вираженою і призводить до високого ступеню перенасичення і, таким чином, до високих швидкостей нуклеації. Це, в свою чергу, дає високий вихід синтезу наночастинок та наноструктур, що є недосяжним при інших методах, включаючи ті, що базовані на нерівноважній плазмі.

Суттєвою вимогою до плазових нанотехнологій є їх детермінованість, тобто можливість керування ними та передбачуваність їх наслідків. І тут ряд беззаперечних переваг (серед них і зазначені вище) мають процеси у нерівноважній плазмі.

Дослідженню можливостей поєднання зазначених властивостей у потужних дугах атмосферного тиску присвячена ця робота. Справа у тому, що в таких системах можуть нестационарно підтримуватись нерівноважні розряди, в яких розвивається іонізаційна нестійкість. Вперше така ситуація була експериментально досліджена в [4] на прикладі ковзаючого дугового розряду. Він запалювався між двома плоскими електродами у місці їх найбільшого зближення і рухався вздовж них під дією газового потоку. Було помічено, що температура газу у квазірівноважній дузі практично не залежить від величини l (довжина розряду), а потужність, що виділяється в розряді, та напруга на ньому зростають практично лінійно з l . Сила струму при наближенні l до критичної $l^* = \mathcal{E}^2/4WR$ (тут W – погонна потужність, а R – опір розряду, тому $\mathcal{E} = JR + \sqrt{Wl/R}$) зменшується практично вдвічі ($J^* = \mathcal{E}/2R$), а потужність P та електричне поле E досягають максимуму. Зрозуміло, що при $l > l^*$ теплові втрати продовжують зростати, а P не може збільшуватись, тому T починає падати. Суттєвою є поведінка провідності плазми, але помітне значення її може бути підтримане процесом ступінчастої іонізації.

Розглянемо вплив повільного зростання E (воно відбувається завдяки збільшенню l) на стійкість такого розряду. Для останнього оберемо т.зв. „каналову” модель [5]. В цій моделі вважається, що в середині каналу провідність висока, а назовні його провідність дорівнює нулю і струму немає. У випадку, коли струм ще не дуже значний, електрон-атомні (молекулярні) зіткнення превалюють над електрон-іонними, тоді провідність σ пропорційна концентрації електронів n_e :

$$\sigma = C \exp\left(-\frac{I}{2T_e}\right), \quad C = \text{const.} \quad (1)$$

При більш значних струмах (ступінь іонізації більший одного відсотка) залежність $\sigma(T)$ вже інша, але, як показано у [5], її можна описати законом (1), підбираючи ефективне значення потенціалу іонізації I_{ef} . Так, у повітрі, азоті та аргоні при тиску 1 атм, в характерному для дуг діапазоні температур:

$$\sigma = 0,83 \cdot 10^{-2} \exp\left(-\frac{36000}{T_e}\right), \quad I_{ef} \approx 6,2 eV. \quad (2)$$

Проте відомо, що одна з основних особливостей плазми полягає в тому, що нелінійні ефекти проявляють себе вже у порівняно невеликих і тому легко досяжних електричних полях. Такою є і нагрівна і пов'язана з нею іонізаційна нестійкість слабо іонізованої плазми. Як показано у багатьох монографіях з цього питання, електронна температура є квадратичною функцією напруженості електричного поля і також зростає при зменшенні долі енергії, що втрачається електронами при зіткненнях:

$$T_e = T \left(1 + \frac{E^2}{E_p^2}\right), \quad (3)$$

де E_p – характерне плазмове поле,

$$E_p^2 = 3mT/e^2 \delta_0 v_{e0}^2, \quad (4)$$

де T – температура плазми в електрон-вольтах, v_{e0} – ефективна частота зіткнень електрона з атомами (молекулами) та іонами, δ_0 – середня доля енергії, що втрачається електроном при одному зіткненні.

При наближенні поля до E_p (але E менше від E_p) провідність квазірівноважної плазми буде визначатись не тільки T , але і E :

$$\sigma = C \exp\left(-\frac{I}{2T(1 + \frac{E^2}{E_p^2})}\right) \sim C \exp\left(-\frac{I}{2T} + \frac{IE^2}{2a}\right), \quad (5)$$

де $a = E_p^2 T$. Як відомо з [5], вимогу для суттєвої зміни провідності у порівняно невеликому інтервалі температур $\Delta T \ll T$, можна подати у загальній диференціальній формі:

$$\zeta_T = (\Delta T/T)^{-1} = \left(\frac{\partial \ln \sigma}{\partial \ln T}\right) \sim \frac{I}{2T}, \quad (6)$$

$$\zeta_E = \left(\frac{\partial \ln \sigma}{\partial \ln E}\right) \sim \frac{IE^2}{a}. \quad (7)$$

Ці вирази отримані з використанням (5), але взагалі можуть бути використані при довільній залежності $\sigma(T)$. Зазначимо, що у [6] стверджується, що до $T \leq 3 \cdot 10^4$ К степенева апроксимація $\sigma(T) = C(T/T_0)^{\zeta_T}$ гарно відповідає даним експериментів.

Звичайний дуговий розряд, де $J = \sigma ES = \text{const}$, S – переріз каналу струму, є достатньо стійким. При невеликому охолодженні його провідність різко падає, що призводить до зростання E та W . В свою чергу, така ситуація веде до зростання T і стабілізації розряду. Інший перебіг подій можливий при $E > E_p$, коли провідність є функцією не тільки T , але і E . В таких умовах невеличке охолодження плазми викликає малу зміну провідності і невелике зростання поля і погонної потужності. Відповідне зменшення енерговнеску, який забезпечує встановлення початкової температури, може при певних умовах привести до втрати стійкості такого розряду.

Думка про можливість розвитку своєрідної іонізаційної нестійкості при великому відносному відриву температур ($\geq 50\%$) висловлена в [5]. Більш детальний аналіз можна провести, використовуючи рівняння балансу теплової енергії в каналі та закон Ома:

$$nc_p \frac{\partial T}{\partial t} = \sigma(T, E) E^2 - 8\pi \frac{\lambda(T) T^2}{lS}, \quad (8)$$

$$\sigma(T, E) R S / I + 1 = \frac{2E^*}{E}, \quad (9)$$

де $E^* = \mathcal{E}/2l$; c_p – теплоємність газу, $\lambda(T)$ – його теплопровідність. Розв'язок системи (8-9) будемо шукати у вигляді $T = T_0 + T_1(t)$; $E = E_0 + E_1(t)$. Використовуючи (6-7), отримаємо:

$$(1 - T_1/T) \frac{\partial T_1}{\partial t} = \omega \zeta_T T_1 \left[\frac{k-2}{k+\zeta_E} + \frac{T_1}{2T} f(\zeta_T, \zeta_E, k) \right], \quad (10)$$

де $\omega = \frac{\sigma E^2}{nc_p T_0}$, $k(E) = E^*/E$, f – поліноміальна

функція вказаних у (10) змінних. В лінійному наближенні з (10) можна отримати вираз для інкременту розвитку нестійкості:

$$\Omega = -\omega \frac{\zeta_T}{1+\zeta_E} \left(1 - \frac{E}{E^*}\right). \quad (11)$$

З (11) випливає, що система є стійкою ($\Omega < 0$) при $E < E^*$ ($I < I^*$), та є нестійкою при зворотніх нерівностях. Тобто, (11) описує нестійкість, яка виникає в наслідок неможливості

компенсувати зростаючі при $I > I^*$ теплові втрати за рахунок джерела. Це і було продемонстровано в [4]. Зауважимо, що при $E_p(T/I)^{1/2} < E < E^*$ ($\Omega < 0$) інкремент Ω за рахунок множника $I/(1 + \zeta_E)$ може зменшитись приблизно у 20 разів.

У звичайних дугових розрядах ($E \ll E^*$; $\zeta_E \ll 1$) збільшення T знижує чутливість провідності (дивись ф-лу (7)) до електричного поля і тому збільшує стійкість системи. Падіння T збільшує чутливість до поля і тим самим знижує стійкість системи. При розвитку нестійкості електричне поле зростає, а провідність, густина струму, температура плазми та питомий енерговнесок - зменшуються. Невелике збільшення поля (в 2-3 рази) при помітному збільшенні концентрації електронів відповідає відносно меншому, ніж T , зниженню E/n і, відповідно T_e в каналі розряду, допомагаючи переходу плазми у нерівноважний стан.

Зміна величини питомого енерговнеску стає від'ємною. Але у відповідності до закону Ома при зменшенні сили струму повний енерговнесок повинен зростати. Ця невідповідність ліквідується швидким ($\sim 1\text{мс}$) зростанням довжини розряду. Електронна температура залишається приблизно сталою (на рівні 10^4 K), а поступальна падає до $2 \cdot 10^3\text{ K}$.

Показано, що квазірівноважний розряд з ковзаючою дугою може мати кілька стадій у відповідності до величини прикладеного електричного поля E :

- I) – рівноважна, $E < E_p(T/I)^{1/2}$;
- II) – перехідна, $E_p(T/I)^{1/2} < E < E^*$;
- III) – нерівноважна, $E > E^*$.

Таким чином, створюються сприятливі умови для виробництва наноструктур в дуговому розряді, включаючи більшу детермінованість процесу.

Список використаних джерел

1. Ostrikov K. and Murphy A.B. Plasma – aided nanofabrication: where is the cutting edge? // J.Phys.D.: Appl. Phys. – 2007. – 40. – P. 2223-2241.
2. Bell M.S., Teo K.B.K. and Milne W.I. Factors determining properties of multi-walled carbon nanotubes deposited by PECVD // J.Phys.D.: Appl. Phys. – 2007. – 40. – P. 2285-2292.
3. Keidar M. Factors affecting synthesis of SWNT in arc discharge // J.Phys.D.: Appl. Phys. – 2007. – 40. – P. 2388-2293.
4. Русанов В.Д., Петрусов А.С., Потанин Б.В. и др. Ионизационная неустойчивость переходного режима дугового разряда // ДАН. – 1993. – 332.- №3.- С. 306-308.
5. Райзер Ю.П. Физика газового разряда: Учеб. руководство. – М., 1987.
6. Yas'ko O.I. Some aspects of the generalization of electric arc characteristics // Pure and Appl. Chem. – 1990. – 62, #9. – P.1817 – 1824.

Надійшла до редколегії 31.03.08.

УДК 621.385.111, 615.014.075

Іванюта О. М., канд. фіз.-мат. наук,
Прокопенко О. В., канд. фіз.-мат. наук

Дослідження водних сольових розчинів за допомогою НВЧ аналізатора на основі резонатора поверхневої хвилі

Розглянуто порівняльні вимірювання діелектричної проникності водяного розчину NaCl з великими втратами в НВЧ діапазоні за допомогою резонатора поверхневої хвилі в прямокутному хвилеводі.

Показано, що високі чутливість і роздільна здатність НВЧ аналізатора дозволяють застосовувати його для дослідження водних сольових розчинів.

Ключові слова: резонатор поверхневої електромагнітної хвилі, водний сольовий розчин, власна добротність.

Ivanyuta O. M., Ph.D.,
Prokopenko O. V., Ph.D.

Investigation of salt water solutions using the microwave analyzer based on surface wave resonator

Comparative measurements of dielectric permittivity of NaCl salt water solution with high losses at microwaves by the use of surface wave resonator in rectangular waveguide are examined.

It is shown that microwave analyzer has enough high sensitivity and discrimination for the investigation of salt water solutions.

Key words: surface electromagnetic wave resonator, salt water solution, unloaded Q-factor.

*E-mail: mga@mail.univ.kiev.ua, wrld@mail.univ.kiev.ua

На сьогоднішній день спостерігається підвищений інтерес до практичних досліджень амплітудних, частотних та температурних залежностей електрофізичних параметрів водних розчинів [2]. Як правило, такі розчини не є стійкими – їх параметри можуть сильно змінюватись у часі залежно від таких факторів як тиск, температура, вологість повітря та ін. Тому, зазвичай, їх консервують, а перед початком досліджень проводять атестацію їх основних параметрів неруйнівними експрес-методами [3].

В попередній роботі [4] було відмічено, що однією з методик експериментального дослідження властивостей таких розчинів може бути методика, що пов'язана з використанням резонатора поверхневої електромагнітної хвилі (РПХ). Відмітимо, що окрім водних розчинів та біоактивних рідин, об'єктами застосування такої методики також можуть бути різноманітні нелінійні інтегральні елементи [5, 8].

В даній роботі експериментально показано можливість використання НВЧ резонатора з добротністю $Q \sim 10^3$ (яким може бути РПХ) для збудження характеристичних резонансів у водних сольових розчинах. На основі отриманих при цьому експериментальних результатів можна визначити показник заломлення водного сольового розчину в НВЧ діапазоні [9]. В даній роботі автори обмежились дослідженнями в 8-мм діапазоні довжин хвиль водного розчину лише

однієї найбільш поширеної солі – хлориду натрію (NaCl).

В роботі використовувався НВЧ аналізатор на основі РПХ, що розміщувався в стандартному прямокутному хвилеводі 8-мм діапазону довжин хвиль [4]. Експериментальні дослідження проводилися за допомогою стандартного панорамного скалярного вимірювача Р2-65. В ході вимірювань використовувалися стандартні методи рефлектометрії, ослаблення та пробного тіла [1]. Детальний опис методики досліджень наведено в роботах [4, 6].

В роботі вимірювались власні добротності півхвильового РПХ з нанесеним на нього водним сольовим розчином. Частота і амплітуда резонансів у такій системі суттєво залежать від показника заломлення (діелектричної проникності, оскільки магнітна проникність $\mu \approx 1$) нанесеної рідини. Наприклад, було зафіксовано, що при заповненні капіляра водним сольовим розчином, виникають резонанси з власною добротністю $\sim 10^2$ і ослабленням НВЧ сигналу до 60 дБ. У той же час при порожньому капілярі ослаблення НВЧ сигналу було досить незначне (до 1 дБ) у всьому діапазоні робочих частот експериментальної установки (25,6 – 37,9 ГГц). Зв'язок РПХ із системою збудження регулювався шляхом зміни кута нахилу α між ним та широкою стінкою прямокутного хвилеводу (див. [4, 6]).

Як еталонні рідини використовувались водні розчини спиртів з відомими діелектричними проникностями [7]. Набір аналогічних експериментальних характеристик використовувався для калібрування та безпосереднього знаходження геометричного фактора для півхвильового РПХ [4]. Попередні експериментальні дослідження показали, що мінімальна зміна концентрації спиртового розчину, яку можна зареєструвати за допомогою НВЧ аналізатора на основі РПХ, складає $0,1 \div 0,5\%$.

Результати експериментальних досліджень наведено на рис. 1–3. З них випливає, що НВЧ аналізатор виявляється чутливим до змін внутрішньої структури води в розчинах на основі електролітів (речовин з різними знаками гідратації). Так, для хлоридів металів, що мають «додатній» знак гідратації, тобто що впорядковують структуру води (Na^+), відповідні резонансні частоти зсуваються в область більш високих частот відносно резонансної частоти для дистильованої води. Відповідно, можна припустити, що для хлоридів металів з «від'ємним» знаком гідратації, які руйнують структуру води (K^+ , Cs^+ , Cr^+ , Rb^+), резонансні частоти навпаки будуть зміщуватись в область більш низьких частот. Виявлений ефект гідратації проявляється для розчину хлориду натрію з концентрацією $0,1 \div 5$ моль/л. При більш високих концентраціях вочевидь стає необхідно враховувати ступінь впливу електропровідності, зміни густини та інших ефектів на параметри НВЧ аналізатора, що значно ускладнює проведення вимірів.

На рис. 1 наведено типову частотну залежність діелектричної сталої водного розчину хлориду натрію для трьох різних значень його концентрації. Видно, що з підвищенням частоти іонне поглинання радикалами різко зменшується. Натомість зростає аномальне поглинання, що викликане вивільненням іонів, які раніше були «зв'язаними» у розчині. Відомо, що максимум втрат відповідає критичній частоті $\omega_{\text{кр}}$ і визначає величину часу релаксації τ , згідно зі співвідношенням $\omega_{\text{кр}}\tau = 1$ [7]. Зазвичай, для визначення часу релаксації τ достатньо виміряти частотну залежність втрат від частоти – відповідну залежність для водного розчину хлориду натрію зображено на рис. 2.

З рис. 2 випливає, що зі зростанням концентрації розчину спостерігається одночасне збільшення його діелектричної проникності та тангенса кута втрат. Відповідно залежність електропровідності розчину хлориду натрію від

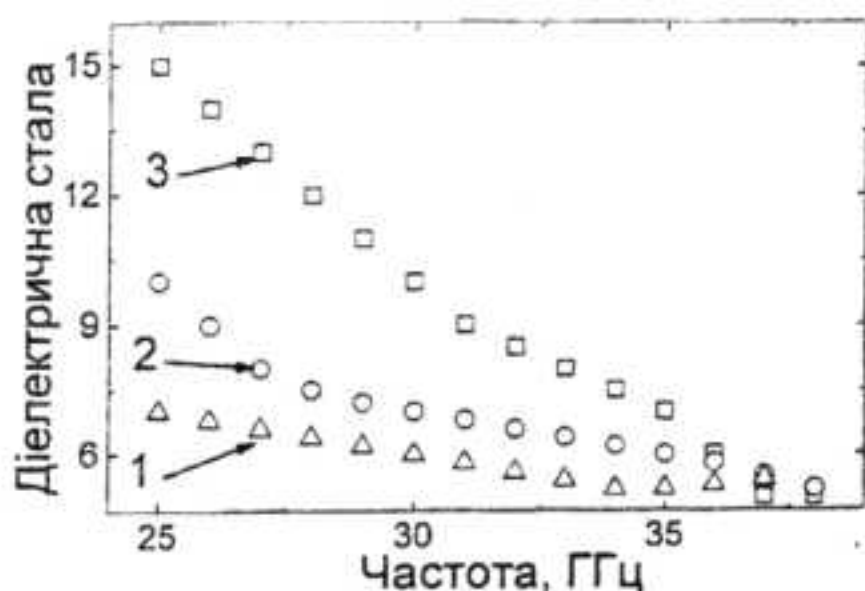


Рис. 1 Частотна залежність діелектричної сталої для водних розчинів хлориду натрію (NaCl) для трьох значень концентрації: 1 – 10^{-4} моль/л; 2 – 10^{-2} моль/л; 3 – 10^{-1} моль/л

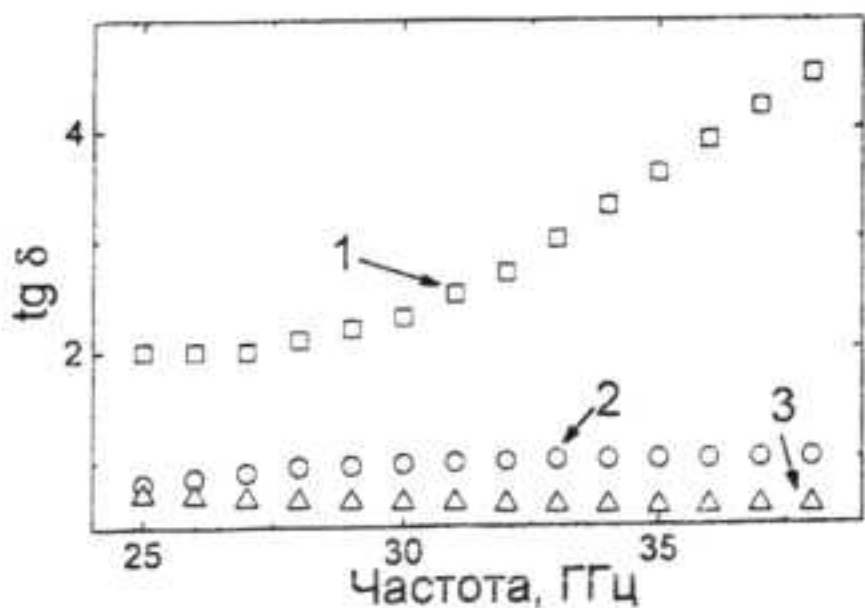


Рис. 2 Частотна залежність тангенса кута втрат для водних розчинів хлориду натрію (NaCl) для трьох значень концентрації: 1 – 10^{-4} моль/л; 2 – 10^{-2} моль/л; 3 – 10^{-1} моль/л

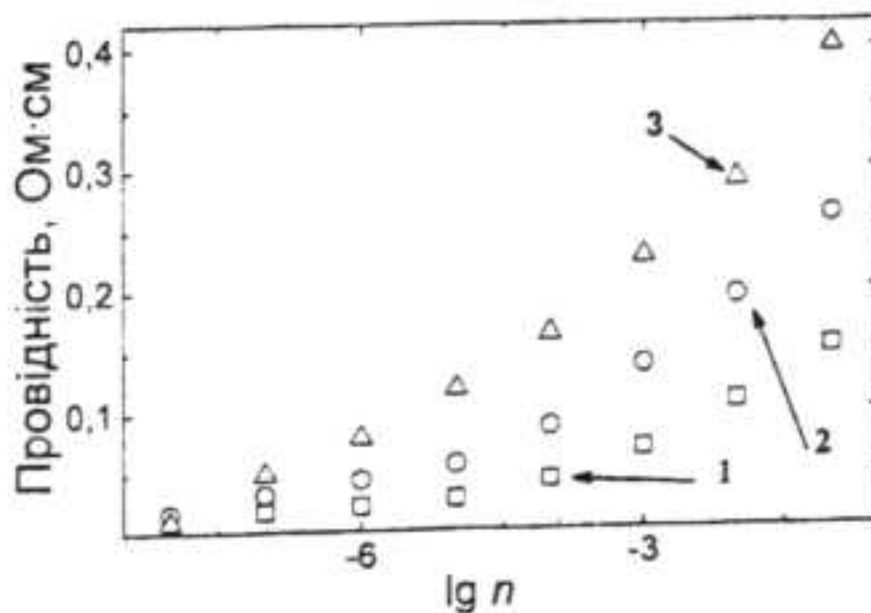


Рис. 3 Залежність електропровідності розчину хлориду натрію (NaCl) як функції логарифму концентрації n (моль/л) при різних частотах: 1 – 26 ГГц; 2 – 30 ГГц; 3 – 35 ГГц

логарифму концентрації (див. рис. 3) також швидко зростає зі збільшенням частоти.

Отримані експериментальні результати узгоджуються з даними, відомими з літератури [2, 7]. Отже, при цих концентраціях хлоридів металів у водних розчинах можна застосувати для їх вивчення НВЧ аналізатор на основі РПХ. Відповідна методика досліджень може бути використана для визначення основних властивостей водних сольових розчинів, наприклад, їх концентрації, провідності та ін.

При збільшенні концентрації водних сольових розчинів необхідно враховувати вплив інших чинників, зокрема тих, що були вже згадані, а також враховувати особливості поширення поверхневої електромагнітної хвилі в нелінійному (багатошаровому) середовищі.

Таким чином, у результаті досліджень водних розчинів встановлено, що розроблений НВЧ аналізатор дозволяє: 1) проводити порівняльні вимірювання діелектричної проникності сильно поглинаючих рідин ($\text{tg} \delta \sim 1$) з похибкою не більше $1 \div 10\%$; 2) визначити знак і інтегральну величину гідратації в водних розчинах малої концентрації ($0,1 \div 5$ моль/л); 3) контролювати зміни концентрації водних розчинів для широкого класу органічних речовин з роздільною здатністю не гірше $0,1 \div 1\%$.

Конструкція резонансної проточної комірки і методика вимірювань вихідного сигналу дозволяють автоматизувати процес вимірювання. Отже, на основі цього пристрою, можуть бути розроблені прилади для швидкого контролю технологічних процесів за участю водних сольових розчинів, а також методи неруйнівного контролю різноманітних матеріалів.

Список використаних джерел

1. Зайцев А. Н., Иващенко П. А., Мильников А. В. Измерения на СВЧ и их метрологическое обеспечение. Учебное пособие. – М.: Издательство стандартов, 1989. – 240 с.
2. Liu C., Pu Y. A Microstrip Resonator With Slotted Ground Plane for Complex Permittivity Measurements of Liquids // IEEE Trans. on Microwaves and Wireless Components Letters. – 2008. – 18, № 4. – P. 257-259.
3. Liquid handling // Nature Biotechnology. – 2005. – 23. – P. 1445-1453.
4. Іванюта О. М., Прокопенко О. В. НВЧ аналізатор біоактивних рідин // Вісник Київського університету. Серія: фізико-математичні науки. – 2007. – № 2. – С. 248-252.
5. Іванюта О. М., Кищенко Я. І., Прокопенко О. В., Ракша В. М. Система збудження джозефсонівських контактів у резонаторі поверхневої хвилі // Український фізичний журнал. – 2002. – 47, № 6. – С. 592-593.
6. Іванюта О. М., Прокопенко О. В., Ракша В. М., Ребіков А. І., Тягульский С. І. Вплив діелектричних підкладок на добротність напівхвильового резонатора поверхневої хвилі // Вісник Київського університету. Серія: фізико-математичні науки. – 2002. – № 2. – С. 328-336.
7. Девятков М. Д. Взаимодействие миллиметрового излучения с биологически активными соединениями и полярными жидкостями // Радиотехника и электроника. – 1978. – 23, № 9. – С. 1982-1985.
8. Prokopenko O. V., Borisenko M. O., Yudina O. S. Perspectives of application of surface wave resonators for creation of spin-transfer devices // Proceedings of the Seventh International Young Scientists' Conference on Applied Physics (June 13-15, 2007). – Kyiv (Ukraine), 2007. – P. 165-166.
9. Melkov G. A., Prokopenko O. V., Malyshev V. Yu. A new method of complex dielectric permittivity determination // Proceedings of the Sixth International Kharkov Symposium on Physics and Engineering of Microwaves, Millimeter, and Submillimeter Waves and Workshop on Terahertz Technologies (June 25-30, 2007). – Kharkov (Ukraine), 2007. – 2. – P. 819-821.

Надійшла до редколегії 20.04.2008

УДК 621.386.834; 622.383

Ільченко В.В., канд. фіз.-мат. наук

Моделювання рекомбінаційного струму в діодах Шотткі через шари квантових точок двох різних типів

В роботі проведено числове моделювання рекомбінаційного струму для структур з бар'єром Шотткі через шари Ge квантових точок в Si та шари InAs квантових точок в GaAs.

Показано, що для реальних значень енергій залягання рівнів квантових точок та інших параметрів просторового їх розміщення у структурах з Ge квантовими точками в Si величина рекомбінаційного струму може на декілька порядків перевищувати надбар'єрний струм при невеликих прямих зміщеннях.

В цей же час, рекомбінаційний струм для структур з шарами InAs квантових точок в GaAs може домінувати над надбар'єрним у цих структурах лише за особливих умов досить глибоких рівнів квантових точок.

Ключові слова: квантова точка, рекомбінація, перетин захоплення, вольт-амперна характеристика.

E-mail: ilch@univ.kiev.ua

Вступ. Дослідженням структур з бар'єром Шотткі з шарами квантових точок (КТ) в останні роки приділяється велика увага. Доволі проста технологія створення металевих контактів, що утворюють бар'єр біля поверхні напівпровідника, дозволяє відносно просто отримати напівпровідникові структури, де шари КТ будуть знаходитись в області вбудованого електричного поля і за рахунок накопичення заряду в шарах КТ можуть впливати на профіль потенціалу в області просторового заряду (ОПЗ). Цей факт є обов'язковою умовою для виникнення таких цікавих, з точки зору практичного застосування, ефектів, якими є ефекти пам'яті, явище негативної диференціальної ємності, ефекти фото- та електролюмінісценції [2-4].

На сьогоднішній день добре відомо, що досить великий електричний заряд, що накопичується в шарах КТ, може впливати на проходження надбар'єрного струму крізь такий діод [5,6]. Але крім цього, квантові стани КТ, що слугують центрами пересадки для носіїв заряду,

Ilchenko V.V., Ph.D.

The modeling of the recombination current of the Schottky diode through two different types of quantum dots layers

The numerical modeling of the recombination current for the structures with Schottky barrier through the Ge quantum dots layer in Si and InAs quantum dots layer in GaAs has been done in present paper.

It was shown, that the value of recombination current in the structures with Ge quantum dots layer in Si can be essentially bigger on several order then current going over the barrier for the typical parameters of the energetic position levels of Ge quantum dots at relatively small forward bias.

At the same time, the recombination current in the structures with InAs quantum dots layer in the GaAs can be dominating on the current going over the barrier for this structures at the special conditions with relatively deep energetic position of the level of quantum dots.

Key Words: quantum dot, recombination, capture cross section, I-V curve.

можуть також брати участь в процесах захоплення електронів та дірок і, таким чином, суттєво змінювати як величину самого рекомбінаційного струму через контакт, так і його відносну частку в сумарному струмі через всю структуру.

В літературі вже існує кілька робіт [1,7], в яких проводилися обчислення рекомбінаційних струмів у контактах метал-напівпровідник з глибокими рівнями, розташованими в області просторового заряду, де процес рекомбінації може відбуватися за схожим механізмом, що і в структурах з шарами КТ. Так, подібна модель, що базувалася на статистичних співвідношеннях Шоклі-Ріда, широко використовувалася в літературі для теоретичних розрахунків рекомбінаційних струмів у контакті метал-напівпровідник та підтвердила свою працездатність добрим узгодженням теоретичних результатів та експериментальних даних [7].

Слід зазначити, що формалізм обчислення рекомбінаційних струмів у термінах перетинів

захоплення практично не залежить від природи походження енергетичних рівнів у забороненій зоні напівпровідника за умов, коли накопиченням заряду, захопленого на цих рівнях, можна знехтувати. Якщо рекомбінаційний струм досить значний, його проходження крізь досліджувану структуру заважає накопиченню заряду. Тож використання узагальненої моделі Шоклі-Ріда для випадку, коли в забороненій зоні напівпровідника знаходиться не один, а досить багато рівнів, розподілених за енергією та координатою, є цілком природним [7].

Наш випадок суттєво відрізняється лише тим, що шари КТ мають певну, доволі невелику, товщину, та енергетичні рівні КТ розміщуються в паралельних металевому контакту площинах на певних технологічно точно витриманих відстанях від міжфазної границі метал-напівпровідник. Кожний з цих рівнів може досить ефективно обмінюватися носіями заряду як із зоною провідності, так і з валентною зоною напівпровідника, чим створює передумови для проходження через структуру певної складової рекомбінаційного струму крізь відповідні квантові стани КТ. Беручи до уваги той факт, що шари КТ мають технологічно задані координати свого розташування, а відповідні енергетичні рівні КТ - типові значення енергій, дані розрахунки було відносно легко модифікувати для нашого випадку, врахувавши при цьому специфіку матеріалу, де розміщуються шари КТ, їх тип та певні особливості процесів захоплення носіїв заряду в КТ.

В даній роботі проводився числовий аналіз величини рекомбінаційного струму для двох типів структур з бар'єром Шотткі з шарами Ge КТ в Si та з шарами InAs КТ в GaAs.

Хід обчислень. На рис. 1,а та 1,б представлені зонні діаграми цих структур з трьома шарами КТ.

З літератури відомо [8], що в структурі з Ge КТ в Si (так звані КТ II типу) мають місце глибокі квантові ями поблизу валентної зони, в яких утворюються квантові рівні, що можуть легко обмінюватися захопленими в них носіями з валентною зоною. Слід зазначити, що в структурах на основі Si потенціальні ями для електронів поблизу зони провідності не утворюються, навпаки, в цих же перерізах, де виникають квантові ями для дірок, за рахунок стрибка потенціалу дна зони провідності між Si та Ge, навпаки, утворюються невеличкі потенціальні бар'єри для електронів (Див. рис.1,а).

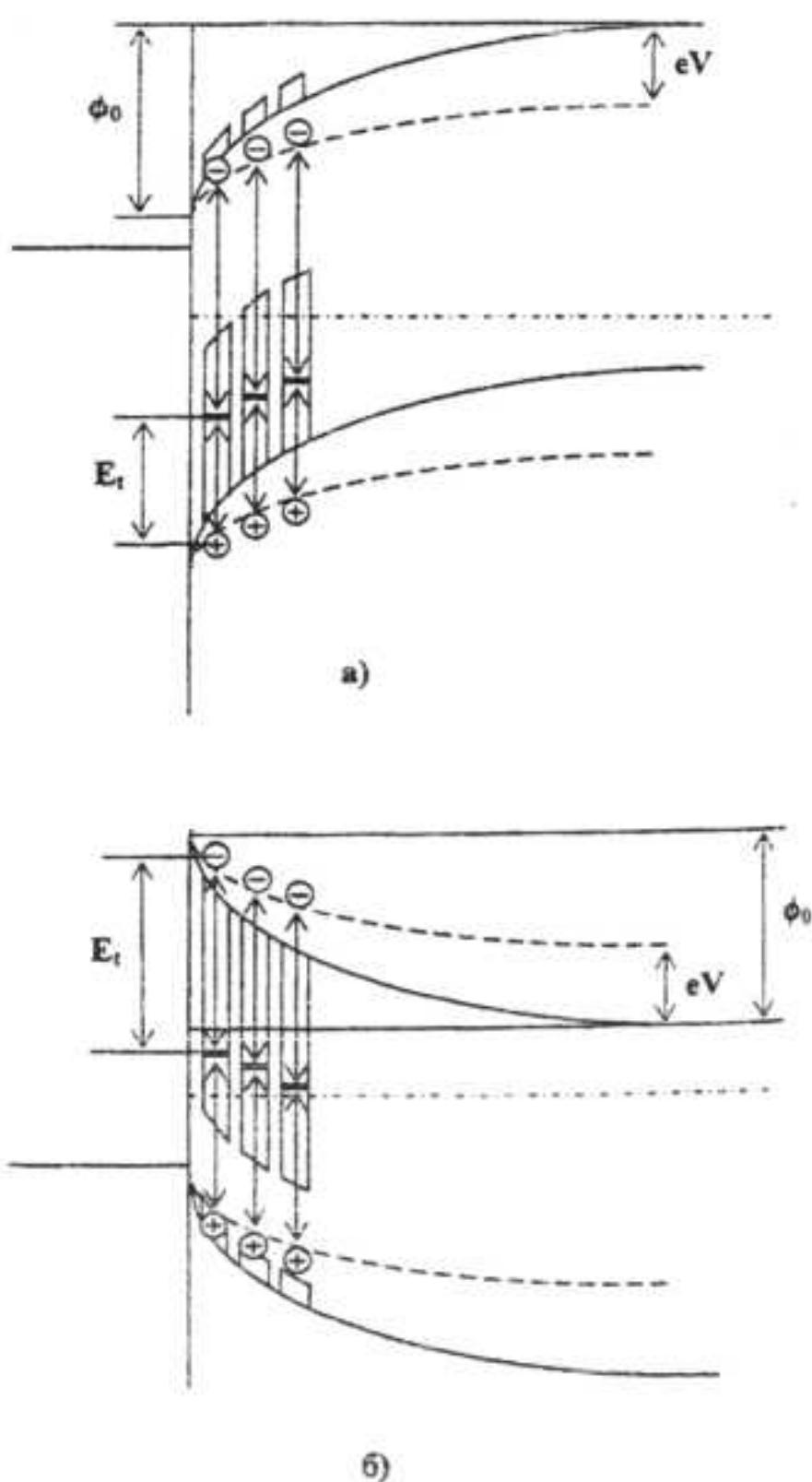


Рис.1 Зонні діаграми структур з бар'єром Шотткі: а) з шарами Ge КТ в Si та б) з шарами InAs КТ в GaAs.

Оскільки висота цих бар'єрів по величині більш як на порядок менше ширини забороненої зони Si та GaAs, їх вплив на концентрації вільних електронів в цих перерізах при кімнатних температурах є незначним і в подальших розрахунках не враховувався. В структурах з InAs КТ в GaAs (так звані КТ I типу) в забороненій зоні утворюються квантові ями як поблизу валентної, так і поблизу зони провідності [9] (Див. рис 1,б). В них теоретично можуть бути локалізовані і електрони, і дірки. Проте внаслідок того, що енергетичний рівень у потенціальній ямі InAs поблизу валентної зони знаходиться дуже близько до межі валентної зони GaAs, вплив цього квантового рівня на рекомбінаційний струм при достатньо високих(кімнатних) температурах у даній задачі теж дуже слабкий, і ним можна знехтувати. Тож

для цього типу структур найбільший вплив на рекомбінаційний струм у контакті обумовлюють квантові рівні, що виникають в більш глибоких потенціальних ямах, що спричинені стрибком дна зони провідності InAs в GaAs.

Таким чином, для обох типів структур, що розглядалися в даній роботі, суттєвий вплив на рекомбінаційний струм створює лише один тип енергетичних рівнів, що виникають за рахунок розмірного квантування в КТ (утворений за рахунок стрибка потенціалу або валентної або зони провідності). Для обох типів досліджуваних у даній роботі структур рекомбінаційний струм може бути представлений у вигляді суми кількох складових, кожна з яких пов'язана із своїм шаром КТ.

При цьому загальний вираз кожної складової рекомбінаційного струму може бути записаний у вигляді [7]:

$$I_{rec} = \frac{qN_{dot}C_nC_p(n(L_{dot},V)p(L_{dot},V) - n_i^2)}{C_n(n(L_{dot},V) + n_1) + C_p(p(L_{dot},V) + p_1)},$$

де N_{dot} - нормована густина КТ в площині шару;
 $C_n = \sigma_n v_{thn}$, $C_p = \sigma_p v_{thp}$ - коефіцієнти

захоплення для електронів та дірок; $n(L_{dot},V)$, $p(L_{dot},V)$ - концентрації вільних електронів та дірок поблизу шарів КТ,

$$n_1 = N_c \exp\left(-\frac{E_g - E_{dot}}{kT}\right), \quad p_1 = N_v \exp\left(-\frac{E_{dot}}{kT}\right) -$$

концентрації електронів та дірок безпосередньо на рівнях КТ, L_{dot} - відстань шару КТ від границі контакту метал-напівпровідник, σ_n , σ_p - перетини захоплення на рівні КТ електронів та дірок; v_{thn} , v_{thp} - теплові швидкості носіїв; E_{dot} - енергетичне положення рівнів КТ; решта позначень - загальновідомі.

В подальшому для розрахунків використовувалися стандартні вирази для форми профілю потенціального бар'єра. Цим самим в роботі нехтувалося впливом ефекту накопичення заряду в шарах КТ на форму профіля потенціального бар'єру, оскільки вважалося, що рекомбінаційний струм, що проходить крізь енергетичні рівні КТ заважає такому накопиченню.

Проте слід зазначити ряд відмінностей у розрахунках рекомбінаційного струму в досліджуваних структурах, що були характерні для різних типів КТ. Так, для структур з Ge КТ в Si концентрації вільних носіїв в тих перерізах структури, де були розміщені шари КТ, можуть

бути записані за допомогою наступних співвідношень: $n(L_{dot}) = n_0 \exp \frac{\varphi(L_{dot},V) + eV}{kT}$

$$\text{та } p(L_{dot}) = p_0 \exp - \frac{\varphi(L_{dot})}{kT}.$$

В той же час для структур з шарами InAs КТ в GaAs аналогічні концентрації вільних носіїв поблизу шарів КТ набуватимуть дещо іншого вигляду: $n(L_{dot}) = n_0 \exp \frac{\varphi(L_{dot},V)}{kT}$ та

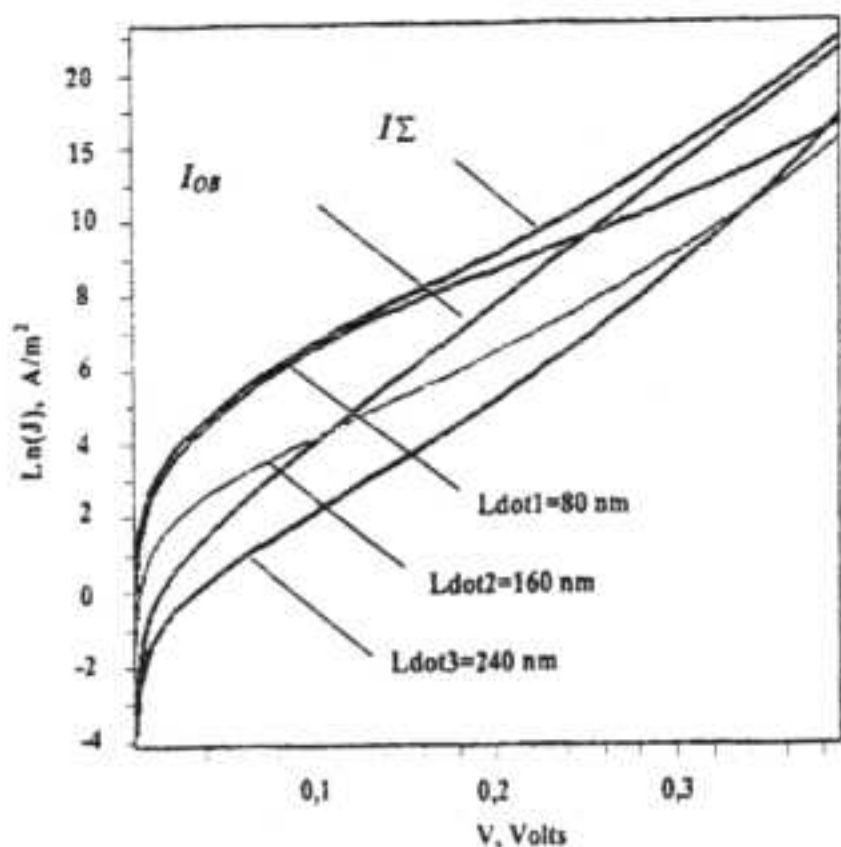
$$p(L_{dot}) = p_0 \exp \frac{\varphi(L_{dot}) + eV}{kT}.$$

Крім того, на результати обчислень будуть суттєво впливати такі параметри структур як ширини заборонених зон матеріалів, в які вбудовуються шари квантових точок. З літератури відомо, що темп рекомбінації через локальний рівень в забороненій зоні тим більший, чим ближче цей рівень знаходиться до середини забороненої зони. Оскільки глибина енергетичних рівнів КТ в обох типах матеріалів приблизно однакова і становить близько 0,25-0,35eV, збільшення ширини забороненої зони для GaAs помітно послаблювало темп рекомбінації через рівні КТ і, відповідно, зменшувало величини рекомбінаційних струмів.

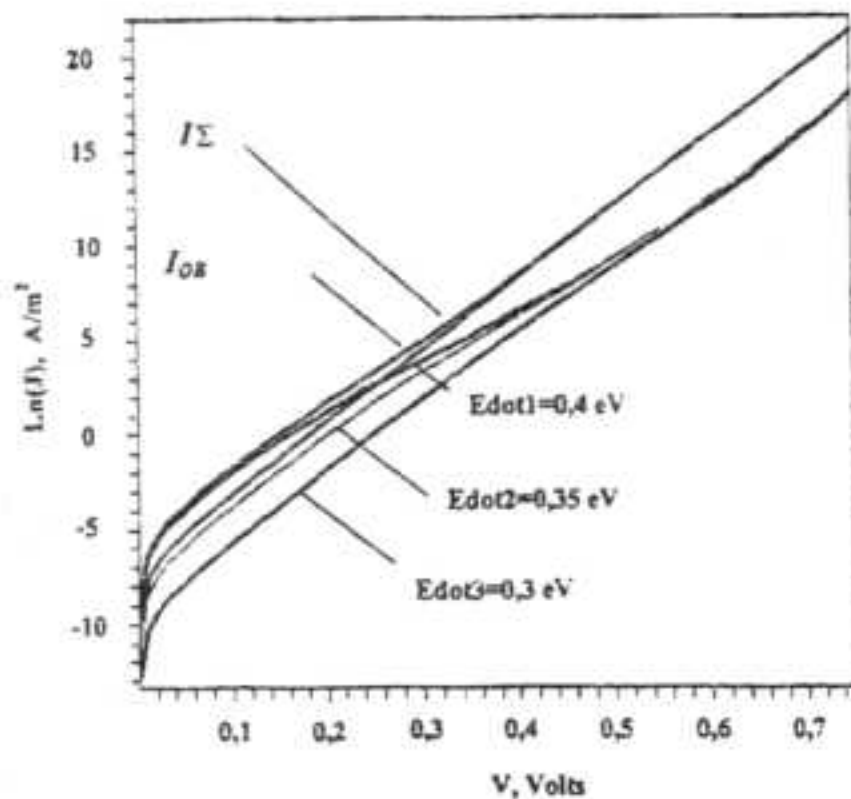
Для порівняння величини рекомбінаційного струму з надбар'єрним струмом, що проходить через контакт метал-напівпровідник, для кожного типу досліджуваних в роботі структур розраховувалася величина надбар'єрного струму для тих самих параметрів напівпровідникової структури, що і при обчисленнях рекомбінаційного струму. Дані обчислення проводились за формулою, яка у випадку діода Шотткі на основі p-кремнію записується у вигляді[1]:

$$I_{ov} = qp_0\mu_pF_0 \exp\left(-\frac{\varphi_0}{kT}\right)\left(\exp \frac{qV}{kT} - 1\right),$$

де F_0 - величина електричного поля на границі контакту метал-напівпровідник; φ_0 - висота потенціального бар'єру на границі метал-напівпровідник для Si ~0,5eV; μ_p - рухливість дірок в Si; решта позначень загальновідомі. Аналогічним чином може бути розрахована величина надбар'єрного струму для діода Шотткі на основі GaAs, але для цього необхідно в виразі для струму використати концентрації вільних електронів n_0 , їх рухливість μ_n та характерну для GaAs дещо більшу, ніж для Si, висоту потенціального бар'єра.



a)



б)

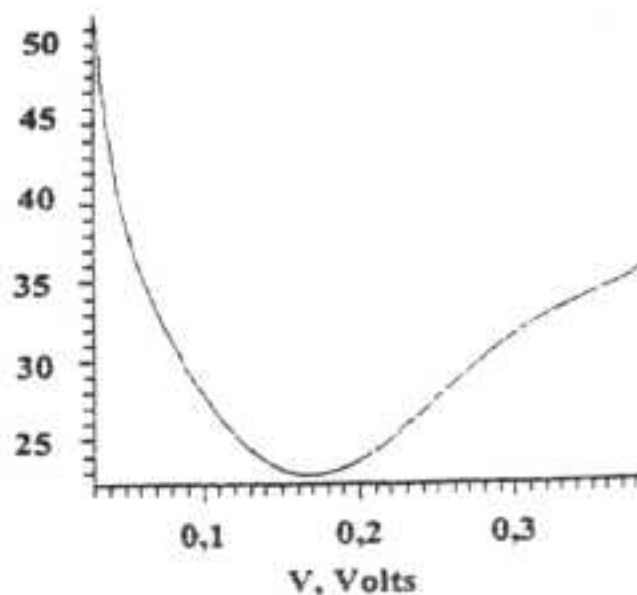
Рис.2 Вольт-амперні характеристики розраховані для двох типів структур з бар'єром Шотткі: а) з шарами Ge КТ в Si для $E_{dot} = 0,4\text{eV}$ та трьох значень координат розташування шарів КТ: $L_{dot1} = 80\text{ nm}$, $L_{dot2} = 160\text{ nm}$, $L_{dot3} = 240\text{ nm}$; б) з шарами InAs КТ в GaAs для однієї координати шару КТ $L_{dot} = 250\text{ nm}$ та трьох значень енергій $E_{dot1} = 0,4\text{eV}$, $E_{dot2} = 0,35\text{eV}$, $E_{dot3} = 0,3\text{eV}$.

Слід зазначити, що рухливість вільних електронів в GaAs майже на півтора порядки перевищує аналогічну величину для дірок в Si. Хоча збільшення рухливості в структурах на основі GaAs значною мірою компенсується більшою висотою бар'єра $\sim 0,95\text{--}1,05\text{eV}$ на

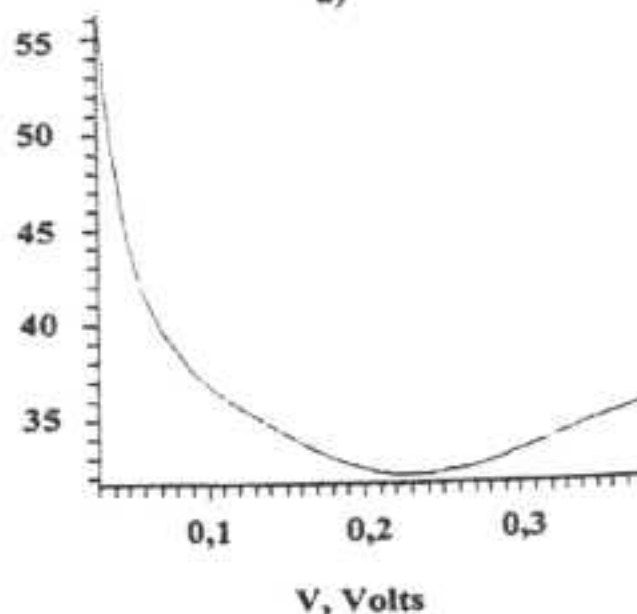
границі метал-напівпровідник. Загалом вираз для надбар'єрного струму діода Шотткі в цьому випадку буде мати вигляд [1]:

$$I_{OB} = qn_0\mu_n F_0 \exp\left(-\frac{\phi_0}{kT}\right) \left(\exp\frac{qV}{kT} - 1\right).$$

Для наочності на рис.2,а та 2,б наведено окремо розраховані рекомбінаційні складові струму, залежно від розміщення шарів КТ, надбар'єрні та сумарні струми, залежно від прикладених до досліджуваних структур напруг.



а)



б)

Рис.3 Залежність коефіцієнтів нелінійності діодів Шотткі від прикладеної напруги для структур з бар'єром Шотткі а) з шарами Ge КТ в Si та б) з шарами InAs КТ в GaAs.

На основі проведених у роботі розрахунків рекомбінаційних та надбар'єрних струмів обчислено залежності коефіцієнтів неідеальності прямих гілок вольт-амперних характеристик

(ВАХ) діодів Шотткі $\alpha = \frac{d \ln I_{\Sigma}}{dV}$ для обох типів

структур. (Див. рис. 3 а, б)).

Аналіз отриманих результатів. На рис.2а подана ВАХ Si діода з бар'єром Шотткі з трьома шарами Ge квантових точок, розташованих на різних відстанях від границі метал-

напівпровідник. З проведеного розрахунку видно, що рекомбінаційний струм через шари КТ може дуже суттєво змінюватись залежно від координат їх розташування, так що при певних положеннях шарів Ge КТ в Si рекомбінаційний струм може досягати значної величини, що на декілька порядків перевищує надбар'єрний струм у цих структурах.

Обчислення коефіцієнта нелінійності для таких діодів, залежно від прямої прикладеної напруги, вказують на існування характерного мінімуму в поведінці цього параметру із зміною напруги. Це відбувається саме в той момент, коли рекомбінаційний струм стає по величині рівним струму надбар'єрному.

Розрахунки рекомбінаційного струму для GaAs діодів з шарами InAs квантових точок дають порівняно невелику величину струму. На рис. 26 наведено результати розрахунків, зроблених для одного шару квантових точок, розмішених у тому перерізі, де темп рекомбінації крізь квантові точки є найбільший. Не дивлячись на це, для мілких рівнів квантових точок рекомбінаційний процес у таких діодах дає зовсім незначний вклад в сумарний струм через діод. Про це ж свідчить значно менший за своєю амплітудою мінімум в залежності коефіцієнту нелінійності від напруги для прямої гілки ВАХ (Див. рис. 36). Як показують чисельні оцінки, рекомбінаційний струм в даних структурах може домінувати над надбар'єрним струмом лише для досить глибоких рівнів квантових точок з енергією залягання більше за 0,4 еВ. Слід також зазначити, що рекомбінаційні струми крізь квантові точки в таких структурах можуть бути більш помітними при досить низьких температурах.

Висновки. Загалом, на підставі проведеного моделювання можна зробити висновок про те, що за рахунок рекомбінації крізь Ge квантові точки в структурах на основі Si можуть виникати значні струми через діод, що по своїй величині здатні значно перевищувати надбар'єрні в певному діапазоні (0-0,25В) прикладених напруг. У той же час більша ширина забороненої зони в GaAs та порівняно більша рухливість електронів не дозволяють рекомбінаційному струму так само легко, як в Si, конкурувати з надбар'єрним струмом за типових значень енергій залягання (0,25-0,35еВ) енергетичних рівнів КТ.

Дана робота була виконана за підтримки українсько-російської наукової програми „Нанофізика та наноелектроніка”, грант МОНУ М/175-2007р.

Список використаних джерел

1. *Стриха В.И.* Теоретические основы работы контакта металл-полупроводник.- Киев: Наукова думка, 1974.- 263 с.
2. *Brunches T., Boucaud P., Sauvage S.* Electroluminescence of Ge/Si self-assembled quantum dots grown by chemical vapor deposition // *Appl. Phys. Lett.*-2000.- 77, No.12.- P.1822-1824.
3. *Peng C.S., Huang Q., Zhang Y.H.* New optical properties of Ge self-organized quantum dots // *Thin Solid Films.*-1998.- No. 323.- P. 174-177.
4. *Леденцов Н.Н., Устинов В.М., Щукин В.А., Копьев П.С., Алферов Ж.И., Бимберг Д.* Гетероструктуры с квантовыми точками: получение, свойства, лазеры.Обзор // *ФТП.*-1998.-32, №4.-С. 385-410.
5. *Chiquito A. J., Pusep Yu. A., Mergulhao S., and Galzerani J. C.* Capacitance-voltage profile in a structure with negative differential capacitance caused by the presence of InAs/GaAs self-assembled quantum dots // *Phys.Rev.* - 2000.- 61.-P. 5499 -5504.
6. *Lin S.D., Ilchenko V.V., Marin V.V.et al.* Observation of the negative differential capacitance in Schottky diodes with InAs quantum dots near room temperature // *Appl. Phys. Lett.*-2007.- 90.- P. 263114(1-3).
7. *Ильченко В.В., Левандовский В.Г., Стриха В.И.* Теоретический анализ генерационно-рекомбинационного тока в контакте металл-аморфный кремний // *УФЖ.*- 1987.- 32, №2.- С.290-295.
8. *Yakimov A I., Dvurechenskij A V., Nikiforov A.I.* Germanium Self-Assembled Quantum Dots on Silicon: Growth, Electric Transport, Optical Phenomena, and Devices // In: *Handbook of Semiconductor Nanostructures and Nanodevices.* Edited by A. A. Balandin and K. I. Wang. American Scientific Publishers 2006.- 1.- P. 33-102.
9. *Lu L., Wang J., Wang Y., Ge W., Yang G., Wang Z.* Conduction-band offset in a pseudomorphic GaAs/In_{0.2}Ga_{0.8} quantum well determined by capacitance-voltage profiling and deep-level transient spectroscopy techniques // *J. Appl. Phys.*- 1998.- 83, No.4.- P. 2093-2097.

Надійшла до редколегії 31.03.08

Косаревич Б.П., студент,
Анісімов І.О.* , д-р. фіз.-мат. наук, проф.

Числовий розрахунок характеристик темного розряду

Наведені результати числового розрахунку характеристик темного розряду на основі розв'язання системи самоузгоджених гідродинамічних рівнянь. Результати числового розрахунку порівнюються з відомими аналітичними розрахунками на основі спрощених моделей. Показано, що при великих довжинах міжелектродного проміжку вольт-амперна характеристика розряду характеризується від'ємним диференціальним опором.

Ключові слова: темний розряд, просторовий електричний заряд, від'ємний диференціальний опір.

E-mail: *ioa@mail.univ.kiev.ua

1. Вступ

Темний таунсендівський розряд серед усіх газових розрядів постійного струму характеризується найменшими струмами і найпростішою структурою – в ньому, на відміну від жевріючого та дугового розрядів, відсутні прикатодний шар і позитивний стовп. Тому його вивчення являє значний методичний інтерес. В той же час роботи, присвячені дослідженню цього розряду, в періодичній літературі практично відсутні.

Відомий аналітичний розрахунок цього розряду [1] є внутрішньо суперечливим: густини зарядів і струмів електронів та іонів знаходяться без урахування просторового заряду, а при розрахунку розподілів електричного поля в міжелектродному проміжку просторовий заряд вважається однорідним, що не відповідає дійсності. Аналітично ж розв'язати самоузгоджену систему рівнянь, що враховує неоднорідність просторового заряду в міжелектродному проміжку, не вдається.

Метою цієї роботи є числовий розрахунок характеристик темного розряду в гідродинамічному наближенні та порівняння отриманих результатів із результатами спрощених аналітичних розрахунків.

2. Вихідні рівняння

Проведемо самоузгоджений розрахунок темного розряду, тобто врахуємо одночасно

Kosarevych B.P., student,
Anisimov I.O.* , Dr. Sci., prof.

Numerical computation of characteristic curves for dark discharge

Characteristic curves for dark discharge were computed via numerical solution of the set of self-consistent hydrodynamic equations. Results of the computation were compared with the well-known analytic solutions obtained from the simplified models. The possibility to obtain the negative differential resistance of the discharge volt-ampere characteristic for long distances between electrodes was demonstrated.

Key Words: dark discharge, space electric charge, negative differential resistance.

просторовий заряд і неоднорідне електричне поле, спричинене цим зарядом. Розглядатимемо одновимірну модель.

Запишемо систему рівнянь у гідродинамічному наближенні, що описують розряд. Вона включає рівняння для повного струму, густина якого в стаціонарному режимі не повинна залежати від координати,

$$[n_i(x)\mu_i + n_e(x)\mu_e]eE(x) = j, \quad (1)$$

рівняння для розподілу густини електронів, обумовленої іонізацією електронним ударом,

$$\frac{dn_e}{dx} = \alpha(E)n_e, \quad (2)$$

а також теорему Гаусса

$$\frac{dE}{dx} = \frac{e}{\epsilon_0} [n_e(x) - n_i(x)]. \quad (3)$$

Тут $n_{e,i}$ – відповідно концентрація електронів та іонів, $\mu_{e,i}$ – відповідно рухливість електронів та іонів, E – напруженість електричного поля, ϵ_0 – діелектрична проникність вакууму, e – елементарний заряд, $\alpha(E)$ – коефіцієнт Таунсенда, що визначається формулою

$$\alpha(E) = pA \exp\left(-\frac{pB}{E}\right), \quad (4)$$

де p – тиск газу, A, B – параметри, що залежать від сорту газу. Вважатимемо тиск по всьому об'єму однаковим, оскільки ступінь іонізації газу в темному розряді дуже малий.

Виключивши за допомогою (1) концентрацію електронів, отримаємо замість рівняння (3) співвідношення

$$\frac{dE}{dx} = \frac{1}{\epsilon_0} \left[\left(1 + \frac{\mu_e}{\mu_i} \right) e n_e - \frac{j}{\mu_i E} \right]. \quad (5)$$

Щоб розв'язати систему (2), (5), її необхідно доповнити граничними умовами.

Одна з граничних умов впливає з того, що єдиним джерелом електронів з катода є вторинна іон-електронна емісія, тобто

$$j_e(x=0) = \gamma j_i(x=0), \quad (6)$$

де γ – коефіцієнт вторинної іон-електронної емісії. Підставивши значення густин електронного та іонного струму і виключивши концентрацію іонів за допомогою (1), рівняння (6) можна переписати у формі:

$$n_e(x=0) = \frac{\gamma j}{(\gamma + 1) e \mu_e E(x=0)}. \quad (7)$$

Таблиця 1. Параметри розрахунку

Параметр	Числове значення	Примітка
P	0.1 Торр	
D	9 см	
A	1360 м ⁻¹	
B	2.35 · 10 ⁴ В/(м·Торр)	для E/p=(10÷60) В/(м·Торр)
γ	0,058	для E·p=(100÷1000) В·Торр/м
μ_e	36 м ² ·Торр/(В·с)	
μ_{Ag}	0,16 м ² ·Торр/(В·с)	

З урахуванням умов на E/p та pE електричне поле E при обраному тиску має знаходитися в інтервалі (1÷6) кВ/м. Ця умова, в свою чергу, приводить до обмеження можливих густин струму згори значенням 5 · 10⁻³ А/м².

Як показують оцінки, для обраних параметрів типові значення довжин вільного пробігу електронів та іонів лежать у межах від частки міліметра до кількох міліметрів, тобто виявляється принаймні на 1-2 порядки меншими від довжини міжелектродного проміжку. Таким чином, застосування гідродинамічної моделі для опису розряду при обраних нами параметрах є цілком припустимим.

4. Результати розрахунку

В результаті розрахунку були отримані розподіли в міжелектродному проміжку електричного поля, концентрації електронів та іонів, густин іонного та електронного струму, а також "вольт-амперна характеристика" розряду

Друга гранична умова випливає з того, що поблизу аноду іонний струм відсутній, а, отже, густина іонів дорівнює нулеві:

$$n_i(x=d) = 0, \quad (8)$$

де d – довжина міжелектродного проміжку.

Для отримання числового розв'язку системи (2), (5) із граничними умовами (7)-(8) необхідно ще задатися числовими значеннями параметрів системи.

3. Числові параметри

За робочий газ оберемо аргон. Електроди вважатимемо мідними. Тиск та віддаль між електродами оберемо типовими для лабораторних експериментів. Параметри, необхідні для розрахунку, подано в таблиці 1 [2].

(точніше, залежність падіння напруги на розряді від густини струму, яка задавалася).

При малих значеннях струму ($j < 10^{-5}$ А/м²), коли просторовий заряд у міжелектродному проміжку є незначним і не створює електричне поле, розподіли концентрації іонів і електронів збігаються з результатами аналітичного розрахунку [1].

При більших густинах струму просторовий заряд, створений переважно іонами Ag⁺, створює однорідне електричне поле. Чим більша густина струму, тим більші концентрації іонів і, відповідно, просторовий заряд.

Розберемо отримані залежності більш детально.

4.1. Міжелектродний розподіл електричного поля

Розподіл електричного поля в міжелектродному проміжку для різних густин струму (в межах від 10⁻⁴ А/м² до 5 · 10⁻³ А/м²) поданий на рис.1. Видно, що зі збільшенням густини стру-

му напруженість поля біля катоду зростає, а біля анода – зменшується. Це обумовлено неоднорідністю концентрації іонів вздовж розрядного проміжку. Слід відзначити, що друга похідна $d^2E(x)/dx^2$ в отриманому розрахунку є додатною, а в аналітичному розрахунку [1] вона є від'ємною.

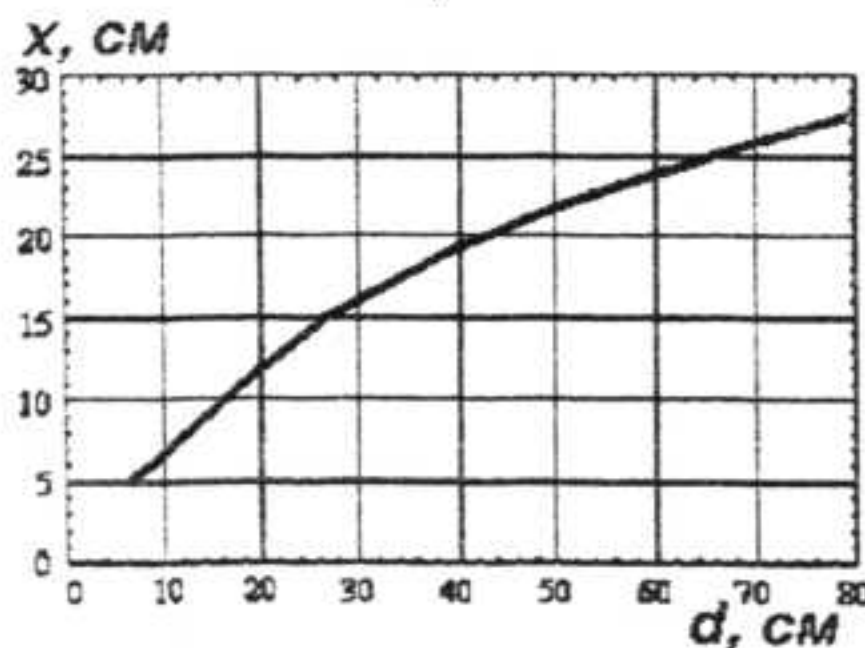
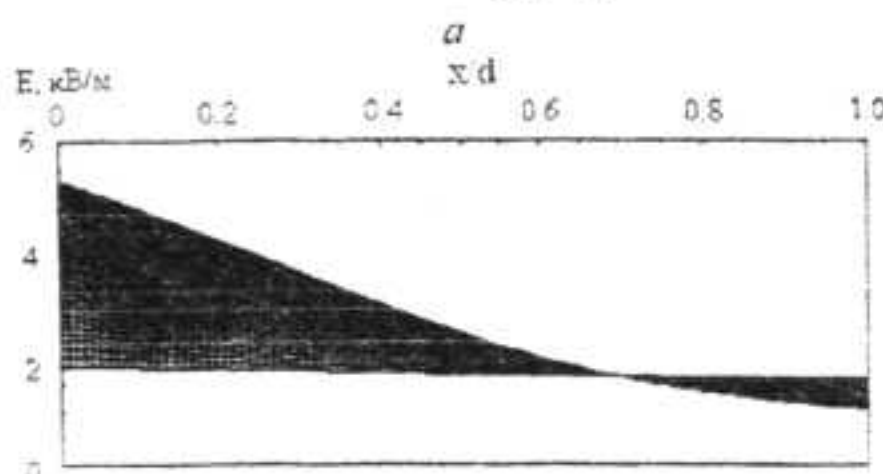
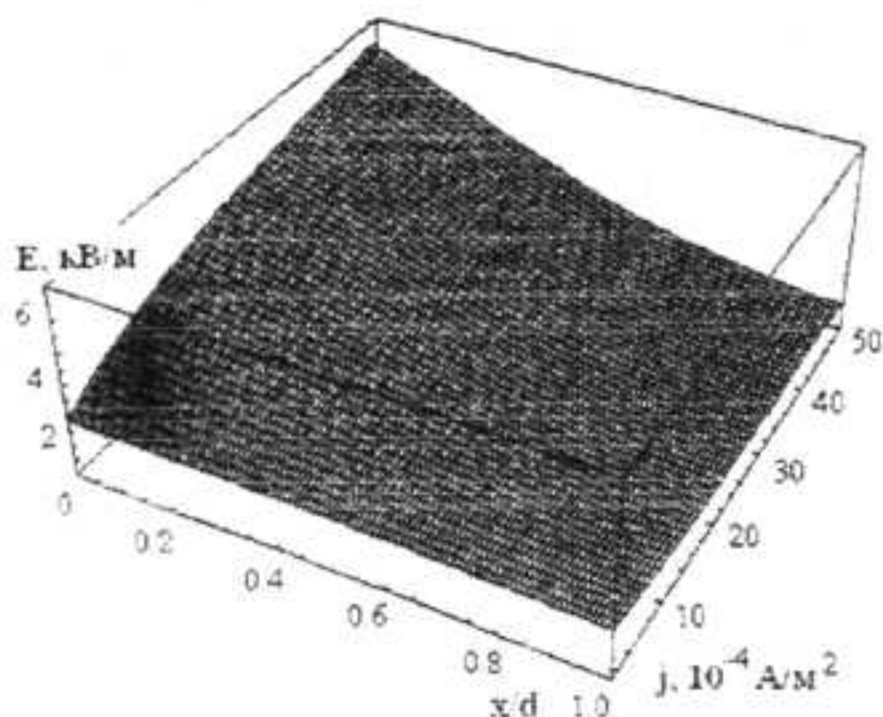


Рис.1: а, б – розподіл електричного поля в міжелектродному проміжку для різних густин струму; в – залежність віддалі від катоду до точки сталого поля від довжини міжелектродного проміжку при $j=10^{-4}$ А/м².

Як видно з рис.1 б, при малих густинах струму в міжелектродному проміжку існує така точка (назвемо її точкою сталого поля), де

електричне поле залишається практично незмінним при зміні густини струму до $5 \cdot 10^{-3}$ А/м². При більших густинах струму вона поступово зсувається до катоду.

На рис.1 в подано залежність координати точки сталого поля від довжини міжелектродного проміжку d . При $d \approx 35$ см точка сталого поля знаходиться посередині між катодом і анодом. При меншій міжелектродній відстані ця точка знаходиться ближче до катоду, а при більшій міжелектродній відстані – ближче до анода.

4.2. Вольт-амперна характеристика

Вольт-амперна характеристика (точніше, залежність падіння напруги на розряді від густини струму), побудована за результатами моделювання, показана на рис.2. Зі зростанням густини струму напруга на розряді також зростає, тобто в аналізованому режимі розрядний проміжок характеризується додатним диференціальним опором.

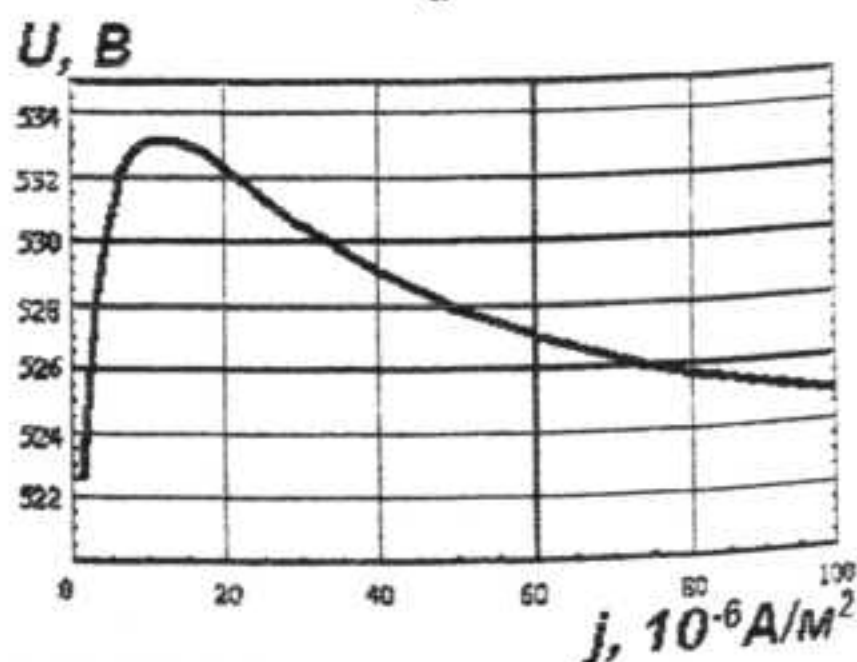
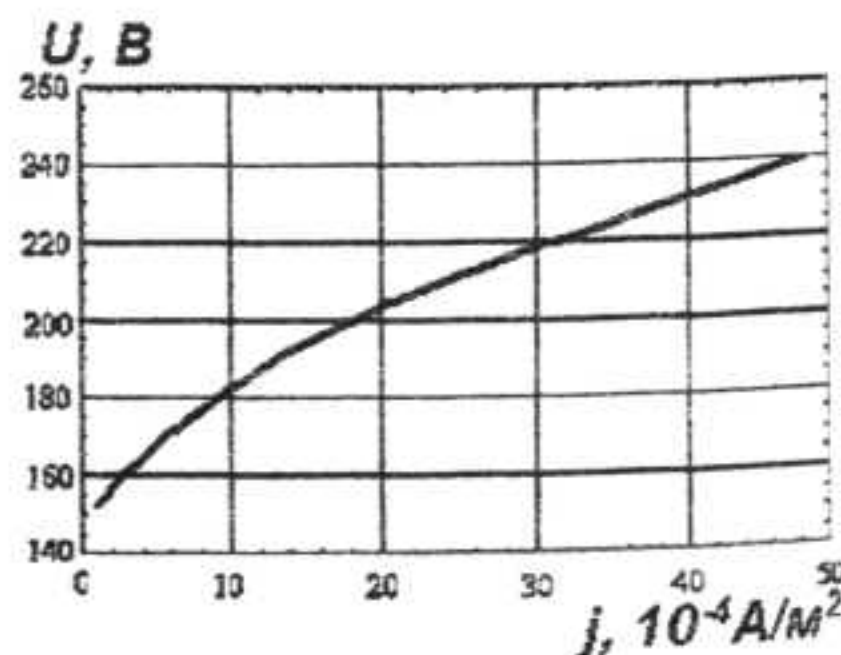


Рис.2. Вольт-амперна характеристика темного розряду для різних довжин міжелектродного проміжку $d=9$ см (а), $d=80$ см (б).

При великій ширині газорозрядного проміжку напруга U може спадати при зростанні густини струму j , в цьому випадку газорозрядний проміжок характеризується від'ємним диференціальним опором. Від'ємний диференціальний опір можна пояснити тим, що точка сталого поля знаходиться поблизу катода, тому зростання електричного поля ліворуч неї може бути перекомпенсоване зменшенням поля праворуч. В цьому випадку вольт-амперна характеристика може мати незначний від'ємний нахил, який спостерігається на рис.2 б.

4.3. Просторові розподіли концентрації іонів і електронів

Розподіли концентрації іонів та електронів уздовж міжелектродного проміжку подані на рис.3. З них видно, що концентрація іонів на більшій частині проміжку значно перевищує концентрацію електронів, тому просторовий заряд визначається переважно розподілом іонів. Відповідно поблизу катода в першому наближенні можна вважати, що $n_i \approx \text{const}$, тоді хід просторової залежності електричного поля буде близьким до лінійного (рис.1 а). Біля анода неоднорідність поля менша, ніж біля катода, оскільки n_i мале (див. граничну умову (8)).

При малих густинах струму концентрація іонів спадає від катода до анода, але при наближенні до значення $j = 1.5 \cdot 10^{-4} \text{ А/м}^2$ на просторовому розподілі концентрації іонів з'являється локальний максимум, розташований поблизу катода. При подальшому зростанні струму максимум концентрації стає чіткішим.

Появу максимуму концентрації іонів можна пояснити таким чином. При великих густинах струму неоднорідність електричного поля зростає, причому найбільший градієнт поля спостерігається біля катода (рис.1). Оскільки біля катода струмом, створеним електронами, можна знехтувати поруч з іонним струмом, то в цій області $j_i = en_i \mu_i E \approx j = \text{const}$. Звідси випливає, що в цій області $n_i E \approx \text{const}$. В результаті зменшення електричного поля при віддаленні від катода приводить до зростання концентрації іонів.

За результатами розрахунку відношення максимальної концентрації іонів до концентрації нейтральних атомів складає величину порядку 10^{-9} .

Відзначимо, що при великих j характер залежностей електронної та іонної концентрацій від координати втрачає подібність до результа-

тів наближеного аналітичного розрахунку [1] і стає схожим на подібні залежності в прикатодній області жевріючого розряду. Справді, названий розрахунок дає експоненціальні залежності $n_{i,e}(x)$, тоді як на відповідних кривих, отриманих у числовому розрахунку, з'являються точки перегину, як у жевріючому розряді поблизу катода [1]. Оскільки наближений розрахунок не враховує впливу зарядів у міжелектродному проміжку на електричне поле, природно припустити, що відзначені особливості якраз пов'язані зі зворотним впливом об'ємного заряду на поле.

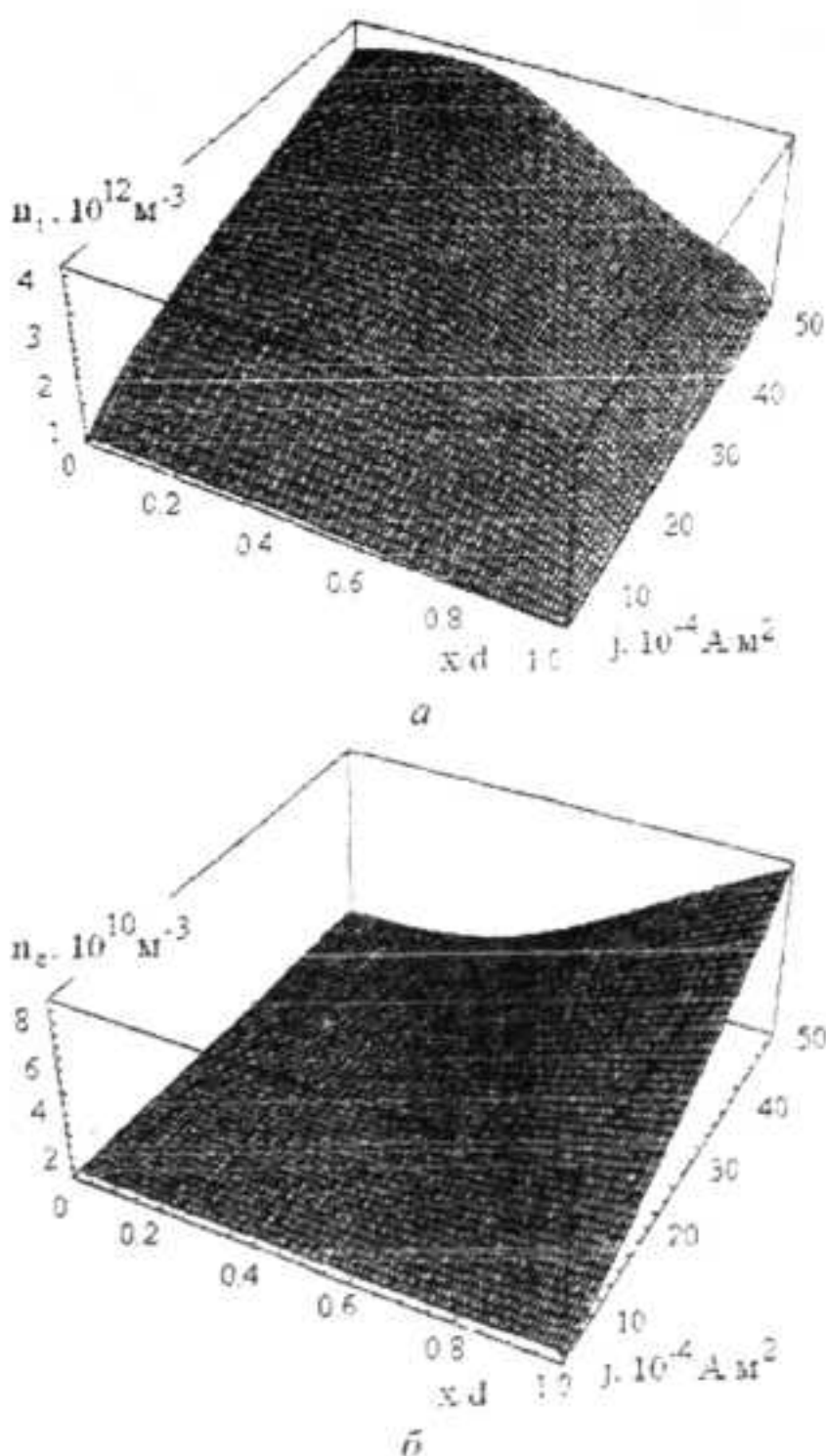


Рис.3. Розподіл концентрації іонів (а) та електронів (б) у міжелектродному проміжку.

4.4. Іонна та електронна густини струму

На рис.4 зображено просторовий розподіл іонної компоненти густини струму, нормованої на повну густину струму. При малих j хід залежності можна описати законом $j_i/j = 1 - \exp(-\alpha x)/(\gamma + 1)$, він збігається з результатами спрощеного аналітичного розрахунку [1].

Однак при переході через значення $j=1.2 \cdot 10^{-3} \text{ А/м}^2$ на графіку з'являється точка перегину, що, як і в попередньому випадку, надає йому подібності до аналогічної залежності в прикатодній області жевріючого розряду.

Зрозуміло, що просторова залежність електронного струму у таких самих відносних одиницях являтиме собою різницю $1-j_e/j$.

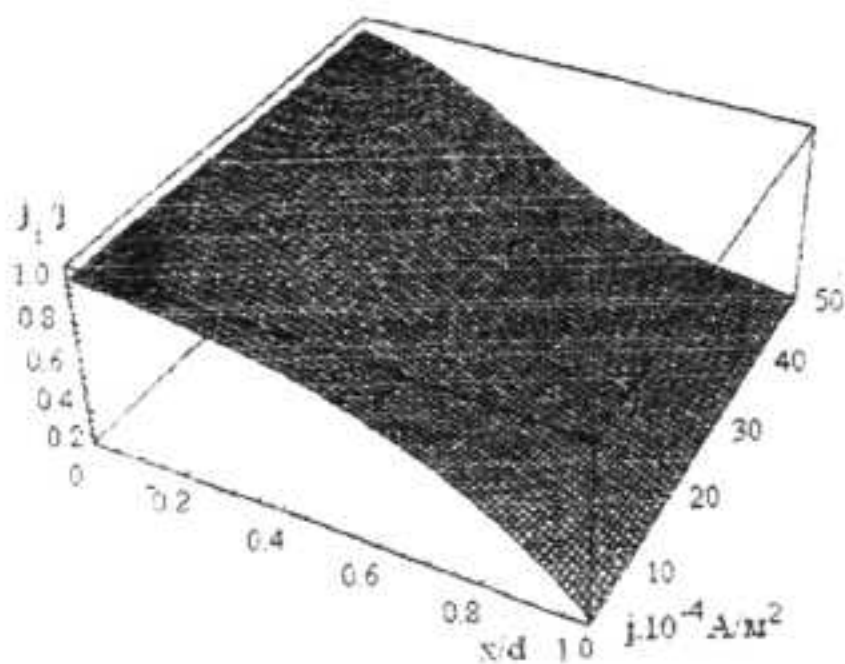


Рис.4. Відношення іонної компоненти густини струму до повної густини струму.

5. Висновки

1. При малих значеннях густини струму ($j \leq 10^{-3} \text{ А/м}^2$) просторовий заряд у розрядному проміжку незначний, тому розподіли концентрацій електронів та іонів і відповідних густин струмів збігаються з відомими результатами аналітичних розрахунків [1].

2. При більших густинах струму стає помітною неоднорідність електричного поля. Напруженість поля на катоді найбільша і спадає майже лінійно на ділянці від катода до середини розрядного проміжку. На ділянці від середини розрядного проміжку до анода друга похідна електричного поля $d^2E(x)/dx^2$ є додатною, що відповідає відповідному вигину електрич-

ного поля. Напруженість поля на аноді мінімальна. Градієнт поля біля аноду близький до нуля, бо там майже відсутній просторовий заряд. Розподіл електричного поля при великих густинах струму ($j \geq 5 \cdot 10^{-3} \text{ А/м}^2$) аналогічний розподілу електричного поля в прикатодних областях тліючого розряду.

3. В міжелектродному проміжку існує точка (так звана точка сталого поля), де зі зростанням густини розрядного струму поле залишається таким самим, як за відсутності струму. При збільшенні густини струму ця точка поступово зсувається до катода. При малих d точка сталого поля зникає. Коли точка сталого поля відсутня або знаходиться біля аноду, нахил вольт-амперної характеристики є додатним. При її наближенні до катода (при великих відстанях між електродами, $d > 50 \text{ см}$) нахил вольт-амперної характеристики може стати від'ємним.

4. При великих густинах розрядного струму на розподілі концентрації іонів виникає максимум, обумовлений градієнтом електричного поля біля катода. На просторових розподілах концентрацій електронів та іонів та густин електронного та іонного струмів виникають точки перегину, що відрізняє їх від результатів спрощеного аналітичного розрахунку та робить подібними до аналогічних залежностей у прикатодній області жевріючого розряду.

Список використаних джерел

1. Ю.П. Райзер. Физика газового разряда. - М.: Наука, 1992.
2. С.М. Левитский. Сборник задач и расчетов по физической электронике. - К.: Издательство Киевского университета, 1964.

Надійшла до редколегії 22.04.08

СУЧАСНА ФІЗИКА

УДК 532.536

Альошин О.Д., д. ф.-м. н., проф.
Абдікарімов Б.Ж., к. ф.-м. н.
Рудніков Є.Г., к. ф.-м. н.

**Взаємозв'язок величин критичних показників
кривої співіснування, теплоємності та
кореляційної функції**

У роботі застосовано новий метод визначення величини критичних показників, оснований на введенні у рівняння флуктуаційної теорії фазових переходів малих параметрів. Це надало можливість запропонувати нові співвідношення між критичними показниками флуктуаційної теорії фазових переходів: вперше визначити величину критичного показника польової залежності теплоємності речовини; уточнити величини критичних показників кривої співіснування, кореляційної функції, температурної залежності теплоємності. Одержані результати підтверджуються даними експериментальних досліджень поблизу критичної точки.

Ключові слова: критична точка, критичні показники, крива співіснування, теплоємність, кореляційна функція

* E-mail: alekhin@univ.kiev.ua

Значний в минулому інтерес до визначення величин критичних показників флуктуаційної теорії фазових переходів (ФТФП) [1] на даний час суттєво зменшився. Численні експериментальні і теоретичні дослідження критичного стану речовини дозволили багаторазово визначити їх величини для великої кількості речовин, зокрема розмірності $d=3$, які відносяться до одного класу універсальності з моделлю Ізінга [2-4]. Однак, на даний час ані теоретично, ані експериментально не визначена величина критичного показника польової залежності теплоємності α_u ($C_v - \Delta \mu^{-\alpha_u} = ((\mu - \mu_c)/\mu_c)^{-\alpha_u}$). Тут μ_c – критичне значення хімічного потенціалу речовини.

Слід вказати також, що розрахунки величин критичних показників [1] різноманітними теоретичними методами: розрахунки тривимірної моделі Ізінга, методу ϵ -розкладів, розв'язку рівнянь ренормалізаційної групи та інші [2,3] приводять до результатів, які

UDK 532.536

Alekhin A.D., Dr. of Sci., Prof.
Abdikarimov B.Zh., Ph.D.
Rudnikov E.G., Ph.D.

**Interrelation among the critical exponents of
coexistence curve, heat capacity and correlation
function**

A new method of critical indices determinate based on small parameters introduction into the equation of fluctuation theory of phase transition have been presented in paper. It allow to proposed the new interrelations among the critical exponents of fluctuation theory of phase transition: to obtained for the first time the critical exponent of field dependence of heat capacity: to define more precisely the critical exponents of coexistence curve, correlation function and temperature dependence of heat capacity. Obtained results are corroborated by experimental studies near the critical point.

Keywords: critical point, critical exponents, coexistence curve, heat capacity, correlation function

дещо, але не дуже сильно, відрізняються між собою. На жаль, в цих розрахунках не визначені конкретні похибки величин цих критичних показників [1]. Тому об'єктивно не можна віддати перевагу тому чи іншому теоретичному методу.

Величини критичних показників, визначені експериментально різними дослідниками також дещо відрізняються між собою [4-6] і від розрахунків теорії [2,3]. В межах похибок експерименту, які значно зростають при підході до критичної точки, не можна віддати перевагу тим чи іншим конкретним величинам цих критичних показників.

Найбільш ретельно за даними експериментальних досліджень визначений критичний показник β , що визначає форму кривої співіснування рідина-пара при наближенні до критичної температури.

$$\Delta \rho = \frac{\rho - \rho_c}{\rho_c} = B_0 t^\beta = B_0 \left(\frac{T - T_c}{T_c} \right)^\beta \quad (1)$$

Тут T_c , ρ_c – критичні значення температури і густини. Вперше розрахунки Гугегейма [2] показали, що величина критичного показника $\beta \approx 1/3$.

Розрахунки показника β за даними сучасних експериментів дозволили більш точно визначити величину цього критичного показника. Однак при цьому з'ясувалось, що різні дослідники схиляються до різних значень показника β . Так, фахівці [4-6], які досліджують криву співіснування просторово однорідних систем після їх ретельного перемішування, відстоюють значення $\beta < 1/3$ ($\beta = 0.32 \pm 0.33$).

Інша група фахівців [7-9], які досліджують поблизу КТ реальну просторово неоднорідну рідинну систему без її перемішування (тобто використовують явище гравітаційного ефекту [10]), стверджують, що величина критичного показника $\beta > 1/3$ ($\beta \geq 0.34$).

Подібна ситуація простежується при порівнянні величин критичного показника теплоємності α_i ($C_v \sim t^{-\alpha_i}$) різних авторів [4,5,6,11,12]. В цих роботах значення α_i змінюються значно сильніше в межах величин $\alpha_i = 0.09 \pm 0.12$. Одержані дані узгоджуються з результатами різних теоретичних розрахунків [2,3].

У зв'язку із викладеним, метою даної роботи було: на основі запропонованого в [13] нового методу визначення величин критичних показників розрахувати не визначену на даний час величину критичного показника польової залежності теплоємності α_μ ($C_v \sim \Delta\mu^{-\alpha_\mu}$). Крім того, уточнити величини малих критичних показників ФТФП [1-3]: α_i , η ($G_2(r) \sim 1/r^{1+\eta}$) і критичного показника кривої співіснування β і проаналізувати їх величини за допомогою відомих співвідношень ФТФП [1]. Похибка теоретичних розрахунків величин критичних показників ФТФП [1-3] в запропонованому методі [13] складає величину не більшу 1%. Слід зазначити, що ця похибка менша реальних похибок експерименту [4-9] чи теоретичних розрахунків [1-4] при визначенні величин критичних показників.

Суть методу полягає у введенні у співвідношення між критичними показниками ФТФП [1] малих параметрів, які можуть бути представлені у вигляді

$$\gamma_0 = 2/3 - \gamma \ll \gamma; \delta_0 = \xi - 2/5 \ll \xi; \beta_0 = \beta - 1/3 \ll \beta; \quad (2)$$

$$\gamma_0 = 4/3 - \gamma \ll \gamma; \delta_0 = 5 - \delta \ll \delta$$

Результати експериментальних [4-9] і теоретичних [2,3] досліджень величин критичних показників дозволяють вишикувати нерівності (2) наступним чином

$$\delta_0 \gg \gamma_0 > \alpha_i > \alpha_\mu \approx \eta \gg \nu_0 \gg \xi_0 \approx \beta_0 \approx (3 \pm 4) \cdot 10^{-3}. \quad (3)$$

Припущення рівності параметрів $\xi_0 \approx \beta_0$ завдяки малості їх величин несуттєво впливає на точність розрахунків величин критичних показників ФТФП [1]. (Реальні експериментальні похибки величин критичних показників такі ж самі і навіть значно більші величин малих параметрів $\xi_0 \approx \beta_0 \approx 3 \cdot 10^{-3}$). В той же час це припущення (3) на основі рівнянь ФТФП [1] і (2) дозволяє одержати для 3-вимірних систем ($d=3$) ряд нових співвідношень між малими критичними показниками α_i , α_μ , η і малими параметрами ν_0 , ξ_0 , наприклад:

$$\alpha_i = 3 \nu_0; \eta = (25/2)\xi_0 + (25/4)\xi_0^2; \nu_0 = 19/3 \xi_0; \quad (4)$$

$$\alpha_\mu = 9/5 \nu_0 + 27/10 \nu_0^2 \approx 34/3 \xi_0;$$

З цих співвідношень безпосередньо слідує, що величина критичного показника теплоємності α_μ дещо менша величини критичного показника кореляційної функції η

$$\alpha_\mu \approx 34/3 \xi_0 < \eta \approx (25/2)\xi_0 \quad (5)$$

Використовуючи (2) і (4) на основі рівнянь ФТФП [1] можна визначити два нових співвідношення між критичними показниками радіуса кореляції, теплоємності, кореляційної функції

$$\nu = \xi^{1/2} \quad (6)$$

$$\alpha_i = 3/4(\alpha_\mu + \eta) \quad (7)$$

Справедливість рівності (6) [13] підтверджується величинами критичних показників радіуса кореляції $\nu = 0.626$ і $\xi = 0.4045$ [13], а також одержаних експериментально [14] і розрахованих теоретично [2,3].

Виходячи із співвідношень (2) і (6), знаходимо взаємозв'язок між малими параметрами ν_0 і ξ_0

$$\nu_0 = 0.0342 - 0.8 \xi_0 \quad (8)$$

$$\xi_0 = 0,0428 - 1,25 v_0$$

З (8) слідує важливий результат, що величина малого параметра v_0 повинна задовольняти нерівності $v_0 < 0,034$. Тобто величина критичного показника теплоємності $\alpha_c = 3v - 2 = 3v_0 < 0,1$.

Розглянемо тепер відомі співвідношення ФТФП [1].

$$\beta = \frac{v}{\xi} \frac{1}{\delta} = \frac{v}{\xi} (3\xi - 1) \quad (9)$$

Тоді на основі (2) і (9) знаходимо рівняння, що визначає малий параметр $\beta_0 = 1/3 - \beta$.

$$\beta_0 = 1/3 - \beta = \frac{25}{6} \xi_0 - \frac{v_0}{2} \quad (10)$$

Як видно, знак малого параметра β_0 залежить від

величин доданків $\frac{25}{6} \xi_0$ і $\frac{v_0}{2}$. Якщо величина

$$\frac{25}{6} \xi_0 > \frac{v_0}{2}, \text{ то } \beta_0 > 0; \text{ якщо ж } \frac{25}{6} \xi_0 < \frac{v_0}{2},$$

то $\beta_0 < 0$. Підставляючи (8) в (10) маємо

$$\beta_0 = 0,178 - 4,7 v_0 \quad (11)$$

Використовуючи нерівність $v_0 < 0,034$ з (11) слідує висновок, що малий параметр $\beta_0 > 0$. Отже величина критичного показника $\beta > 1/3$.

Перейдемо тепер до кількісного визначення величини критичного показника кривої співіснування β (1). Для цього використаємо умову $\beta_0 \approx \xi_0$ і співвідношення $v_0 = 19/3 \xi_0$. Тоді з (8) і (11) знаходимо величини малих параметрів $v_0 = 0,0295$, $\xi_0 = 0,0045$ і нарешті $\beta_0 = 0,005 > 0$. Отже, величина критичного показника кривої співіснування (1) $\beta = 1/3 + 0,005 = 0,338$. Цей результат підтверджується даними експериментів [7-9], де використовується явище гравітаційного ефекту. В той же час цей результат не узгоджується з даними експериментів [4-6], в яких речовина перед вимірюваннями густини $\Delta\rho(t)$ (1) ретельно перемішувалась, що виводило систему із стану рівноваги.

Використовуючи одержані результати для величин малих параметрів: $v_0 = 0,0305$; $\xi_0 = 0,0045$; $\beta_0 = 0,05$, на основі (2) знаходимо величини критичних показників температурної і польової залежностей теплоємності ($C_v \sim t^{-\alpha_c}$, $C_v \sim \Delta\mu^{-\alpha_\mu}$) $\alpha_c = 3v_0 = 0,091$; $\alpha_\mu = 3\xi/v - 3\xi = 0,058$ і кореляційної функції $\eta = 2 - 3(\delta-1)/(\delta+1) = 5 - 2/\xi = (25/2)\xi_0 + (25/4)\xi_0^2 = 0,062$. радіусу кореляції $v = 0,638$, $\xi = 0,405$. Ці результати підтверджують також справедливості запропонованого в [13] рівняння (7).

Слід наголосити, що величина критичного показника польової залежності ізохорної теплоємності $\alpha_\mu = 0,058$ [13] визначена вперше. Як видно, за своєю величиною вона дещо менша величини критичного показника кореляційної функції $\eta = 0,062$ ($\alpha_\mu \leq \eta$), що підтверджує висновок, який слідує із (5).

Одержані на основі [13] і (8), (11) результати дозволили вперше показати, що вони задовольняють також відомому співвідношенню між критичними показниками ФТФП [1]

$$\frac{\alpha_c}{\alpha_\mu} = \frac{v}{\xi} \quad (12)$$

Дійсно, відношення $\alpha_c/\alpha_\mu = 1,57$, а $v/\xi = (2/3 - 0,034)/(2/5 + 0,045) = 1,574$.

Далі для порівняння проаналізуємо, чи відповідають цьому фундаментальному співвідношенню (12) величини малих критичних показників $\alpha_c = 0,07 \pm 0,12$ ($\langle \alpha_c \rangle = 0,095$) і $\eta = 0,03 \pm 0,04$ ($\langle \eta \rangle = 0,035$), розрахованих на основі теоретичних підходів [2,3] і експериментальних даних [5,11,12].

На жаль, на даний час теоретично [2,3] величина критичного показника α_μ не визначена. Не визначена вона і експериментально.

Однак вище на основі (5) було вже показано, що $\alpha_\mu < \eta$. Тоді згідно [2,3] величина критичного показника α_μ повинна задовольняти нерівності $\langle \alpha_\mu \rangle < \langle \eta \rangle = 0,035$. Отже і відношення α_c/α_μ повинно задовольняти нерівності $\langle \alpha_c \rangle / \langle \alpha_\mu \rangle > \langle \alpha_c \rangle / \langle \eta \rangle = 2,71$ ($\alpha_c/\alpha_\mu > v/\xi$). Як бачимо, величина цього відношення значно перевищує величину $v/\xi = 1,574$, одержану на основі [2,3]. Тобто при цих величинах критичних показників $\langle \alpha_c \rangle$, $\langle \eta \rangle$ [2,3] співвідношення (12) ФТФП [1] не виконується. З нерівності $\alpha_c/\alpha_\mu >$

v/ξ слідує, що величина $\langle \alpha_1 \rangle = 0,095$ – завищена, а $\langle \eta \rangle = 0,035$ – занижена.

В той же час розрахунки величин критичних показників ФТФП [1], проведені на основі нового методу [13], дозволили одержати наступні результати.

1. Показано, що величина критичного показника кривої співіснування $\beta = 0,338 > 1/3$, що підтверджується результатами експериментів [7-9], в яких явище гравітаційного ефекту не усувається ретельним перемішуванням речовини поблизу КТ.

2. Вперше визначено величину критичного показника пологої залежності теплоємності вздовж напрямку критичної ізотерми $\alpha_\mu = 0,058$;

3. Зроблено уточнення величин малих критичних показників: теплоємності $\alpha_1 = 0,091$ і кореляційної функції $\eta = 0,062$ ($\alpha_\mu < \eta$).

4. Одержані на основі рівнянь ФТФП [1] величини критичних показників: $v = 0,638$, $\xi = 0,405$, $\alpha_1 = 0,091$, $\alpha_\mu = 0,058$ задовольняють з фундаментальному співвідношенню ФТФП [1]: $\alpha_1/\alpha_\mu = v/\xi$.

5. Вперше знайдено прямий взаємозв'язок між критичними показниками теплоємності $\alpha_1 = 3v_0 = 19\xi_0$, $\alpha_\mu = 9/5 v_0 = 34/3 \xi_0$ і кореляційної функції $\eta = 25/2 \xi_0$.

З цих співвідношень слідує, що у випадку $\eta = 0$ ($\xi_0 = 0$) величини критичних показників α_1 і α_μ також приймають нульові значення ($\alpha_1 = 0$, $\alpha_\mu = 0$). Тобто при $\eta = 0$ теплоємність речовини $C_v \sim t^{-\alpha_1}$ і $C_v \sim \Delta\mu^{-\alpha_\mu}$ а також похідна $d^2P/dt^2 \sim d^2/dt^2(t^{2-\alpha_1})$ [1] не мають степеневі розбіжності в критичній точці. Отже припущення про нульове значення критичного показника η ($\eta = 0$), протирічить основним положенням ФТФП [1] і свідчить про те, що застосування поблизу КТ кореляційної функції Орнштейна-Церніке $G(r) \sim 1/r$ [15] може носити лише якісний характер.

Список використаних джерел

1. Паташинский А.З., Покровский В.Л. Флуктуационная теория фазовых переходов. – М.: Наука, 2-е изд., перераб. 1982. – 382 с.
2. Стенли Г. Фазовые переходы и критические явления. Пер. с англ. – М.: Мир, 1973. – 419 с.
3. Ма Ш. Современная теория критических явлений. Пер. с англ. – М.: Мир, 1980. – 298 с.

4. Анисимов М.А. Критические явления в жидкостях и жидких кристаллах. – М.: Наука. – 1987. – 271 с.
5. Анисимов М.А., Рабинович В.А., Сычев В.В. Термодинамика критического состояния индивидуальных веществ. – М.: Энергоиздат, 1990. – 190 с.
6. Cheng Hongyuan, M.A. Anisimov, J.V. Sengers, Fluid Phase Equilibria 128, 67 (1997).
7. Артюховская Л.М., Шиманская Е.Т., Шиманский Ю.И. Исследование кривой сосуществования и изотермической сжимаемости бензола вблизи критической точки жидкость-пар // УФЖ – 1970. – Т. 15, № 2. – С. 1974–1981.
8. Артюховская Л.М. Исследование поведения некоторых термодинамических функций индивидуальных веществ вблизи критической точки жидкость-пар на основе измерений гравитационного эффекта: Дис... канд. физ.-мат. наук – Киев, 1973. – 220 с.
9. Булавин Л.А. Критичні явища у рідинах. – Київ: “Київський університет”, 1997. – 174 с.
10. Голик А.З., Шиманский Ю.И., Алехин А.Д. и др. Исследование гравитационного эффекта вблизи критической точки индивидуальных веществ и растворов // В сб.: Уравнение состояния газов и жидкостей / К 100-летию уравнения Ван-дер-Ваальса / – М.: Наука, 1975. – С. 189–217.
11. Анисимов М.А., Берестов А.Т., Воронов В.П. и др. // ЖЭТФ. – 1979. – т. 76. – С. 1661.
12. Анисимов М.А., Воронель А.В., Оводова Т.М. // ЖЭТФ. – 1971. – т. 61. – С. 1092.
13. Alekhin A.D. // Journal of Molecular Liquids. – 2005. – 120 – P. 43–45.
14. Алехин А.Д. Поведение радиуса корреляции в гравитационном поле // ЖЭТФ – 1977. – Т. 72, N 5. – С. 1880–1884.
15. Ландау Л.Д., Лифшиц Е.М. Статистическая физика. 3-е изд., доп. – М.: Наука, 1976. – 584 с.

Надійшла до редколегії 14.03.2008

УДК 512

Булавін Л.А., акад. НАН України, д. ф.-м. н., проф., Гаврюшенко Д.А., к. ф.-м. н., Сисоєв В.М., д. ф.-м. н., проф., Якунов* П.А., студ.

Розрахунок просторового розподілу компонентів бінарної суміші в плоскій порі з несиметричним пристінковим потенціалом

Для розрахунку концентрації бінарної суміші розглянуто гратчасту модель рідини. Знайдено розв'язок диференціального рівняння для плоскопаралельної пори з несиметричним експоненціальним пристінковим потенціалом. Одержаний розв'язок ізопериметричної задачі дозволяє досліджувати зміну термодинамічних властивостей речовини у пристінковому шарі у порівнянні із властивостями рідини у об'ємній фазі.

Ключові слова: гратчаста модель бінарного розчину, неоднорідна система, критична точка розшарування, пристінковий потенціал.

*E-mail: y9@yandex.ru

Важливим результатом розвитку статистичної фізики є можливість дати кількісний опис бінарного розчину однорідної рідини в широкому інтервалі зміни термодинамічних параметрів, включаючи окіл можливих критичних точок. Стан такої системи можна повністю задати набором інтенсивних параметрів, жодний з яких не описує розмірів системи або форми поверхні, що її обмежує, та загальним числом частинок $N = N_A + N_B$, де N_A та N_B – кількість частинок сорту "А" та "В" відповідно.

Традиційні методи опису властивостей неоднорідної системи полягає так званому локальному наближенні [1]. Для того, щоб дати послідовну термодинамічну теорію, яка описує поведінку однокомпонентної неоднорідної системи, раніше нами був запропонований фундаментальний підхід, що базується на обчисленні внеску від кожного шару між еквіпотенційними поверхнями не у відповідні термодинамічні потенціали, а в гамільтоніан системи [4]. Базуючись на цьому підході для хімічного потенціалу однокомпонентної системи $\mu(\vec{r})$ у зовнішньому полі можна отримати вираз:

L.A. Bulavin, UNAC Acad.,
D.A. Gavryushenko, Ph.D.,
V.M. Sysoyev, Sc.D.,
P.A. Yakunov*, stud.

Calculation of spatial distribution of binary solution component concentration in flat pore with asymmetrical wall potential

To calculate a concentration profile of a binary solution the lattice model of liquid has been revised. The solution of differential equation for flat pore with an asymmetrical exponential wall potential was found.

The obtained solution allows to investigate deviations of thermodynamic properties of substance in near-wall layer.

Key words: binary solution lattice model, nonuniform system, critical point of lamination, wall potential.

$$u(\vec{r}) = \mu_0 - \mu(\vec{r}) + \Delta\mu_{cor}(\vec{r}), \quad (1)$$

де $\Delta\mu_{cor}$ – внесок від кореляційних ефектів, для якого було отримано вираз у вигляді нескінченного ряду.

У роботі [5] запропонований в [66] формалізм було узагальнено на випадок двокомпонентної системи. Зокрема, для опису термодинамічної поведінки неоднорідного бінарного флюїду в обмеженій системі було застосовано гратчасту модель рідини, в рамках якої молекули розчину знаходяться у вузлах гратки, причому загальне число частинок N співпадає з кількістю вузлів гратки N' :

$$N = N' = N_A + N_B \quad (2)$$

В роботах [5] та [7] було отримано вираз для $u_{BA}(z)$, аналогічний виразу (1):

$$u_{BA}(z) = \mu_0 - \mu(z) + \Delta\mu_{cor}(z). \quad (3)$$

Для випадку достатньо плавної просторової зміни густини $\Delta x(z)$ на відстанях порядку радіусу дії міжмолекулярних сил далеко від

критичної точки розшарування $\Delta x(z_1) - \Delta x(z) \sim (z_1 - z)$, і тоді вираз для $\Delta \mu_{crit}$ можна подати у вигляді розкладу в ряд Тейлора за степенями вигляді:

$$\begin{aligned} \Delta \mu_{crit}(z) = & \frac{1}{2} \frac{d^2 \Delta x(z)}{dz^2} \int_{-L}^L dz_1 C_2(z, z_1) (z_1 - z)^2 + \frac{1}{2} \left(\frac{d \Delta x(z)}{dz} \right)^2 \int_{-L}^L dz_1 \int_{-L}^L dz_2 C_3(z, z_1, z_2) (z_1 - z)(z_2 - z) + \\ & + \frac{d \Delta x(z)}{dz} \left[\int_{-L}^L dz_1 C_2(z, z_1) (z_1 - z) + \Delta x(z) \int_{-L}^L dz_1 \int_{-L}^L dz_2 C_3(z, z_1, z_2) (z_1 - z) \right] + \dots \end{aligned} \quad (4)$$

Оскільки в околі точки розшарування на критичній ізотермі, згідно з теорією масштабних перетворень [2], $(\mu - \mu_c) \sim |x - x_c|^\delta$, де δ – відповідний критичний індекс, то зрозуміло, що нульові моменти прямих кореляційних функцій порядку п'ять і більше розбігаються в самій критичній точці, тому рід (4) розбігається.

Для оцінки внеску кореляційних доданків $\Delta \mu_{crit}(z)$ розглянемо випадок, коли в системі поблизу стінок діють сили притягання (відштовхування), потенціал яких промодельємо наступним виразом:

$$\begin{aligned} u_A(z) = & N_l (B_{A1} + B_{A2}) e^{-kz} \operatorname{ch} skz + \\ & + N_l (B_{A1} - B_{A2}) e^{-kz} \operatorname{sh} kz, \\ u_B(z) = & N_l (B_{B1} + B_{B2}) e^{-kz} \operatorname{ch} skz + \\ & + N_l (B_{B1} - B_{B2}) e^{-kz} \operatorname{sh} kz, \end{aligned} \quad (5)$$

де $B_{\theta 1}$ ($B_{\theta 2}$) та $B_{\theta 2}$ ($B_{\theta 1}$) – амплітуди пристінкових потенціалів на лівій та правій межі пори для речовин А та В відповідно, k^{-1} – ефективний радіус дії цих сил. Необхідно відмітити, що значення цих сталих залежать як від природи рідини (і від наявності домішок) та матеріалу стінки, що обмежує об'єм, так і від стану поверхні – якості її механічної обробки, чистоти поверхні, ступені створення дефектів внаслідок радіаційного опромінення тощо [3]. Завдяки останнім факторам вона може змінюватися в дуже широких межах навіть для тієї ж самої речовини, тому точних даних щодо значення цих сталих для даної речовини не існує.

Тоді в наближенні плавної неоднорідності [8], враховуючи лише перший доданок ряду (4), можна отримати диференціальне рівняння відносно $\Delta x(z)$ [9]

$$\begin{aligned} \frac{d^2 \Delta x}{dz^2} - \kappa^2 \Delta x = & A \kappa^2 \left[\lambda + \right. \\ & + (B_{BA1} + B_{BA2}) e^{-kz} \operatorname{ch} skz + \\ & \left. + (B_{BA1} - B_{BA2}) e^{-kz} \operatorname{sh} kz \right], \end{aligned} \quad (6)$$

побудова розв'язку якого має певні особливості [8]. Зокрема, рівняння (6) необхідно розв'язувати за умов трансверсальності

$$\left. \frac{d \Delta x(z)}{dz} \right|_{z=L} = 0, \quad (7)$$

та ізопериметризму

$$\int_{-L}^L \Delta x(z) dz = 0 \quad (8)$$

У рівнянні (6) $\kappa = \frac{1}{\sqrt{2\xi}}$, ξ – радіус кореляції,

$B_{BA1} = (B_{B1} - B_{A1}) N_l$, $B_{BA2} = (B_{B2} - B_{A2}) N_l$, $A = A(\xi)$ – стала, що залежить від параметрів критичної точки даного розчину, λ – стала, значення якої визначається з умови.

Зауважимо, що накладання на розв'язок рівняння (6) додаткової умови пов'язано з необхідністю завдання хімічного потенціалу однорідної системи. Розв'язок рівняння (6) за умов (7) та (8) має вигляд:

$$\Delta x(z) = A\kappa^2 e^{-\kappa L} \left[(B_{BA1} + B_{BA2}) \left(-\frac{k}{\kappa} \frac{1}{k^2 - \kappa^2} \frac{\text{sh } kL}{\text{sh } \kappa L} \text{ch } \kappa z + \frac{1}{k^2 - \kappa^2} \text{ch } \kappa z + \frac{1}{k\kappa^2 L} \text{sh } kL \right) + \right. \\ \left. + (B_{BA1} - B_{BA2}) \left(-\frac{k}{\kappa} \frac{1}{k^2 - \kappa^2} \frac{\text{ch } kL}{\text{ch } \kappa L} \text{sh } \kappa z + \frac{1}{k^2 - \kappa^2} \text{sh } \kappa z \right) \right]. \quad (9)$$

На рис.1. наведено просторовий розподіл концентрації модельного бінарного розчину $\Delta x(z)$.

Аналіз одержаних виразів вказує на те, що далеко від критичної точки профілі концентрації рідини збігаються з розв'язком, який відповідає локальному наближенню, тобто лише на відстанях порядку радіуса дії поверхневих сил мають місце відхилення концентрації від середнього значення, у всьому ж об'ємі системи концентрація розчину дорівнює середній. Але при наближенні до критичної точки сингулярно збурений доданок починає відігравати усе більшу роль, і зміна концентрації розчину відбувається в більшій частині об'єму, тобто характерною відстанню, на якій відбуваються суттєві зміни густини, є вже не радіус дії поверхневих сил, а радіус кореляції системи ξ .

Одержаний розв'язок ізопериметричної задачі дозволяє досліджувати зміну термодинамічних властивостей речовини у пристінковому шарі у порівнянні із властивостями рідини у об'ємній фазі. Суттєва зміна повинна призвести до відповідних явищ, що спостерігаються експериментально, навіть на

молекулярному рівні. Зокрема, результати експерименту з пропускання повільних нейтронів [10], демонструють суттєву залежність гравітаційного ефекту в розчині етану від розмірів системи. Аналіз отриманих в [10] результатів вказує на те, що критичні індекси, які обраховуються за дослідними даними, не змінюються при врахуванні кореляційних ефектів.

Таким чином, з проведеного аналізу випливає, що при розрахунках критичних амплітуд за даними з розсіяння нейтронів та світла необхідно враховувати кореляційний внесок у розподіл густини, і лише потім проводити врахування інших членів рівняння стану речовини. Отримані результати можуть використовуватись при експериментальному дослідженні термодинамічних властивостей багатокомпонентних флюїдів поблизу критичної точки.

Список використаних джерел

1. Монстер А.. Химическая термодинамика. – М.: Мир, 1971. – 296 с.
2. Покровский А.З., Потошинский В.Л.. Флуктуационная теория фазовых переходов. – М.: Наука, 1982. – 382 с.
3. Дерягин Б.В., Чураев Н.В., Муалев В.М., Поверхностные силы. – М.: Наука, 1987. – 400с.
4. Булавин Л.А., Гаврюшенко Д.А., Сысоев В.М. Профиль плотности флюида в плоскопараллельной поре с неидеальными стенками в гравитационном поле // Журнал физической химии. – 2004. – Т. 78 (11). – С. 2039-2042.
5. Булавин Л.А., Гаврюшенко Д.А., Сысоев В.М. Розрахунок розподілу компонентів бінарної суміші в обмеженій системі // Український фізичний журнал. – 2007. – Т. 52 (10). – С. 933-938.
6. Булавин Л.А., Гаврюшенко Д.А., Сысоев В.М. Плотность неоднородной жидкости во внешнем поле // Доповіді Академії наук України. – 1997. – № 7.
7. Булавин Л.А., Гаврюшенко Д.А., Сысоев В.М. Розрахунок просторового розподілу концентрації бінарної суміші в зовнішньому полі // Доповіді Академії наук України. – 2008. – №1. – С. 68-71.
8. Булавин Л.А., Гаврюшенко Д.А., Сысоев В.М. Расчет профиля плотности жидкости в плоских несмачиваемых порах // Журнал физической химии. – 1996. – Т. 70 (3). – С. 559-561.
9. Bulavin L.A., Gavryushenko D.A., Sysyoyev V.M. Density profile of liquid in finite-size systems // J. Mol. Liquids. – 2003. – Vol. 105(2). – P. 127-130.
10. Булавин Л.А., Мельниченко Ю.Б. Парциальная численная плотность C_2H_6 в системе $CO_2 - C_2H_6$ // Журн. эксп. и теор. физ. – 1985. – Т. 81 (85). – С. 910-920.

Надійшла до редакції 24.01.2008р

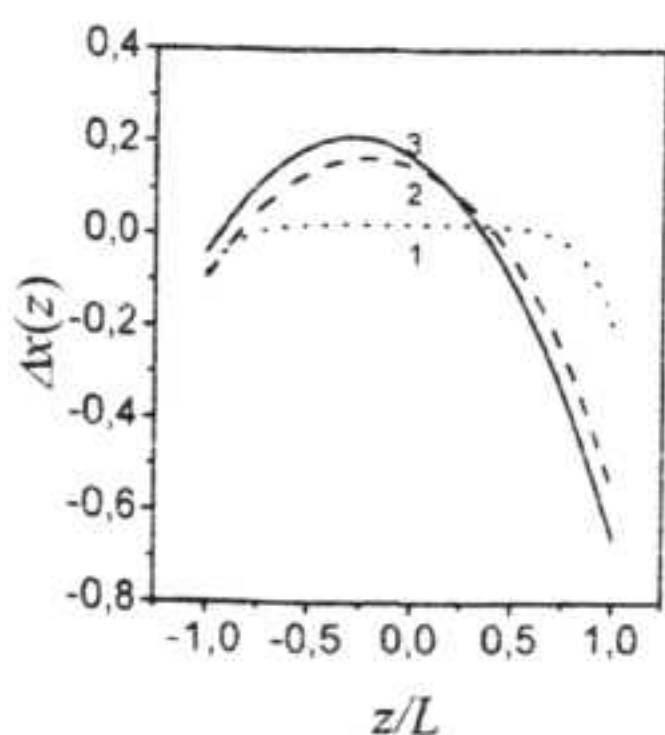


Рис.1. Залежності концентрації $\Delta x(z)$ бінарного розчину для різних значень ξ : 1 – 10^{-6} м, 2 – $5 \cdot 10^{-6}$ м, 3 – 10^{-5} м ($L = 10^{-5}$ м, $k^{-1} = 3 \cdot 10^{-10}$ м, $B_{BA1} = 10^{-27}$ Дж, $B_{BA2} = 0.5 \cdot 10^{-27}$ Дж).

УДК 519.9

Заболотний М.А., к. ф.-м.н., ст.н.с.,
Куліш М.П., д. ф.-м.н., проф.,
Дмитренко О.П., к. ф.-м.н., доц.,
Харкянен В.М., д. ф.-м.н., проф.,
Барабаш Ю.М., к. ф.-м.н., ст.н.с.,
Кобус* О.С., асп.,
Стасюк* К.В., інж.

Спектральна залежність фотогенерації в фулеренмістких аморфних молекулярних напівпровідниках

В роботі досліджуються залежності квантового виходу фотогенерації носіїв струму в аморфних молекулярних напівпровідниках. В експериментах для вимірювань спектральної чутливості використовуються зразки з тонких шарів карбазолмістких матеріалів, отриманих за допомогою поливу толуольних розчинів. Установлено зменшення величини квантового виходу фотогенерації при збільшенні довжини світлової хвилі. Запропонована феноменологічна модель процесу термалізації, що пояснює цю закономірність, яка заснована на припущенні ньютонівського характеру залежності швидкості енергетичних втрат нерівноважним електроном. В рамках моделі встановлено залежності довжини і часу термалізації від частоти збуджуючого світла, температури аморфних молекулярних напівпровідників (АМН) та параметра, що визначає швидкість втрати надлишкової енергії нерівноважним носієм електричного заряду.

Ключові слова: аморфні молекулярні напівпровідники, квантовий вихід, фотогенерація.

Zabolotny M.A., Cand.Sci (Phys.-Math.),
Kulish M.P., Dr of Sci., Prof.,
Dmitrenko O.P., Cand.Sci (Phys.-Math.),
Kharkyanen V.M., Dr. Sci., Prof.,
Barabash Y.M., Cand.Sci (Phys.-Math.),
Kobus* O.S., post-graduate stud.,
Stasiuk* K.V., eng.

Photogeneration spectral dependence of the fullerenecontaining amorphous molecular semiconductors

Quantum output of the photogeneration dependences of current transmitters at the amorphous molecular semiconductors are explored in these work. For experimental measurements of the spectral sensitiveness carbazolecontaining materials thin layer samples were prepared by pouring on of toluol solutions and used. Value reduction of the quantum output of the photogeneration were determined with light wave length increasing. We offer a phenomenological model of the thermalization process, that could explain this conformity to the law, which is based on the supposition of newtonian character of the speed power losses dependence by a nonequilibrium electron. Within model frameworks the thermalisation dependences of length and time are set on excitant light frequency, amorphous molecular semiconductors (AMS) temperature and parameter, that determines speed losses of surplus energy by the nonequilibrium electric charge transmitter.

Key Words: amorphous molecular semiconductors, quantum output, photogeneration.

E-mail*: kobus_olena@ukr.net, Stasiuk_Katerina@univ.kiev.ua

1. Вступ

Аморфні молекулярні напівпровідники (АМН) на основі сенсibilізованих фотопровідних полімерів мають фотопровідні властивості у видимому та близькому до інфрачервоного хвильових діапазонах [1]. Причиною цього є те, що молекули сенсibilізатора діють не лише як центри поглинання світла, але також як центри (або їх складові) фотогенерації носіїв заряду. Ці властивості АМН дають можливість використовувати їх для розвитку мікроелектронних і інформаційних систем [1]. Добре вивчені і широко використовуються АМН

на основі полі-N-вінілкарбазолу (ПВК), полі-N-епоксипропілкарбазолу (ПЕПК), поліантрацелінгліциділового ефіру (ПАГЕ), полі-N-гліциділкарбазолу (ПГК), карбазолвмісного поліорганосилоксану (КСПО). Фоточутливість цих АМН обумовлена наявністю в них карбазольних ядер, що мають велику замкнену спряжену π – електронну систему, регулярністю їх структури і густиною упаковки. Відомо [1], що у ряді структур ПВК (ПЕПК) – ПГК – КСПО реалізується поступове збільшення відстані між карбазольними ядрами за рахунок зростання кількості атомів в ланках основного полімерного ланцюга, якими з'єднуються сусідні мономерні

Збільшення відстані між карбазольними ядрами в полімерному ланцюзі сприяє зростанню його гнучкості і одночасно покращанню умов щодо утворення комплексів із переносом заряду (КПЗ). В АМН, утворених на основі ПГК і ПАГЕ, має місце аналогічна структура ланцюга, і змінюється лише вигляд донорного включення. Перехід від донорних ядер карбазолу до замінників із більш протяжною системою зв'язків зменшує енергію іонізації донора, отже приводить до розширення спектральної чутливості АМН.

В якості сенсibilізаторів звичайно використовуються сполуки 2,4,7 – тринітро флуорену (ТНФ), 2,4,5,7 – тетранітро флуорену (ТЕНФ), кислоти 2,4,7 –тринітро-9-дицианометилефлуорен-4-карбоксиліку (КТНФДЦМК), тетраціаноквінодиметану (ТЦНК); андециліка ефіру кислоти 2,7 – динітро (АЕКДДМФК), фулерени C_{60} і C_{70} . Використання перерахованих АМН дозволяє синтезувати світлочутливі матеріали, область спектральної чутливості яких охоплює практично весь видимий діапазон.

Відомо [1, 2], що для цього типу АМН процес фотогенерації складається з декількох етапів, із яких можна виділити два головних. Перший триває близько 10^{-11} – 10^{-12} с [3, 4]. На протязі цього етапу після поглинання кванту світла формується нейтральний екситоноподібний стан, який завдяки автоіонізації переходить в іон, тобто у «гарячу» пару носіїв заряду. Утворена "гаряча" пара втрачає надлишкову енергію завдяки непружній взаємодії із сусідніми атомами і протягом часу термалізації t_f розділяється на відстань термалізації r_f .

Процес термалізації можна вважати завершеним [4], коли надлишкова енергія нерівноважної пари носіїв заряду зменшується до такого значення ΔW , що взаємодія із оточенням стає пружним. Термалізація приводить до переходу більш рухомого носія заряду від фотогенеруючого центру до молекули карбазолу АМН, що завершується формуванням зв'язаної пари електрон-дірка, яка перебуває в тепловій рівновазі із навколишнім середовищем.

В подальшому розпочинається другий етап. На протязі цього етапу відбувається дисоціація пари електрон-дірка за рахунок процесу дифузійно – дрейфового руху носіїв заряду. Цей етап може бути задовільно описаний моделлю Онзагера [1, 2, 5], яка враховує можливість процесів тунелювання. Однак, модель Онзагера не придатна для опису першого етапу процесу

генерації заряду в АМН, оскільки вона не може бути застосована для опису термодинамічно нерівноважних систем. Така багатоступеневість процесу фотогенерації носіїв електричного заряду в АМН створює значні труднощі при його експериментальному дослідженні, оскільки вимірюється лише результат реалізації всіх етапів (наприклад, квантовий вихід фотогенерації або люмінесценції). Водночас теоретичне дослідження дозволяє аналізувати кожний етап окремо і дає можливість передбачати зміни кінцевих параметрів процесу фотогенерації, обумовлених цими станами.

Метою даної роботи є експериментальне дослідження залежності довжини термалізації у карбазолмістких АМН від довжини хвилі збуджуючого світла і розробка адекватної моделі процесу термалізації носіїв заряду.

2. Експериментальне дослідження

Вільні носії електричного заряду в АМН утворюються після проходження ланцюгу окремих елементарних процесів, кожний з яких впливає на закономірності процесу фотогенерації. Вже відмічалось, що експериментально реєструється результат проходження всіх цих окремих елементарних процесів, кожний з яких визначається своїми параметрами (що обумовлюються характеристиками АМН) і тому по різному залежать від експериментальних умов.

При вивченні процесу фотогенерації носіїв електричного заряду в АМН особливий інтерес представляє залежність квантового виходу фотогенерації від довжини хвилі збуджуючого світла. Це пояснюється тим, що ця залежність в основному визначається лише термалізацією, а не всією сукупністю процесів, що визначають процес фотогенерації.

Залежність квантового виходу фотогенерації η в антрацені і пентацені, а також в сполуках карбазолмістких сполуках досліджувалась в роботах [1, 2, 4, 6, 7]. В них було показано, що квантовий вихід фотогенерації носіїв заряду зростає при збільшенні частоти поглинаючого світла. Із одержаних експериментальних даних по вивченню фотопроектів в пентацені і при використанні моделі Онзагера для опису механізму дисоціації електронно – діркової пари було зроблено [2, 6] висновок про зростання довжини термалізації носіїв заряду при збільшенні частоти збуджуючого світла. Такий висновок ілюструє Рис.1, на якому представлено експериментальну криву одержану при використанні даних роботи [6].

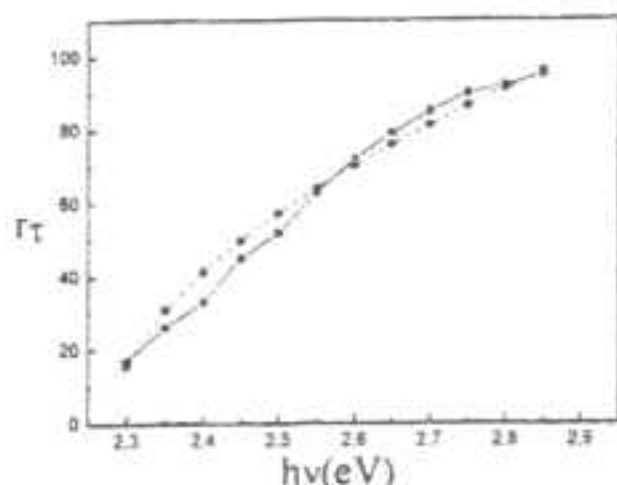


Рис. 1. Експериментальна (суцільна) [6] та теоретична (розривна) криві залежності довжини термалізації в пентацені від енергії кванту збудження.

В роботі [1] було показано, що в полімерних напівпровідниках в околі границі власного поглинання світла спектральна залежність квантового виходу фотогенерації носіїв заряду може бути описана наближенням співвідношенням

$$\eta(h\nu) \approx C \cdot (h\nu - E_G^{ad})^n, n \approx \frac{2}{5} \quad (1)$$

Тут величина C – деяка функція температури і напруженості електричного поля, E_G^{ad} – ширина адіабатичної енергетичної щілини. Відмітимо, що задовільної моделі яка б обґрунтовувала вираз (1) не існує.

В даній роботі були проведені додаткові дослідження залежності квантового виходу фотогенерації носіїв $\eta(\lambda)$ для більш широкого числа АМН. В цих дослідженнях використовувалися зразки товщиною 1,4-1,5 мкм, одержані на скляній основі із електропровідним підшаром SnO_2 за допомогою поливу із толуольного розчину стандартним методом [9]. Поверхня зразка в коронному розряді була заряджена позитивними іонами. Поверхневий потенціал вимірювали динамічним зондовим методом. Спектральну чутливість S_λ визначали по 20% - ному спаду потенціалу вільної поверхні АМН під дією випромінювання інтенсивністю 0,1 мВт/см². Виміри проводили після зарядки вільної поверхні АМН до потенціалу 190 В. Спектри поглинання зразків визначали за допомогою спектрофотометра С – 26. Результати експериментів представлені в таблиці 1. В ній приведено результати дослідження спектральних залежностей S_λ (по рівню спаду 0,2 потенціалу ($S_{dv}/v=0.2$)), коефіцієнт поглинання зразка $\kappa(\lambda)$, квантового виходу фотогенерації носіїв $\eta(\lambda)$ та енергії спорідненості до електрону E_A . Згідно з

наведеними в таблиці 1 результатами, залежність квантового виходу фотогенерації від λ (або енергії кванту світла W_ϕ) посилюється при зменшенні енергії спорідненості до електрону E_A сенсibilізатора. Такий характер залежності $\eta(\lambda)$ можна якісно зрозуміти, враховуючи, що початкову енергію гарячого електрону W_e (при нехтуванні вкладом поляризаційних ефектів) можна оцінити як $W_e = W_\phi - I_d + E_A + U$. (Тут I_d – енергія іонізації, U – потенціальна енергія гарячого електрону.) У цьому випадку відносна зміна $\Delta W_e/W_e$ при зміні W_ϕ визначається відношенням

$\Delta W_e / (W_e - I_d + E_A + U)$, величина якого зростає при зменшенні E_A . Відмітимо, що якщо такий спрощений аналіз є справедливим то при використанні у якості акцептора фулерена C_{60} , для якого є характерне значення $E_A = 2,65$ еВ, маємо одержати наноккомпозит квантовий вихід фотогенерації носіїв зарядів у якому в області поглинання майже не залежить від довжини хвилі світла.

3. Теоретична модель

Незважаючи на суттєві результати досягнуті при довготривалих дослідженнях процесу термалізації фотозбуджених носіїв електричних зарядів в АМН [1, 2, 6, 7, 10, 11, 12] (визначення порядку величин часу t_r та довжини термалізації L_T , співвідношень між загальними властивостями фононних спектрів та значеннями t_r і L_T , зв'язок між t_r та коефіцієнтом дифузії носіїв заряду тощо), закономірності процесу термалізації залишається нез'ясованими. Насамперед такий стан пояснюється тим, що процес термалізації є нерівноважним. Це приводить, наприклад, до того, що носій заряду під час термалізації не можливо характеризувати такими рівноважними параметрами як коефіцієнтом дифузії або рухливістю. Крім того при створенні моделей термалізації з використанням квантово-механічних характеристик виникають труднощі в дослідженнях відповідних гамільтоніанів внаслідок неповноти інформації про значення перерізів взаємодії.

В даній роботі запропонована феноменологічна модель процесу термалізації гарячих носіїв заряду, яка дозволяє описати залежність характеристик цього процесу від довжини хвилі збуджуючого світла та температури середовища. Модель заснована на

припущенні, що надлишок енергії "гарячого" нерівноважного носія заряду зменшується завдяки її переходу до сусідніх молекул із швидкістю переходу, що залежить від різниці температур гарячих електронів і оточення. Передбачається [6], що процес термалізації припиняється, коли енергія гарячих носіїв заряду зменшується до величини, коли стає неможливими процеси непружної взаємодії із оточенням. Це стає можливим, коли температура гарячих носіїв на малу величину ΔT перевищує температуру навколишнього середовища T_0 .

Водночас припускається, що потенціальна енергія взаємодії гарячої частинки із оточенням в процесі термалізації не змінюється, а зовнішні поля відсутні або малі і тому не впливають на процес термалізації.

При такому розгляді надлишкова енергія нерівноважного носія заряду визначається його кінетичною енергією або температурою. Це положення дозволяє визначити швидкість переміщення нерівноважного електрона $v(t)$ наступним виразом:

$$v(t) = \sqrt{\frac{3k_b T(t)}{m}} \quad (2)$$

тут k_b - це стала Больцмана; m - маса (взагалі кажучи ефективна) нерівноважного носія заряду.

Використовуючи припущення, що рух носія заряду під час термалізації прямолінійний, довжина термалізації може бути визначена за допомогою співвідношення:

$$r_t = \int_0^{t_t} v(t) dt = \sqrt{\frac{3k_b}{m}} \cdot \int_0^{t_t} \sqrt{T(t)} dt \quad (3)$$

Процес втрати надлишкової енергії нерівноважним електроном визначається його взаємодією із фононами. Фононний спектр визначається структурою АМН і часто невідомий. В таких випадках може бути виправдано використання феноменологічних

моделей [9, 10], які дозволяють пояснити залежності, що спостерігаються на експерименті, а також отримати інформацію про мікропроцеси в АМН. Припустимо, що процес теплового переходу - ньютонівський, тобто коефіцієнт теплового переходу пропорційний різниці температур $(T(t) - T_0)$. Тоді температурна кінетика нерівноважного носія заряду описується наступним рівнянням:

$$\frac{dT(t)}{dt} = -\frac{2\chi}{3k_b} (T(t) - T_0) \quad (4)$$

де χ - параметр, що визначає швидкість втрати надлишкової енергії нерівноважним гарячим електроном. Його величина, в загальному випадку, повинна залежати як від форми фононного спектру, так і від особливостей взаємодії нерівноважного електрона із молекулами зразка. Незважаючи на вказану залежність, при подальшому аналізі виберемо її константою.

Звернемо увагу на те, що співвідношення (4) є частинним випадком більш загального виразу $\frac{dT(t)}{dt} = -\alpha \cdot (T(t) - T(0))^n$, де параметр n визначається структурою фононного спектру, а розмірний коефіцієнт α - визначає величину теплового потоку.

Початковою умовою розв'язку цього рівняння використаємо співвідношення, з якого витікає, що $T(0)$ визначається енергією $\hbar\omega$ поглинутого кванту світла і роботою A необхідною, щоб створити збуджений стан [2]. Параметр A має залежати від значень I_d , E_A та енергій поляризаційних перебудов.

$$T(0) = \frac{2}{3 \cdot k_b} \cdot (\hbar \cdot \omega - A). \quad (5)$$

Розв'язок рівняння (4) при початковій умові має вигляд:

$$T(t) = T_0 \cdot (1 - \exp(-\frac{2\chi}{3k_b} t)) + \frac{2}{3k_b} \cdot (\hbar \cdot \omega - A) \cdot \exp(-\frac{2\chi}{3k_b} t) \quad (6)$$

Врахуємо, що:

$$T(t_t) = T_0 + \Delta T. \quad (7)$$

Використовуючи рівняння (6, 7) отримаємо вираз для часу термалізації t_t в наступному вигляді:

$$t_t = \frac{3k_b}{2\chi} \cdot \ln \left[\frac{1}{\Delta T} \left(\frac{2(\hbar \cdot \omega - A)}{3k_b} - T_0 \right) \right] \quad (8)$$

Із рівнянь (3) і (8) витікає, що радіус термалізації дорівнює:

$$r_T = \sqrt{\frac{2}{m}(\hbar \cdot \omega - A) \cdot \frac{3 \cdot k_h}{\chi}} \cdot \left[1 - \sqrt{\frac{\Delta T}{\frac{2}{3 \cdot k_h} \cdot (\hbar \cdot \omega - A) - T_0}} \right] \quad (9)$$

Відзначимо, що вирази (8) і (9) дозволяють знайти співвідношення між l_T і r_T незалежно від невідомого параметра χ :

$$r_T = \frac{2 \sqrt{\frac{2}{m}(\hbar \cdot \omega - A)}}{\ln \left(\frac{1}{\Delta T} \left(\frac{2 \cdot (\hbar \cdot \omega - A)}{3 \cdot k_h} - T_0 \right) \right)} \cdot \left[1 - \sqrt{\frac{\Delta T}{\frac{2}{3 \cdot k_h} \cdot (\hbar \cdot \omega - A) - T_0}} \right] \cdot l_T \quad (10)$$

4. Обговорення результатів

При аналізі отриманих теоретичних результатів (вирази (8), (9) (10)) необхідно врахувати, що довжина термалізації не є величиною, яка безпосередньо вимірюється на експерименті. Значення r_T можна встановити, використовуючи інформацію про величину

$$\eta(\hbar\omega) = \frac{\Phi_0}{2\gamma\langle r_T \rangle} \exp\left[-\frac{r_c}{\langle r_T \rangle}\right] \sum_{m=0}^{\infty} \frac{(r_c/\langle r_T \rangle)^m}{m!} \sum_{n=0}^{\infty} \left[1 - \exp(-2\gamma\langle r_T \rangle) \sum_{k=0}^{m+n} \frac{(2\gamma\langle r_T \rangle)^k}{k!} \right] \quad (11)$$

де Φ_0 – квантовий вихід первинного процесу автоіонізації, а r_c і γ відповідно рівні:

$$r_c = \frac{e^2}{4\pi\epsilon\epsilon_0 k_h T}, \quad \gamma = \frac{eE}{2k_h T} \quad (12)$$

де ϵ – діелектрична проникність середовища, ϵ_0 – електрична стала, E – напруженість зовнішнього електричного поля.

Аналітичний аналіз залежності r_T як функції $\eta(\hbar\omega)$, що впливає із співвідношення (11), складний, і описується в роботі [6]. Проведений в цій роботі чисельний аналіз показує, що за всіх інших рівних умов (температурі, напруженості електричного поля) більшому значенню квантового виходу відповідає більше значення початкової відстані між зарядами, що входять до складу близнюкової пари зарядів. Це означає, що експериментальне збільшення квантового виходу фотогенерації при зростанні частоти поглинаючого світла вказує на підвищення довжини термалізації носія електричного заряду r_T .

Із формули (9) випливає, що r_T зростає з енергією кванта світла $\hbar\omega$ (або зменшенні λ): що

квантового виходу фотогенерації η при використанні моделі, яка описує процес дисоціації пари дірка – електрон. Так, при застосуванні традиційної моделі Онзагера інформацію про r_T можна отримати із даних, одержаних в роботах [6, 11] за допомогою наступного співвідношення:

добре узгоджується із раніше опублікованими результатами [1, 7]. Співставлення результатів розрахунку при використанні виразу (9) із експериментальними даними, одержаними при дослідженнях пентацену [6] приведено на Рис.1. Добре узгоджена поведінка теоретичних і експериментальних кривих вказує на адекватність запропонованої моделі процесу термалізації. Відзначимо, що залежність $r_T(\lambda)$ добре узгоджується із результатами моделі, яка припускає втрату надлишкової енергії нерівноважного носія заряду завдяки зіткненню із навколишніми молекулами [4,8]. Із виразу (9) також витікає, що довжина термалізації зростає при зменшенні температури навколишнього середовища, що спостерігається експериментально [2, 6].

Отже, запропонована модель процесу термалізації носіїв заряду якісно пояснює спектральну і температурну залежності довжини термалізації і квантового виходу фотогенерації носіїв електричного заряду і дає співвідношення між r_T і l_T , яке знаходиться в задовільній відповідності із експериментальними результатами.

5. Висновки

1. Проведені експериментальні дослідження квантового виходу фотогенерації носіїв електричного заряду η в ПВК, ПЕПК, ПАГЕ, ПГК, сенсібілізованих акцепторами електронів (ТНФ, ТЕНФ, Ud-ДДФК, ТЕНФДЦМ, ТЦХДН) показують істотну залежність квантового виходу фотогенерації від довжини хвилі збуджуючого світла, що вказує на значний внесок в процеси фотогенерації термалізації носіїв заряду.

2. Запропонована феноменологічна модель процесу термалізації носіїв заряду в органічних молекулярних напівпровідниках. Параметрами моделі є показник степені, який визначає залежність швидкості "теплого потоку гарячого носія заряду" від різниці температур гарячого носія заряду і оточення, а також

коефіцієнт пропорційності, який від величини різниці температур вказує на значення теплового потоку. Запропонована модель має ряд припущень:

а). процес термалізації є завершеним, коли надлишковий енергія нерівноважної пари носіїв заряду зменшується до такого значення, коли взаємодія із оточенням стає пружним.

б). швидкість втрати надлишкового енергії гарячими носіями заряду пропорційна різниці температур нерівноважних носіїв заряду і навколишнього середовища.

3. Приведена модель термалізації при використанні моделі Онзагера (або її модифікацій) дозволяє якісно пояснити спектральну і температурну залежності квантового виходу фотогенерації носія заряду.

Список використаних джерел

1. Е.Л.Александрова. Светочувствительные полимерные полупроводники // *Физика и техника полупроводников*, 38(10), ст.1153 - 1194 (2004).
2. М.В.Курик, Э.А.Силнныш. Электронные состояния органических молекулярных кристаллов. Зинатне, Рига. 329 с. (1988).
3. L.Valkunas, V.Gubinas, A.Undzenas. Charge separation and recombination in polymers and aggregates // *Proc. Int Conf. Electronic Processes in Organic Materials*, 22B, с. 54, (1998).
4. Yu. Barabash, V. Kharkyanen, M. Kulish, O. Dmitrenko, M. Zabolotny, E. Kobus, and N. Sokolov. Process of Thermalization of Charges Carriers in Organics Molecular Semiconductors. // *Nanosystems. Nanomaterials, Nanotechnologies*. 5(2), 641-648 (2007).
5. М.А.Заболотный, А.Г.Загородный, Н.П.Боролина. Влияние тунелирования на образование носителей заряда в полимерных полупроводниках под воздействием электронного поля // *Физика твердого тела*, 33 (8), ст. 2345-2349(1991).
6. Э.А.Силинныш, В.А.Колесников, И.Я.Музиканте, Д.Р.Балоде, А.К.Гайлис. О механизмах фотогенерации носителей заряда в органических молекулярных кристаллах // *Известия Академии Наук Латвийской ССР серия физических и технических наук*, (5), ст. 14-28(1981).
7. E.L. Alexandrova. Spectral dependence quantum yield in complexes with carry a charge PVK- TNDZM // *Proc. conf. Silverlos and unusual photographic processes*, 1(1), ст. 95 - 96 (1984).
8. Yu.M. Barabash, M.A. Zabolotny, N.I. Sokolov. Photogeneration charge carriers in photosensitive organic semiconductors // *Semiconductors physics, quantum electronics and optoelectronics*, 2 (4), ст. 51 - 54(1999).
9. Н.А. Давиденко, М.А. Заболотный, А.А. Ищенко. Влияние электрического поля на фотопроводимость и спектры электронного поглощения центров фотогенерации аморфных молекулярных полупроводников // *Химия высоких энергий*, 38(1), ст.23(2004).
10. А.И.Назаров, В.В.Сергеев, Расчет квантового выхода неравновесных носителей в диэлектрике при действии ионизирующего излучения // *Журнал Технической Физики*, 67(6), ст.127-130(1997).
11. М.Поуп, Ч.Свенберг. *Электронные процессы в органических кристаллах*. Т2. Мир, М., 462 с.(1985).
12. Н. Sano, A. Mozumder Model of thermalization of quasi - free electrons in high - mobility liquids and its relationship with electrons mobility // *J. Chem. Phys.*, 66 (2), p. 689 - 698 ,(1977).

Надійшла до редколегії 28.03.2008

УДК 538.22

Беженар А.А.^{*}, асп.,
Копань В.С.^{**}, д.ф.-м. н., проф.

Вплив внутрішніх границь на електроопір порошкового фторопласту

Показано, що питомий електроопір границь між частинками порошкового фторопласту (ФП) менший за питомий електроопір самого ФП. Тому порошковий ФП схожий на композиційний матеріал, в якому струм проходить переважно по електропровідному каркасу з границь.

Ключові слова: електроопір, фторопласт, границя.

^{*}E-mail: R10@univ.kiev.ua, ^{**}E-mail: kopa@pulsary.com.ua

В мікроелектроніці в якості ізоляторів використовують кераміку [1]. В ролі діелектрика в транзисторах давно слугує діоксид кремнію (SiO_2), який має товщину всього декількох атомних шарів, тому важливою проблемою стає тунельний струм та струм протікання [4]. В останні роки були створені діелектрики на основі гафнію [6] з високою діелектричною проникливістю, які є матеріалом для реалізації 45-нанометрового технологічного процесу [4]. В конденсаторах також має місце струм протікання з першої пластини на другу, що здійснюється через ізолятор. ФП використовуються для виготовлення термостійких конденсаторів, які експлуатуються при підвищених температурах ($\sim 200^\circ\text{C}$), а також в конденсаторах, що працюють в ланцюгах струму високої частоти [2].

Слабкими місцями ізолятора являються зовнішня поверхня, границі між зернами (кристалітами, бездефектними областями), а також треки [5]. В дефектах будови порушені валентні міжатомні зв'язки, потрібна менша енергія для руху носіїв заряду в порівнянні з їх рухом в об'ємі ізолятора. Тому по цим місцям протікає небажаний струм та здійснюється пробій ізолятора.

Мета роботи – дослідження електропровідності границь фторопласту, який являється одним з найкращих діелектриків, що використовується в електротехніці [3].

Оскільки струм проходить по поверхні частинок, то пресований порошок ФП можна розглядати як композиційний матеріал (КМ), де

A.A. Bezhenar^{*}, post-graduate stud.,
V.S. Kopan^{**}, Dr of Sci. Prof

Influence of internal boundaries on the electrical resistance of powder fluoroplastic

It is shown that resistivity of boundaries between particles of powder fluoroplastic less than resistivity of fluoroplastic. Therefore powder fluoroplastic similar to composite material where current flow preferably through the electroconductive frame of boundaries.

Key Words: electrical resistance, fluoroplastic, boundary.

поверхня частинок ФП відіграє роль провідника, а об'єм частинки – ізолятором.

1. Електроопір границь порошкового ФП

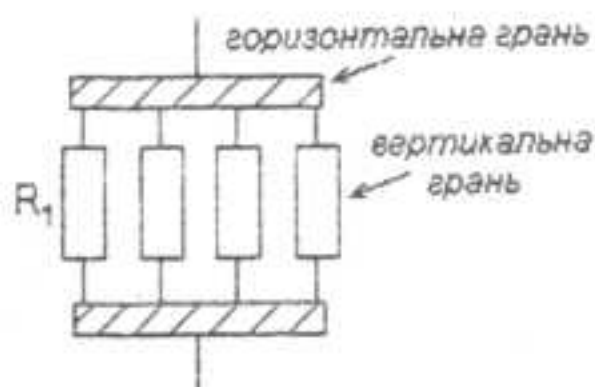


Рис.1. Еквівалентна схема електроопорів граней частинки ФП.

Модель частинки порошкового ФП у формі паралелепіпеда (рис.1) має дві горизонтальні грані і чотири вертикальні. Нехай струм протікає згори вниз. Опір однієї вертикальної грані позначимо через R_1 . Опори R_1 паралельно з'єднані. Вздовж поверхні горизонтальної грані струм не тече, бо в усіх її точках приблизно однаковий потенціал. Тоді загальний опір R однієї частинки ФП буде:

$$\frac{1}{R} = \frac{4}{R_1} \Rightarrow R = \frac{R_1}{4}. \quad (1)$$

Опір стовпчика з таких ФП частинок R_{cm} буде визначатись як

$$R_{cm} = k \cdot R, \quad (2)$$

де $k = l/l_0 = 20$ – кількість частинок (шарів) у стовпчику ($l = 0,5$ мм – товщина всього зразка; $l_0 = 25$ мкм – товщина однієї частинки ФП).

Підставимо (2) в (1)

$$R_1 = 4 \cdot \frac{R_{cm}}{k} = \frac{R_{cm}}{5} \quad (3)$$

У ФП порошковому зразку є n стовчиків з фторопластових частинок, які є паралельно з'єднані між собою. Тоді опір ФП зразка $R_{фп}$ буде визначатись так:

$$\frac{1}{R_{фп}} = \frac{n}{R_{cm}} \Rightarrow R_{фп} = \frac{R_{cm}}{n} \quad (4)$$

Отже, опір границі

$$R_1 = \frac{nR_{фп}}{5} \quad (5)$$

В нашому зразку кількість стовчиків

$$n = \frac{S}{S'} = \frac{5.675 \cdot 10^{-5} \text{ м}^2}{10^{-6} \text{ м}^2} = 57,$$

де S – площа зразка, S' – площа частинки ФП.

Згідно експерименту опір всього зразка $R_{фп} = 2.17 \cdot 10^6 \text{ Ом}$ (при напрузі на зразку $U = 500 \text{ В}$ та напруженості $E = 10 \text{ кВ/см}$).

Тоді опір однієї грані частинки вибраного розміру (паралелепіпед висотою 25 мкм із квадратним перерізом $10^{-3} \times 10^{-3} \text{ м}^2$)

$$R_1 = 11.4 \cdot R_{фп} = 2.48 \cdot 10^7 \text{ Ом}.$$

Оскільки дві грані сусідніх частинок дотикаються і R_1 опір одної, то опір границі:

$$R_{гп} = \frac{R_1}{2} = 1.24 \cdot 10^7 \text{ Ом} \quad (6)$$

Цей опір границі між частинками порошкового ФП визначений експериментально. Такі границі утворюють електропровідний каркас, по якому і протікає основна частина струму. Тому порошок ФП можна вважати композиційним матеріалом, в якому практично неелектропровідна матриця армована електропровідниками – границями.

2. Оцінка питомого електроопору границь порошкового ФП

Опір однієї границі між частинками ФП визначається як

$$R_{гп} = \rho_{гп} \frac{l_0}{2 \cdot S_0} \Rightarrow \rho_{гп} = R_{гп} \frac{2 \cdot S_0}{l_0} \quad (7)$$

де $l_0 = 25 \cdot 10^{-6} \text{ м}$ – висота однієї частинки ФП; S_0 – площа поперечного перерізу грані частинки; $2 \cdot S_0$ – площа поперечного перерізу границі між частинками порошкового ФП.

Будемо вважати, що товщина границі (провідний шар) рівна діаметру молекули ФП.

Молекула ФП має продовгувату циліндричну форму (рис.2) з розмірами $d_1 = 0.565 \cdot 10^{-9} \text{ м}$ та $d_2 = 1.688 \cdot 10^{-9} \text{ м}$ [3].

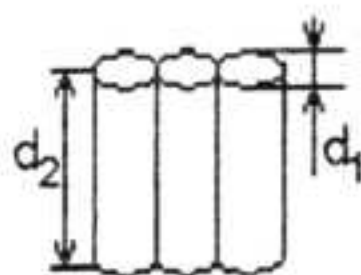


Рис.2. Схематичне зображення молекул ФП.

Розглянемо деякі варіанти розташування молекул ФП в границі:

1) Молекули розташовані паралельно напрямку проходження електричного струму, тобто вертикально.

Можливі варіанти розрахунку площі грані частинки S_0 :

а) Розраховуємо S_0 як площу прямокутника. Якщо a – горизонтальна довжина сторони частинки ФП, то

$$S_0 = a \cdot d_1 = 10^{-3} \cdot 0.565 \cdot 10^{-9} = 0.565 \cdot 10^{-12} \text{ м}^2.$$

Тоді згідно (7) маємо питомий електроопір

$$\rho_{гп}^* = 1.24 \cdot 10^7 \frac{2 \cdot 0.565 \cdot 10^{-12}}{25 \cdot 10^{-6}} = 0.56 \text{ Ом} \cdot \text{м} \quad (8)$$

б) Розраховуємо S_0 як площу поперечного перерізу однієї молекули помножену на кількість молекул, що розташовані на довжині сторони a частинки ФП. Площа поперечного перерізу однієї молекули

$$S_{мол} = \pi \frac{d_1^2}{4}$$

Кількість молекул m визначаємо як відношення довжини сторони частинки ФП до діаметру однієї молекули, тобто $m = a/d_1$. Тоді

$$S_0 = m \cdot S_{мол} = \frac{a}{d_1} \cdot \frac{\pi d_1^2}{4} = \frac{\pi a d_1}{4} = 0.444 \cdot 10^{-12} \text{ м}^2.$$

І, згідно (7), маємо питомий електроопір

$$\rho_{гп}^* = 1.24 \cdot 10^7 \frac{2 \cdot 0.444 \cdot 10^{-12}}{25 \cdot 10^{-6}} = 0.44 \text{ Ом} \cdot \text{м} \quad (9)$$

2) Молекули розташовані перпендикулярно напрямку проходження електричного струму.

Можливі два варіанти розташування молекул: коли молекули лягають одна за одною, горизонтально і паралельно площині грані та ребру частинки (розрахунок питомого електроопору такий самий як в 1.а), і коли молекули дотикаються одна до одної і розташовані перпендикулярно до площини грані. Тоді

$$S_0 = a \cdot d_2 = 10^{-3} \cdot 1.688 \cdot 10^{-9} = 1.688 \cdot 10^{-12} \text{ м}^2.$$

І, згідно (7), маємо питомий електроопір

$$\rho_{гп}^* = 1.24 \cdot 10^7 \frac{2 \cdot 1.688 \cdot 10^{-12}}{25 \cdot 10^{-6}} = 1.67 \text{ Ом} \cdot \text{м} \quad (10)$$

3) Оскільки невідомо як саме розташовані молекули ФП, а розміри молекул d_1 та d_2 одного порядку, то можна оперувати поняттям середнього розміру молекули:

$$\bar{d} = \frac{d_1 + d_2}{2} = 1.127 \cdot 10^{-9} \text{ м.} \quad (11)$$

Тоді середня площа грані частинки:

$$\bar{S}_0 = a \cdot \bar{d} = 1.127 \cdot 10^{-12} \text{ м}^2. \quad (12)$$

Середній питомий електроопір границі, згідно (7)

$$\bar{\rho}_{\text{г}} = R_{\text{г}} \frac{2 \cdot \bar{S}_0}{l_0} = 1.12 \text{ Ом} \cdot \text{м}. \quad (13)$$

Це значення питомого електроопору границі лежить десь посередині серед отриманих з інших міркувань питомих електроопорів. Тому можна вважати, що саме таке значення питомого електроопору границі є найбільш достовірним.

3. Залежність питомого електроопору границі ФП від товщини границі

Таку залежність можна отримати, якщо у формулу (7) підставити $S_0 = a \cdot d$:

$$\rho_{\text{г}} = R_{\text{г}} \frac{2 \cdot a}{l_0} \cdot d = 4.95 \cdot 10^8 \cdot 2 \cdot d, \quad (14)$$

де $2 \cdot d$ – товщина границі між частинками порошкового ФП.

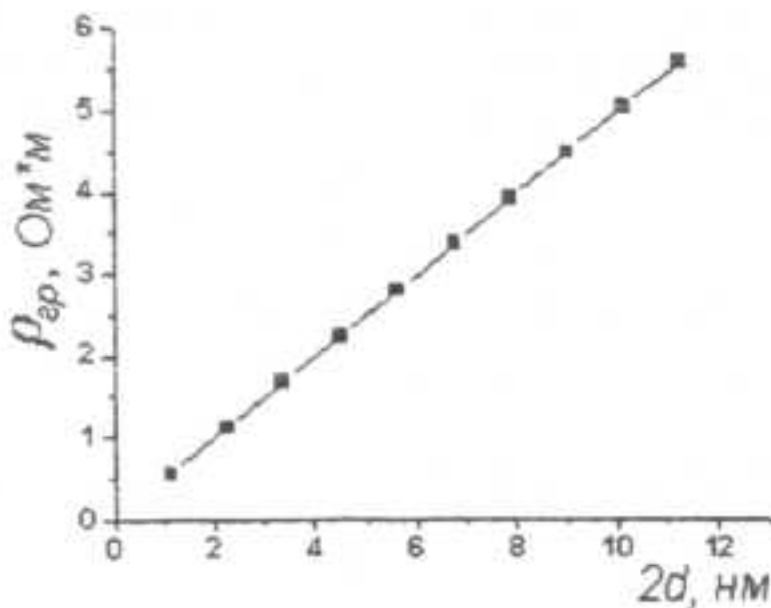


Рис.3. Залежність питомого електроопору границі $\rho_{\text{г}}$ від її товщини $2 \cdot d$.

Границя може мати різну товщину, якщо струм проходить не по одній молекулі (як ми припускали раніше), а по двох, трьох чи більше молекулам. Таким чином, розраховуючи товщину границі як добуток середнього діаметру молекули ФП та кількості молекул у границі, ми отримаємо певний набір значень питомого електроопору границі в залежності від її товщини (рис.3).

4. Питомий електроопір ламаної границі ФП

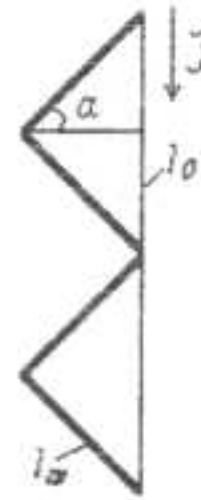


Рис.4. Схематичне зображення ламаної границі частинки порошкового ФП.

Раніше ми вважали, що границя між частинками ФП плоска. Насправді вона може мати ламану форму, наприклад пилкоподібну, таку, схема якої наведена на рис.4. Довжина границі $l_{\text{г}}$ – це довжина ламаної. Питомий електроопір ламаної границі

$$\rho_{\text{г}} = R_{\text{г}} \frac{2 \cdot \bar{S}_0}{l_{\text{г}}}. \quad (15)$$

Довжина границі, очевидно, буде залежати від кута α :

$$l_{\text{г}} = \frac{l_0}{\sin \alpha}. \quad (16)$$

де l_0 – довжина плоскої границі.

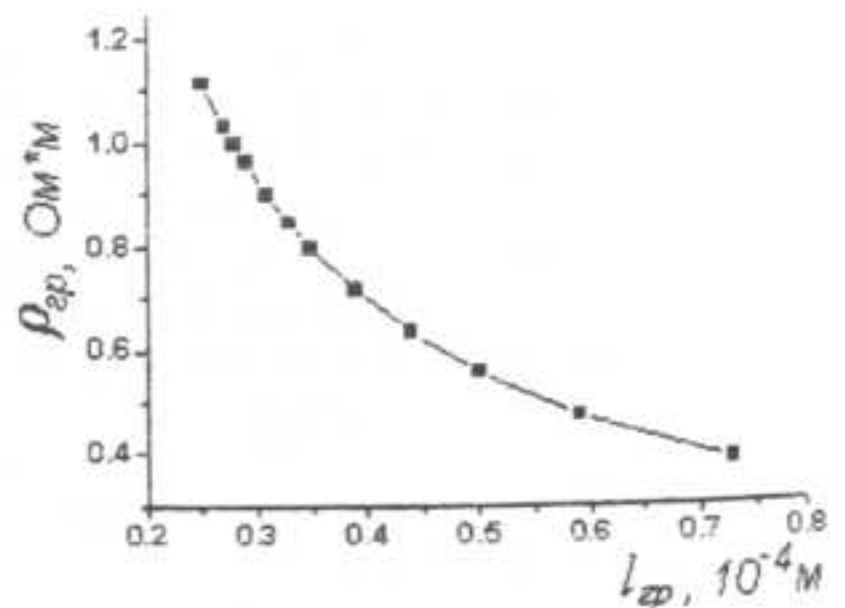


Рис.5. Залежність питомого електроопору границі $\rho_{\text{г}}$ від її довжини $l_{\text{г}}$.

Підставляємо (16), значення l_0 , а також значення опору границі ФП $R_{\text{г}}$ з (6) та значення середньої площі грані з (12) в (15):

$$\rho_{\text{г}} = R_{\text{г}} \frac{2 \cdot \bar{S}_0}{l_0} \cdot \sin \alpha = 1.12 \cdot \sin \alpha. \quad (17)$$

Якщо $\alpha = 90^\circ$, то $l_0 = l_{\text{г}}$ і $\rho_{\text{г}} = 1.12 \text{ Ом} \cdot \text{м}$. Для випадку, коли струм проходить перпендикулярно "ребрам" границь (рис.4), отримуємо залежність питомого електроопору границі ФП від її довжини, наведену на рис.5. При $l_{\text{г}} \rightarrow \infty$ ($d \rightarrow 0$)

$\rho_{\text{гп}} \rightarrow 0$, бо ФП складається з одних електропровідних границь. Цей математичний висновок не реалізується, бо $l_{\text{гп}}$ обмежена розмірами молекул.

Перевірка розрахунку електроопору границь на модельних зразках

Модельні зразки складаються з шарів ФП, в яких є декілька паралельних розрізів довжиною $a_{\text{п}} = 0.8 \cdot 10^{-2}$ м.

Розрахунок проводимо по аналогії з розрахунком електроопору однієї границі порошкового ФП. Оскільки тут буде дві вертикальні грані опором R_2 , то їхній загальний опір буде $R = R_2/2$. Але кожна границя є спільною для граней двох сусідніх частинок, які дотикаються, тому електроопір границі буде $R'_{\text{гп}} = R_2/2 = R$. Використовуючи вирази (2) і (4) отримуємо вираз для R , а отже і для електроопору границі:

$$R'_{\text{гп}} = R = \frac{n \cdot R_{\text{фп}}}{k}. \quad (18)$$

Для визначення питомого електроопору скористаємось формулами (7) та (11), а також поняттям середньої площі грані частинки. По аналогії з (12) маємо:

$$\bar{S}_0 = a \cdot \bar{d} = 0.8 \cdot 10^{-2} \cdot 1.127 \cdot 10^{-9} = 9.012 \cdot 10^{-12} \text{ м}^2.$$

Тоді питомий електроопір границі

$$\rho'_{\text{гп}} = R'_{\text{гп}} \cdot \frac{2 \cdot \bar{S}_0}{l_0} = 0.721 \cdot 10^{-6} \cdot R'_{\text{гп}}. \quad (19)$$

Розглянемо декілька типів зразків:

1 зразок: $k = 5$; $n = 5$. З експерименту визначаємо опір всього зразка $R_{\text{фп}} = 1.3 \cdot 10^6$ Ом (при $U = 150$ В; $E = 12$ кВ/см). Підставляємо ці значення в (18) і отримуємо

$$R'_{\text{гп}} = \frac{5 \cdot 1.3 \cdot 10^6}{5} = 1.3 \cdot 10^6 \text{ Ом}.$$

Підставляючи це значення в (19), отримуємо величину питомого електроопору границі $\rho'_{\text{гп}}$ першого модельного зразка

$$\rho'_{\text{гп}} = 0.721 \cdot 10^{-6} \cdot 1.3 \cdot 10^6 = 0.94 \text{ Ом} \cdot \text{м}. \quad (20)$$

2 зразок: $k = 5$; $n = 10$. З експерименту визначаємо опір всього зразка $R_{\text{фп}} = 1.11 \cdot 10^6$ Ом (при $U = 150$ В; $E = 12$ кВ/см). Підставляючи ці значення в (18), маємо

$$R'_{\text{гп}} = \frac{10 \cdot 1.11 \cdot 10^6}{5} = 2.22 \cdot 10^6 \text{ Ом}.$$

Підставляючи це значення в (19), отримуємо $\rho'_{\text{гп}}$ другого модельного зразка

$$\rho'_{\text{гп}} = 0.721 \cdot 10^{-6} \cdot 2.22 \cdot 10^6 = 1.6 \text{ Ом} \cdot \text{м}. \quad (21)$$

3 зразок: $k = 3$; $n = 5$. З експерименту визначаємо опір всього зразка $R_{\text{фп}} = 0.93 \cdot 10^6$ Ом (при $U = 90$ В; $E = 12$ кВ/см). Підставляючи ці значення в (18), маємо

$$R'_{\text{гп}} = \frac{5 \cdot 0.93 \cdot 10^6}{3} = 1.55 \cdot 10^6 \text{ Ом}.$$

Підставляючи це значення в (19), отримуємо $\rho'_{\text{гп}}$ третього модельного зразка

$$\rho'_{\text{гп}} = 0.721 \cdot 10^{-6} \cdot 1.55 \cdot 10^6 = 1.12 \text{ Ом} \cdot \text{м}. \quad (22)$$

Значення $\rho'_{\text{гп}}$ ФП модельних зразків (20), (21) та (22) близькі до $\rho'_{\text{гп}}$ порошкового ФП (13). Це говорить про те, що розрахунки проведені вірно.

Висновки

1. В порошковому фторопластовому зразку струм проходить переважно по границям. Тому порошковий ФП можна розглядати як композиційний матеріал з провідників у формі границь, вкраплених в діелектричну матрицю.
2. Значення питомого електроопору границь лежить в інтервалі $0.44 \div 1.67$ Ом·м. Питомий електроопір фторопластової плівки в цілому 10^{10} Ом·м [5].

Список використаних джерел

1. Тареев Б.М. Физика диэлектрических материалов. – М., Энергоатомиздат, 1982. – 320 с.
2. Ануфриев Ю.А., Гусев В.Н., Смирнов В.Ф. Эксплуатационные характеристики и надежность электрических конденсаторов. – Москва, 1976. – 224 с.
3. Панин Ю.А., Малкевич С.Г., Дунаевская Ц.С. Фторопласты. – Л., 1978. – 232 с.
4. Чоу Р. Революционная технология затворов на основе диэлектрика с высокой диэлектрической проницаемостью корпорации Intel открывает путь для дальнейшей реализации закона Мура // Новинки и тенденции технологий. – 01, 2004.
5. Беженар А.А., Копань В.С., Півень Н.І. Вісник Київського університету. Серія: фіз.-мат. науки. – 2006. – №3. – с. 410 - 414.
6. Балсанова Л.В., Базаров Б.Г., Базарова Ж.Г. Оксидные керамические материалы, содержащие гафний, как твердые электролиты // Всероссийская конференция "Керамика и композиционные материалы". – Сыктывкар, 2004. – 236 с.

Надійшла до редколегії 12.02.2008

УДК 621.382

Ніколенко* А.С., студ., Кондратенко С.В., к.ф.-м.н., доц., Вакуленко О.В., д.ф.-м.н., проф., Юхимчук В.О., к.ф.-м.н.

Вплив термічних відпалів на фотопровідність Si/Ge гетероструктур з квантовими точками

В роботі методом спектроскопії фотоструму досліджено вплив термічних відпалів Si/Ge структур з КТ на енергію переходів дірок з локалізованих в КТ станів в делокалізовані стани змочувального шару та Si оточення. Компонентний склад та пружні деформації в структурах оцінено методом спектроскопії КРС. Зменшення енергій переходів на спектрах фотопровідності з відпалом пояснено впливом інтердифузії, що призводить до розмиття гетерограниці та зменшення глибини потенціальної ями КТ.

Ключові слова: гетероструктури, квантові точки, фотопровідність.

A. S. Nikolenko*, stud., S. V. Kondratenko, Ph. D., O. V. Vakulenko, Dr., V. O. Yukhymchuk, Ph. D.

Effect of thermal annealing on photoconductivity of Si/Ge heterostructures with quantum dots

Effect of thermal annealing of Si/Ge heterostructures with quantum dots on the energy of holes' bound-to-continuum transitions from quantum dot to wetting layer and Si surrounding was investigated by photocurrent spectroscopy. Composition and mechanical stresses of investigated structures were estimated by Raman scattering spectroscopy. The decrease of transition energies in photocurrent spectra with annealing was explained by effect of interdiffusion which causes heteroboundary spreading and decreasing of quantum dot potential well.

Key Words: heterostructures, quantum dots, photoconductivity.

*E-mail: Nikolenko_mail@ukr.net

Вступ

Гетероструктури Si/Ge з квантовими точками (КТ) є об'єктом інтенсивних досліджень в останні декілька років. Такий інтерес, по-перше, пов'язаний з перспективою створення на основі таких структур для формування активних елементів принципово нових електронних приладів з використанням квантово-розмірних ефектів на базі добре розвинутої кремнієвої технології. По-друге, необхідністю розуміння фізичних процесів, які є відповідальними за ефекти самоформування КТ в режимі росту Странського-Крастанова та їх енергетичну структуру.

Особливістю таких структур є наявність неоднорідних механічних напружень. Всередині КТ деформація має характер розтягу в напрямку їх росту та стиску в площині росту. Найбільш напруженою є область наноострівців по контуру їх основи в площині росту, а найбільш відрелаксованою – їх вершини. В процесі зарощування острівців кремнієм відбувається їх стиск як в площині росту, так і в напрямку росту. В кремнієвому оточенні найбільш напружена область знаходиться поблизу вершини КТ [1]. Наявні неоднорідні механічні напруження сприяють процесу інтердифузії за участю атомів

кремнієвої підкладки. В свою чергу компонентне перемішування призводить до зменшення напружень. Особливо цей процес актуальний при високих температурах епітаксії (600-750°C) [2]. При відносно низьких температурах формування острівців (300°C) процес компонентного перемішування стає помітним при вирощуванні багатошарових структур. В цьому випадку зростає час, протягом якого відбувається інтердифузія, обумовлена неоднорідними напруженнями, крім того в острівці вбудовуються атоми Si не тільки з підкладки, але й з кремнієвих спейсерів. Внаслідок компонентного перемішування в процесі росту, наноострівці являють собою Si_{1-x}Ge_x твердий розчин зі змінним об'ємним компонентним складом. Відпал структур при високих температурах (600-700°C) також призводить до значної додаткової релаксації напружень за рахунок компонентного перемішування, що в свою чергу обумовлює збільшення розмірів КТ [3].

Si/Ge Гетероструктури з квантовими точками відноситься до другого типу, в яких просторове обмеження існує лише для носіїв заряду одного знаку. В даному випадку в GeSi острівцях реалізується квантова яма для дірок. Наявність неоднорідних напружень призводить до вигину

зон на границі гетеропереходу та утворення зв'язаних станів для електронів в області біля КТ. Накопичення дірок в КТ призводить до збільшення вигину зон та, відповідно, до збільшення енергії зв'язку електронів [4].

Оптичні переходи в SiGe гетероструктурах з КТ досліджувались методами спектроскопії поглинання, фотолюмінісценції та фотоструму [4-10]. В області 0.7 – 0.8 еВ як в спектрах люмінесценції, так і в спектрах фотоструму спостерігаються непрямі в просторі переходи електронів з локалізованих станів КТ в делокалізовані стани зони провідності кремнієвого оточення [4,9,11-13]. При збільшенні рівня заселеності КТ дірками спостерігається зсув енергії такого переходу в високоенергетичний бік [14]. В структурах р-типу в області 0.3 – 0.4 еВ спостерігаються переходи дірок між підзонами валентної зони та рівнями розмірного квантування наноострівців [5,11,15]. Умовою спостереження останніх є наявність дірок в КТ, що реалізується в структурах з р-типом провідності. Енергія переходів обох типів залежить від розривів зон гетеропереходу та впливу розмірного квантування.

Метою даної роботи було встановлення взаємозв'язку між енергетичним спектром наноострівців, який залежить від компонентного складу, та напружень, з одного боку та енергією відповідних оптичних переходів з іншого.

Експериментальна частина

Структури, що досліджувалися в даній роботі, були отримані методом молекулярно-променевої епітаксії на підкладці з кремнію КДБ-20 орієнтації (001), з вирощеним на ній спочатку буферним кремнієвим шаром товщиною 50 нм. На буфер осаджувалось 8 моношарів (МШ) германію, після чого, сформовані острівці покривалися шаром кремнію 12 нм. Така процедура повторювалася 10 раз. Температура підкладки при епітаксії складала 300°C.

Для дослідження повздовжньої фотопровідності на поверхні з епітаксіальними шарами вплавліялися два омичні контакти із золота діаметром 1 мм на відстані 10 мм один від одного. Вимірювання спектральних залежностей фотопровідності проводились на базі стандартного інфрачервоного спектрометра. Модульоване випромінювання від джерела світла (глобара) після проходження монохроматора фокусувалось на поверхні з епітаксіальними шарами в області між контактами. Вимірювання фотоструму проводились з використанням

підсилювача струму та стандартної техніки синхронного детектування. Виміряні спектральні залежності приводилися до сталої кількості квантів збуджувального випромінювання за допомогою реєстрації випромінювання піроелектричним приймачем.

Спектри комбінаційного розсіювання світла (КРС) реєструвалися при кімнатній температурі на подвійному дифракційному спектрометрі ДФС-24. Для збудження спектрів КРС використовувалося випромінювання Ag⁺-лазера з довжиною хвилі $\lambda=514.5$ нм. Розсіяне світло реєструвалося охолодженим фотоелектронним помножувачем, який працював в режимі лічби фотонів. Реєстрація смуг КРС здійснювалося в геометрії $z(x,y)-z$, де осі x , y , z , паралельні до кристалографічних напрямків [100], [010], [001], відповідно.

Відпал зразків проводився в атмосфері азоту в діапазоні температур 500-700°C протягом 15 хв

Результати та обговорення

Спектри КРС досліджуваних зразків до відпалу та після відпалу наведено на Рис.1. Використавши відомі співвідношення для $Si_{1-x}Ge_x$ твердого розчину острівців:

$$\omega_{GeSi}(cm^{-1}) = 387 + 81 \cdot (1-x) - 78 \cdot (1-x)^2 - 575 \cdot \epsilon \quad (1)$$

$$\omega_{GeGe}(cm^{-1}) = 282.5 + 16 \cdot x - 385 \cdot \epsilon \quad (2)$$

що пов'язують частоту відповідних смуг з їх компонентних складом (x) та величиною пружної деформації (ϵ), було знайдено величини x , ϵ (табл.1).

Табл.1. Температура відпалу, частоти Ge-Ge та Ge-Si смуг, вміст Ge та пружна деформація.

T, °C	ω_{Ge-Ge} cm ⁻¹	ω_{Ge-Si} cm ⁻¹	x	ϵ
до відпалу	313.2	419.8	0.91	-0.040
600°C	309.3	421.3	0.82	-0.035
700°C	302.0	421.8	0.61	-0.025

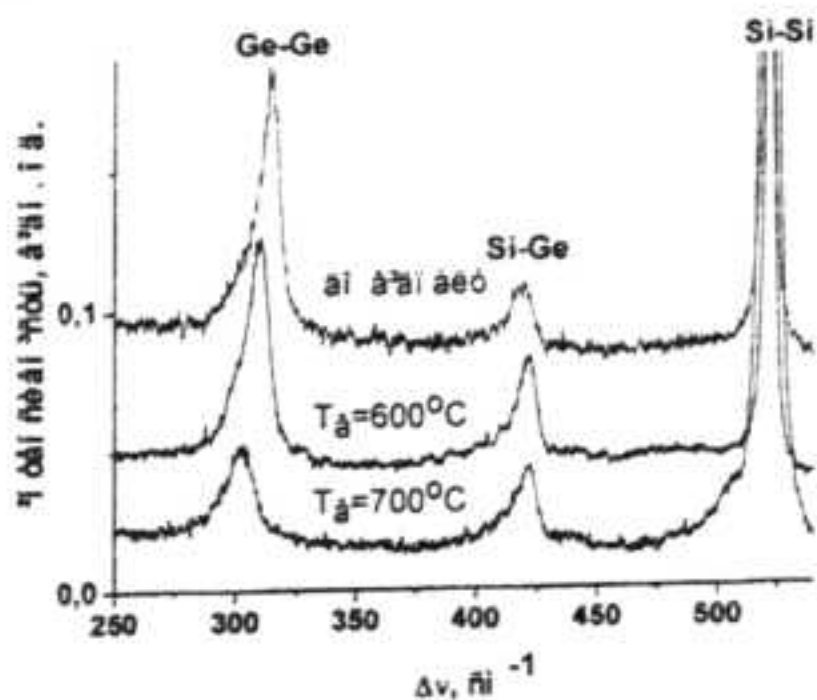


Рис.1. Спектр КРС досліджуваних Si/Ge гетероструктур до відпалу та після відпалу при температурах 600 та 700°C.

На рис.2 наведено спектральні залежності поздовжньої фотопровідності досліджуваної структури, виміряні при 77 К до відпалу та після відпалів при температурах 600 та 700°C. Для всіх спектрів фотопровідності в області від 0.3 до 0.4 eV спостерігалось два максимуми (T1, T2), частотне положення та форма яких змінювались з відпалом. Зокрема, після відпалу при 700°C, спостерігався зсув смуги T2 в інфрачервону область на величину порядку 10 меВ. Також на спектрах фотопровідності в області 0.38 eV спостерігалось наростання сигналу (T3), початок якого з відпалом теж зазнавав зсуву в низькочастотний бік. Для пояснення природи фоточутливості в цій області слід розглянути зонну діаграму досліджуваної структури. Відомо, що в системі SiGe КТ реалізується другий тип гетероструктур, в яких потенційні ями для носіїв заряду протилежних знаків знаходяться по різні боки гетеропереходу. Дірки у валентній зоні наноострівців виявляються локалізованими в потенціальній ямі і не можуть приймати участь у процесах переносу. Фотострум в середній інфрачервоній області в структурах р-типу може бути зумовлений переходами дірок з локалізованих станів валентної зони наноострівців в делокалізовані стани валентної зони кремнієвого оточення та/або у стани валентної зони змочувального шару, де вони є вільними і можуть давати внесок у фотострум [5, 11, 15, 16]. Можна припустити, що у нашому випадку походження максимумів T1 та T2 пов'язано з переходами дірок з квантоворозмірних станів валентної зони наноострівців в стани змочувального шару, наростання сигналу в області від 0.38 eV пов'язане з переходами дірок з локалізованих станів валентної зони наноострівців в континуум делокалізованих станів валентної зони кремнієвого оточення (рис.4а). Енергія таких переходів залежить від величини розриву валентної зони гетеропереходу та величини квантового обмеження.

Величина енергетичного розриву валентної зони в таких структурах залежить від компонентного складу острівців та їх оточення, наявності механічних напружень та впливу квантоворозмірних ефектів. Найкраще узгодження між теоретичними розрахунками та експериментальними даними для ненапруженого Si/Si_{1-x}Ge_x гетеропереходу дає середня величина розриву валентної зони $\Delta E_V^0 = 0.58 \cdot x$ [17].

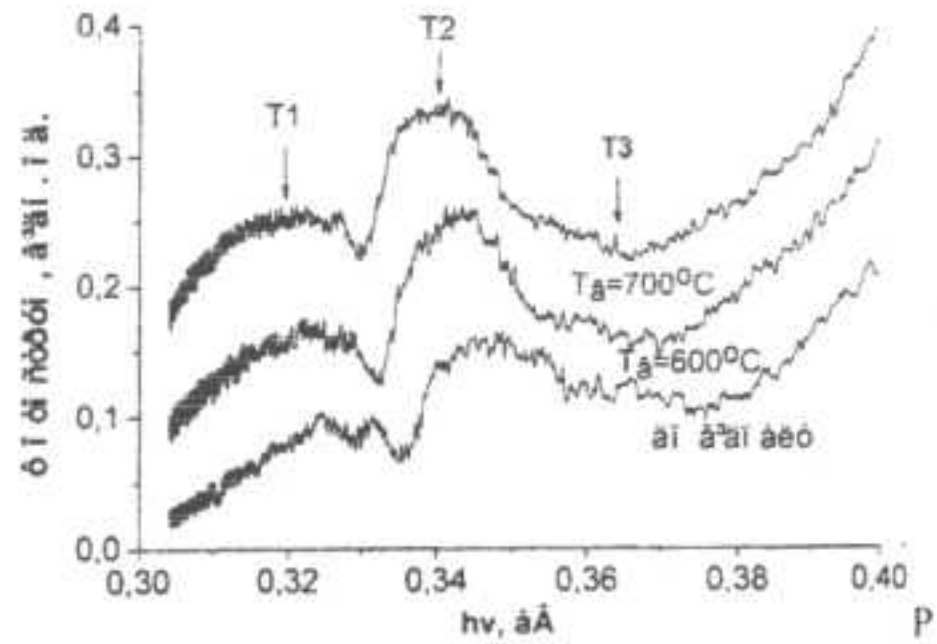


Рис. 2. Спектральні залежності поздовжньої фотопровідності багат шарової гетероструктури з GeSi КТ при 77 К до відпалу та після відпалу при температурах 600 та 700°C.

Використовуючи визначені зі спектрів КРС величини пружної деформації (ϵ_{xx} , ϵ_{zz}), можна оцінити величину розриву валентної зони на інтерфейсах з врахуванням деформацій [18]:

$$\Delta E_V = \Delta E_V^0 + a_V \cdot (2 \cdot \epsilon_{xx} + \epsilon_{zz}) \quad (6)$$

де a_V - деформаційний потенціал Ge. Таким чином величина енергетичного розриву валентної зони гетеропереходу Si_{1-x}Ge_x/Si може бути апроксимована залежністю від вмісту x наступним виразом: $\Delta E_V(x) = \Delta E_V^0 \cdot x$ [eV] [19].

Наявність двовісної деформації стиску в наноострівців призводить до зняття виродження валентної зони, при якому відбувається її розщеплення та зсув зони важких дірок в бік менших енергій по відношенню до зони легких дірок. Величини зсувів зон для легких та важких дірок відносно середнього значення ΔE_V можуть бути визначені з врахуванням даних КРС для досліджуваних структур наступним чином [17]:

$$\Delta E_{lh} = -\frac{1}{6} \Delta_0 + \frac{1}{4} \delta E + \frac{1}{2} \left[\Delta_0^2 + \Delta_0 \delta E + \frac{9}{4} (\delta E)^2 \right]^{1/2} \quad (7)$$

$$\Delta E_{hh} = -\frac{1}{3} \Delta_0 - \frac{1}{2} \delta E. \quad (8)$$

де величина δE у випадку деформації в напрямку [001] має вигляд $\delta E_{001} = 2a_V(\epsilon_{zz} - \epsilon_{xx})$. Величини спін-орбітального розщеплення валентної зони Δ_0 , деформаційного потенціалу a_V та пружних модулів c_{12} , c_{11} для Si_{1-x}Ge_x наноострівців визначаються лінійною інтерполяцією даних для Si та Ge. Як показали розрахунки величини енергетичних розривів валентної зони для важких ΔE_{hh} та легких дірок ΔE_{lh} до відпалу становлять 0.769 та 0.617 eV, відповідно.

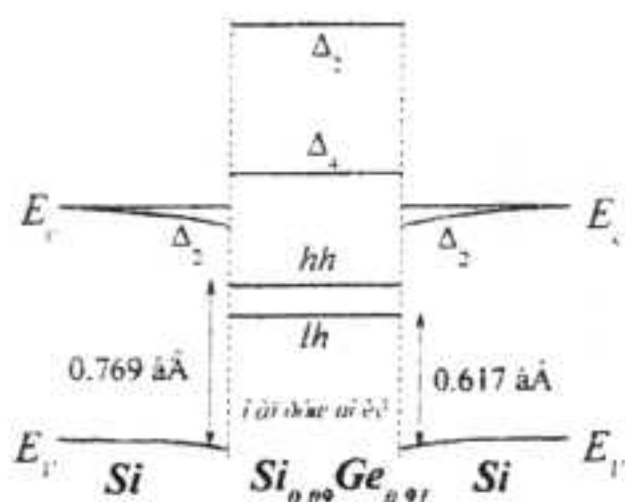


Рис.3. Зонна структура гетеропереходу GeSi-KT/Si-оточення досліджуваної структури до відпалу.

Для оцінки ефектів квантового обмеження розглянемо КТ як тривимірну прямокутну потенціальну яму зі стінками кінцевої висоти. При цьому висота стінок буде визначатися розрахованими вище розривами зон ΔE_{lh} та ΔE_{hh} для підзон легких та важких дірок відповідно. Оскільки підзона важких дірок забезпечує домінуючий внесок ($\sim 85\%$) у стани дірок [20], тому в подальшому обмежимося розглядом задачі лише для підзони важких дірок. Ефективна маса важких дірок m_{SiGe}^* визначалася як $1/m_{SiGe}^* = x/m_{Ge}^* + (1-x)/m_{Si}^*$, де m_{Ge}^* , m_{Si}^* - ефективні маси важких дірок для чистого германію та кремнію, які становлять $0.537m_0$ та $0.33m_0$, відповідно. Для оцінки дозволених енергетичних станів було проведено чисельний розв'язок стаціонарного рівняння Шредінгера для важких дірок в прямокутній потенціальній ямі. Енергія перших станів в потенціальній ямі КТ становить 0.372 eV (111), 0.385 eV (211, 121), 0.399 eV (221). Енергія переходу з основного стану КТ (111) в континуум станів кремнієвого оточення становить 0.397 eV. Тому наростання сигналу фотопровідності (Т3) пов'язане з переходами важких дірок з квантоворозмірних станів КТ в валентну зону Si оточення (рис.4а). При переході з квантоворозмірних станів КТ в стани валентної зони змочувального шару, енергія переходів є меншою за енергію відповідних переходів в валентну зону кремнієвого оточення, тому максимум фотопровідності Т2 може бути зумовлений переходами зі станів з квантовими числами (111) та (211, 121) в стани валентної зони змочувального шару, а максимум Т1 з переходами зі стану (221) в стани валентної зони змочувального шару (рис.4а). Ймовірність переходів з квантоворозмірних станів КТ з вищими індексами є набагато меншою, тому їх впливом можна знехтувати [21, 22].

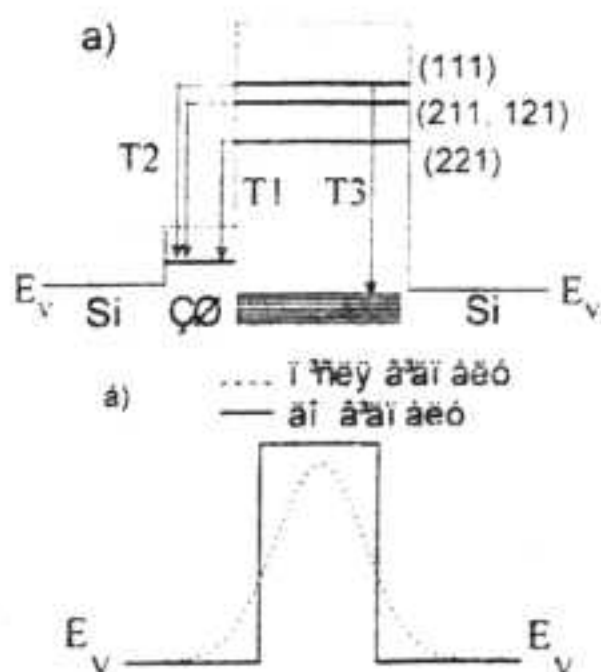


Рис.4. Схема досліджуваних переходів (Т1, Т2, Т3) з локалізованих станів в делокалізовані (а) та схематичне зображення валентної зони гетеропереходу до та після відпалу структури (б).

Як випливає з даних спектроскопії КРС при відпалі структур при 700°C концентрація германію в КТ зменшується з 0.91 до 0.61. Проте зміщення енергій переходів Т1, Т2 та Т3 є незначним (на 2% у порівнянні зі значенням для невідпаленої структури). При цьому, визначення глибини потенціальної ями використовуючи формули (7) та (8) вже не є коректним. Відпал зразків разом зі зменшенням величин механічних напружень та зменшенням вмісту Ge внаслідок процесів інтердифузії, призводить також до розмиття границі гетеропереходу між КТ та Si оточенням (рис.4б). В результаті гетероперехід стає більш плавним і, відповідно, змінюється форма потенціальної ями і зменшується її глибина [3]. Як наслідок, має зменшитись енергія переходів з локалізованих станів КТ в делокалізовані стани змочувального шару та кремнієвого оточення. Зміщення енергій переходів Т1, Т2 та Т3 в бік менших енергій з відпалом можна пояснити зменшенням енергії переходів з локалізованих станів в делокалізовані. Слід зазначити, що на енергію переходу з основного стану локалізованого в КТ в делокалізовані стани змочувального шару та кремнієвого оточення ефект розмиття гетерограниці впливає несуттєво. Це можна пояснити тим, що хвильова функція основного стану є локалізованою всередині об'єму КТ.

Таким чином, відпал структур з SiGe КТ призводить до релаксації напружень, зменшення середнього вмісту Ge в КТ, розмиття гетерограниці між КТ та Si оточенням та, як результат, зменшення енергії переходів з локалізованих станів валентної зони КТ в делокалізовані стани змочувального шару та Si.

Висновки

В роботі було досліджено вплив термічних відпалів SiGe структур з КТ на енергію переходів дірок з локалізованих станів КТ в делокалізовані стани змочувального шару та Si оточення. Компонентний склад та пружні деформації в структурах визначались з допомогою спектроскопії комбінаційного розсіяння світла. Розриви валентної зони гетеропереходу досліджуваних структур оцінювались методом деформаційних потенціалів використовуючи данні КРС. Зміщення енергій переходів на спектрах фотопровідності в бік менших енергій з відпалом пояснено впливом інтердифузії, що призводить до розмиття гетерограниці та зменшення глибини потенціальної ями КТ.

Подяка

Автори роботи вдячні доктору Нікіфорову А.І. за надані зразки.

Список використаних джерел

1. А.В. Двуреченский, А.И. Якимов Квантовые точки 2 типа в системе Ge/Si // ФТП – 2001. - 35. - с. 1143-1153.
2. Z.F. Krasilnik, P.M. Lytvyn, D.N. Lobanov et al. Microscopic and optical investigation of Ge nanoislands on silicon substrates // Nanotechnology – 2002. – 13. - p. 81-85.
3. Song Tong, Fei Liu, A. Khitun et al. Tunable normal incidence Ge quantum dot midinfrared detectors // J.Appl.Phys. – 2004. - 96, - p. 773.
4. A.I. Yakimov, N.P. Stepina, A.V. Dvurechenskii Interband absorption in charged Ge/Si type-II quantum dots // Phys.Rev.B.- 2001. – 63. - p. 045312-045318.
5. T. Fromherz, W. Mac, et al. Intraband absorption and photocurrent spectroscopy of self-assembled p-type Si/SiGe quantum dots // Appl. Phys. Lett. – 2002. – 80. – p. 2093-2095.
6. P. Boucaud, V. Le Thanh et. al. Intraband absorption in Ge/Si self-assembled quantum dots // Appl. Phys. Lett. – 1999. – 74. - p. 401.
7. A.I. Yakimov, N.P. Stepina, et al. Depolarization shift of the in-plane polarized interlevel resonance in a dense array of quantum dots // Phys.Rev.B. – 2000. – 62. - p. 9939-9942.
8. M. El kurdi et. al. Electroabsorption spectroscopy of Ge/Si self-assembled islands // J.Appl.Phys. – 2005. - 97. - p. 083525-083532.

9. M. Larsson, A. Elfving, P. O. Holtz and W.-X. Ni Spatially direct and indirect transitions observed for Si/Ge quantum dots // Appl. Phys. Lett. – 2003. - 82, - p. 4785-4787.
10. O. G. Schmidt, U. Denker, and K. Eberl et. al. Effect of overgrowth temperature on the photoluminescence of Ge/Si islands // Appl. Phys. Lett. -2000. – 77. - p. 2509-2511.
11. A.I. Yakimov, A.V. Dvurechenskii, Yu.Yu. Proskuryakov Normal-incidence infrared photoconductivity in Si p-i-n diode with embedded Ge self-assembled quantum dots // Appl. Phys. Lett. – 1999. – 75. p. 1413-1415.
12. S. Tong, J.L. Liu, J. Wan, and Kang L. Wang Normal-incidence Ge quantum-dot photodetectors at 1.5 mm based on Si substrate // Appl. Phys. Lett. – 2002. – 80. - p. 1189.
13. A.I. Yakimov, A.V. Dvurechenskii, A.I. Nikiforov Stark effect in type-II Ge/Si quantum dots // Phys.Rev.B – 2003. – 67. – p. 125318.
14. P. Boucaud, S. Sauvage, M. Elkurdi et al. Optical recombination from excited states in Ge/Si self-assembled quantum dots // Phys.Rev.B. – 2001. - 64, - p. 155310-155316.
15. C.H. Lin et al. Broadband SiGe/Si quantum dot infrared photodetectors // J.Appl.Phys. – 2007. – 101. – p. 033117-033121.
16. O.V. Vakulenko, S.V. Kondratenko, A.S. Nikolenko et.al. Optical absorption in Ge/Si heterostructures with Ge QDs // Nanotechnology - 2007. - 18, - p. 185401.
17. Chris G.Van de Walle Band lineups and deformation potentials in the model-solid theory // Phys. Rev. B. – 1989. – 39. - p. 1871.
18. E. Kasper Properties of Silicon Germanium and SiGe: Carbon. – London: INSPEC, 2000.
19. M.Ya. Valakh, V.O. Yukhymchuk, A.S. Nikolenko Optical and photoelectrical properties of GeSi nanoislands // Semicond Sci. Technol. – 2007. - 22, - p.326-329.
20. A.V. Nenashev, A.V. Dvurechenskii and A.F. Zinovieva Wave functions and g factor of holes in Ge/Si quantum dots // Phys.Rev.B. – 2003. – 67. – p. 205301-205311.
21. А.Я. Шик Физика низкоразмерных систем. - Санкт-Петербург, Наука, 2001.
22. Л.Е. Воробьев Оптические свойства наноструктур. - Санкт-Петербург, Наука, 2001.

Надійшла до редколегії 10.03.2008

УДК 548:539.32

Онанко А.П., к. ф.-м. н., с.н.с.

Вказівна поверхня непружно-пружного тіла сплаву Ti_3Al

Досліджувалась температурна залежність внутрішнього тертя і модуля пружності (вказівної поверхні непружно-пружного тіла) сплаву Ti_3Al . Виявлено релаксацийний максимум внутрішнього тертя при ~ 390 К, обумовлений процесом переорієнтації втілених атомів H в гантельних конфігураціях.

Ключові слова: модуль пружності, внутрішнє тертя, дефекти.

E-mail: onanko@mail.kiev.ua

Вимір внутрішнього тертя (ВТ) при різних температурах, амплітудах ультразвукової деформації ε та на різних частотах f дозволяє одержувати додаткову інформацію про не тільки про властивості та симетрію дефектів структури, але й про динаміку їх взаємодії. Неруйнуючий метод ВТ дозволяє притулити спектр структурних дефектів: розмішуючи їх по тривалості релаксації τ і по їх внеску в затухання пружних коливань. Вимірювання фону ВТ Q^{-1}_f після різних термообробок дає інформацію про зміни полів термопружних напруг σ_i в сплав Ti_3Al . Зростання висот максимумів ВТ Q^{-1}_m свідчить про зростання концентрації дефектів, а уширення максимумів ΔQ^{-1}_m при цьому відображує процес релаксації нових типів структурних дефектів. Наступна робота є продовженням [1].

Експеримент

Для виміру температурних залежностей модуля пружності E і ВТ сплаву Ti_3Al після механічного різання і шліфування використовувалися методики згінних резонансних коливань на частоті $f \approx 1$ кГц і чотирьохскладового п'єзоелектричного вібратора на частоті $f \approx 118$ кГц при знакозмінній деформації $\varepsilon \approx 10^{-6}$ у вакуумі не гірше $P \approx 10^{-3}$ Па [1]. Погрішність виміру відносної зміни модуля пружності ΔE

Onanko A.P., Sen. Sci. Res., PhD

Indicator surface of inelasticity-elasticity state of Ti_3Al alloy

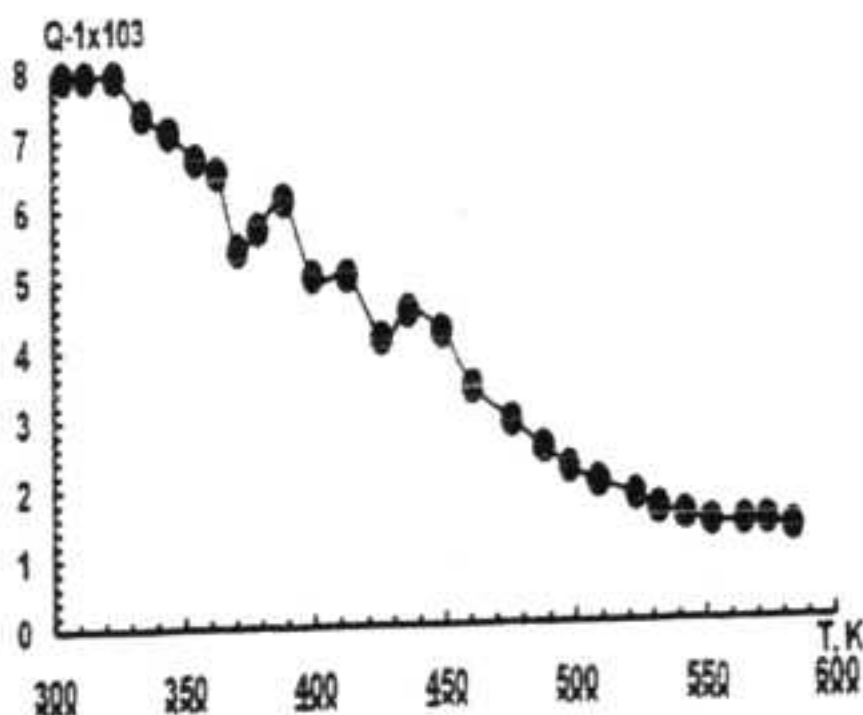
Temperature dependence of internal friction and elastic module (indicator surface of inelasticity-elasticity state) of Ti_3Al alloy were explored. There was the relaxation maximum of internal friction at ~ 390 K, conditioned by the mechanism caused by reorientation interstitial atoms of H in dumbbell configurations.

Key words: elastic module, internal friction, defects.

дорівнювала $\sim 0,1\%$, а виміру ВТ $\sim 2\%$. Для визначення температурного положення максимумів ВТ по релаксації модуля пружності $\Delta E/E$ одночасно вимірювалась температурна залежність E . Насичення зразків воднем відбувалося електролітичним методом. В якості аноду використовувалася платинова фольга, зразок був катодом.

Результати і їх обговорення

На рис.1 наведено температурну залежність внутрішнього тертя сплаву Ti_3Al у початковому стані після механічної обробки.



Представлено: член-кор. НАН України,
д.ф.-м.н., проф. Макара В.А.

Рис. 1. Температурна залежність внутрішнього тертя Q^{-1} сплаву Ti_3Al у початковому стані після механічної обробки.

Спостерігалася висока початкова величина фону ВТ $Q^{-1}_f \approx 8,0 \cdot 10^{-3}$. Значний високий початковий фон ВТ при першому нагріванні порівнюючи з фоном при повторному нагріванні свідчить про наявність полів напруг σ_i , які виникають внаслідок термомеханічної обробки, в сплав Ti_3Al . Аномальний хід фону ВТ на рис. 1 свідчить про зменшення величин цих напружень σ_i у процесі нагріву. Відпал структурних дефектів викривляє вигляд температурного спектру ВТ. При відпалі має місце ухід домішок, вакансій.

Виявлено максимум ВТ в сплав Ti_3Al при $T_{M1} \approx 390$ К. Для максимуму ВТ умова його існування записується у вигляді [2]:

$$\omega \tau_0 \exp(H/k_B T_m) = 1. \quad (1)$$

З наведеного співвідношення випливає, що $H = k_B T_m \ln(1/f_m \tau_0)$. Підставивши $\tau_0 = 1/f_0 = h/k_B T_m$, одержуємо формулу для визначення енергії активації максимуму ВТ

$$H = k_B T_m \ln(k_B T_m / h f_m) + \Delta S. \quad (2)$$

де T_m і f_m – температура і частота максимуму відповідно. Ентропійна добавка ΔS лежить в межах погрішності експерименту; точність оцінки H залежить в основному від точності підтримки температури нагріву зразка (2 К) $\approx (2\%)$. Використання для визначення H температури максимуму ВТ T_m і частоти коливань зразка f_m не вимагає виконання спеціальних умов. При накладенні процесів релаксації температурне положення T_m кожного максимуму легко вимірюється, а частота f_m задається експериментально.

Енергію активації H максимуму ВТ оцінювалася за формулою Верта-Маркса (2). Була отримана величина енергії активації $H_1 \approx 0,77 \pm 0,1$ еВ, з умови існування релаксацийного максимуму ВТ формули (1) $\omega \tau = 1$ визначена постійна часу релаксації $\tau_{01} \approx 1,9 \cdot 10^{-14}$ сек та частотний фактор релаксації $f_{01} \approx 5,3 \cdot 10^{13}$ Гц

Динамічний модуль пружності E є більш чутливою характеристикою процесів фізико-хімічної взаємодії по границям розділу, ніж ВТ. Результати вимірювання температурної залежності модуля пружності E сплаву Ti_3Al у початковому стані після механічної обробки представлені на рис. 2.

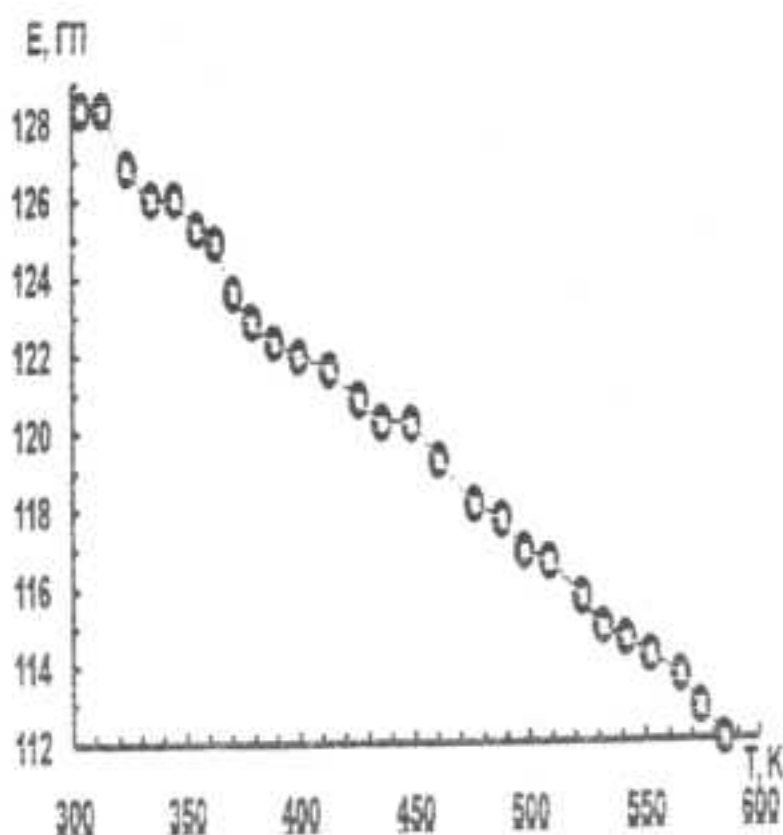


Рис. 2. Температурна залежність модуля пружності E сплаву Ti_3Al у початковому стані після механічної обробки.

Модуль пружності E має спадний лінійний характер, що пояснюється теорією ангармонізму кристалів з врахуванням квадратичних членів [2]. Зменшення модуля пружності не було пов'язане з гістерезисним дислокаційним процесом, так як ВТ було амплітудне незалежним. На температурній залежності E сплаву на рис. 2 спостерігалась релаксація модуля пружності $\Delta E/E$ в області існування максимуму ВТ $T_{M1} \approx 390$ К. Після відпалу спостерігалось збільшення абсолютних значень модуля пружності E , що свідчило про зміцнення сплаву.

Визначено малий максимум ВТ при $T_{M2} \approx 440$ К з енергією активації $H_2 \approx 0,85 \pm 0,1$ еВ, з постійної часу релаксації $\tau_{02} \approx 1,8 \cdot 10^{-14}$ сек і частотним фактором релаксації $f_{02} \approx 5,5 \cdot 10^{13}$ Гц. Цій максимум ВТ можливо обумовлений процесом зміни числа вакансій в вакансійних комплексах під впливом змінної деформації.

Виявлено також малий максимум ВТ при $T_{M3} \approx 530$ К з енергією активації $H_3 \approx 1,0 \pm 0,1$ еВ, з

постійної часу релаксації $\tau_{03} \approx 3,1 \cdot 10^{-14}$ сек і частотним фактором релаксації $f_{03} \approx 3,2 \cdot 10^{13}$ Гц. Релаксація модуля пружності $\Delta E/E$ в температурному інтервалі існування максимуму ВТ в сплаві дозволяє припустити відповідний механізм, обумовлений релаксацийним процесом переорієнтації комплексів вакансія – водень V-H під впливом ультразвукової змінної деформації.

На рис.3 представлена температурна залежність вказівної поверхні непружно-пружного тіла сплаву Ti_3Al .

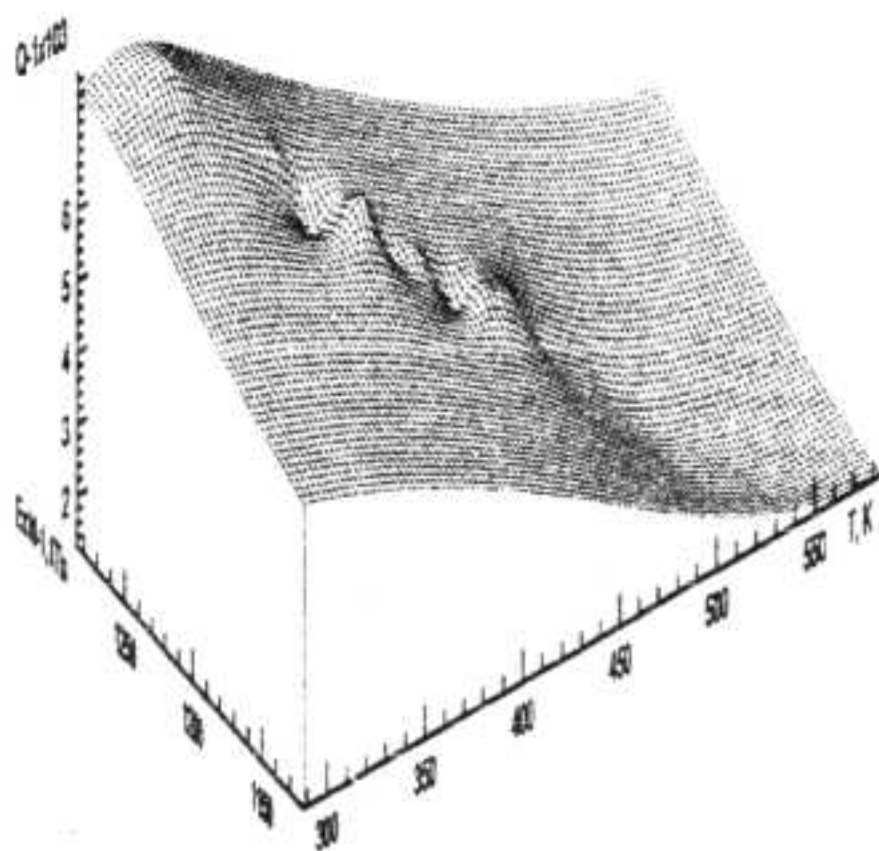


Рис. 3. Температурна залежність внутрішнього тертя Q^{-1} і модуля пружності E (вказівної поверхні непружно-пружного тіла) сплаву Ti_3Al у початковому стані після механічної обробки і відпалу.

Причиною індукованої воднем H локалізованої пластичності є підвищена концентрація електронів провідності – підсилений воднем металічний характер міжатомних зв'язків. Локальне збільшення концентрації електронів провідності в водневих хмарах навколо дислокацій призводить до зменшення модулю зсуву G і напруги активації джерел дислокацій σ_d [3]. H підсилює металічний характер міжатомних зв'язків, що не може бути причиною крихкого руйнування. Зменшення відстані між дислокаціями L_d в скупченнях завдяки H призводить до збільшення дислокацій в них і до збільшення напруження на

провідну дислокацію σ_d , що є причиною полегшеного формування мікротріщини.

Поблизу вершин мікротріщин, що утворюються при механічній обробці поверхні завжди існує певна початкова надлишкова (над середньою по кристалу) локальна щільність дислокацій N_d . Миттєвий розподіл цих дислокацій у випадку ненапруженої тріщини та при впливі зовнішньої знакозмінної (ультразвукової) механічної напруги може суттєво відрізнятись. Відповідно сила, що діє на верхівку мікротріщини, порушує умову термодинамічної стабільності початкового положення цієї верхівки, що призводить до її стрибкоподібного руху до нового метастабільного рівноважного положення. Цей процес супроводжується збільшенням радіусу верхівки мікротріщини та утворенням комплексу замкнутих дислокаційних петель навколо верхівки, що додатково гальмують розвиток мікротріщини.

Релаксація модуля пружності $\Delta E/E$ в області існування максимуму ВТ в сплаві Ti_3Al при $T_{M1} \approx 390$ К дозволяє припустити відповідний механізм, обумовлений релаксацийним процесом переорієнтації втілених атомів H в гантельних конфігураціях під впливом ультразвукової змінної деформації.

Список використаних джерел

1. Онанко А.П. Кінетика пружно-непружного тіла сплаву $Al_{0.5}Be_{0.5}$ // Вісник Київського університету. Сер.: Фізико-математичні науки. – 2007. – № 1. – с.394–395.
2. Новик А., Берри Б. Релаксационные явления в кристаллах. – М.: Атомиздат, 1975. – 472 с.
3. Колачев Б.А. Физическое металловедение титана. – М.: Металлургия, 1976. – 276 с.

Надійшла до редколегії 18.02.08

В.Є. Погорелов, проф., д.ф.-м.н., П.О. Голуб, студ., Н.І. Школьнікова, к.ф.-м.-н.

Спектр комбінаційного розсіювання 4-(н-гексилосі)феніл-4'-(н-октилокси)бензоат в кристалічній фазі.

Досліджено та інтерпретовано спектр комбінаційного розсіювання світла нещодавно синтезованого рідкого кристалу

Виміряно і проаналізовано повний коливний спектр 4-(н-гексилосі)феніл - 4'-(н-октилокси)бензоат в кристалічній фазі.

Ключові слова: рідкий, кристал, спектр, комбінаційне розсіювання світла

Останнім часом рідкокристалічні дисплеї знайшли широке застосування в різноманітних інформаційних технологіях. Це монітори персональних комп'ютерів, телевізійних приймачів, різноманітних інших засобів відображення інформації. В свою чергу, це зумовило синтезування нових речовин, здатних утворювати рідкокристалічні структури з передбачуваними властивостями. До таких речовин відносяться два ізомери складних ефірів—4-(н-гексилосі)феніл-4'-(н-октилокси)бензоат (далі – зразок X_1) та 4-(н-октилокси)феніл-4'-(н-гексилосі)бензоат (далі – зразок X_2).

Незважаючи на близькість хімічної структури цих двох зразків, вони відрізняються характером фазових перетворень. Так, в зразку X_1 перехід з стану ізотропної рідини до нематичної мезофази відбувається за температури 362.5 K, охолодження цього зразка до температури 338 K зумовлює мезофазний перехід нематик – смектик С, подальше охолодження перехід з стану ізотропної рідини до нематичної мезофази відбувається за температури 362 K, мезофазний перехід нематик – смектик С спостерігається за температури 319 K.

Зразок X_2 переходить до кристалічного стану за температури 307 K. Нагрів зразка з кристалічного стану до температури 324 K зумовлює перехід до нематичного стану. При цьому в процесі нагріву смектична фаза не утворюється (на відміну від зразка X_1).

Відомо, що коливальні спектри суттєво пов'язані з будовою молекул, з особливостями внутрішньо молекулярної динаміки. Тому доцільно дослідити особливості коливальних спектрів визначених вище двох ізомерів складного ефіру, трансформацію коливальних

V.Ye. Pogorelov, Prof., Dr. Sci., P.O. Golub, N.I. Shcolnicova, PhD

The Raman spectrum of 4-(n-hexyloxy)phenyl-4'-(n-octyloxy)benzoate in crystal phase.

The spectrum of recently created liquid crystal has been investigated and interpreted.

It measured and analyzed the full vibrational spectrum of 4-(n-hexyloxy)-phenyl-4-(n-octyloxy)-benzoate in crystal phase.

Key words: liquid, crystal, spectrum, Raman scattering.

спектрів при фазових перетвореннях цих зразків, й, врешті рещт, визначити зв'язок між специфікою фазових перетворень і особливостями структури цих двох ізомерів.

В цій публікації аналізується спектр комбінаційного розсіювання 4-(н-гексилосі)феніл-4'-(н-октилокси)бензоат (далі – зразок X_1)

В роботі використовувався автоматизований спектрально-обчислювальний комплекс на базі подвійного монохроматора ДФС-24.

Джерело збуджуючого комбінаційне розсіювання (КР) світла випромінювання - аргонний лазер з довжиною хвилі випромінювання 488 нм, з потужністю 50–100 мВт. Така потужність не приводить до суттєвого нагрівання зразка під час реєстрації спектрів КР.

Розсіяне світло спостерігається під кутом 90° відносно напрямку збуджуючого лазерного променя. Всі спектри було записано при ширинах вхідної та вихідної щілин, що не перевищували 100 мк. При таких щілинах ширина апаратної функції монохроматора не перевищує 5 см^{-1} .

Спектри КР зразка X_1 в кристалічній фазі було зареєстровано за температури 293 K.

Зразок було синтезовано в Харківському "Інституті монокристалів" НАН України у відділі "Оптично активних органічних сполук".

Структурну формулу зразка X_1 наведено на рис. 1

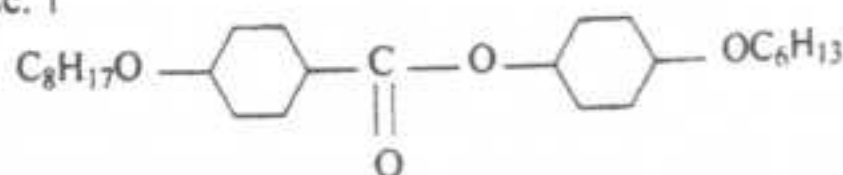


Рис. 1 Структурна формула X_1

Видно, що в молекулі наявні два бензольні кільця, два хімічні зв'язки CH_2 , дванадцять

зв'язків CH_2 , вісім зв'язків CH , три зв'язки $\text{C}-\text{O}$ і один зв'язок $\text{C}=\text{O}$. Всі ці хімічні групи атомів будуть давати свій внесок у спектр даної молекули.

На рисунках 2а і 2б показано оглядовий спектр

Спектральна область $1800 - 2800 \text{ cm}^{-1}$ була вирізана тому, що на цьому частотному інтервалі не було виявлено жодної спектральної лінії.

Тепер поділимо цей спектр на частини і детально розглянемо кожну.

Розглянемо частину спектра в інтервалі $600 - 1000 \text{ cm}^{-1}$ (рис.3).

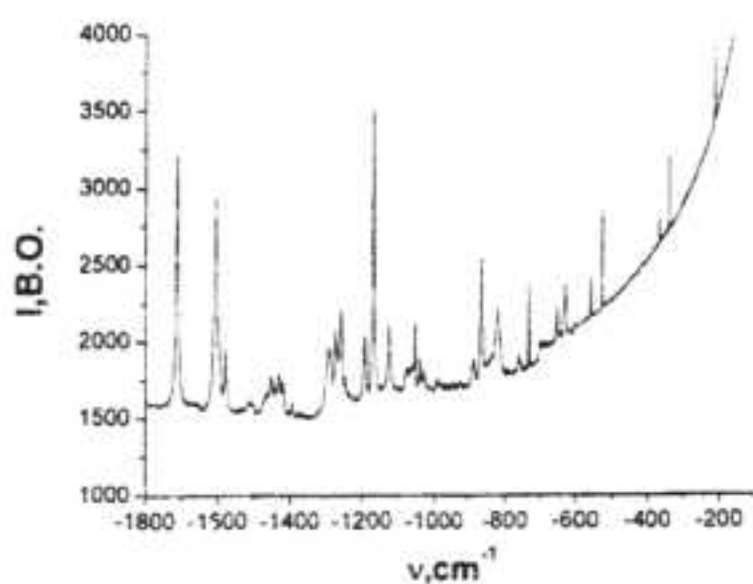


Рис. 2а Коливальний спектр X_1 на ділянці частот $200 - 1800 \text{ cm}^{-1}$

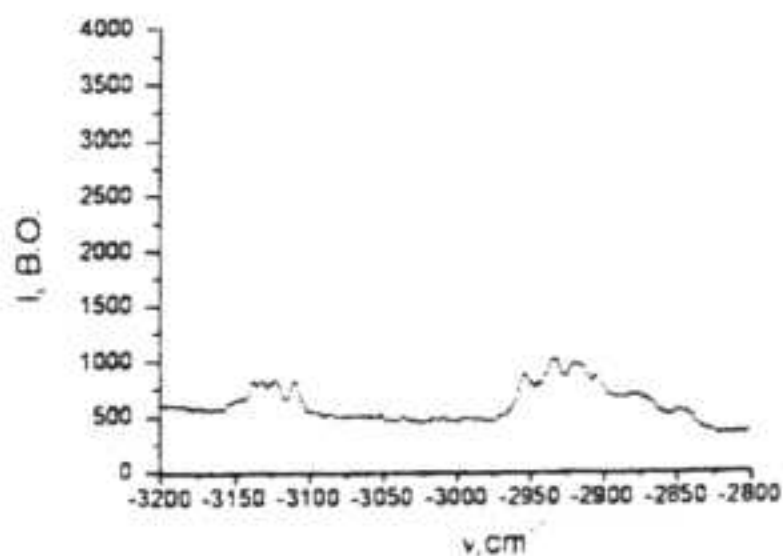


Рис. 2б Коливальний спектр X_1 на ділянці частот $2800 - 3200 \text{ cm}^{-1}$

Тут виділяємо спектральні смуги на таких частотах: 630 cm^{-1} , 653 cm^{-1} , 701 cm^{-1} , 734 cm^{-1} , 820 cm^{-1} , 870 cm^{-1} і 892 cm^{-1} . Відразу треба зазначити, що спектральна смуга 734 cm^{-1} також належить власне до спектра самого лазера (у лазерному спектрі зафіксовано лінію на частоті 735 cm^{-1}). Всі інші смуги вочевидь пов'язані з коливаннями

двох бензольних кілець. Так смугу на частоті 630 cm^{-1} можна пов'язати з таким типом коливання:



, де зірочками позначено зв'язок бензольного кільця з іншими компонентами молекули; характерна частота для цього типу зв'язку становить 629 cm^{-1} . Інші лінії можна інтерпретувати як деформаційні коливання самого бензольного кільця.

На рис.4 показано частотний інтервал $1000 - 1200 \text{ cm}^{-1}$, де можна виділити такі смуги: 1031 cm^{-1} , 1043 cm^{-1} , 1056 cm^{-1} , 1126 cm^{-1} і 1171 cm^{-1} . Смугу 1031 cm^{-1} можна пов'язати з коливанням хімічного зв'язку $\text{C}-\text{O}$, а 1043 cm^{-1} , 1056 cm^{-1} і 1126 cm^{-1} – з колаваннями бензольних кілець, а неоднорідності на інтервалі $1150 - 1190 \text{ cm}^{-1}$ вочевидь виникли в результаті їхньої взаємодії.

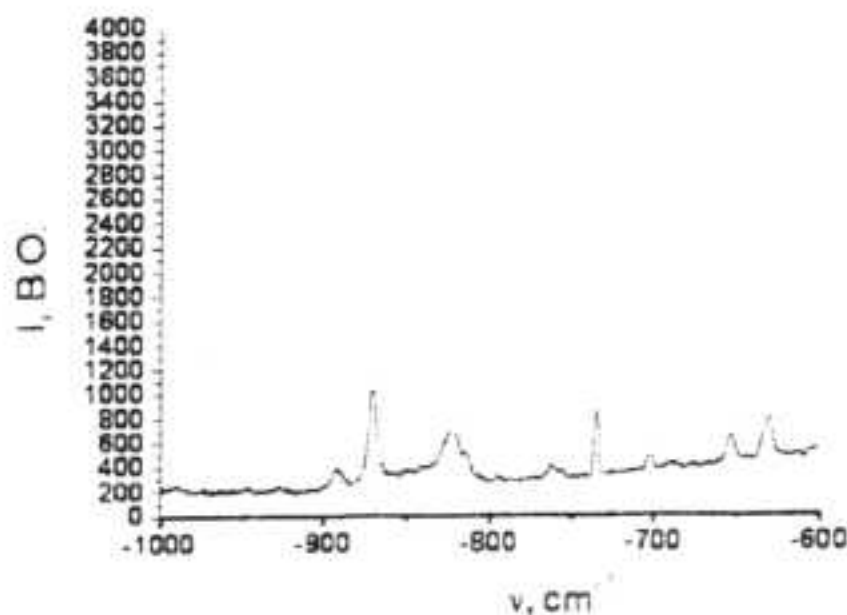


Рис. 3 Коливальний спектр X_1 на ділянці частот $600 - 1000 \text{ cm}^{-1}$

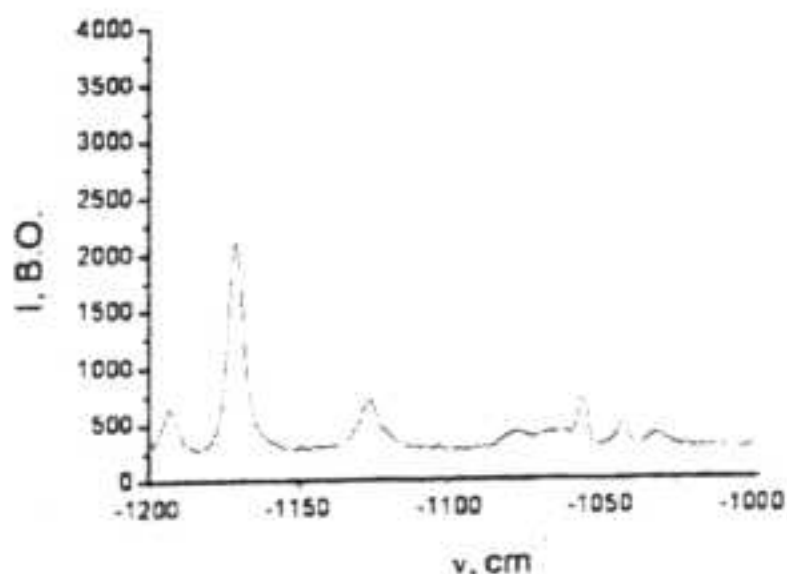


Рис. 4 Коливальний спектр X_1 на ділянці частот $1000 - 1200 \text{ cm}^{-1}$

На рис.5 видно комплекс смуг в інтервалі $1200 - 1350 \text{ cm}^{-1}$, які накладаються одна на одну. Його можна інтерпретувати як коливання груп RCH_3 , де радикал R позначає будь-який хімічний елемент (в нашому випадку це буде ще один атом вуглецю).

В області $1350 - 1800 \text{ cm}^{-1}$ (рис.6) можна виділити такі лінії: 1422 cm^{-1} , 1431 cm^{-1} , 1454 cm^{-1} , 1509 cm^{-1} , дуплет 1581 cm^{-1} і 1607 cm^{-1} , а також 1713 cm^{-1} . В інтервалі $1400 - 1500 \text{ cm}^{-1}$ наявні коливні смуги як груп атомів CH_2 , так і RCH_3 . Причому значення характеристичної частоти 1450 cm^{-1} співпадає і для CH_2 , і для RCH_3 . Група атомів RCH_3 розташована на самих кінцях молекули, тому можна очікувати, що саме вона коливається з найближчою до характеристичної частотою, а саме 1454 cm^{-1} . Смуга 1509 cm^{-1} і дублет пов'язані з коливанням бензольного кільця, а найбільш інтенсивна смуга КР 1713 cm^{-1} виникає внаслідок коливання $\text{C}=\text{O}$. Її велику інтенсивність можна пояснити тим, що атом оксигену приєднується до молекули збоку, а тому може коливатися з великою амплітудою.

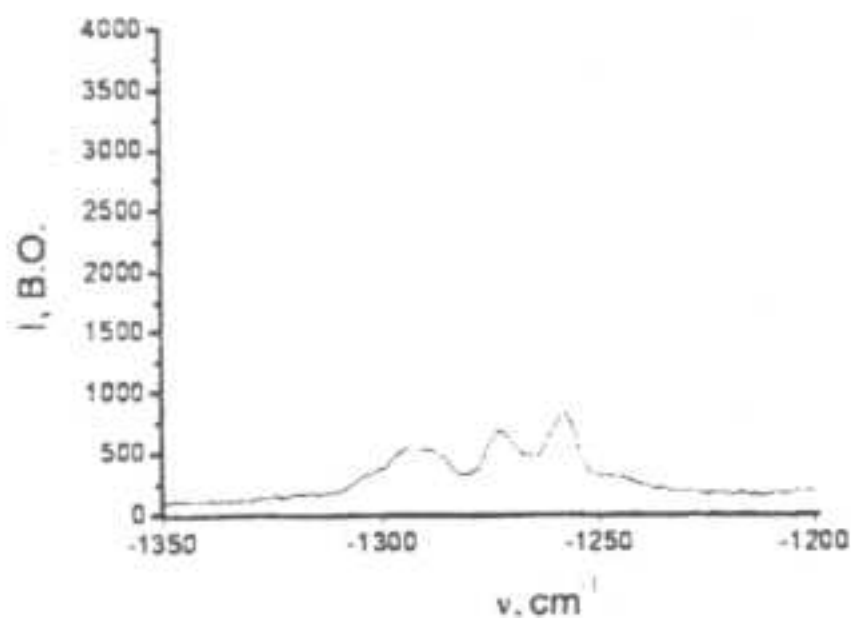


Рис. 5 Коливальний спектр X_1 на ділянці частот $1200 - 1350 \text{ cm}^{-1}$

Останню частину спектра проаналізуємо, розглянувши мал. 26. Аналізуючи комплекс спектрів в інтервалі $2850 - 3000 \text{ cm}^{-1}$, можна виділити смуги 2853 cm^{-1} , 2879 cm^{-1} , 2908 cm^{-1} , 2917 cm^{-1} , 2934 cm^{-1} і 2954 cm^{-1} . Порівнюючи зі значеннями характеристичних частот для груп RCH_3 і CH_2 , можна зробити висновок, що коливання з частотою 2853 cm^{-1} і 2879 cm^{-1}

належать RCH_3 , а всі інші – CH_2 . В області $3100 - 3150 \text{ cm}^{-1}$ виділяємо смуги 3109 cm^{-1} , 3123 cm^{-1} , 3132 cm^{-1} і 3137 cm^{-1} . Всі вони належать до ко-

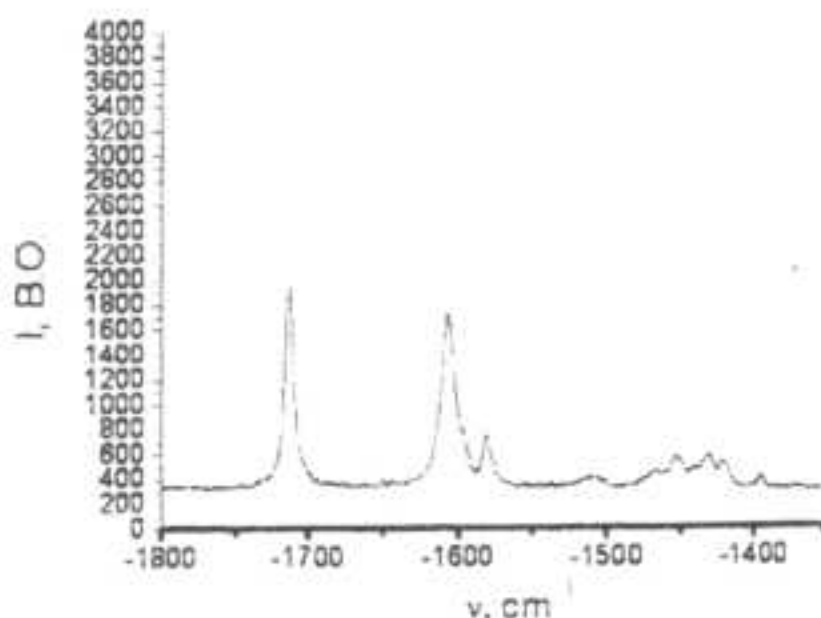


Рис. 6 Коливальний спектр X_1 на ділянці частот $1350 - 1800 \text{ cm}^{-1}$

ливань групи $\text{C} - \text{H}$, оскільки вона має найменшу атомну масу, а отже і її смуги в спектрі КР повинні розташовуватися найдалі від лінії збудження.

Висновок. В даній статті було виміряно і проаналізовано повний колильний спектр рідкого кристалу 4-(н-гексилосі)феніл - 4'-(н-октилокси)бензоат в кристалічній фазі.

Ця робота виконана за підтримки Державного фонду фундаментальних досліджень МОН України, Угода № М/174-2007.

Список використаних джерел

1. Волькенштейн М.В., Ельяшевич М.А., Степанов Б.И. «Колебания молекул», Т.1, Москва – Ленинград, ГИТТЛ, 1949. – с.413 – 426.
2. Волькенштейн М.В., Ельяшевич М.А., Степанов Б.И. «Колебания молекул», Т.2, Москва – Ленинград, ГИТТЛ, 1949. – с.

Надійшла до редколегії 24.01.2008р

УДК 620.187:539/5: 669.268

Рогуль Т.Г. к. ф.-м. н.

Термічна стабільність осадженого магнетронним розпиленням наддрібнозернистого хрому

Методами просвічуючої електронної мікроскопії досліджено структуру хромових покриттів, одержаних магнетронним розпиленням. Вивчена еволюція дислокаційної структури покриттів при відпалі в інтервалі температур 1073 – 1873 К, встановлена відповідність між структурними перетвореннями, мікротвердістю та пластичністю покриттів.

Ключові слова: хром покриття, структура, мікротвердість, пластичність.

E-mail: rogul@ipms.kiev.ua

Захищені, зносостійкі хромові покриття давно і успішно застосовуються в сучасній промисловості. Розвиток вакуумної техніки та розробка нових конструкційних сплавів на основі хрому з підвищеною пластичністю, що можуть бути використані для виготовлення мішеней для розпилення, відкривають нові можливості для створення хромових покриттів із заданими характеристиками. Особливу цікавість викликає дослідження можливості отримання покриттів з нанокристалічною структурою, яку пов'язують з появою цілого ряду нових або екстремальних властивостей, зокрема з високою твердістю, міцністю. В той же час, беручи до уваги те, що ресурс роботи наноструктурованих покриттів залежить від термічної стабільності їх структури, дослідження еволюції дислокаційної структури покриттів при відпалі є актуальною задачею. У даній роботі приведені результати дослідження тонкої структури хрому, осадженого методом магнетронного розпилення, встановлена відповідність між структурними особливостями осаджених шарів та їх твердістю і пластичністю при відпалі.

Магнетронне розпилення здійснювали за допомогою магнетронної системи, що була встановлена на установці УКР-77П-1. В якості вихідного сплаву для одержання покриттів використовували малолегований сплав хрому (99 мас.% Cr та La, Ta, V). Попередній вакуум у камері становив $5 \cdot 10^{-6}$ мм рт. ст., робочий тиск після напуску аргону – $3 \cdot 10^{-3}$ мм рт. ст., струм мішені – 1,5 А, швидкість осадження досягала 0,25 мкм/хв. Осадження проводили в циклічному режимі (розпилення здійснювали дискретними

T. G. Rogul, Cand. Sci

Thermal stability of produced by magnetron sputtering ultrafine-grain chromium

The structure of chromium coatings, produced by magnetron sputtering, is investigated by the methods of the transmission electron microscopy. The evolution of dislocation structure with the annealing in the range of temperatures 1073 – 1873 K is studied, correlation between the structural transformations, the microhardness and the plasticity of coatings is established.

Key Words: chromium, coatings, structure, microhardness, plasticity.

інтервалами протяжністю в 3 хв. і з перервою між ними в 10 хв.) на підкладки із полікристалічної міді та свіжо сколотого монокристалічного NaCl. Відпал відокремлених від підкладок покриттів проводили в аргоні, при температурах відпалу вищих від 1273 К для зменшення газонасичення в робочий простір печі вміщали стружку титано – цирконієвого сплаву. Дослідження тонкої структури покриттів проводили на електронних мікроскопах ЕМВ-100БР, ЕММА-4, JEM-100СХ (прискорююча напруга 100 кВ), JEM-1000 (прискорююча напруга 1000 кВ). Виміри мікротвердості з поперечних шліфів покриттів проводили за стандартною методикою на приладі PMT-3.

Розпилення за магнетронною технологією в циклічному режимі дало можливість одержати хромове покриття товщиною 42 мкм з розміром зерен 40–60 нм (рис. 1 а). Дослідження показали, що на мікродифракційних знімках покриття, крім ліній від ОЦК ґратки хрому, зафіксовано рефлекси від оксиду Cr_2O_3 . Мікротвердість одержаного покриття сягає 18,7 ГПа, в той час як характерне її значення для масивного хрому становить 1,6 ГПа.

Відпал покриття в інтервалі температур 1073 – 1473 К (1 год.) не призводить до зміни розміру зерен, хоча вже при температурі відпалу 1173 К спостерігаються окремі аномально вирослі зерна, кількість яких надзвичайно мала в порівнянні з кількістю зерен малих розмірів. Активний ріст зерна спостерігається лише після відпалу при 1573 К (рис. 1 б), при цьому в об'ємі зерен чітко розрізняються ланцюжки дисперсних

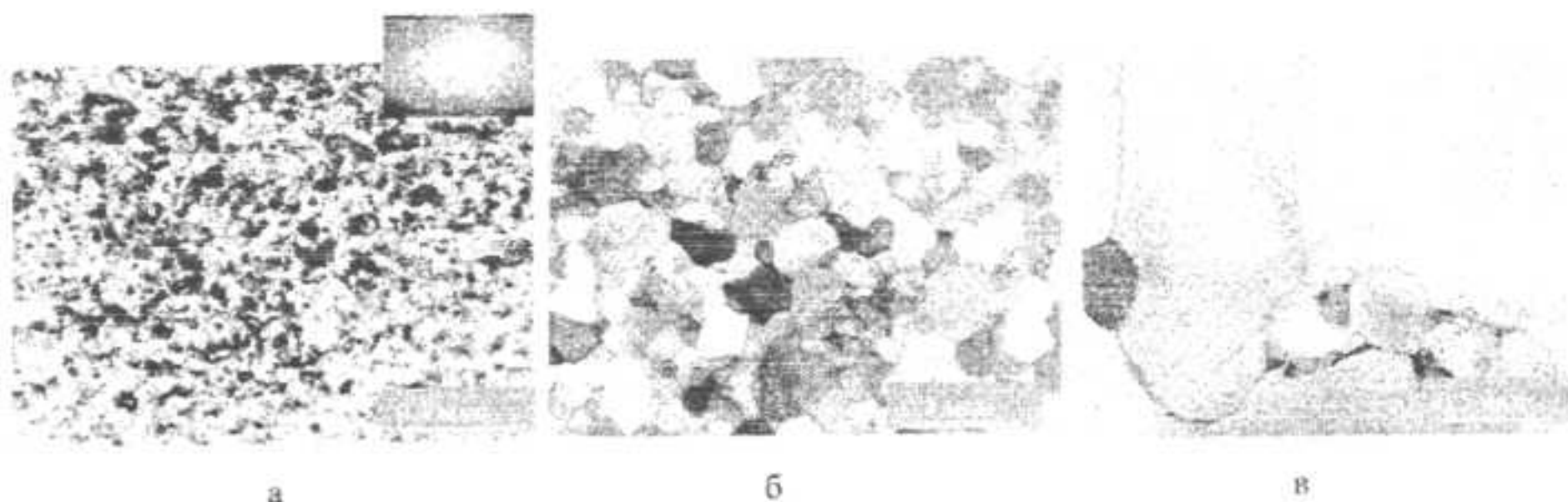


Рис. 1. Структура одержаних магнетронним розпиленням хромових покриттів товщиною 42 мкм у вихідному стані (а) і після відпалу при 1573 (б) та 1873 К (в) на протязі 1 години.

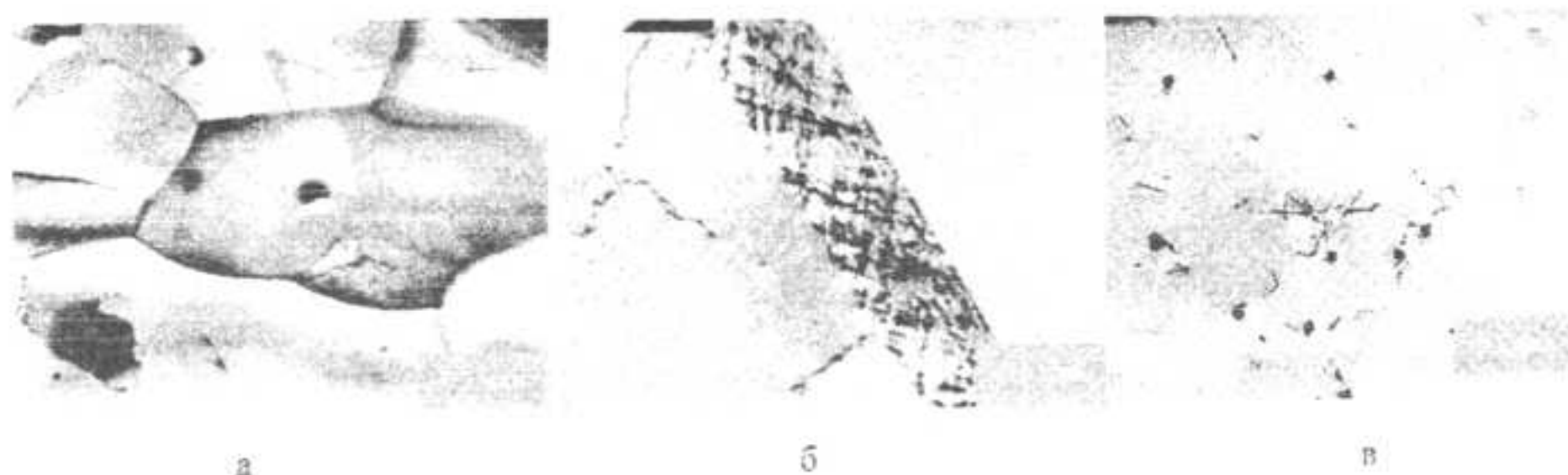


Рис. 2. Структура деформованої хромової мішені у вихідному стані (а) і після відпалу при 1573 (б) та 1873 К (в) на протязі 1 години.

включень Cr_2O_3 , які повторюють конфігурацію меж зерен у вихідному покритті. У той же час, при температурі відпалу 1573 К деформований сплав хрому, з якого була виготовлена мішень для розпилення, був повністю рекристалізований (рис. 2 б), розмір зерен становив 40 – 80 мкм. При відпалі в інтервалі температур 1673 – 1873 К швидкість росту зерен в покритті значно перевищила значення, зафіксовані в деформованому розпилюваному хромі при таких же температурах відпалу. І вже при температурі відпалу 1873 К, як і в деформованому хромі після відпалу при такій же температурі, переважна кількість зерен в покритті мала розмір 450 – 500 мкм (рис. 1 в, 2 в). Однак, навіть після відпалу покриття при 1873 К в ньому збереглися окремі зерна розміром 5 – 7 мкм. Розмір, кількість і конфігурація термічно стабільних дисперсних включень оксиду хрому Cr_2O_3 в осадженому хромі не змінилися навіть після відпалу при 1873

К. Розрахунки показують, що об'ємна доля таких включень не перевищує 3%.

Співставлення і аналіз одержаних результатів показує, що зміщення температури збиральної рекристалізації осадженого наддрібнозернистого хрому в область більш високих температур пояснюється закріпленням меж зерен дисперсними виділеннями Cr_2O_3 . Термічно активований відрив меж зерен від закріплюючої їх фази відбувається тільки при температурах відпалу вищих від 1573 К і супроводжується активним ростом зерен. Розмір зерен осадженого і розпилюваного хрому до відпалу (вихідний стан), а також залежність зміни розміру зерна від температури відпалу приведені на рис. 3.

Необхідно зазначити, що стабілізуючий вплив дисперсних частинок другої фази на розорієнтовану коміркову або дрібнозернисту структуру з ОЦК ґраткою за умов високотемпературного відпалу розглянуто в

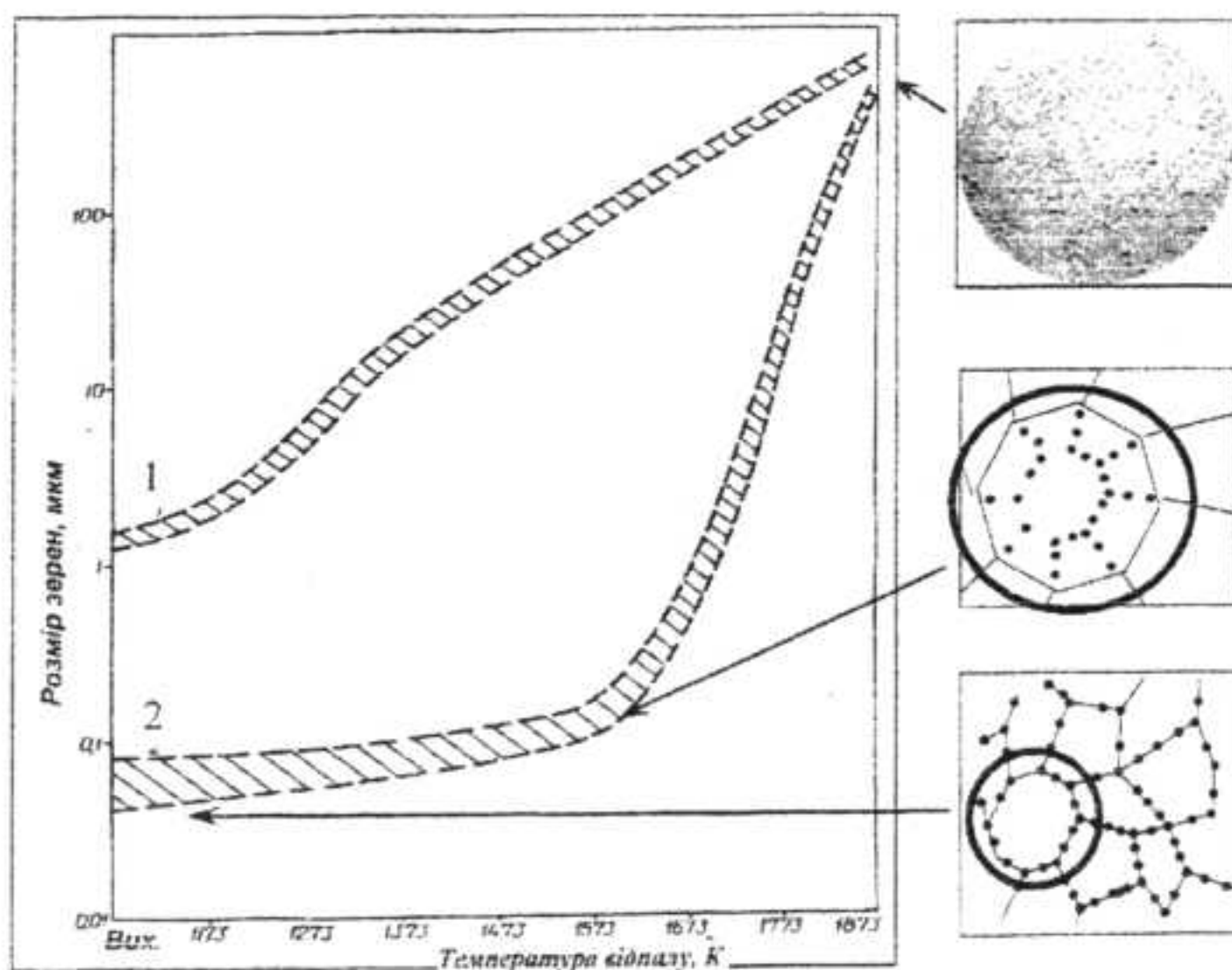


Рис. 3. Вплив температури відпалу на розмір зерен розпилюваного (1) та осадженого (2) хрому.

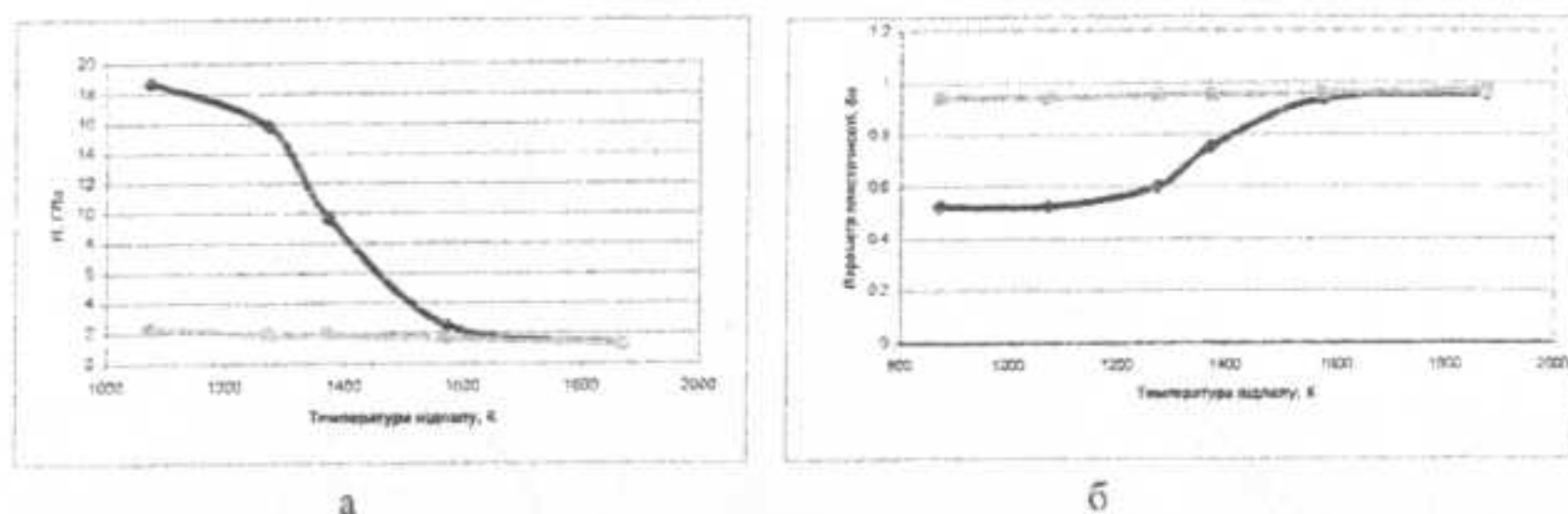
роботах Ю. В. Мільмана [1,2]. У роботах приведені рівняння, що описують зміну розмірів зерен (комірок) за умов „сповільненої” рекристалізації (межі зерен переміщуються разом з частинками), а також знайдено вираз для розрахунку критичного розміру коагульованої другої фази, за якого окремі межі зерен відриваються від частинок (перехід до „швидкої” рекристалізації та аномального росту зерен). У роботі [1], також зазначено, що у випадку, коли частинки другої фази стабільні і не коагулюють, а кількість їх достатньо велика, то перехід до „швидкої” рекристалізації не повинен відбуватися аж до температур, близьких до температури плавлення. У випадку ж осадженого наддрібнозернистого хрому частинки другої фази стабільні і не коагулюють навіть при температурі відпалу 1873 К, але так, як їх вміст невеликий (об’ємна доля менша від 3%), то активний ріст зерен фіксується вже при температурах вищих від 1573 К (температура плавлення хрому – 2170 К [1]).

Значно уповільнений в порівнянні з масивним хромом ріст зерен при відпалі в осадженому наддрібнозернистому хромі призводить до збереження його надвисоких значень твердості аж до температури відпалу 1573 К (рис. 4). Зниження мікротвердості

осадженого хрому до значень, характерних для масивного хрому (1.6 ГПа), має місце лише при температурі відпалу 1873 К, коли розміри зерен у покритті і в масивному хромі однакові. Зауважимо, що мікротвердість електролітичних хромових покриттів, осаджених у стандартному електроліті, хоча і може сягати 9.8 – 11.1 ГПа, втім різке її зниження фіксується вже при нагріві вище 873 – 1073 К [3].

Детально природа надвисокої твердості наддрібнозернистого осадженого хрому, яка не може бути пов’язаною лише з сумарним ефектом наявності наддрібного зерна, фази Cr_2O_3 , механізмами дисперсного та твердорозчинного твердіння, розглянута в роботі [4]. Показано, що надвисока твердість осадженого наддрібнозернистого хрому є наслідком впливу атомів кисню, що проникли у межі зерен і внаслідок більш сильного хімічного зв’язку з атомами хрому у порівнянні зі зв’язком між самими атомами хрому, спричинили часткову ліквідацію („заліковування”) слабких, розрихлених, несучільних ділянок у межах зерен. При цьому зменшується частка „недосконалого” матеріалу, підвищується когезія зерен, і, відповідно, різко зростає твердість наддрібнозернистого матеріалу.

Для оцінки пластичності матеріалу при



а

б

Рис. 4. Залежність твердості (а) і параметра пластичності (б) осажденного за магнетронною технологією (◆) і розпильованого (мішені) (■) хрому від температури відпалу.

індентуванні в роботі [5] запропоновано аналітичний вираз для розрахунку параметра пластичності δ_n , що характеризує частку пластичної деформації в загальній деформації зразка при проникненні в нього індентора:

$$\delta_n = 1 - 14,3 (1 - \nu - 2 \nu^2) H/E,$$

де H – твердість; ν і E – коефіцієнт Пуассона і модуль Юнга відповідно (для хрому $\nu = 0,31$, $E = 279$ ГПа [5]).

На рис. 4, б продемонстровано залежність параметра пластичності δ_n від температури відпалу для розпильованого хрому і одержаних за магнетронною технологією наддрібнозернистих хромових покриттів. Видно, що δ_n покриття наближається до значень 0,9, характерних для металів, лише після відпалу при температурах вищих від 1573 К.

Таким чином, виконані дослідження показали, що сформована в хромових покриттях за умов магнетронного розпилення наддрібнозерниста структура (розмір зерен 40–60 нм) є стабільною аж до температур відпалу 1573 К (1 год). Характер зміни твердості покриттів при відпалі цілком узгоджується зі структурними перетвореннями і зниження мікротвердості осажденного хрому до значень, характерних для масивного хрому (1,6 ГПа), має місце лише при температурі відпалу 1873 К, коли розміри зерен у покритті і в масивному хромі однакові. При цьому пластичність хромових наддрібнозернистих покриттів значно нижча від пластичності розпильованого хрому і наближається до характерних для металів

значень ($\delta_n > 0,9$) лише після відпалу при температурах вищих від 1573 К.

Список використаних джерел

1. Мильман Ю.В., Рябошапка К.П. О рекристаллизации дисперсноупрочненных сплавов с ОЦК решеткой // Фізика металлов и металловедение. – 1971. – Том 32, вып. 5. – С. 998–1006.
2. Мильман Ю.В. Развитие представлений о рекристаллизации дисперсноупрочненных материалов на основе теории диффузионного движения включений М.А. Кривоглаза // Металлофизика, новейшие технологии. – 2005. – 27. – № 1. – С. 59 – 74.
3. Вейнер Р. Гальваническое хромирование. Блестящее и твердое. Пер. с нем. -М.: Машиностроение, 1979. – 296 с.
4. Фирстов С.А., Роголь Т.Г., Свечников В.Л. и др. Структура, механическое поведение и нанотвердость поликристаллических хромовых и молибденовых покрытий, полученных методом магнетронного распыления // Металлофизика, новейшие технологии. – 2003. – 25. – № 9. – С. 1153 – 1164.
5. Мильман Ю.В., Галанов Б.А., Чугунова С.И. Характеристика пластичности, получаемая при измерении твердости. –АН Украины ИПМ Препр-т. – 1992. – 26 с.

Надійшла до редколегії 18.02.2008

УДК 539.15.04

Розуван К.П., н.с., Розуван* С.Г., н.с.,
Шайкевич І.А., д.ф.м.н., проф.

Лінійна і нелінійна спектроскопія π - зв'язків у політиофенів.

У роботі приведені результати спектральних досліджень політиофенів за допомогою еліпсометричних вимірювань. Теоретично досліджено вплив двохфотонних переходів при нелінійних процесах третього роду. Показано, що зареєстрований різкий спад тензора третього рангу у видимій області для широкого класу сполук з розподіленими π -зв'язками атомів вуглецю може являтися штучним ефектом внаслідок існування сильних смуг поглинання в близькій ультрафіолетовій області.

Ключові слова: політиофени, спектральна еліпсометрія, чотирьохпучкова взаємодія, фемтосекунди, двохфотонні переходи.

*E-mail: sgri@univ.kiev.ua

Гібридизація електронних орбіталей виникає в широких класах хімічних сполук і є відповідальною за цілий ряд фізико-хімічних властивостей цих речовин [1]. За останню декаду наукові інтереси великої кількості дослідницьких груп були сфокусовані на дослідженні sp^2 гібридизованих зв'язків вуглецю. Це обумовлено декількома досить важливими причинами. По-перше, це пов'язано з різким зростанням кількості наукових досліджень над фулеренами та вуглецевими нанотрубками як потенційно перспективними матеріалами для використання в галузі нанотехнологій. На нашу думку діапазон можливого застосування нанотехнологій охоплює в свою чергу надзвичайно широкий спектр наукових дисциплін починаючи від біофізики і закінчуючи оптоелектронікою [2]. Це й обумовлює широкий фронт досліджень в даному напрямку. Наступною важливою причиною зростання інтересу до вуглецевих сполук являється прогрес останнього десятиріччя в інтегральній та нелінійній оптиці. Оптичні мікросхеми стали невід'ємною частиною оптичної обробки сигналу в оптоволоконних лініях зв'язку. Так, наприклад, оптичні чіпи на основі інтерферометра Маха-

Rozouvan K.P., Rozouvan* S.G.,
Shaykevich I.A., Dr Sci

Linear and nonlinear spectroscopy of π - bonds in polythiophenes.

In this work are presented results of spectral ellipsometrical measurements of polythiophenes. Influence of two-photon transitions as third order nonlinear processes has been studied theoretically. Registered sharp drop of third order nonlinear tensor spectral curve in visible spectral region (for a variety of materials with π -conjugated bonds of carbon atoms) may be an artifact due to strong absorption bands in near UV spectral region.

Key words: polythiophenes, spectral ellipsometry, DFWM, two-photon transitions.

Цандера з кристалом мікропроцесора вже серійно випускаються для вузлів спектрального ущільнення передачі інформації [3] (wavelength division multiplexing). Водночас, ступінь інтеграції оптичних мікросхем залишається порівняно низькою. Однією з перешкод на шляху побудови мініатюрних оптичних логічних вузлів являється відсутність функціональних матеріалів з високою оптичною нелінійністю. Перспективними матеріалами які б підходили на роль такого матеріалу являються полімери з розподіленим π -зв'язком. Такий зв'язок відповідає за нелінійність третього рангу в полімерах і може досліджуватись за допомогою чотирьохпучкової взаємодії. Спектральна чотирьохпучкова взаємодія [4] являє собою фактично різновид фемтосекундної оптичної спектроскопії яка досліджує спектральну залежність виродженого тензора оптичної нелінійності третього рангу. Поряд з такими методиками як зет-скан (Z-scan), затвор Кера (Kerr-gate technique) та генерація третьої гармоніки [5] спектральна чотирьохпучкова взаємодія дозволяє знаходити параметри виродженого тензора, а не згортку його декількох спектральних компонент. При дослідженнях широкого класу вуглецевих

сполук спостерігався дуже різкий спад їх нелінійно оптичних характеристик в короткохвильовій області спектра. Задачею даної публікації являється дослідження даного ефекту з використанням останніх теоретичних результатів для дослідження $sp2$ зв'язків [6]. Дана робота являється логічним продовженням досліджень роботи [7].

Вироджена чотирьохпучкова взаємодія з геометрією пучків уперед (degenerate four-wave mixing measurements in folded box CARS geometry) дозволяє експериментально отримувати релаксацийні криві, які представляють собою суперпозицію когерентного максимуму (coherent artifact) та повільно затухаючої компоненти. Максимум являє собою результат чотирьохпучкової взаємодії і є функцією згортки трьох взаємодіючих через нелінійне середовище імпульсів світла. Повільно затухаюча компонента являється дифрагованим світлом на динамічних ґратках які виникають в результаті прямих оптичних переходів. Час затухання повільної компоненти визначається напівшириною збудженого рівня. Для того щоб дати уявлення про систему енергетичних рівнів типової сполуки з розподіленим π - зв'язком нами були проведені спектральні еліпсометричні вимірювання політіофенових плівок

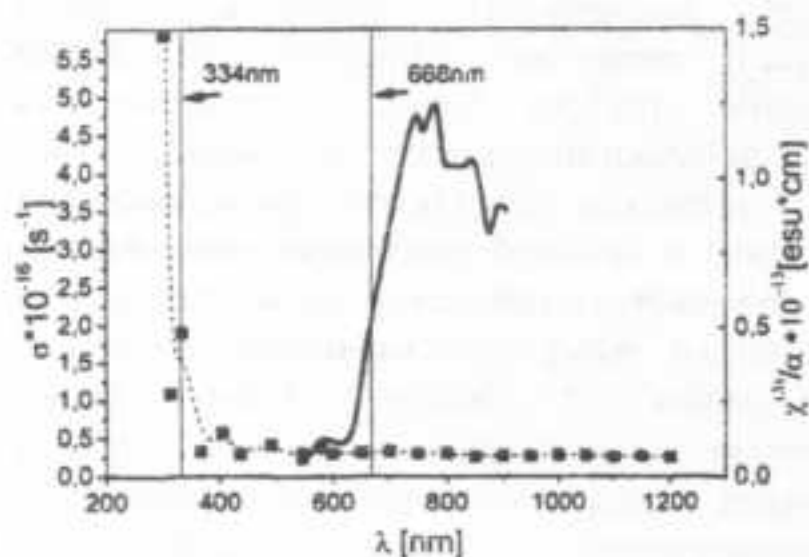


Рис.1 Спектральна залежність оптичної провідності (---■---) та величин $\chi_{(111)}^{(3)}/\alpha$ (—) для політіофенової плівки.

аналогічних до досліджуваних раніше [8] із застосуванням устаткування і методик, які детально описані нами в [7]. Типовий хід нелінійної спектральної залежності для політіофену характеризується сильним спадом компонент тензора для області спектра менше 600 нанометрів [8].

Результати вимірювань спектральної залежності оптичної провідності представлені на Рис.1. Для порівняння на цьому ж рисунку також приведена спектральна крива компоненти тензора нелінійності третього рангу нормалізована до коефіцієнта поглинання (figure of merit) яка була опублікована раніше [8]. Видно, що подвоєна частота різкого короткохвильового спаду на кривій нелінійного тензора майже співпадає з границею сильної полоси поглинання політіофену в області 350 нанометрів. Виникає закономірне питання про роль двохфотонних переходів при чотирьохпучковій взаємодії. Розглянемо систему основного рівня 0 і двох збуджених рівнів 1 і 2 які відповідають переходам з основного рівня з

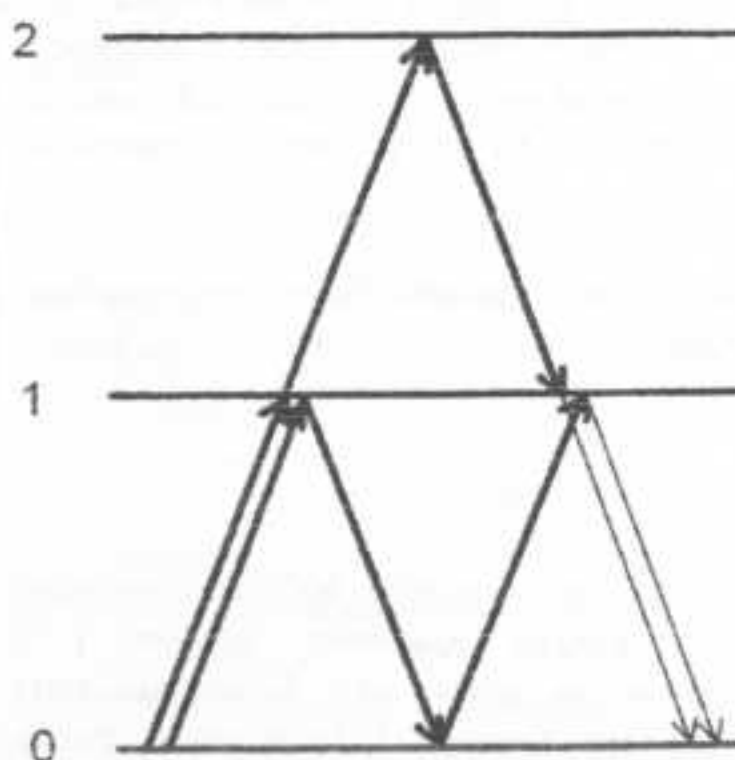


Рис.2 Схеми виродженої чотирьохпучкової взаємодії для дворівневої та трьохрівневої системи.

відповідними частотами ω і 2ω . Класична („вироджена“) схема чотирьохпучкової взаємодії для переходів між рівнями 0 і 1 а також схема чотирьохпучкової взаємодії з участю двохфотонних переходів зображені на рис.2 Нас цікавлять когерентні процеси які проходять за часи порядку десятка фемтосекунд. За цей час спонтанні переходи не в змозі суттєво вплинути на динаміку міжрівневих переходів оскільки вони мають істотну ефективність в значно більшому часовому інтервалі (більше десятка пікосекунд). Коефіцієнти Ейнштейна які визначають ймовірності вимушених переходів порівнюють:

$$B_{01} = B_{10} = (2\pi |\vec{d}_{01}| / \hbar)^2 / 3 \quad (1)$$

$$B_{12} = B_{21} = (2\pi |\vec{d}_{12}| / \hbar)^2 / 3.$$

де величини $\vec{d}_{01}, \vec{d}_{12}$ позначають оператори дипольних моментів:

$$\vec{d}_{ij} = \int \vec{r} \Psi_i^* \Psi_j d^3r \quad (2)$$

Верхній індекс (*) позначає операцію комплексного спряження. У випадку чотирьохпучкової взаємодії електричне поле, яке взаємодіє з речовиною являється суперпозицією трьох плоских хвиль:

$$E(r, t) = \sum_m \sum_{k=1}^3 \begin{pmatrix} E_k \\ 0 \\ 0 \end{pmatrix} \exp(i \begin{pmatrix} (-1)^{k+1} \\ (-1+2\delta_{1k})^k \\ 5 \end{pmatrix} * \frac{2\pi}{3\sqrt{3}\lambda_m} - i\omega_m t), \quad (3)$$

де δ_{1k} - символ Кронекера, а індекс m позначає спектральні компоненти хвильового пакета. Співвідношення (3) записане для випадку сходження трьох лазерних пучків до повздовжньої осі з тангенсом кута, який дорівнює 5. Взагалі наближення плоских хвиль в рівнянні (3) являється не зовсім точним оскільки фемтосекундні лазерні системи генерують Гаусові пучки світла. Однак у нашому випадку ми розглядаємо сфокусовані пучки (перетяжку Гаусових пучків) в дуже маленькому об'ємі тонких полімерних плівок, де з високою точністю можна вважати, що вони мають еквіфазні поверхні у вигляді паралельних площин.

Четвертий пучок, який реєструвався в експерименті генерувався просторово розподіленим ансамблем диполів $\vec{d}_{01}$ та $\vec{d}_{12}$. Нами були проведені комп'ютерні розрахунки чотирьохпучкової взаємодії. Для того щоб обчислити величину електричного поля, яке випромінюється у напрямку $\vec{K}_m = \frac{2\pi}{\lambda_m} \hat{r}$ ($|\hat{r}|=1$)

необхідно просумувати всі хвильові пакети, які генеруються просторово розподіленими диполями в інтервалі довжин хвиль, який задається спектральним профілем лазерної лінії:

$$E_{total} = \frac{1}{4\pi\epsilon_0} \sum_m \iiint \left[\frac{\omega_m^2}{c^2 r} \hat{r} \times \vec{P}(x, y, z) \times \hat{r} + \left(\frac{1}{r^3} - \frac{i\omega_m}{cr^2} \right) * (3\hat{r}(\hat{r} \cdot \vec{P}(x, y, z)) - \vec{P}(x, y, z)) \right] \exp(i \begin{pmatrix} x \\ y \\ z \end{pmatrix} \vec{K} - i\omega_m t) dxdydz, \quad (4)$$

де індекс m позначає спектральні компоненти хвильового пакету, а дипольні моменти одиниці об'єму $\vec{P}$ можна виразити через тензори нелінійності середовища:

$$\vec{P}(x, y, z, t) = (\chi^{(3)}(\omega, -\omega, \omega, -\omega) + \chi^{(3)}(\omega, \omega, -\omega, -\omega)) \vec{E}^3(x, y, z, t) \quad (5)$$

Два тензори нелінійності $\chi^{(3)}$ в співвідношенні (5) визначають нелінійність в дворівневій і трьохрівневій схемах і виражаються через дипольні моменти переходів за співвідношеннями приведеними в роботі [7].

Результати комп'ютерних обчислень кутового розподілу інтенсивності сигналу, що реєструється приведені на рис.3. Інтервал зміни кожного з двох кутів на рисунку складає 22° . Кількісні величини четвертого пучка досить сильно залежать від параметрів системи (геометричних розмірів об'єму речовини де розповсюджується електромагнітна хвиля, спектральної ширини хвильового пакету тощо). Наприклад, об'єм збудженої речовини співрозмірний з довжиною хвилі світла не в змозі продукувати світловий промінь в дуже малому тілесному куті із-за очевидних дифракційних обмежень. При наших розрахунках світло з довжиною хвилі 800нм (спектральна напівширина 1нм) фокусувалося в точці з поперечними розмірами 10 мікрон на тонкій плівці товщиною 0.2 мікрона. З рисунка видно, що чотирьохпучкова взаємодія надзвичайно сильно залежить від коефіцієнтів Ейнштейна переходів в трьохрівневій системі. При відносно невеликих варіаціях перерізів вимушених переходів (до одного порядку) основний-збуджений і збуджений-збуджений станів, величина інтенсивності четвертого (сигнального) пучка змінюється на порядки. З чисто фізичної точки зору цей ефект якісно може бути пояснений параметричною взаємодією різних електромагнітних хвиль внаслідок

нелінійності середовища. Якщо, наприклад, ми аналізуємо дві системи динамічних ґраток, які виникають на переходах основний-збуджений та збуджений-збуджений станів, то ми можемо розглядати інтерференцію світлового пучка на кожній з ґраток незалежно [9]. У випадку нелінійної поляризації середовища, зміна інтенсивності

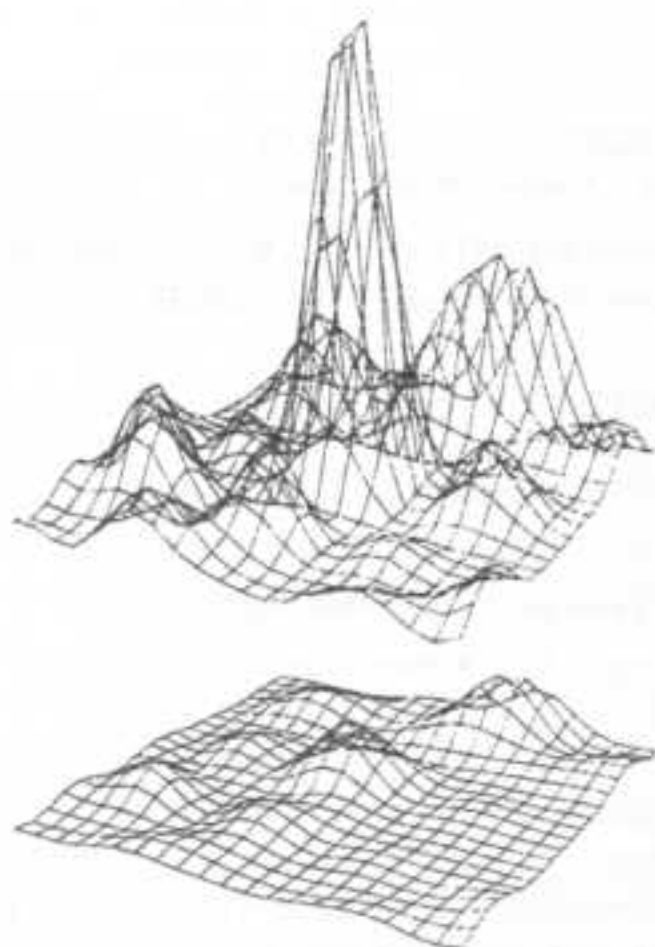


Рис.3 Кутовий розподіл інтенсивності четвертого (сигнального) пучка для $B_{12}/B_{01}=0.25$ (верхня крива) і для $B_{12}/B_{01}=0.64$ (нижня крива).

електромагнітного поля що формує, наприклад, першу ґратку приводить до зміни параметрів другої ґратки. Це може призвести в тому числі і до зменшення видимості інтерференційної картини бо профіль тієї ґратки, що формує цю картину, складним чином деформувався.

Зрозуміло, що даний ефект зниження нелінійності третього порядку може бути присутнім у всіх речовинах зі співрозмірними величинами ймовірностей вимушених переходів основний-збуджений та збуджений-збуджений станів. Так, наприклад, ефект реєструвався і при нелінійно оптичних експериментах з фулеренами [7], де також спостерігалось відносне зменшення інтенсивності нелінійних процесів у короткохвильовій області. На практиці це означає, що сполуки вуглецю з гібридизацією s - та p - орбіталей і сильними полосами поглинання в видимій області будуть демонструвати зменшення ефективності процесів нелінійності третього порядку

внаслідок параметричної взаємодії з нелінійними процесами другого порядку [7]. З точки зору практичних висновків даний результат може мати досить суттєве значення. Справа в тому, що однією з основних ідей створення функціонального матеріалу з високими нелінійностями третього порядку було збільшення величин компонент тензора за рахунок збільшення коефіцієнта поглинання у видимій області. Технічно це може бути досягнуто або варіаціями хімічної формули полімеру або імплантацією в полімер металевих наночастинок з сильними плазмонними смугами поглинання в червоній області [10]. Порівняно невеликі успіхи, які були досягнуті на цьому шляху тепер можна інтерпретувати з точки зору систематичного штучного заниження величин нелінійності досліджуваних матеріалів внаслідок ефектів, які були розглянуті в даній публікації.

Список використаних джерел.

1. Y.K. Syrkin, M.E.Dyatkina, Structure of molecules and the chemical bond. -London: Butterworths Scientific Publications, 1950. -309 pages.
2. Nanomaterials Handbook by Yury Gogotsi (Editor) New York: CRC, 2006. -800 pages.
3. Arora, A.; Subramaniam, S. Wavelength Conversion Placement in WDM Mesh Optical Networks// Photonic Network Communications, -2002. -4(2). -P.78-91.
4. Rozouvan S. Broad band degenerate four wave mixing measurements // Journal of the Optical Society of America B. -2000. -17(8). -P.1354-1359.
5. Mills D.L. Nonlinear Optics -Berlin: Springer- Verlag, 1991. -184 pages.
6. Rozouvan S. Polarization dependent electrostrictive grating measurements // The Journal of Chemical Physics -2007. -127. -P.84501- 84507.
7. Розуван К.П., Розуван С.Г., Шайкевич І.А. 'Нелінійна спектроскопія фулеренів', Вісник Київського університету. Сер. фіз.-мат., вип.2, С.349-353 (2007).
8. W. Schrof, S.Rozouvan, T.Hartmann, V.Belov, H. Mohwald and E. van Keuren, 'Nonlinear optical properties of novel low-bandgap polythiophenes', Journal of the Optical Society of America B., 15, No 2, 889-894 (1998).
9. H.J.Eichler P.Guenter D.W. Pohl Laser-Induced Dynamic Gratings Springer-Verlag Berlin, 1986. -256 pages.
10. W.Schrof, S.Rozouvan, E. van Keuren, D. Horn, J. Schmitt, G. Decher, 'Nonlinear optics of polyelectrolyte thin films containing gold nanoparticles investigated by wavelength dispersive femtosecond degenerate four wave mixing', Advanced Materials, 10, No 4, 338-341 (1998).

Надійшла до редколегії 24.01.2008р

УДК 539.2:539.12.04

¹Селищев П.О., д. ф.-м. н., проф., ²Діденко Т.П.,
інж.

Перехідна повзучість матеріалу під опроміненням при умові генерації вакансій ковзаючими дислокаціями

В межах математичної моделі
"ковзання – переповзання" розглянуто вплив
генерації вакансій ковзаючою дислокацією на
повзучість матеріалів під опроміненням.

Розраховано стаціонарні значення для
концентрацій міжвузлових атомів та вакансій,
густини дислокацій та швидкості радіаційної
повзучості. Досліджено стійкість
стаціонарного розв'язку та побудовано фазові
портрети моделі.

Ключові слова: кристал, опромінення,
дефекти, повзучість.

¹ E-mail: Selyshchev@mail.univ.kiev.ua

² E-mail: Didenko_T@mail.univ.kiev.ua

Вступ

Особливе значення для розуміння
розвитку структури радіаційних дефектів має
вивчення механізмів її утворення: ролі
нелінійних взаємозв'язків між її елементами
(точковими дефектами, дислокаціями). Вид,
кількість та розподіл радіаційних дефектів, їх
взаємодія суттєво впливають на повзучість
матеріалу під опроміненням. Оскільки
повзучість значно обмежує експлуатаційний
ресурс конструкційних матеріалів, то пошук
шляхів її мінімізації є надзвичайно актуальним
на сьогоднішній день.

Дослідження дозової залежності
повзучості опромінених матеріалів дозволили
встановити, що швидкість повзучості як функція
часу чи дози в залежності від параметрів
опромінення може мати якісно різний вигляд. А
саме, при наближенні до встановленого режиму
величина деформації зразка може змінюватись
немонотонно, спостерігаються різкі стрибки
похідної дозової залежності повзучості;
характерні особливості залежності швидкості
повзучості від умов опромінення і властивостей
матеріалу [2,3,4].

Існують теоретичні моделі, які
пояснюють поведінку повзучості матеріалу під
опроміненням [1]. Встановлено, що різним
особливостям перехідної повзучості

¹P.O. Selyshchev, Prof.

²T.P. Didenko, Eng.

Transitional creep of material under the irradiation at the condition of vacancies generation by sliding dislocations

Within the limits of mathematical model of
"sliding – crawling" over is considered influence of
generation of vacancies by sliding dislocation on
the creep of materials under the irradiation.

The stationary values were expected for
concentrations of interstitial atoms and vacancies,
density of dislocation and radiation creep speed.
Stability of stationary stay was explored and the
phase pictures of model were built.

Key words: crystal, irradiation, defects,
creep.

відповідають різні механізми, які ретельно
вивчаються.

Але, незважаючи на це, до сих пір не
встановлена роль генерації точкових дефектів
ковзаючою дислокацією на рух дислокацій.
Вирішенню цього питання і присвячена дана
робота.

Постановка задачі

Зміну повзучості опромінених матеріалів
будемо розглядати в межах математичної моделі
ковзання-переповзання. Модель ковзання
дислокацій, лімітованого їх переповзанням є
найбільш перспективною. На відміну від інших
моделей, вона дає значення швидкості
повзучості, котре по порядку величини
співпадає зі значенням, отриманим
експериментально.

Швидкість повзучості, отримана в
рамках даної моделі, є нелінійною функцією
прикладеної напруги і швидкості утворення
змішень. Причиною повзучості навантаженого
матеріалу є ковзання дислокацій, викликане
прикладеною напругою. Перешкоджаючі
ковзанню бар'єри (домішки, дислокації лісу)
дислокації долають шляхом переповзання,
завдяки чому дислокації переходять в нові
площини ковзання і ковзають під дією
зовнішньої напруги до зупинки на новому
бар'єрі. Останній знову долається

переповзанням і процес повторюється. Швидкість подолання бар'єру залежить від модуля різниці потоків вакансій і міжвузлових атомів.

В роботі враховано, що ковзаюча дислокація здатна генерувати точкові дефекти, а також захоплювати точкові дефекти і їх малі скупчення.

Швидкість повзучості опроміненого зразка описується рівнянням

$$\dot{\epsilon} = v |J_i - J_v| \quad (1)$$

де $J_i = \alpha_i n_i$ і $J_v = \alpha_v n_v$ - густини потоків міжвузлових атомів і вакансій (їх концентрації n_i і n_v) на дислокації з вектором Бюргерса, паралельним прикладеному навантаженню (їх густина $\rho_d/3$, а коефіцієнт переважного поглинання ними міжвузлових атомів - z_d). $\alpha_i = z_d \rho_d D_i/3$, $\alpha_v = \rho_d D_v/3$. ρ_d - повна густина дислокацій в кристалі. Припускалось, що в середньому одна третина дислокацій має вектор Бюргерса паралельний прикладеному навантаженню. $D_i = D_i^0 \exp(-E_m^i/kT)$ і $D_v = D_v^0 \exp(-E_m^v/kT)$ - коефіцієнти дифузії міжвузлових атомів і вакансій, E_m^i і E_m^v - відповідні енергії міграції, T - температура зразка, k - стала Больцмана.

Параметр v представимо у вигляді $v = L(\sigma)/l$, де L - середня довжина ковзання дислокацій є функцією прикладеної напруги σ , l - середній розмір бар'єрів.

Концентрацію точкових дефектів і густину дислокацій визначимо з рівнянь

$$\begin{aligned} \frac{dn_v}{dt} &= K - (\rho_d + \rho_i) D_v (n_v - n_v^e) + \beta \rho_d n_v^0 n_i \\ \frac{d\rho_d}{dt} &= k - \alpha |D_v n_v - z_d D_i n_i| \rho_d \\ \frac{dn_i}{dt} &= K - (z_d \rho_d + z_i \rho_i) D_i n_i \end{aligned} \quad (2)$$

K - швидкість генерації точкових дефектів, k - швидкість генерації дислокацій, α - константа, пов'язана з відходом дислокацій на границі зерен;

$\alpha = \frac{v}{Rb}$, де R - радіус зерна, b - вектор Бюргерса, ρ_i - густина інших дефектів на стоках, ρ_d - густина дислокацій, n_v^e -

концентрація термічних вакансій, останній доданок в першому рівнянні системи (2) описує генерацію дефектів ковзаючою дислокацією, де n_v^0 - концентрація вакансій утворена однією сходинкою, n_i - кількість сходинок на дислокації, D_i, D_v - коефіцієнти дифузії, z_d, z_i - коефіцієнти переважного поглинання міжвузлових атомів і дислокацій.

Припустимо, що стоками є лише дислокації, і їх розподіл є ізотропним, тобто

$$\rho_i = 2\rho_d$$

$$z_i = z_d$$

$$\gamma = \beta n_v^0 n_i$$

Тоді, враховуючи це перепишемо систему рівнянь (2) наступним чином

$$\begin{aligned} \frac{dn_v}{dt} &= K - 3\rho_d D_v (n_v - n_v^e) + \gamma \rho_d \\ \frac{d\rho_d}{dt} &= k - \alpha |\rho_d D_v n_v - K/3| \\ 0 &= K - 3z_d \rho_d n_i D_i \end{aligned} \quad (3)$$

Час характерної зміни концентрації міжвузлових атомів набагато менший, ніж для концентрації вакансій, тому вважатимемо, що для міжвузлових атомів рівноважний стан встановився і тому останнє рівняння системи (3) записано в алгебраїчному вигляді.

В даній системі врахована генерація точкових дефектів опроміненням, поглинання дефектів дислокаціями, утворення дислокацій зовнішньою напругою, вихід дислокацій на границі зерен і генерація вакансій ковзаючою дислокацією. Рівняння (3) описують нелінійні взаємозв'язки між елементами системи. Це конкуруючий вплив термічних і радіаційних потоків точкових дефектів на рух дислокацій. Повзучість, що визначається ковзанням дислокацій, але лімітована їх переповзанням залежить від різниці потоків. Потоки дефектів, які мають радіаційну і термічну природу, по-різному залежать від температури: перший з її зростанням зменшується, другий - зростає. При певних значеннях параметрів, потоки можуть зрівнятися, і швидкість повзучості буде мати мінімум. Визначення величини мінімального значення потребує врахування флуктуацій. Справа в тому, що система (3) записана для середніх величин. Проте, будь-яка флуктуація рівного нуля в середньому сумарного потоку призводить до руху дислокацій і появи

повзучості. Тому середня швидкість повзучості буде додатною.

Система може мати декілька класів різних розв'язків, які необхідно встановити. Таким чином задача зводиться до необхідності побудувати всі якісно різні фазові портрети.

Стационарні розв'язки. Стійкість стационарних розв'язків

Для побудови фазового портрету необхідним є знаходження стационарних розв'язків. Стационарний розв'язок швидкості повзучості опроміненого матеріалу визначається системою рівнянь (3) при умові рівності нулю часових похідних концентрації точкових дефектів та густини дислокацій.

Розв'язок являє собою точку у фазовому просторі, яка називається особливою або стационарною точкою.

Вирази для стационарних концентрацій міжвузлових атомів та вакансій, а також густини дислокацій мають наступний вигляд

$$n_v^s = K / 3z_d \rho_d^s D_i \quad (4)$$

$$\rho_d^s = 3k / \alpha (\gamma + 3D_v n_v^s) \quad (5)$$

$$n_v^s = (n_v^e + \gamma / 3D_v) (1 + \alpha K / 3k) \quad (6)$$

Підставляючи, (4), (5), (6) в (1) отримаємо стационарне значення для швидкості повзучості

$$\dot{\epsilon} = vk / \alpha \quad (7)$$

Зі зміною того чи іншого параметру стационарні значення концентрації дефектів та швидкості повзучості змінюються.

Так з рівняння (4) видно, що стационарне значення концентрації міжвузлових атомів залежить, насамперед, від коефіцієнта дифузії та значення густини дислокацій: чим більша ρ_d , тим менша концентрація міжвузлових атомів. Теж саме відбувається зі зростанням температури (а значить і із збільшенням D_i).

З рівняння (5) випливає, що густина дислокацій безпосередньо залежить від концентрації термічних вакансій, коефіцієнта дифузії (залежать від температури), виходу дислокації на границі зерен та генерації вакансій рухомою дислокацією. Коли вище вказані величини мінімальні, стационарне значення густини дислокацій максимальне, при ж їх збільшенні ρ_d^s прямує до мінімальної позначки.

Аналізуючи вираз (6) можна сказати, що в залежності від температури домінуючий вплив на стационарне значення концентрації вакансій здійснюють або термічні вакансії, або вакансії генеровані при ковзанні дислокації. А саме, при збільшенні температури стає більше термічних вакансій, зростає коефіцієнт дифузії, що призводить до зменшення γ ; також має вплив швидкість генерації дислокацій.

З вигляду рівняння (7) бачимо, що для розглянутої моделі стационарне значення повзучості не залежить від рухливості міжвузлових атомів. Суттєво впливає радіус зерна матеріалу та швидкість генерації дислокацій.

Отже, маємо єдиний стационарний розв'язок системи (3), який описує стационарно однорідний розподіл концентрацій точкових дефектів, густини дислокацій і стационарну швидкість повзучості. Фізично цей розподіл реалізується при тих умовах, при яких він стійкий.

Для дослідження стійкості стационарних станів і топологічної структури фазових траєкторій в їх околі прослідкуємо еволюцію малих збурень δn_v і $\delta \rho_d$. Для цього у відповідні рівняння системи (3) підставимо $n_v = n_v^s + \delta n_v$ і $\rho_d = \rho_d^s + \delta \rho_d$. Лінеаризуємо їх відносно малих збурень і, використовуючи стационарність ρ_d^s і n_v^s , отримаємо систему лінійних рівнянь відносно $\delta \rho_d$ і δn_v . Її розв'язок експотенційно залежить від часу: $\delta \rho_d$, δn_v пропорційно $\exp(\lambda t)$. Декремент затухання малої флуктуації λ задовольняє рівнянню

$$\lambda^2 + p\lambda + q = 0 \quad (8)$$

де

$$p = D_v (\alpha n_v^s + 3\rho_d^s) > 0$$

$$q = \alpha \gamma D_v \rho_d^s + D_v^2 \rho_d^s n_v^s > 0$$

Значення p і q визначають поведінку розв'язку в околі особливої точки і класифікацію особливої точки. p і q – дійсні числа, при додатному дискримінанті мають однаковий знак. В цьому випадку точка фазової траєкторії з часом наближається чи віддаляється від особливої точки. При $p > 0$ і $q > 0$ всі фазові траєкторії в околі стационарної точки наближаються до неї, а точка називається стійким вузлом.

Фазовий портрет моделі

Для якісного аналізу фазової траєкторії використаємо метод ізоклін. Ізокліна являє собою лінію, яку фазові траєкторії перетинають під однаковим кутом. Вони задовольняють рівнянню $d\rho_d / dn_v = \text{const}$

Головні ізокліни є розв'язками рівнянь

$$\begin{aligned} \frac{d\rho_d}{dn_v} &= 0 \\ \frac{d\rho_d}{dn_v} &= \infty \end{aligned} \quad \text{або} \quad (9)$$

На перетині головних ізоклін знаходяться особливі точки.

Для ізокліни $dn_v = 0$,

$$n_v^1 = n_v^e + \frac{\gamma}{3D_v} + \frac{K}{3\rho_d D_v} > n_v^e \quad (10)$$

Для ізокліни $d\rho_d = 0$

а) Коли вираз під модулем менший за нуль

$$n_v^{2a} = \frac{\alpha K - 3k}{\alpha D_v \rho_d} \wedge n_v^{2a} < n_v^e \quad (11)$$

в) Коли вираз під модулем більший за нуль

$$n_v^{2b} = \frac{3k + \alpha K}{\alpha D_v \rho_d} \wedge n_v^{2b} > n_v^e \quad (12)$$

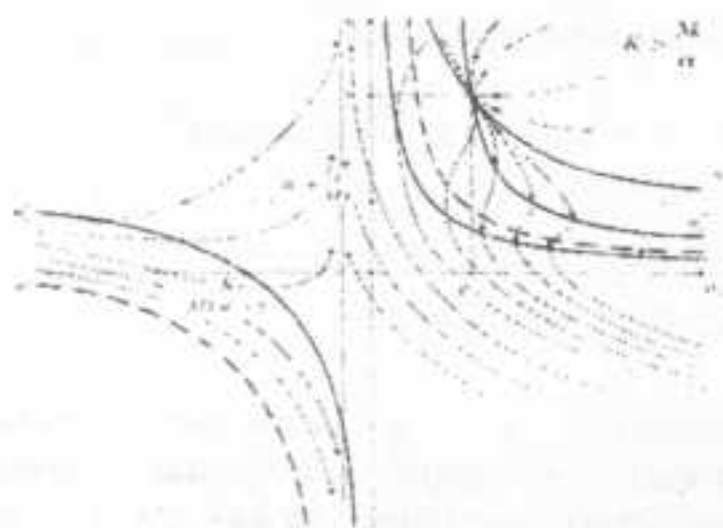


Рис.1 Фазовий портрет системи (3). Випадок, при якому швидкість генерації дислокацій (k) менша за швидкість генерації вакансій (K), з урахуванням концентрації термічних вакансій. $n_v^1, n_v^{2a}, n_v^{2b}$ - головні ізокліни, n_v^e - крива, де $\dot{\epsilon} = 0$

Більша частина деформації опроміненних матеріалів накопичується на перехідній стадії повзучості, тому є цікавим дослідження особливостей дозової (часової) залежності повзучості опроміненого зразка, динаміки її зміни. Динаміку розвитку повзучості зручніше за все представити за допомогою фазових портретів. В залежності від значень параметрів фазові портрети можуть мати різну топологічну структуру.

Для даної моделі, котрій відповідає вакансійний механізм повзучості, керуючими параметрами зручно вибрати концентрацію вакансій і густину дислокацій

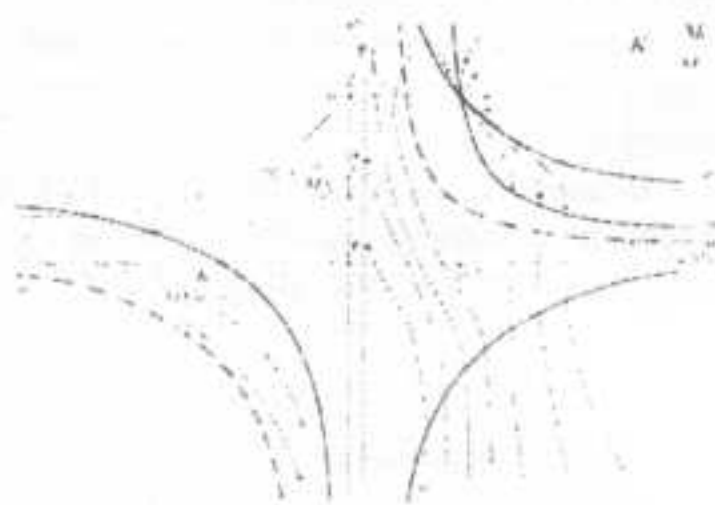


Рис.2 Фазовий портрет системи (3). Випадок, при якому швидкість генерації дислокацій (k) перевищує швидкість генерації вакансій (K) з урахуванням концентрації термічних вакансій. $n_v^1, n_v^{2a}, n_v^{2b}$ - головні ізокліни, n_v^e - крива, де $\dot{\epsilon} = 0$

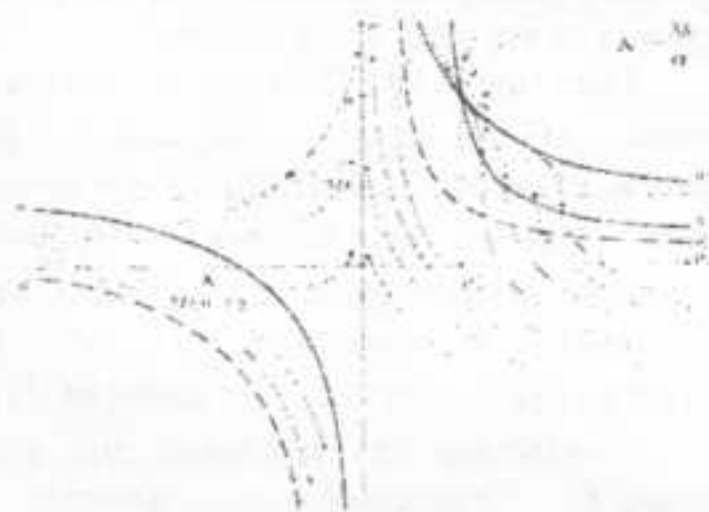


Рис.3 Фазовий портрет системи (3). Граничний випадок, при якому швидкість генерації вакансій (K) дорівнює швидкості генерації дислокацій (k) з урахуванням концентрації термічних вакансій. $n_v^1, n_v^{2a}, n_v^{2b}$ - головні ізокліни, n_v^e - крива, де $\dot{\epsilon} = 0$

Обговорення результатів

Фазові портрети мають якісно різний вигляд в залежності від того, що домінує: швидкість генерації вакансій (K) чи швидкість генерації дислокацій (k). Проаналізуємо фазовий портрет представлений на рис.1 для випадку, коли $\alpha K > k$. Фізичний зміст мають лише додатні значення концентрації вакансій та густини дислокацій, однак, щоб зрозуміти еволюцію фазового портрету при зміні параметрів, побудовано фазові траєкторії і головні ізокліни в усьому фазовому просторі. Рух точки по фазовій траєкторії відповідає еволюції системи, тобто зміні концентрації дефектів та густини дислокацій. В залежності від початкових умов можна виділити наступні області:

1). При великих значеннях ρ_d, n_v спостерігається монотонний рух до стаціонарної точки. При цьому швидкість повзучості не приймає нульового значення.

2). При невеликих концентраціях вакансій і великих густинах дислокацій спостерігається наступна картина: ρ_d досягає спочатку мінімального значення, потім максимального і при монотонному зростанні n_v прямує до стаціонару. При певному значенні швидкість повзучості може бути рівною нулю.

3). ρ_d досягає мінімуму, збільшується, при цьому n_v має максимум і монотонно рухається до особливої точки.

4). При дуже малих значеннях густини дислокацій ρ_d , величина концентрації вакансій n_v наближається до деякого граничного значення. Модель в цій області працює погано, так як не враховані інші стоки, крім дислокацій.

Вцілому фазові портрети подібні, різницю складає розташування вертикальної асимптоти. Для даного випадку вона зсунута вліво, для випадку, коли $\alpha K < k$, вона зсувається вправо, і таким чином зменшується область, де траєкторії прямують до стаціонарної точки; також для випадку, де $\alpha K < k$ головна ізокліна n_v^* разом з мінімумом переходить вниз.

Висновки

Генерація вакансій рухомими дислокаціями призводить до появи додаткової концентрації вакансій. Завдяки цьому підвищується середня швидкість руху дислокацій і зменшується їх густина. В межах досліджуваної моделі існує єдиний режим стаціонарної повзучості. Стаціонарне значення швидкості повзучості додатне (нульового значення не досягає), воно зростає зі збільшенням швидкості генерації дислокацій і з ростом зерна матеріалу. Часова (дозова) залежність перехідної повзучості в залежності від початкового стану опроміненого зразка має якісно різні види.

Список використаних джерел

- 1 Селищев П.А Самоорганизация в радиационной физике. Киев Аспект-Полиграф. 2004 .-239с.
- 2 Карасев В.С. Деформация аустенитной стали ОХ18Н10Т при циклическом действии нейтронного облучения // ВАНТ. Сер.: ФРП и РМ.-1981.- Вып.5(19). С.84-87
- 3 Карасев В.С., Тоцкий А.Ю. Смена механизма переходной ползучести в металлах под облучением // ВАНТ. Сер.: ФРП и РМ.-1990.-Вып 1(52).С.37-40
- 4 Карасев В.С., Тоцкий А.Ю. Влияние примеси на переходную радиационную ползучесть никеля // ВАНТ. Сер.: ФРП и РМ.-1994.- Вып.1(61).-с. 3-8

Надійшла до редколегії 12.02.08

Підписано до друку 16.03.2008. Формат 70x100¹/16.

Папір офсетний. Гарнітура Times New Roman.

Ум. друк. арк. 19,02.

Наклад 350 прим. Зам. № 66.

ДП «Інформаційно-аналітичне агентство»
вул. Еспланадна 4–6, м. Київ, 01601, Україна
тел/факс 287-03-79, 289-77-62
E-mail: iaa@dstati.kiev.ua

Наукова бібліотека
ім. М. Максимовича
КНУ
ім. ТАРАСА ШЕВЧЕНКА



13366JB

45 - чит зал періодики та дисерт. Ц. 25.00

