Розділ 11

ЧИСЕЛЬНЕ ІНТЕГРУВАННЯ ЗВИЧАЙНИХ ДИФЕРЕНЦІЙНИХ РІВНЯНЬ ТА ЇХНИХ СИСТЕМ

11.1. Загальні положення

Існує клас задач, в яких рівняння крайової задачі мають тільки прості похідні від одного параметра (позначимо його як $\tau > 0$). Наприклад, це задачі про згин стрижня чи вісесиметричної оболонки (τ є просторовим параметром) або еволюційні чи динамічні (τ є часовим параметром), для яких основне рівняння можна записати як

$$du(\tau)/d\tau = u'(\tau) = \varphi(\tau, u)$$
 also $d^2u(\tau)/d\tau^2 = u''(\tau) = \psi(\tau, u)$. (11.1)

В останньому випадку задачу можна представити у вигляді системи з двох диференційних рівнянь першого порядку:

$$du(\tau)/d\tau = u'(\tau) = \varphi(\tau, u); \quad d\varphi(\tau, u)/d\tau = \varphi'(\tau, u) = \psi(\tau, u).$$
(11.2)

Більш того, будь-яке рівняння з похідною *N*-го порядку можна представити у вигляді *N* рівнянь з похідними першого порядку. Тобто завжди можна привести таку задачу до системи рівнянь з похідними першого порядку:

$$du_i(\tau)/d\tau = \varphi_i(\tau, u_1, ..., u_N); \quad \tau > 0; \quad i = 1, ..., N,$$
(11.3)

де *N* – одночасно кількість диференційних рівнянь першого порядку та функцій, які необхідно знайти. В задачі обов'язково задаються вихідні умови:

$$u_i(0) = (u_i)_0; \quad i = 1, ..., N.$$
 (11.4)

Задача для системи диференційних рівнянь (11.3) з вихідними умовами (11.4) називається задачею Коші. Її розв'язок існує та є одиничним, якщо одночасно виконуються наступні умови:

• функції $\varphi_i(\tau, u_1, ..., u_N)$ є безперервними в замкненій області

$$D = \{ |\tau| \le a; |u_i - (u_i)_0| \le b; i = 1, ..., N \};$$
(11.5)

• функції $\varphi_i(\tau, u_1, ..., u_N)$ задовольняють умовам Ліпшіца

$$|\varphi_{i}(\tau, u_{1}', ..., u_{N}') - \varphi_{i}(\tau, u_{1}'', ..., u_{N}'')| \leq A \cdot \{|u_{1}' - u_{1}''| + ... + |u_{N}' - u_{N}''|\}$$
(11.6)

для будь-яких точок $(\tau, u'_1, ..., u'_N)$ та $(\tau, u''_1, ..., u''_N)$ області визначення D. Тут $A \in$ деякою константою, яка не залежить від τ .

11.2. Методи розв'язування задачі Коші для одного рівняння

Для розв'язування задачі Коші розроблено дві групи методів різного порядку наближення: Рунге-Кутта (Ейлера, Рунге-Кутта) і багатокрокові різницеві методи (Адамса, "прогнозу-корекції").

Ідеї всіх методів розглянемо у випадку N = 1 у рівняннях (11.3), (11.4). Рівняння (11.3) інтегрується згідно з теоремою про середнє при інтегруванні:

$$(u)_{n+1} = (u)_n + \int_{\tau_n}^{\tau_{n+1}} \varphi(\tau, u) d\tau = (u)_n + \tilde{\varphi}_n(\tilde{\tau}, \tilde{u}) \cdot \Delta \tau , \qquad (11.7)$$

де $\Delta \tau = \tau_{n+1} - \tau_n$; індекс n = 0, 1, ... є номером часового шару або кроку, але параметри $\tilde{\tau}$ й \tilde{u} функції $\tilde{\varphi}_n(\tilde{\tau}, \tilde{u})$, при яких інтегрування дає точне значення – невідомі, тому функція $\tilde{\varphi}_n(\tilde{\tau}, \tilde{u})$ утворюється у різних методах по-різному.

11.2.1. Методи Ейлера, Рунге-Кутта

У методі Рунге-Кутта функція $\tilde{\varphi}_n(\tilde{\tau}, \tilde{u})$ утворюється як комбінація декількох зважених функцій k_i , які обчислюються завдяки *явної m-етапної процедури*:

$$(u)_{n+1} \approx (u)_n + \tilde{\varphi}_n(\tilde{\tau}, \tilde{u}) \cdot \Delta \tau = (u)_n + \left(\sum_{j=1}^m c_j k_j\right) \cdot \Delta \tau ; \qquad (11.8)$$

$$\begin{cases} k_{1} = \varphi(\tau_{n}, (u)_{n}) = \varphi_{n}; \\ k_{2} = \varphi(\tau_{n} + a_{2}\Delta\tau, (u)_{n} + b_{21}\Delta\tau k_{1}); \\ k_{3} = \varphi(\tau_{n} + a_{3}\Delta\tau, (u)_{n} + b_{31}\Delta\tau k_{1} + b_{32}\Delta\tau k_{2}); \\ \dots; \\ k_{m} = \varphi(\tau_{n} + a_{m}\Delta\tau, (u)_{n} + b_{m1}\Delta\tau k_{1} + b_{m2}\Delta\tau k_{2} + \dots + b_{m,m-1}\Delta\tau k_{m-1}). \end{cases}$$
(11.9)

Коефіцієнти c_j, a_j, b_{jq} утворюються з системи алгебраїчних рівнянь (САР), яка збирається для обраного значення *m* з умов точності, причому приймається, що $\sum_{j=1}^{m} c_j = 1$, а значення *m* повинно співпадати з *p* – порядком наближення формули. Навіть при цих умовах, при *m* > 1 САР є недостатньо визначеною, тобто має більше невідомих, ніж рівнянь, а тому має багато варіантів розв'язків, причому не всі з них є прийнятними. Основні (класичні) варіанти функції $\tilde{\varphi}_n(\tilde{\tau}, \tilde{u})$ методу Рунге-Кутта наведено в таблиці 11.1.

Доведено, що в наслідок того, що метод Рунге-Кутта наближує розв'язок вихідного рівняння, він збігається при $\Delta \tau \rightarrow 0$, причому порядок його точності співпадає з порядком апроксимації.

При *m* > 5 формули методу Рунге-Кутта майже не використовуються.

11.2.2. Метод Адамса, багатокрокові методи

Метод Адамса – один з так званих багатокрокових методів з однаковим кроком $\Delta \tau$. Для початку процесу такі методи потребують попередніх підрахунків декількох значень функцій (φ)_n, (n = 0, 1, ...), причому ці значення потрібно одержати зі значно вищою точністю, ніж це бажано для усього розв'язку: на порядок або (краще) на два порядки. Якщо для цього буде застосовано метод Рунге-Кутта відповідного порядку наближення, то крок $\Delta \tau$ теж потрібно брати меншим, ніж у багатокроковому методі, який буде застосовуватися у подальшому. Це є недоліком багатокрокових методів.

TT		
Назва методу, порядок наближення	Функції $\tilde{\varphi}_n(\tilde{\tau},\tilde{u})$. Всюди $k_1 = \varphi(\tau_n,(u)_n) = \varphi_n$	
Ейлера, 1-й	k_1	
Ейлера	k_{2} , де $k_{2} = \varphi(\tau_{n} + \Delta \tau / 2, (u)_{n} + k_{1} \Delta \tau / 2)$	
(два варіанти), 2-й	$(k_1 + k_2)/2$, $\exists e \ k_2 = \varphi(\tau_n + \Delta \tau, (u)_n + k_1 \Delta \tau)$	
	$(k_1 + 4k_2 + k_3)/6,$	
	де $k_2 = \varphi(\tau_n + \Delta \tau / 2, (u)_n + k_1 \Delta \tau / 2);$	
Рунге-Кутта	$k_{3} = \varphi \left(\tau_{n} + \Delta \tau, (u)_{n} - (k_{1} - 2k_{2})\Delta \tau \right)$	
(два варіанти), 3-й	$(k_1 + 3k_3)/4$,	
	де $k_2 = \varphi(\tau_n + \Delta \tau / 3, (u)_n + k_1 \Delta \tau / 3);$	
	$k_{3} = \varphi\left(\tau_{n} + 2\Delta\tau/3, (u)_{n} + 2k_{2}\Delta\tau/3\right)$	
	$(k_1 + 2k_2 + 2k_3 + k_4)/6,$	
	де $k_2 = \varphi(\tau_n + \Delta \tau/2, (u)_n + k_1 \Delta \tau/2);$	
	$k_{3} = \varphi \left(\tau_{n} + \Delta \tau / 2, (u)_{n} + k_{2} \Delta \tau / 2 \right);$	
	$k_4 = \varphi \left(\tau_n + \Delta \tau, (u)_n + k_3 \Delta \tau \right)$	
	$(k_1+3k_2+3k_3+k_4)/8,$	
Рунге-Кутта	де $k_2 = \varphi(\tau_n + \Delta \tau / 3, (u)_n + k_1 \Delta \tau / 3);$	
(три варіанти),	$k_{3} = \varphi(\tau_{n} + 2\Delta\tau/3, (u)_{n} - (k_{1}/3 - k_{2})\Delta\tau);$	
4-й	$k_4 = \varphi \left(\tau_n + \Delta \tau, (u)_n + (k_1 - k_2 + k_3) \Delta \tau \right)$	
	$\left[k_{1}+(2-\sqrt{2})k_{2}+(2+\sqrt{2})k_{3}+k_{4}\right]/6,$	
	μe $k_2 = \varphi(\tau_n + \Delta \tau / 2, (u)_n + k_1 \Delta \tau / 2);$	
	$k_{3} = \varphi \Big(\tau_{n} + \Delta \tau / 2, (u)_{n} - \Big[(1/2 - 1/\sqrt{2})k_{1} - (1 - 1/\sqrt{2})k_{2} \Big] \Delta \tau \Big);$	
	$k_4 = \varphi \left(\tau_n + \Delta \tau, (u)_n - \left[k_2 / \sqrt{2} - (1 + 1 / \sqrt{2}) k_3 \right] \Delta \tau \right)$	

Таблиця 11.1. Функції $\tilde{\varphi}_n(\tilde{\tau},\tilde{u})$ у методах Ейлера і Рунге-Кутта

Для отримання формул багатокрокових методів у формулі (11.7) замість підінтегральної функції $\varphi(\tau, u)$ застосовують один з відомих інтерполяційних многочленів, який її наближує через декілька відомих значень функцій φ_n , (n = 0, 1, ...). У разі використання другої формули Ньютона (див. п.9.3.2) $\varphi_n \approx \varphi_{n-1} + q \nabla \varphi_{n-1} + \frac{q(1+q)}{2!} \nabla^2 \varphi_{n-1} + ...,$ де $q = (\tau - \tau_0) / \Delta \tau$, отримують явну формулу Адамса, в якій функція

$$\tilde{\varphi}_{n}(\tilde{\tau},\tilde{u}) \approx \varphi_{n} + \nabla \varphi_{n}/2 + 5\nabla^{2} \varphi_{n}/12 + 3\nabla^{3} \varphi_{n}/8 + 251\nabla^{4} \varphi_{n}/720 + 95\nabla^{5} \varphi_{n}/288 + \dots$$
(11.10)

(∇ – оператор диференціювання вказаного порядку), для застосування якої ще потрібно провести диференціювання.

Узагальнюючи цей та інші подібні методи, було розроблено теорію лінійних *m*-крокових (багатокрокових) різницевих методів, в яких розв'язок дається формулою

$$a_{0} \cdot (u)_{n+1} + a_{1} \cdot (u)_{n} + \dots + a_{m} \cdot (u)_{n-m+1} \approx (b_{0} \cdot \varphi_{n+1} + b_{1} \cdot \varphi_{n} + \dots + b_{m} \cdot \varphi_{n+1-m}) \cdot \Delta \tau, \qquad (11.11)$$

де $a_0 \neq 0$; n = m - 1, m, m + 1, ..., а також для усунення невизначеності прийнято, що $\sum_{i=0}^{m} b_i = 1$. Розрахунки починаються з n + 1 = m, тобто з рівняння

$$a_{0} \cdot (u)_{m} + a_{1} \cdot (u)_{m-1} + \dots + a_{m} \cdot (u)_{0} \approx (b_{0} \cdot \varphi_{m} + b_{1} \cdot \varphi_{m-1} + \dots + b_{m} \cdot \varphi_{0}) \cdot \Delta \tau.$$
(11.12)

Якщо в (11.11) прийняти $b_0 = 0$, то утворюється група явних схем, при $b_0 \neq 0$ – неявних. Доведено, що повинна виконуватися умова $\sum_{j=0}^{m} a_j = 0$, а також що завжди порядок наближення $p \leq 2m$ для неявних схем та $p \leq 2m-1$ для явних схем.

Всі схеми методу Адамса можна одержати з (11.11), якщо прийняти $a_0 = -a_1 = 1$, а інші $a_j = 0$; j = 2,...,m. Тоді повертаємося до формули (11.7) з

різними виразами для функції $\tilde{\varphi}_n(\tilde{\tau},\tilde{u}) = \sum_{j=0}^m b_j \varphi_{n+1-j}$:

$$(u)_{n+1} \approx (u)_n + (b_0 \cdot \varphi_{n+1} + b_1 \cdot \varphi_n + \dots + b_m \cdot \varphi_{n+1-m}) \cdot \Delta \tau = (u)_n + \left(\sum_{j=0}^m b_j \varphi_{n+1-j}\right) \cdot \Delta \tau .$$
(11.13)

У таблиці 11.2 наведено декілька перших виразів цієї функції. Порядок наближення методу Адамса $p \le m+1$ для неявних схем та $p \le m$ для явних. При виконанні деяких додаткових умов для неявних схем і при парних *m* порядок наближення методу Адамса $p \le m+2$.

Схема	т	р	Функції $ ilde{arphi}_n(ilde{ au}, ilde{u})$
явна	1	1	φ_n (метод Ейлера)
	2	2	$(3\varphi_n-\varphi_{n-1})/2$
	3	3	$(23\varphi_n - 16\varphi_{n-1} + 5\varphi_{n-2})/12$
	4	4	$(55\varphi_n - 59\varphi_{n-1} + 37\varphi_{n-2} - 9\varphi_{n-3})/24$
	5	5	$(1901\varphi_n - 2774\varphi_{n-1} + 2616\varphi_{n-2} - 1274\varphi_{n-3} + 251\varphi_{n-4})/720$
неявна	1	2	$(\varphi_{n+1}+\varphi_n)/2$
	2	3	$(5\varphi_{n+1}+8\varphi_n-\varphi_{n-1})/12$
	3	4	$(9\varphi_{n+1}+19\varphi_n-5\varphi_{n-1}+\varphi_{n-2})/24$
	4	5	$(251\varphi_{n+1} + 646\varphi_n - 2646\varphi_{n-1} + 106\varphi_{n-2} - 19\varphi_{n-3})/720$

Таблиця 11.2. Функції $\tilde{\varphi}_n(\tilde{\tau},\tilde{u})$ у методі Адамса

Відомо, що при m > 5 формули методу Адамса є нестійкими, тому й не використовуються.

Неявні схеми мають підвищений порядок наближення, але потребують ітераційного процесу для одержання $(u)_{n+1}$.

У разі використовування методу простих ітерацій вираз (11.13) записують як

$$(u)_{n+1}^{(k+1)} \approx (u)_n + \left[b_0 \cdot \varphi_{n+1}(\tau_n + \Delta \tau, (u)_{n+1}^{(k)}) + b_1 \cdot \varphi_n + \dots + b_m \cdot \varphi_{n+1-m} \right] \cdot \Delta \tau , \qquad (11.14)$$

де $k = 0, 1, ... \epsilon$ номером ітерації.

Якщо всі відомі члени позначити як

$$F_n = (u)_n + \left[b_1 \cdot \varphi_n + \dots + b_m \cdot \varphi_{n+1-m}\right] \cdot \Delta \tau, \qquad (11.15)$$

то (11.14) запишеться як

$$(u)_{n+1}^{(k+1)} \approx b_0 \cdot \varphi_{n+1}(\tau_n + \Delta \tau, (u)_{n+1}^{(k)}) \cdot \Delta \tau + F_n.$$
(11.16)

Визначимо число *M* таке, що $|\partial \varphi(\tau, u)/\partial u| \leq M$. Тоді умовою збіжності ітераційного процесу (11.16) буде нерівність

$$\Delta \tau < 1/(b_0 \cdot M), \tag{11.17}$$

яка є досить жорсткою при значних M (вимагає дуже малих кроків $\Delta \tau$).

Для інших багатокрокових схем рекомендують проводити аналіз характеристичного рівняння

$$\sum_{j=0}^{m} a_j q^{m-j} = 0, \qquad (11.18)$$

а саме: визначають, чи задовольняють всі його корені умові стійкості та збіжності | *q* |≤1, причому корені, для яких | *q* |=1, не повинні бути кратними.

11.2.3. Методи типу "прогноз-корекція"

Методами типу "прогноз-корекція" (або "предіктор-коректор") називають методи, в яких розв'язок утворюється послідовним виконанням таких дій: попереднього "прогнозу" (з недостатньою точністю) та наступною "корекцією" для отримання достатньої точності.

Багато варіантів такого методу можна одержати з неявних багатокрокових методів, якщо в ітераційному процесі обмежити кількість ітерацій. Зокрема, з (11.14) при *k* = 0 отримаємо "прогнозне" значення

 $(u)_{n+1}^{progn} = (u)_{n+1}^{(1)} \approx (u)_n + \left[b_0 \cdot \varphi_{n+1}(\tau_n + \Delta \tau, (u)_{n+1}^{(0)}) + b_1 \cdot \varphi_n + \dots + b_m \cdot \varphi_{n+1-m} \right] \cdot \Delta \tau , \quad (11.19)$ а потім "кореговане" (див. таблицю 11.3):

$$(u)_{n+1}^{corr} = (u)_n + \left[b_0 \cdot \varphi_{n+1}(\tau_n + \Delta \tau, (u)_{n+1}^{progn}) + b_1 \cdot \varphi_n + \dots + b_m \cdot \varphi_{n+1-m} \right] \cdot \Delta \tau .$$
(11.20)

Для визначення отриманої точності розглядається різниця $(u)_{n+1}^{corr} - (u)_{n+1}^{progn}$. Якщо вона значна — зменшують $\Delta \tau$ або обирають схему з більшим порядком наближення p.

Відомо ще багато інших формул методу. Зокрема, в них використовується такий підмічений факт, що різниця $(u)_{n+1}^{corr} - (u)_{n+1}^{progn}$ майже пропорційна локальній похибці усікання. Її помножують на ваговий коефіцієнт і додають до раніше отриманого результату.

У методі Дімсдейла 4-го порядку наближення (див. таблицю 11.3) можна без проблем міняти крок $\Delta \tau$, просто застосовуючи нове значення.

Назва методу	Формули методу. Всюди $\varphi_{n+1}^{mod} = \varphi(\tau_{n+1}, (u)_{n+1}^{mod})$
Прогноз за Адамсом- Башфортом та корек- ція за Адамсом- Мултоном	$(u)_{n+1}^{progn} = (u)_{n} + (55\varphi_{n} - 59\varphi_{n-1} + 37\varphi_{n-2} - 9\varphi_{n-3}) \cdot \Delta\tau;$ $(u)_{n+1}^{mod} = (u)_{n+1}^{progn} + 251 [(u)_{n}^{corr} - (u)_{n}^{progn}] / 270;$ $(u)_{n+1}^{corr} = (u)_{n} + (9\varphi_{n+1}^{mod} + 19\varphi_{n} - 5\varphi_{n-1} + \varphi_{n-2}) \cdot \Delta\tau;$ $(u)_{n+1} = (u)_{n+1}^{corr} - 19 [(u)_{n+1}^{corr} - (u)_{n+1}^{progn}] / 270$
Хеммінга	$(u)_{n+1}^{progn} = (u)_{n-3} + 4(2\varphi_n - \varphi_{n-1} + 2\varphi_{n-2}) \cdot \Delta \tau / 3;$ $(u)_{n+1}^{mod} = (u)_{n+1}^{progn} + 112 [(u)_n^{corr} - (u)_n^{progn}] / 121;$ $(u)_{n+1}^{corr} = \left\{ [9(u)_n - (u)_{n-2}] + 3(\varphi_{n+1}^{mod} + 2\varphi_n - \varphi_{n-1}) \cdot \Delta \tau \right\} / 8;$ $(u)_{n+1} = (u)_{n+1}^{corr} - 9 [(u)_{n+1}^{corr} - (u)_{n+1}^{progn}] / 121$
Мілна	$(u)_{n+1}^{progn} = (u)_{n-3} + 4(2\varphi_n - \varphi_{n-1} + 2\varphi_{n-2}) \cdot \Delta \tau / 3;$ $(u)_{n+1}^{mod} = (u)_{n+1}^{progn} + 28[(u)_n^{corr} - (u)_n^{progn}] / 29;$ $(u)_{n+1}^{corr} = (u)_{n-1} + (\varphi_{n+1}^{mod} + 4\varphi_n + \varphi_{n-1}) \cdot \Delta \tau / 3;$ $(u)_{n+1} = (u)_{n+1}^{corr} - [(u)_{n+1}^{corr} - (u)_{n+1}^{progn}] / 29$
	Примітка . Метод Мілна є нестійким, якщо $\partial \varphi / \partial u < 0$.
Дімсдейла	$(u)_{n+2} = (u)_n + 2\varphi_n \cdot \Delta \tau \text{ (або інший вираз для } (u)_{n+2});$ $(u)_{n+1} = 0.5[(u)_n + (u)_{n+2}] + 0.25(\varphi_n - \varphi_{n+2}) \cdot \Delta \tau;$ $(u)_{n+2} = (u)_n + [\varphi_n + 4\varphi_{n+1} + \varphi_{n+2}] \cdot \Delta \tau / 3$

Таблиця 11.3. Модифіковані методи "прогноз-корекція" 4-го порядку наближення

11.3. Методи інтегрування систем звичайних диференційних рівнянь

Методи, що розглянуті вище для одного рівняння, дуже просто пристосовуються до системи рівнянь (11.3) з похідними першого порядку та початковими умовами (11.4). Аналогічно (11.7)

$$(u_i)_{n+1} = (u_i)_n + \int_{\tau_n}^{\tau_{n+1}} \varphi_i(\tau, u_1, ..., u_N) d\tau \approx (u_i)_n + \tilde{\varphi}_i(\tilde{\tau}, \tilde{u}_1, ..., \tilde{u}_N) \cdot \Delta \tau; \quad i = 1, ..., N . (11.21)$$

Функції $\tilde{\varphi}_i(\tilde{\tau}, \tilde{u}_1, ..., \tilde{u}_N)$ для методу Рунге-Кутта аналогічні наведеним у таблиці 11.1. Зокрема, для першого варіанта 4-го порядку наближення:

$$\tilde{\varphi}_{i}(\tilde{\tau},\tilde{u}_{1},...,\tilde{u}_{N}) = \left[(k_{1})_{i} + 2(k_{2})_{i} + 2(k_{3})_{i} + (k_{4})_{i} \right] / 6; \quad i = 1,...,N,$$
(11.22)

157

$$\begin{aligned} &(k_{1})_{i} = \varphi_{i} \left(\tau_{n}, (u_{1})_{n}, ..., (u_{N})_{n} \right) = (\varphi_{i})_{n}; \\ &(k_{2})_{i} = \varphi_{i} \left(\tau_{n} + \Delta \tau / 2, (u_{1})_{n} + (k_{1})_{1} \Delta \tau / 2, ..., (u_{N})_{n} + (k_{1})_{N} \Delta \tau / 2 \right); \\ &(k_{3})_{i} = \varphi_{i} \left(\tau_{n} + \Delta \tau / 2, (u_{i})_{n} + (k_{2})_{1} \Delta \tau / 2, ..., (u_{N})_{n} + (k_{2})_{N} \Delta \tau / 2 \right); \\ &(k_{4})_{i} = \varphi_{i} \left(\tau_{n} + \Delta \tau, (u_{1})_{n} + (k_{3})_{1} \Delta \tau, ..., (u_{N})_{n} + (k_{3})_{N} \Delta \tau \right). \end{aligned}$$
(11.23)

Проблеми виникають, якщо процеси, які описуються диференційними рівняннями, мають різні масштаби (швидкості зміни). Такі системи диференційних рівнянь мають назву *жорстких* систем.

Для оцінювання ситуації зазвичай використовують частотний аналіз.

Позначимо якийсь розв'язок системи (11.3) як $v_i(\tau)$ та будемо розглядати різниці $z_i(\tau) = u_i(\tau) - v_i(\tau)$. Вони задовольняють системі диференційних рівнянь

$$dz_{i}(\tau)/d\tau = \varphi_{i}(\tau, v_{1} + z_{1}, ..., v_{N} + z_{N}) - \varphi_{i}(\tau, v_{1}, ..., v_{N}); \quad \tau > 0; \quad i = 1, ..., N.$$
(11.24)

Будемо вважати $z_i(\tau)$ малим збуренням, яке внесено в основний розв'язок $v_i(\tau)$. Праву частину системи (11.24) завжди можна розкласти в ряд Тейлора з достатньою точністю та одержати запис системи із застосуванням матриці:

 $dz_i(\tau)/d\tau = K(\tau, v_1, ..., v_N) \cdot \{z_1, ..., z_N\}^T + O(|z|); \quad \tau > 0; \quad i = 1, ..., N,$ (11.25) де O(|z|) є величина вищого порядку малості, якою можна знехтувати. Матриця $K(\tau, v_1, ..., v_N)$ має компоненти

$$K_{ii} = \partial \varphi_i(\tau, v_1, ..., v_N) / \partial(u_i); \quad i, j = 1, ..., N.$$
(11.26)

Її компоненти або стаціонарні, або залежать від τ . Проводиться аналіз власних чисел λ_i цієї матриці. Якщо відношення (число жорсткості)

$$s = \max_{1 \le i \le N} |\operatorname{Re} \lambda_i| / \min_{1 \le i \le N} |\operatorname{Re} \lambda_i|, \qquad (11.27)$$

є великим (конкретного граничного значення, на жаль, немає), а також якщо

$$\operatorname{Re}\lambda_i < 0; \quad i = 1, ..., N,$$
 (11.28)

то система диференційних рівнянь вважається *жорсткою* на розглядуваному інтервалі інтегрування (якщо власні числа не залежать від τ , то – на будь-якому інтервалі).

Складові системи з більш жорсткими характеристиками швидше затухають і з часом майже не впливають на розв'язок інших складових, але саме вони "контролюють" величину кроку $\Delta \tau$, суттєво обмежуючи його "зверху": перевищення кроком деякого критичного значення призводить до втрати стійкості розв'язку.

Доведено, що всі явні схеми мають лише умовну стійкість зі значним обмеженням $\Delta \tau$ "зверху". А серед неявних лінійних багатокрокових методів немає абсолютно стійких схем, які б мали порядок точності вище другого.

Кардинальним рішенням цієї проблеми є застосування для жорстких систем *чисто неявних* схем, для яких розмір кроку $\Delta \tau$ впливає лише на точність розв'язку. Поширеним є метод Гіра, в якому використовується багатокрокова

схема типу (11.11), в якій $b_0 = 1$, а всі інші $b_j = 0$; j = 1, ..., m. Її записують у вигляді

$$a_0 \cdot (u)_{n+1} - \varphi \left(\tau_n + \Delta \tau, (u)_{n+1}\right) \cdot \Delta \tau \approx -\sum_{j=1}^m a_j \cdot (u)_{n+1-j}$$
(11.29)

для одного рівняння та у вигляді

$$a_{0} \cdot (u_{i})_{n+1} - \varphi_{i} \left(\tau_{n} + \Delta \tau, (u_{1})_{n+1}, \dots, (u_{N})_{n+1} \right) \cdot \Delta \tau \approx -\sum_{j=1}^{m} a_{j} \cdot (u_{i})_{n+1-j}$$
(11.30)

для системи рівнянь. Із умови досягнення p-го порядку наближення записується система рівнянь відносно a_i ; j = 1, ..., m:

$$a_0 = -\sum_{j=1}^m a_j; \quad \sum_{j=1}^m j \cdot a_j = -1; \quad \sum_{j=1}^m j^k \cdot a_j = 0; \quad k = 2, ..., p.$$
(11.31)

При фіксованому *m* вона має лише один розв'язок, причому найвищий порядок наближення p = m.

Декілька схем методу Гіра наведено у таблиці 11.4 (використовуються схеми до m = 10 включно).

Природно, що ці формули можна застосовувати й для одного рівняння, тобто при N = 1.

Неявні схеми методу Гіра мають дуже добру, але все ж умовну стійкість.

Назва методу	т	Формула. Всюди $(F_i)_{n+1} = \varphi_i (\tau_n + \Delta \tau, (u_1)_{n+1},, (u_N)_{n+1}) \cdot \Delta \tau$
Ейлера неявний	1	$(u_i)_{n+1} - (u_i)_n = (F_i)_{n+1}$
	2	$[3(u_i)_{n+1} - 4(u_i)_n + (u_i)_{n-1}]/2 = (F_i)_{n+1}$
Гіра	3	$\left[11(u_i)_{n+1} - 18(u_i)_n + 9(u_i)_{n-1} - 2(u_i)_{n-2}\right]/6 = (F_i)_{n+1}$
	4	$[25(u_i)_{n+1} - 48(u_i)_n + 36(u_i)_{n-1} - 16(u_i)_{n-2} + 3(u_i)_{n-3}]/12 = (F_i)_{n+1}$

Таблиця 11.4. Формули методу Гіра

11.4. Загальні рекомендації

Який метод застосовувати? Це складна проблема, яка потребує компромісу між величиною похибки, часом розрахунку та стійкістю розв'язку, а також врахування особливостей задачі. При однаковій теоретичній точності кращими є формули, в яких коефіцієнти мають однакові знаки та не дуже відрізняються між собою за абсолютними значеннями (це зменшує похибки округлення), які мають меншу кількість дій, не потребують додаткового алгоритму для початку розрахунків, стійкі при більших значеннях $\Delta \tau$. Тому для окремих рівнянь та нежорстких систем найбільш поширеними є методи Рунге-Кутта, а з них перший варіант 4-го порядку наближення (див. таблицю 11.1). Але й він все ж має недоліки: потребує відносно значну кількість дій та не дуже зручний у контролі потрібної величини кроку $\Delta \tau$.

При практичному застосуванні описаних у цьому Розділі методів та схем, а також багатьох інших, після обирання порядку наближення схеми, тобто p, основним питанням є проблема визначення розміру кроку $\Delta \tau$ та пов'язана з ним проблема стійкості схеми.

Часто для оцінювання розміру кроку $\Delta \tau$ користуються *правилом Рунге*. Відповідно до нього, спочатку знаходять розв'язки задачі з двома різними кроками $\Delta \tau$ (позначимо останні як h_1 та h_2), досягаючи одного й того же значення τ . Потім обчислюються наближені значення функцій

$$\rho_i(\tau) \approx \left[(u_i)_{h_1} - (u_i)_{h_2} \right] / \left(h_2^p - h_1^p \right); \quad i = 1, \dots, N,$$
(11.32)

за допомогою яких для кожної *i*-ої функції визначаються основні члени похибки розв'язку:

$$(\varepsilon_{h_k})_i = (h_k)^p \cdot \rho_i(\tau) + O((h_k)^{p+1}); \quad k = 1, 2; \quad i = 1, ..., N.$$
(11.33)

В цих формулах p, як і раніше, є порядком наближення схеми. Якщо такі похибки не задовольняють, потрібно ще зменшувати крок $\Delta \tau$. Позначимо як ε бажану точність розв'язку всієї системи, тоді з (11.33) і (11.32)

$$\Delta \tau \le \min\{(h_{\varepsilon})_i\}, \quad \text{de } (h_{\varepsilon})_i \approx \sqrt[p]{\varepsilon \cdot \left| (h_2^p - h_1^p) / \left[(u_i)_{h_1} - (u_i)_{h_2} \right] \right|}; \quad i = 1, \dots, N.$$
(11.34)

Зазвичай приймають $h_2 = h_1/2$. Тоді з (11.32) і (11.34)

$$\rho_i(\tau) \approx 2^p \Big[(u_i)_{h_1/2} - (u_i)_{h_1} \Big] / \Big(h_1^p \cdot (2^p - 1) \Big); \quad i = 1, ..., N,$$
(11.35)

$$\Delta \tau \le \min\{(h_{\varepsilon})_i\}, \quad \text{de } (h_{\varepsilon})_i \approx \frac{h_1}{2} \sqrt[p]{\varepsilon \cdot (2^p - 1)/|(u_i)_{h_1} - (u_i)_{h_1/2}|}; \quad i = 1, ..., N.$$
 (11.36)

Формулу (11.36) можна використовувати й для підвищення точності отриманого розв'язку за методом Річардсона:

$$u_{i}(\tau) \approx (u_{i})_{h_{1}/2} + \left[(u_{i})_{h_{1}/2} - (u_{i})_{h_{1}} \right] / (2^{p} - 1); \quad i = 1, ..., N.$$
(11.37)

На жаль, єдиних для всіх схем розв'язування задачі Коші формул та умов немає. Це дуже складна та велика тема, яка лише частково розглянута у цьому Розділі. Для отримання додаткової інформації можна рекомендувати, наприклад, книгу [2], в якій велика увага приділена впливу на результати похибок усікання при розрахунках на ЕОМ.

Контрольні питання до підрозділу 11.1

1. Яку крайову задачу називають задачею Коші?

2. Які умови потрібно виконати, щоб розв'язок задачі Коші існував та був одиничним? Контрольні питання до підрозділу 11.2

1. В яких діапазонах значень знаходяться параметри $\tilde{\tau}$ й \tilde{u} функції $\tilde{\varphi}_n(\tilde{\tau},\tilde{u})$ для рівняння (11.7)?

2. Якою є кількість можливих розв'язків системи рівнянь Рунге-Кутта (11.9)?

3. Якою функцією у формулі (11.7) замінюється функція $\tilde{\varphi}_n(\tilde{\tau},\tilde{u})$ у багатокрокових методах (у методі Адамса)?

4. Як в багатокрокових методах (у методі Адамса) вирішується проблема виконання умов стійкості та збіжності?

5. Якім чином будуються методи типу "прогноз-корекція"?

Контрольні питання до підрозділу 11.3

1. Які проблеми виникають при інтегруванні систем звичайних диференційних рівнянь?

2. Коли виникають так звані жорсткі системи рівнянь?

3. Який максимальний порядок точності мають абсолютно стійкі багатокрокові методи інтегрування систем звичайних диференційних рівнянь?

Контрольні питання до підрозділу 11.4

1. Якими є загальні рекомендації щодо вибору методу отримання розв'язку задачі Коші?

Частина III КРАЙОВІ ЗАДАЧІ МЕХАНІКИ ДЕФОРМІВНОГО ТВЕРДОГО ТІЛА

Розділ 12

ЗАГАЛЬНІ СПІВВІДНОШЕННЯ МЕХАНІКИ СУЦІЛЬНИХ СЕРЕДОВИЩ

12.1. Системи координат. Тензор метрики простору

Якщо середовище має в будь-якому елементарному об'ємі $d\Omega$ масу $dm = \overline{\rho} d\Omega$, де його густина $\overline{\rho}$ є безперервною функцією координат, то середовище вважається суцільним.

У загальному випадку для описування стану тіла з об'ємом Ω_0 та поверхнею S_0 зазвичай вводять початковий час t_0 та відповідний йому початковий стан; поточний час t з відповідними об'ємом Ω та поверхнею S. Ще вводять декілька координатних систем. Введемо такі системи (рис.12.1):

• глобальну нерухому декартову систему координат з основним базисом $\vec{k_i}$, i = 1, 2, 3 (не показана);

• глобальну нерухому, в загальному випадку криволінійну не ортогональну, систему координат з базисом \vec{e}_i та координатними лініями x^i ; i = 1,2,3. Вона вводиться для опису руху реальних об'єктів. Базис \vec{e}_i може співпадати з базисом \vec{k}_i . Положення початку базису \vec{e}_i може бути довільним. Умова $x^i = const$; i = 1,2,3 задає координатну, в загальному випадку криволінійну, поверхню. Дві координатні поверхні x^i та x^j пересікаються



Рис.12.1 Координатні системи

(при $i \neq j$). Лінія їх пересічення називається координатною лінією, вона відповідає третій координаті x^k .

У цій системі *початкове* положення P_0 будь-якої матеріальної точки задається радіус-вектором $\vec{r}(t_0) = a^1 \vec{e}_1 + a^2 \vec{e}_2 + a^3 \vec{e}_3 = a^i \vec{e}_i = \vec{a}$, а *поточне* положення Pцієї ж точки – радіус-вектором $\vec{r}(t) = x^1(t)\vec{e}_1 + x^2(t)\vec{e}_2 + x^3(t)\vec{e}_3 = x^i(t)\vec{e}_i = \vec{x}(t)$. Співвідношення $a^i = a^i(x^j(t)); x^i(t) = x^i(a^j,t); i, j = 1,2,3$ є взаємо-однозначними зв'язками. Коваріантні вектори базису \vec{e}_i , який називається основним, задаються як

$$\vec{e}_i = \lim_{\Delta a^i \to 0} \frac{\Delta \vec{r}_0}{\Delta a^i} = \frac{\partial \vec{r}_0}{\partial a^i} = \frac{\partial \vec{a}}{\partial a^i}; \quad i = 1, 2, 3.$$
(12.1)

Ці три вектора основного базису є дотичними до координатних ліній у будь-якій точці P_0 недеформованого тіла. Найчастіше застосовують дві системи координат (ортогональні): декартову (ДСК) та циліндричну (ЦСК). У ДСК зазвичай позначують $x^i = x, y, z$; всі модулі $|\vec{e}_i| = 1$; у ЦСК $x^i = \rho, \vartheta, z$; причому $|\vec{e}_i| = |\vec{e}_3| = 1$, $|\vec{e}_2| = \rho$. Якщо початки координатних систем збігаються, то $\vec{r} = \rho \cos \vartheta \vec{e}_1 + \rho \sin \vartheta \vec{e}_2 + z \vec{e}_3$;

• локальну рухому ("вморожену"), в загальному випадку криволінійну не ортогональну, систему координат з базисом \vec{E}_i , яка супроводжує кожну матеріальну точку *P* тіла при його деформуванні. На поточне положення кожної такої точки та координатної системи одночасно вказує радіус-вектор $\vec{r} = \vec{r}(a^j, t)$.

У системі координат з базисом \vec{e}_i дев'ять величин

$$g_{ij} = (\vec{e}_i, \vec{e}_j) = \vec{e}_i \cdot \vec{e}_j = \frac{\partial \vec{a}}{\partial a^i} \cdot \frac{\partial \vec{a}}{\partial a^j}$$
(12.2)

називають коваріантними компонентами симетричного метричного тензора. У ДСК і ЦСК відповідно

$$(g_{ij})_{\mathcal{ACK}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \delta_{ij}; \quad (g_{ij})_{\mathcal{UCK}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
 (12.3)

Тут і нижче верхній цифровий індекс при ρ – показник степені; δ_{ij} – символ Кронекера.

Дійсним *рімановим простором* називається *n*-вимірний простір з координатами точок $a^1, a^2, ..., a^n$, якщо в ньому задано коваріантний тензор другого рангу з компонентами $g_{ij}(a^1,...,a^n)$, який задовольняє наступним умовам:

• компоненти $g_{ij}(a^1, ..., a^n)$ є дійсними функціями координат, мають безперервні частинні похідні;

- тензор з компонентами $g_{ij}(a^1,...,a^n) \in$ симетричним, тобто $g_{ij} = g_{ji}$;
- $g = \det |g_{ii}| \neq 0$.

ДСК і ЦСК – окремі випадки ріманова простору.

Для проведення деяких операцій, поряд з основним, вводиться взаємний базис \vec{e}^i у такий спосіб, щоб $\vec{e}^i \cdot \vec{e}_j = \delta^i_j = \delta_{ij}$ (вектори \vec{e}_i вказують напрям координатних ліній, а вектори \vec{e}^i – перпендикулярні до координатних поверхонь). Аналогічно g_{ij} вводиться контраваріантний симетричний метричний тензор з компонентами $g^{ij} = \vec{e}^i \cdot \vec{e}^j$. Тоді

Частина	II. Розд)іл 12
---------	----------	--------

$$g_{ij}g^{jk} = \delta_i^k; \quad g^{ij} = (g_{ij})^{-1}; \quad \vec{e}^i = g^{ij}\vec{e}_j; \quad \vec{e}_i = g_{ij}\vec{e}^j.$$
(12.4)

3 урахуванням (12.3) і (12.4)

$$(g^{ij})_{\mathcal{A}CK} = (g_{ij})_{\mathcal{A}CK} = \delta_{ij}; \quad (g^{ij})_{\mathcal{U}CK} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\rho^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
 (12.5)

Згідно з (12.1), нескінченно малий приріст $d\vec{a}$ вектору \vec{a} представляється у вигляді $d\vec{a} = da^i \vec{e}_i$, з використанням (12.4) – аналогічно: $d\vec{a} = da_i \vec{e}^i$.

Для зручності, зокрема, щоб метрика простору не була присутня у фізичних рівняннях, може вводитися локальний ортонормований "фізичний" (місцевий) базис \vec{b}_i , тобто такий, щоб усі три $|\vec{b}_i|=1$. Цього можна досягти, якщо взяти \vec{b}_i наступного вигляду (під радикалом тут не додавати):

$$\vec{b}_i = \vec{e}_i / \sqrt{g_{ii}} = \sqrt{g^{ii}} \vec{e}_i$$
 (12.6)

163

З (12.3), (12.5) і (12.6) випливає, що в ДСК основний і "фізичний" базиси співпадають, чого не можна сказати про інші, де базиси зазвичай різні. У "фізичному" базисі та ДСК часто всі компонентні індекси записують внизу, хоча для цих базисів немає значення, де записані компонентні індекси.

Примітка 12.1. Величина $\sqrt{g_{ii}} = |\vec{e}_i|$, тобто є модулем відповідного базисного вектора. Тому ще один варіант визначення базисних векторів $\vec{e}_i = \sqrt{g_{ii}}\vec{b}_i$.

Будь-який вектор \vec{q} в основному, взаємному і "фізичному" базисах має компоненти q^i, q_i, \vec{q}_i відповідно. Він виражається як

$$\vec{q} = q^i \vec{e}_i = q_i \vec{e}^i = \breve{q}_i \vec{b}_i.$$
(12.7)

3 (12.7) з урахуванням перетворень (12.4) і (12.6)

$$q_{j} = q^{i}g_{ij}; \quad q^{j} = q_{i}g^{ij}; \quad \breve{q}_{j} = q^{j}\sqrt{g_{jj}}.$$
 (12.8)

Тому і координатні вектори \vec{a} точки P_0 , \vec{r} точки P, проведені до них з початку координат, можна записати як

$$\vec{a} = a^i \vec{e}_i = a_i \vec{e}^i = \breve{a}_i \vec{b}_i; \quad \vec{r} = x^i \vec{e}_i = x_i \vec{e}^i = \breve{x}_i \vec{b}_i.$$
 (12.9)

Будь-який тензор другого рангу с в основному, взаємному і "фізичному" базисах визначається як:

$$\sigma = \sigma^{ij} \vec{e}_i \otimes \vec{e}_j = \sigma_{ij} \vec{e}^i \otimes \vec{e}^j = \breve{\sigma}_{ij} \vec{b}_i \otimes \vec{b}_j, \qquad (12.10)$$

де знак \otimes вказує на подвійне перемноження векторів; σ_{ij} , σ^{ij} – компоненти тензора σ (строго кажучи, компоненти $\breve{\sigma}_{ij}$ не створюють тензор, оскільки не підкоряються правилам перерахунку компонент тензора при повороті системи координат). З (12.10) з урахуванням перетворень (12.4) і (12.6) отримаємо (під радикалом тут не додавати):

$$\sigma^{ij} = \sigma_{mn} g^{im} g^{jn}; \quad \sigma_{ij} = \sigma^{mn} g_{im} g_{jn}; \quad \breve{\sigma}_{ij} = \sigma^{ij} \sqrt{g_{ii} / g^{jj}} = \sigma_{mn} g^{im} g^{jn} \sqrt{g_{ii} / g^{jj}} . \quad (12.11)$$

За визначенням, квадрат лінійного елементу $(ds_0)^2 = d\vec{r}_0 \cdot d\vec{r}_0 = da^i \vec{e}_i \cdot da^j \vec{e}_j = g_{ij} da^i da^j$. Ще одна метрична характеристика:

$$\sqrt{g} = \sqrt{\det |g_{ij}|} = \vec{e}_1 \cdot (\vec{e}_2 \times \vec{e}_3) = \vec{e}_2 \cdot (\vec{e}_3 \times \vec{e}_1) = \vec{e}_3 \cdot (\vec{e}_1 \times \vec{e}_2).$$
(12.12)
Зокрема, для ДСК $\sqrt{g} = 1$, а для ЦСК $\sqrt{g} = \rho$.

12.2. Кінематичні співвідношення

Кінематика визначає просторово-часові поняття, що описують рух. Далі всюди, де спеціально не обумовлено, застосовується тривимірний Евклідовий простір, декартова система координат і поняття часу, засноване на існуванні абсолютної одночасності (змінна $t \ge 0$). У механіці деформівного твердого тіла (МДТТ) розглядають системи (тіла) об'ємом Ω , обмежені поверхнею S, що міняють свою форму з плином часу, водночас зберігаючи свій склад. Тому в МДТТ зазвичай використовується так звана *вморожена* або *конвекційна* система координат: трійка чисел (координат) нерозривно зв'язана з *конкретнюю* точкою тіла, тому значення чисел змінюються в часі (підхід Лагранжа). Натомість, коли розглядають стан рідкого середовища або газу поблизу твердого тіла, трійка чисел (координат) частіше постійна в часі і вказує *місце*, для якого обчислюються параметри середовища (підхід Ейлера).

Трансформація точки тіла в часі може бути описана оператором однозначного відображення $\Pi(t_o,t)$ при будь-яких t_o та t (зі стану t_o у стан t через стан t_*) з наступними груповими властивостями (абстрактна алгебра):

 $\Pi(t,t) = I;$ $\Pi(t_o,t) = \Pi(t_*,t) \odot \Pi(t_o,t_*);$ $\Pi(t_o,t) \odot \Pi(t,t_o) = I,$ (12.13) де I – тотожне відображення. Друге співвідношення відбиває транзитивність, третє стверджує, що відображення $\Pi(t_o,t)$ є обернене до $\Pi(t,t_o)$.

У системі відліку \Re точка *P* в момент часу *t* визначається координатами x^i , тому відображення П задається співвідношенням:

$$x^{i} = \Im^{i}((x^{j})_{o}, t_{o}, t), \text{ afo } x^{i} = \Im^{i}(\vec{x}_{o}, t_{o}, t), \text{ afo } \vec{r} = \Im(\vec{r}_{o}, t_{o}, t).$$
 (12.14)

Функції \mathfrak{T}^i не можуть бути довільними, вони зв'язані груповими властивостями відображення (12.13). Функції \mathfrak{T}^i мають бути безперервними і такими, що безперервно диференціюються необхідну кількість разів.

Траєкторія точки P тіла — це геометричне місце положень P_t у системі відліку \Re при змінному t. Параметрично її визначає співвідношення:

$$x^{i} = \mathfrak{I}^{i}(\tilde{x}^{j}, \tilde{t}, t), \qquad (12.15)$$

де $\tilde{x}^{j}, \tilde{t} - \phi$ іксовані.

Поле переміщень за відрізок часу від t_o до t визначається як різниця положень

$$u^{i} = \Im^{i}(\vec{r}_{o}, t_{o}, t) - \Im^{i}(\vec{r}_{o}, t_{o}, t_{o}).$$
(12.16)

Швидкість в момент t точки P, яка знаходилася в момент t_o в \vec{r}_o , представляється вектором \vec{V} з компонентами в \Re

$$V^{i} = \dot{x}^{i} = \frac{\partial \mathfrak{I}^{i}(\vec{r}_{o}, t_{o}, t)}{\partial t}.$$
(12.17)

Аналогічно прискорення

Частина III. Розділ 12	165

$$w^{i} = \ddot{x}^{i} = \dot{V}^{i} = \frac{\partial^{2} \mathfrak{I}^{i}(\vec{r}_{o}, t_{o}, t)}{\partial t^{2}}.$$
 (12.18)

Рух є стаціонарним, якщо поле швидкостей V^i не залежить від часу t.

Згідно з підходом Лагранжа достатньо розглядати поточну конфігурацію порівняно з деякою відомою конфігурацією з часом t_* . Якщо час t_* збігається з початковим t_0 (частіше приймають $t_0 = 0$), то такий варіант називають "*повне формулювання Лагранжа*" (TL – від Total Lagrangian). Інакше, якщо $t_* > t_0$ та використовується малий часовий крок $\Delta t = t - t_*$, то такий варіант називають "*модифіковане формулювання Лагранжа*" (UL – від Updated Lagrangian). Останній варіант відображення руху громіздкий у використанні, оскільки необхідно знати багато функції З^{*i*} на всьому часовому просторі.

Всі три варіанти: TL, UL, Ейлера, зрештою, еквівалентні.

Координати точки P_0 в системі відліку \Re позначимо як a^i , i = 1, 2, 3. Їх називають лагранжевими змінними. Також позначимо:

$$\hat{\mathfrak{I}}(\vec{a},0,t) = \tilde{\Phi}(\vec{a},t); \quad \hat{\mathfrak{I}}(\vec{r},t,0) = \Psi(\vec{r},t).$$
 (12.19)

Тоді

$$\vec{r} = \vec{\Phi}(\vec{a},t); \quad \vec{a} = \vec{\Psi}(\vec{r},t); \quad \vec{a} = \vec{\Phi}(\vec{a},0); \quad \vec{r} = \vec{\Psi}(\vec{r},0).$$
 (12.20)

При цьому зовсім не обов'язково, щоб лагранжеві змінні були координатами точок тіла саме в системі відліку \Re . Така абстрактна конфігурація може бути задана достатньо довільно, що часто спрощує розгляд. Головне, що існує взаємно-однозначний зв'язок між \vec{r} та $\vec{\alpha}$ (див. (12.20)), тобто

$$\partial x^{i} / \partial a^{j} = \left| \partial \Phi^{i}(\vec{a},t) / \partial a^{j} \right| > 0.$$
(12.21)

Інші характеристики руху можна отримати, знаючи залежності (12.20):

• поле переміщень за час від t_0 до t

$$u^{i} = u^{i}(\vec{a},t) = \Phi^{i}(\vec{a},t) - \Phi^{i}(\vec{a},t^{o}) = x^{i} - \alpha^{i}; \qquad (12.22)$$

• поле швидкостей

$$V^{i} = \dot{x}^{i} = \dot{u}^{i} = \frac{\partial u^{i}(\vec{a},t)}{\partial t} = \frac{\partial x^{i}(\vec{a},t)}{\partial t}; \qquad (12.23)$$

• поле прискорень

$$w^{i} = \dot{V}^{i} = \ddot{x}^{i} = \ddot{u}^{i} = \frac{\partial^{2} u^{i}(\vec{a},t)}{\partial t^{2}} = \frac{\partial^{2} x^{i}(\vec{a},t)}{\partial t^{2}}.$$
 (12.24)

Також можна описати зміни будь-яких фізичних полів, що супроводжують матеріальну точку в її русі, наприклад, температури

$$\theta = \theta(\vec{r}, t) = \theta(\vec{r}(\vec{a}, t), t) \tag{12.25}$$

або густини середовища

$$\overline{\rho} = \overline{\rho}(\vec{r}, t) = \overline{\rho}(\vec{r}(\vec{a}, t), t).$$
(12.26)

Формулювання Лагранжа дозволяє стежити за частками середовища, траєкторії яких утворюються автоматично.

У формулюванні Ейлера чотири "незалежні ейлерові змінні" x^1, x^2, x^3, t визначають положення частки M^t , а "ейлеровими змінними, які необхідно

знайти", є компоненти *швидкості* частки M в той же момент. При цьому зручно ввести таку змінну τ , для якої $d\tau = dt$ і позначити:

$$\hat{x}^{i} = x^{i}; \quad \hat{V}^{i} = V^{i}; \quad i = 1, 2, 3; \quad \hat{x}^{4} = t; \quad \hat{V}^{4} = 1.$$
 (12.27)

Тоді замість (12.17) маємо:

$$\frac{d\hat{x}^{m}}{d\tau} = \hat{V}^{m}(\hat{x}^{1}, \hat{x}^{2}, \hat{x}^{3}, \hat{x}^{4}); \quad m = 1, 2, 3, 4; \qquad \text{afo} \quad \frac{d\hat{x}}{d\tau} = \dot{\bar{x}} = \vec{V}(\vec{x}).$$
(12.28)

Однією з важливих властивостей формулювання Ейлера є те, що відразу виявляється інваріантність розв'язку відносно змінної τ (періодичність, стаціонарність), а також аналогічність між групами траєкторій, лініями течії тощо.

Перехід від Ейлерових змінних до Лагранжевих координат здійснюється шляхом інтегрування за часом ейлерових змінних з подальшим розв'язуванням отриманої системи рівнянь відносно трьох постійних інтегрування. Звільнившись від постійних інтегрування, одержуємо Лагранжеві координати, які потрібно підставити в усі описи фізичних полів, що супроводжують матеріальну точку в її русі, наприклад: температури або густини середовища.

12.3. Тензор деформацій Гріна-Лагранжа

Розглянемо початкову конфігурацію тіла з координатним вектором $\vec{r}_0 = \vec{a}$ до точки P_0 , а також поточну з координатним вектором \vec{r} до точки P. Використовуємо в загальному випадку криволінійну систему координат з базисом \vec{e}_i . Відповідно до (12.1) у початковому стані:

$$\vec{e}_i = \frac{\partial \vec{r}_0}{\partial a^i}; \quad \vec{e}^j = \frac{\partial \vec{r}_0}{\partial a_j}; \quad d\vec{r}_0 = da^i \vec{e}_i = da_j \vec{e}^j.$$
(12.29)

У поточному стані маємо локальний ("вморожений") здеформований базис

$$\vec{E}_{i} = \frac{\partial \vec{r}}{\partial a^{i}}; \quad \vec{E}^{j} = \frac{\partial \vec{r}}{\partial a_{j}} = C^{ji}\vec{E}_{i}; \quad d\vec{r} = da^{i}\vec{E}_{i} = da_{j}\vec{E}^{j}$$
(12.30)

(про *С^{ji}* – нижче). Тоді квадрати вектора, що зв'язує дві нескінченно близькі матеріальні точки, до і після деформування є, відповідно:

$$(ds_0)^2 = d\vec{r}_0 \cdot d\vec{r}_0 = (da^i \vec{e}_i) \cdot (da_j \vec{e}^j) = da^i da_j \delta^j_i = da^i da_i = da^i g_{ij} da^j = g_{ij} da^i da^j;$$
(12.31)

$$ds^{2} = d\vec{r} \cdot d\vec{r} = (da^{i}\vec{E}_{i}) \cdot (da_{j}\vec{E}^{j}) = C^{j}_{i}da^{i}da_{j} = C_{ij}da^{i}da^{j} = C^{ij}da_{i}da_{j}.$$
(12.32)

Примітка 12.2. Деформацію можна розглядати і щодо будь-якої іншої конфігурації, але, наприклад, для **TL**-формулювання потрібна деформація саме відносно початкової конфігурації.

За визначенням, інваріант

$$ds^{2} - (ds_{0})^{2} \equiv 2 \in_{ii} da^{i} da^{j}$$
(12.33)

є об'єктивною мірою деформації Гріна-Лагранжа в точці тіла щодо *початкової* конфігурації, де \in_{ij} – компоненти цього тензора деформацій. Об'єктивними мірами деформацій є такі, що не залежать від жорстких повороту та зміщення.

Якщо в усьому тілі $ds^2 - (ds_0)^2 = 0$, то рух тіла – абсолютно жорсткий.

Рух точки тіла як абсолютно жорсткого в лагранжевих змінних описують виразом

$$\vec{r}^* = \mathbf{R}^T(t) \cdot \vec{r} + \vec{c}(t), \qquad (12.34)$$

де \vec{r} визначається лагранжевими змінними; $\vec{c}(t)$ – вектор жорсткого плоскопаралельного зміщення; $\mathbf{R}^{T}(t)$ – транспонована матриця жорсткого повороту з компонентами R^{ij} , тобто $\mathbf{R}\mathbf{R}^{T} = \mathbf{R}^{T}\mathbf{R} = \mathbf{I}$ (I– одинична матриця). Якщо система відліку жорстко зв'язана з тілом, то

$$\vec{c}(t) \equiv \vec{0}; \quad \vec{r} = \hat{\vec{x}}; \quad \vec{r}^* = \hat{\vec{x}}^*; \quad \hat{\vec{x}}^* = \mathbf{R}^T(t)\hat{\vec{x}},$$
(12.35)

причому компоненти вектора \vec{x} не залежать від t. Тому

$$\frac{d\vec{x}^*}{dt} = \left(\frac{d\mathbf{R}}{dt}\right)^T \vec{\hat{x}} = \left|\vec{\hat{x}} = \mathbf{R}\vec{\hat{x}}^*\right| = \frac{d\mathbf{R}^T}{dt} \mathbf{R}\vec{\hat{x}}^* = \mathbf{\Xi}\vec{\hat{x}}^* ; \quad \frac{(d\hat{x}^i)^*}{dt} = \Xi^{ij}(\hat{x}^j)^*, \quad (12.36)$$

де матриця Ξ з компонентами $\Xi^{ij} = \frac{dR^{ik}}{dt}R^{kj}$ є кососиметричною (див. підрозділ 4.1).

Якщо в тілі є точки з $ds^2 - (ds_0)^2 \neq 0$, то вони перебувають (і тіло також) у деформованому стані.

Підставимо (12.31) і (12.32) у (12.33):

$$2 \in_{ij} da^{i} da^{j} = C_{ij} da^{i} da^{j} - g_{ij} da^{i} da^{j} = (C_{ij} - g_{ij}) da^{i} da^{j}.$$
(12.37)

З (12.37) випливає, що *поточні* компоненти симетричного тензора деформації Гріна–Лагранжа (Гріна-Сен-Венана) відносно *початкової* конфігурації:

$$=_{ij} = 0.5(C_{ij} - g_{ij}).$$
(12.38-a)

Виразимо компоненти C_{ij} через компоненти вектора переміщень \vec{u} . Оскільки $\vec{r} = \vec{r}_0 + \vec{u}$, а $C_{ij} = \vec{E}_i \cdot \vec{E}_j$, то з урахуванням (12.1) і першого виразу (12.30) визначимо спочатку, що

$$\vec{E}_{i} = \frac{\partial \vec{r}}{\partial a^{i}} = \frac{\partial (\vec{r}_{0} + \vec{u})}{\partial a^{i}} = \vec{e}_{i} + \frac{\partial \vec{u}}{\partial a^{i}}; \quad \vec{E}_{j} = \vec{e}_{j} + \frac{\partial \vec{u}}{\partial a^{j}}.$$
(12.39)

Після цього (C_{ij} є аналогом g_{ij} з (12.2), але у здеформованому стані)

$$C_{ij} = \vec{E}_i \cdot \vec{E}_j = \left(\vec{e}_i + \frac{\partial \vec{u}}{\partial a^i}\right) \cdot \left(\vec{e}_j + \frac{\partial \vec{u}}{\partial a^j}\right).$$
(12.40)

Підставимо (12.40) в (12.38-а):

$$\epsilon_{ij} = \frac{1}{2} \left[\left(\vec{e}_i + \frac{\partial \vec{u}}{\partial a^i} \right) \cdot \left(\vec{e}_j + \frac{\partial \vec{u}}{\partial a^j} \right) - g_{ij} \right].$$
(12.41)

Після розкриття дужок і врахування $\vec{e}_i \cdot \vec{e}_j = g_{ij}$, отримаємо, що

$$\epsilon_{ij} = \frac{1}{2} \left(\vec{e}_i \cdot \frac{\partial \vec{u}}{\partial a^j} + \vec{e}_j \cdot \frac{\partial \vec{u}}{\partial a^i} + \frac{\partial \vec{u}}{\partial a^i} \cdot \frac{\partial \vec{u}}{\partial a^j} \right).$$
(12.42)

Позбудемося векторів. Для цього врахуємо, що відповідно до (12.7) $\vec{u} = u^n \vec{e}_n$. Для (12.42) послідовно отримуємо:

$$\frac{\partial \vec{u}}{\partial a^{i}} = \frac{\partial (u^{n} \vec{e}_{n})}{\partial a^{i}} = \frac{\partial u^{n}}{\partial a^{i}} \vec{e}_{n} + u^{k} \frac{\partial \vec{e}_{k}}{\partial a^{i}} = \left(\frac{\partial u^{n}}{\partial a^{i}} + u^{k} \Gamma_{ki}^{n}\right) \vec{e}_{n} = \nabla_{i} u^{n} \vec{e}_{n}; \qquad (12.43)$$

$$\frac{\partial \vec{u}}{\partial a^{j}} = \left(\frac{\partial u^{n}}{\partial a^{j}} + u^{k} \Gamma_{kj}^{n}\right) \vec{e}_{n} = \nabla_{j} u^{n} \vec{e}_{n}; \qquad (12.44)$$

$$\frac{\partial \vec{u}}{\partial a^{i}} \cdot \frac{\partial \vec{u}}{\partial a^{j}} = (\nabla_{i} u^{m} \vec{e}_{m}) \cdot (\nabla_{j} u^{n} \vec{e}_{n}) = \nabla_{i} u^{m} \nabla_{j} u^{n} (\vec{e}_{m} \cdot \vec{e}_{n}) = \nabla_{i} u^{m} \nabla_{j} u^{n} g_{mn} = \nabla_{i} u^{m} \nabla_{j} u_{m}, \quad (12.45)$$

де позначені коваріантні похідні компонент вектора:

$$\nabla_i u^n = \frac{\partial u^n}{\partial a^i} + u^k \Gamma_{ki}^n; \qquad (12.46)$$

а Γ_{ki}^{n} – символи Крістофеля другого роду, що симетричні за нижніми індексами і є компонентами розкладу $\partial \vec{e}_{k} / \partial a^{i}$ по вихідному базису \vec{e}_{n} :

$$\Gamma_{ik}^{n} = \Gamma_{ki}^{n} = \frac{1}{2} g^{nj} \left(\frac{\partial g_{kj}}{\partial a^{i}} + \frac{\partial g_{ij}}{\partial a^{k}} - \frac{\partial g_{ik}}{\partial a^{j}} \right); \quad i, j, m, n = 1, 2, 3.$$
(12.47)

Підставивши (12.43) ... (12.45) у (12.42), з урахуванням $\vec{e}_i \cdot \nabla_j u^n \vec{e}_n = \nabla_j u_i;$ $\vec{e}_j \cdot \nabla_i u^n \vec{e}_n = \nabla_i u_j$, маємо остаточно, що в довільній системі координат

$$\in_{ij} = 0.5(\nabla_{i} u_{j} + \nabla_{j} u_{i} + \nabla_{i} u^{m} \nabla_{j} u_{m}); \quad i, j, m = 1, 2, 3.$$
(12.38-6)

Звернемо особливу увагу на те, що в (12.47) та (12.38-б) наявне згортання по індексам n = 1, 2, 3 та m = 1, 2, 3 відповідно.

Вектор, що фігурує в (12.39), а саме

$$\vec{E}_i = \frac{\partial \vec{r}}{\partial a^i} = \vec{e}_i + \frac{\partial \vec{u}}{\partial a^i} = (\delta_i^m + \nabla_i u^m) \vec{e}_m = \nabla_i x^m \vec{e}_m = X_{mi} \vec{e}_m, \qquad (12.48)$$

називається градієнтом руху Гріна (Коші-Гріна). Функції $C_{ij} = \vec{E}_i \cdot \vec{E}_j = \vec{E}_j \cdot \vec{E}_i = C_{ji} = X_{mi}X_{mj}$ – коваріантні компоненти симетричного метричного тензора для конфігурації в конвективній (вмороженій) системі координат – компоненти тензора *міри деформації Гріна* (Коші-Гріна).

Компоненти тензора деформацій не є незалежними. Для суцільного твердого тіла, що деформується, виходячи з вимоги збереження суцільності, виводяться *рівняння сумісності деформацій* (їх шість):

$$\frac{\partial^{2} \in_{ij}}{\partial a^{m} \partial a^{n}} + \frac{\partial^{2} \in_{mn}}{\partial a^{i} \partial a^{j}} - \frac{\partial^{2} \in_{jm}}{\partial a^{i} \partial a^{n}} - \frac{\partial^{2} \in_{in}}{\partial a^{m} \partial a^{j}} - 2 \in_{kl} \left(\Gamma_{ij}^{k} \Gamma_{mn}^{l} - \Gamma_{in}^{k} \Gamma_{jm}^{l} \right) + 2\Gamma_{mn}^{k} \in_{ijk} + 2\Gamma_{ij}^{k} \in_{mnk} - 2\Gamma_{mj}^{k} \in_{ink} - 2\Gamma_{in}^{k} \in_{mjk} = 0,$$
(12.49)

де

 $\epsilon_{mnk} = \frac{1}{2} \left(\frac{\partial \epsilon_{nk}}{\partial a^m} + \frac{\partial \epsilon_{km}}{\partial a^n} - \frac{\partial \epsilon_{mn}}{\partial a^k} \right); \quad imnj: \ 1212, \ 1313, \ 2323, \ 1213, \ 2123, \ 3132. \ (12.50)$

Якщо визначити деформацію в координатах $\vec{a} = \vec{a}(\vec{x},t)$, то отримаємо міру деформації Альмансі-Гамеля (тут не викладаємо).

При малих деформаціях та поворотах вирази (12.38-б) лінеаризують:

$$\epsilon_{ij} \approx \varepsilon_{ij} = 0.5(\nabla_i u_j + \nabla_j u_i).$$
(12.51)

Оскільки в ДСК всі символи Крістофеля дорівнюють нулю, то з (12.38-б) та (12.51):

$$(\epsilon_{ij})_{\mathcal{ACK}} = \frac{1}{2} \left(\frac{\partial u_j}{\partial a^i} + \frac{\partial u_i}{\partial a^j} + \frac{\partial u_m}{\partial a^i} \frac{\partial u_m}{\partial a^j} \right); \quad (\epsilon_{ij})_{\mathcal{ACK}} \approx (\varepsilon_{ij})_{\mathcal{ACK}} = \frac{1}{2} \left(\frac{\partial u_j}{\partial a^i} + \frac{\partial u_i}{\partial a^j} \right).$$
(12.52)

Другу формулу з (12.52) зазвичай називають формулою для *тензора деформацій Коші*.

У ЦСК ненульовими є лише $\Gamma_{22}^1 = -\rho$; $\Gamma_{21}^2 = 1/\rho$. З урахуванням (12.3) ... (12.8) з (12.51) можна отримати, що в локальній ФСК при глобальній ЦСК:

$$\vec{\varepsilon}_{11} = \frac{\partial \vec{u}_1}{\partial \rho}; \quad \vec{\varepsilon}_{22} = \frac{1}{\rho} \frac{\partial \vec{u}_2}{\partial \vartheta} + \frac{\vec{u}_1}{\rho}; \quad \vec{\varepsilon}_{33} = \frac{\partial \vec{u}_3}{\partial z}; \quad \vec{\varepsilon}_{31} = \frac{1}{2} \left(\frac{\partial \vec{u}_1}{\partial z} + \frac{\partial \vec{u}_3}{\partial \rho} \right);$$

$$\vec{\varepsilon}_{12} = \frac{1}{2} \left(\frac{\partial \vec{u}_2}{\partial \rho} + \frac{1}{\rho} \frac{\partial \vec{u}_1}{\partial \vartheta} - \frac{\vec{u}_2}{\rho} \right); \quad \vec{\varepsilon}_{23} = \frac{1}{2} \left(\frac{1}{\rho} \frac{\partial \vec{u}_3}{\partial \vartheta} + \frac{\partial \vec{u}_2}{\partial z} \right).$$
(12.53)

При збереженні в (12.38-б) нелінійних членів явні вирази для ЦСК стають значно складнішими, тому є сенс користуватися виразом (12.38-а).

Приклад 12.1. Знайдемо $(\bar{\varepsilon}_{12})_{llCK}$ для малих деформацій, тобто компоненту тензора деформацій у "фізичній" системі координат. Згідно з (12.51) та (12.46) $(\varepsilon_{12})_{llCK} = \frac{1}{2} (\nabla_1 u_2 + \nabla_2 u_1) = \frac{1}{2} (\frac{\partial u_2}{\partial a^1} - u_k \Gamma_{12}^k + \frac{\partial u_1}{\partial a^2} - u_k \Gamma_{21}^k)$. Оскільки у ЦСК ненульові тільки $\Gamma_{12}^1 = -\rho$ та $\Gamma_{12}^2 = \Gamma_{21}^2 = 1/\rho$, то $u_k \Gamma_{12}^k = u_k \Gamma_{21}^k = u_2 \Gamma_{21}^2 = u_2 / \rho$. Відповідно до (12.8), (12.3) та (12.5) у ЦСК $\breve{u}_1 = u^1 \sqrt{g_{11}} = (u_1 g^{11}) \sqrt{g_{11}} = u_1$, а $\breve{u}_2 = u^2 \sqrt{g_{22}} = (u_2 g^{22}) \sqrt{g_{22}} = (u_2 / \rho^2) \rho = u_2 / \rho$, тобто $u_1 = \breve{u}_1$ та $u_2 = \breve{u}_2 \rho$. З урахуванням того, що $a^1 = \rho$, $a^2 = \vartheta$, маємо: $(\varepsilon_{12})_{llCK} = \frac{1}{2} (\frac{\partial (\breve{u}_2 \rho)}{\partial \rho} - 2 \frac{\breve{u}_2 \rho}{\rho} + \frac{\partial \breve{u}_1}{\partial \vartheta}) = \frac{1}{2} (\frac{\partial \breve{u}_2}{\partial \rho} \rho - \breve{U}_2 + \frac{\partial \breve{u}_1}{\partial \vartheta})$. Згідно з (12.11) та (12.5) $(\breve{\varepsilon}_{12})_{llCK} = \varepsilon_{nm} g^{1m} g^{2n} \sqrt{g_{11} g_{22}} = (\varepsilon_{12})_{llCK} g^{11} g^{22} \rho = (\varepsilon_{12})_{llCK} (1/\rho^2) \rho = (\varepsilon_{12})_{llCK} / \rho$. Тобто отримали формулу, яка фігурує у (12.53): при глобальній ЦСК у локальній ФСК.

Далі деформація буде розглядатися щодо початкової конфігурації, тому будемо застосовувати більш звичний запис координат: замість a^i будемо писати x^i , якщо це не є принциповим.

12.4. Рівняння балансу. Аксіоми Нолла

12.4.1. Рівняння балансу

Рівняннями балансу називають загальний закон збереження і його окремі випадки: закони збереження маси, імпульсу, моменту імпульсу та інші.

Загальний закон збереження виражає деякий баланс у зв'язній підобласті Ω_1 з границею S_1 , що є частиною загальної області Ω . Може бути записаний як

$$\frac{d}{d\tau} \int_{\Omega_1} A_i d\Omega + \int_{S_1} B_i dS = \int_{\Omega_1} W_i d\Omega , \qquad (12.54)$$

де $d/d\tau$ – оператор повної похідної; A, B та W – відповідні величини (скаляри, вектори або тензори; в даному випадку – компоненти векторних величин). Їхня сутність наступна: $A = A(\vec{x}, \tau)$ – об'ємна густина величини, що розглядається (енергія, маса, імпульс та ін.); $B = B(\vec{x}, \tau, \vec{v})$ – швидкість густини потоку через границю області S_1 ; $W = W(\vec{x}, \tau)$ – приріст об'ємної густини за одиницю часу, зумовлений припливом ззовні (задається); \vec{x}, τ – Ейлерові координати; \vec{v} – вектор зовнішньої нормалі до границі області S_1 ; індекс *i* вказує напрямок (для скалярних величин його немає).

Суть закону наступна: кількість деякої субстанції W_i , що надходить у підобласть Ω_1 , витрачається на:

а) її поповнення в самій підобласті $\Omega_1(A_i)$;

б) поповнення її втрат через поверхню (B_i).

Нагадаємо, що повні (субстанціональні) похідні від скалярної функції $f(\vec{x},\tau)$ і від вектора \vec{u} з компонентами u^i будуть мати вигляд відповідно

$$\frac{df(\vec{x},\tau)}{d\tau} = \frac{\partial f(\vec{x},\tau)}{\partial \tau} + \frac{\partial f(\vec{x},\tau)}{\partial x^{i}} \frac{dx^{i}}{d\tau} = \frac{\partial f(\vec{x},\tau)}{\partial \tau} + \vec{V} \cdot grad(f(\vec{x},\tau));$$

$$\frac{d\vec{u}(\vec{x},\tau)}{d\tau} = \frac{\partial \vec{u}(\vec{x},\tau)}{\partial \tau} + \frac{\partial \vec{u}(\vec{x},\tau)}{\partial x^{i}} \frac{dx^{i}}{d\tau} = \frac{\partial \vec{u}(\vec{x},\tau)}{\partial \tau} + \vec{V} \cdot \nabla \vec{u}.$$
(12.55)

Наслідки із закону збереження.

Наслідок 1. Закон поверхневих взаємодій: величина *B* є непарною функцією аргументу \vec{v} , тобто $B(\vec{x}, \tau, \vec{v}) = B(\vec{x}, \tau, -\vec{v})$. Інакше кажучи, потік субстанції по обидві сторони поверхні є однаковим за величиною.

Наслідок 2. Спроєктуємо B_i на напрямки базисних векторів \vec{e}_j . Введемо позначення: $B_i(\vec{x},\tau,\vec{v}) = b_i^j(\vec{x},\tau) \cdot v_j = \nabla_j b_i^j(\vec{x},\tau)$. Тоді функції (скалярні, векторні або тензорні) A, B, W в усіх точках області Ω , де вони мають безперервні похідні, задовольняють системі диференційних рівнянь у частинних похідних

$$\frac{\partial A_i}{\partial \tau} + \nabla_j (A_i V^j + b_i^j) = W_i$$
(12.56)

(рівняння в частинних похідних закону збереження за ейлеровими змінними).

Природною граничною умовою на поверхні (границі) S в точці, що визначається \vec{x} , називається рівність

$$\nabla_{j} b_{i}^{j}(\vec{\hat{x}},\tau) = b_{i}^{j}(\vec{\hat{x}},\tau) \cdot v_{j} = \hat{B}_{i}(\vec{\hat{x}},\tau),$$
 (12.57)

де $\hat{B}_i(\vec{x},\tau)$ – задана величина.

12.4.2. Окремі випадки рівняння балансу

12.4.2.1. Закон збереження маси

Маса речовини в області Ω визначається як

$$m = \int_{\Omega} \overline{\rho}(\vec{\hat{x}}, \tau) d\Omega, \qquad (12.58)$$

де $\overline{\rho}(\vec{x},\tau)$ – густина речовини. В законі (12.54) вона відповідає величині $A(\vec{x},\tau)$. Приймемо B = W = 0 (маса в ейлеревому об'ємі є незмінною). Отримаємо закон збереження маси, що має вигляд

$$\frac{d}{d\tau} \int_{\Omega} \overline{\rho}(\vec{\hat{x}}, \tau) d\Omega = 0.$$
(12.59)

Якщо використати рівняння збереження в частинних похідних за змінними Ейлера (12.56), то з (12.59) отримаємо рівняння нерозривності

$$\frac{\partial \overline{\rho}(\widehat{x},\tau)}{\partial \tau} + \nabla_j \left(\overline{\rho}(\overline{\widehat{x}},\tau) \cdot V^j \right) = 0.$$
(12.60)

Використовуючи перший з виразів (12.55), перепишемо (12.60) у вигляді

$$\frac{d\overline{\rho}(\bar{\hat{x}},\tau)}{d\tau} + \overline{\rho}(\bar{\hat{x}},\tau)div\vec{V} = 0 \quad \text{afo} \ \frac{d}{d\tau}(\ln(\overline{\rho}(\bar{\hat{x}},\ \tau))) = -div\vec{V}, \qquad (12.61)$$

де $div\vec{V}$ – швидкість об'ємного деформування.

При використанні лагранжевих змінних з інтегральної умови

$$m = \int_{\Omega(t)} \overline{\rho}(\vec{r}, t) d\Omega = \int_{\Omega(t_0)} \overline{\rho}(\vec{a}, t_0) d\Omega = m_0$$
(12.62)

і відомого співвідношення

$$d\Omega = \det \left| \frac{\partial x^i}{\partial a^j} \right| d\Omega_0$$
(12.63)

випливає закон збереження маси в локальній формі

$$\frac{\overline{\rho}(\vec{a},t_0)}{\overline{\rho}(\vec{r},t)} = \frac{\overline{\rho}_0}{\overline{\rho}} = \det\left|\frac{\partial x^i}{\partial a^j}\right| = \det\left|\delta_j^i + \frac{\partial u^i}{\partial a^j}\right| = \det\left|X_{ij}\right| = \sqrt{G} = J.$$
(12.64)

12.4.2.2. Рівняння руху

Закон збереження імпульсу: швидкість зміни кількості руху тіла об'єму Ω дорівнює імпульсу прикладених до нього сил. У загальному законі збереження (12.56) при цьому: $A^i = \overline{\rho}V^i$ – кількість руху в *i*-му напрямку; $B^i = -P^i = -\sigma^{ij}v_j$ – компоненти вектора поверхневих сил; $W^i = O^i = \overline{\rho}F^i$ – компоненти вектора об'ємних сил. Враховуючи (12.55) і проводячи лінеаризацію виразів, отримуємо в ейлерових змінних

$$\nabla_{j}(\sigma^{ij}) + O^{i} = \overline{\rho}_{0} \frac{\partial^{2} u^{i}}{\partial \tau^{2}}, \qquad (12.65)$$

де σ^{ii} – компоненти симетричного тензора напружень. В лагранжевих координатах лінеаризований вираз закону збереження імпульсу має вигляд:

$$(\nabla_{j})_{0}(\sigma^{ij}) + \overline{\rho}_{0}F^{i} = \overline{\rho}_{0}\frac{\partial^{2}u^{i}}{\partial t^{2}}, \qquad (12.66)$$

де оператор $(\nabla_i)_0$ містить похідні відносно початкового стану.

Інша назва виразів (12.65), (12.66) – рівняння руху.

У теорії коливань та динамічного навантаження в рівняння руху може вводитися дисипативний член. Це макропредставник процесів тертя і поглинання енергії у мікрооб'ємах. Зазвичай приймають, що цей член є пропорційним швидкості руху частки тіла. Рівняння (12.66) після введення до нього дисипативного члена ($A^i = \overline{\rho}V^i + \alpha u^i$) приймає вигляд:

$$(\nabla_{j})_{0}(\sigma^{ij}) + O^{i} = \overline{\rho}_{0} \frac{\partial^{2} u^{i}}{\partial t^{2}} + \alpha \frac{\partial u^{i}}{\partial t}, \qquad (12.67)$$

де α – коефіцієнт дисипації.

Для отримання рівняння руху у Лагранжевих координатах часто використовують дещо інший шлях.

Кількість руху деякого поточного (здеформованого) об'єму Ω_* , кожна точка якого має швидкість \vec{V} та густину матеріалу $\bar{\rho}$, визначається як

$$\Im = \int_{\Omega_*} \vec{V} \, \vec{\rho} d\Omega \,. \tag{12.68}$$

Рівнянням кількості руху (другим законом Ньютона) називають рівняння

$$d\Im/dt = \vec{R}, \qquad (12.69)$$

де результуюча всіх сил, що діють на об'єм

$$\vec{R} = \int_{\Omega_*} \vec{O} d\Omega + \int_{S_*} \vec{P} dS , \qquad (12.70)$$

 $\vec{O} = O^{j}\vec{E}_{j}$ та $\vec{P} = P^{j}\vec{E}_{j}$ – відповідно щільність об'ємних (масових) та поверхневих сил; а вектори \vec{E}_{j} , як і раніше, визначають здеформований базис.

З поверхневими силами пов'язують компоненти тензора напружень, які діють у тій же точці поверхні (природні граничні умови):

$$\vec{P} = \boldsymbol{\sigma} \cdot \vec{v} \Big|_{S^*}, \qquad (12.71)$$

де \vec{v} – зовнішня нормаль до поверхні тіла в актуальній точці. Відмітимо, що скалярний добуток $\mathcal{O}\vec{v}$, де σ є тензором, а \vec{v} – вектором, має результатом вектор.

Застосуємо формулу Остроградського-Гаусса:

$$\int_{S^*} \boldsymbol{\sigma} \cdot \vec{\boldsymbol{\nu}} \, dS = \int_{\Omega^*} \vec{\nabla} \cdot \boldsymbol{\sigma} \, d\Omega \,, \tag{12.72}$$

де $\vec{\nabla} \cdot \boldsymbol{\sigma} = div(\boldsymbol{\sigma})$ є дивергенцією тензора другого рангу $\boldsymbol{\sigma}$, причому вектороператор коваріантного диференціювання $\vec{\nabla} = \vec{e}^i \nabla_i$.

Вважаємо, що всі фізичні величини, що розглядаються, у тілі є безперервними, а також виконується закон збереження маси, тобто $d(\bar{\rho}d\Omega)/dt = 0$. Тоді з використанням (12.68) для лівої частини (12.69) можна записати, що:

$$d\mathfrak{J}/dt = \int_{\Omega_*} \frac{dV}{dt} \,\overline{\rho} d\Omega = \int_{\Omega_*} \vec{V} \,\overline{\rho} d\Omega = \int_{\Omega_*} \vec{u} \,\overline{\rho} d\Omega \,, \qquad (12.73)$$

де $\vec{V} = d\vec{V} / dt = d^2\vec{u} / dt^2 = \vec{u}$ – вектор прискорення. Тепер рівняння (12.69):

$$\int_{\Omega_*} \overline{\rho} \vec{\ddot{u}} \, d\Omega = \int_{\Omega_*} \vec{O} \, d\Omega + \int_{\Omega_*} \vec{\nabla} \cdot \boldsymbol{\sigma} \, d\Omega \tag{12.74}$$

або у "зібраному" вигляді:

$$\int_{\Omega_*} (\vec{\nabla} \cdot \boldsymbol{\sigma} + \vec{O} - \vec{\rho} \vec{u}) d\Omega = 0.$$
(12.75)

Оскільки інтеграл (12.75) повинен дорівнювати нулю для будь-якого об'єму Ω_{*}, то вираз під інтегралом теж повинен дорівнювати нулю (основна лема механіки), тобто рівняння руху елементарного об'єму тіла має вигляд:

$$\vec{\nabla} \cdot \boldsymbol{\sigma} + \vec{O} = \overline{\rho} \vec{\ddot{u}} . \tag{12.76}$$

Це рівняння руху таке ж саме, як і (12.66), тільки у поточній конфігурації та векторній формі запису.

Якщо прискорення відсутнє або ним можна знехтувати, то рівняння руху (12.76) перетворюється у рівняння рівноваги елементарного об'єму тіла:

$$\vec{\nabla} \cdot \boldsymbol{\sigma} + \vec{O} = 0. \tag{12.77}$$

12.4.3. Аксіоми Нолла

Для тіла, що деформується, сформульовано наступні основні принципи:

• *принцип детермінізму* (причинності): тензор напружень у момент часу *t* повністю визначається історією руху тіла до цього моменту часу. Принцип може не виконуватися, якщо у середовищі є так звані "внутрішні зв'язки", коли при будь-яких зовнішніх впливах можливі не всі деформації;

• *принцип локальної дії* (просторової локалізації): на напружений стан в точці впливають процеси, що відбуваються лише в нескінченно близьких точках тіла;

• *принцип матеріальної незалежності від системи відліку*: тензор напружень є інваріантним відносно до будь-яких безперервних і таких, що безперервно диференціюються, перетворень системи координат. Основний наслідок: в залежності напружень від історії руху тіла час не може входити явно.

Сукупність цих трьох принципів називають *аксіомами Нолла*. Вони дозволяють, виходячи із загальних позицій, формулювати визначальні рівняння для середовищ, зокрема, вводити поняття середовищ з пам'яттю, без пам'яті та з "обмеженою" пам'яттю; однорідних і неоднорідних, ізотропних і анізотропних середовищ; пружних, пластичних, пружно-пластичних, в'язко-пружних, в'язкопластичних і в'язко-пружно-пластичних середовищ, а також визначати конкретні залежності, загальну кількість характеристик середовища та їх основні властивості тощо.

12.5. Принцип можливих переміщень

Візьмемо рівняння рівноваги (12.77) та у скалярний спосіб помножимо його на вектор варіації переміщень $\delta \vec{u}$:

$$(\vec{\nabla} \cdot \boldsymbol{\sigma} + \vec{O}) \cdot \delta \vec{u} = 0. \tag{12.78}$$

Після інтегрування у всьому об'ємі тіла:

$$\int_{\Omega} (\vec{\nabla} \cdot \boldsymbol{\sigma} + \vec{O}) \cdot \delta \vec{u} \, d\Omega = 0 \,. \tag{12.79}$$

За правилами диференціювання $\vec{\nabla} \cdot (\boldsymbol{\sigma} \cdot \delta \vec{u}) = (\vec{\nabla} \cdot \boldsymbol{\sigma}) \cdot \delta \vec{u} + \boldsymbol{\sigma} \cdot (\vec{\nabla} \cdot \delta \vec{u})$, тому

$$\vec{\nabla} \cdot \boldsymbol{\sigma}) \cdot \delta \vec{u} = \vec{\nabla} \cdot (\boldsymbol{\sigma} \cdot \delta \vec{u}) - \boldsymbol{\sigma} \cdot (\vec{\nabla} \cdot \delta \vec{u}).$$
(12.80)

Після підставлення (12.80) у (12.79) отримуємо вираз:

$$\int_{\Omega} \vec{\nabla} \cdot (\boldsymbol{\sigma} \cdot \delta \vec{u}) d\Omega - \int_{\Omega} \boldsymbol{\sigma} \cdot (\vec{\nabla} \cdot \delta \vec{u}) d\Omega + \int_{\Omega} \vec{O} \cdot \delta \vec{u} d\Omega = 0.$$
(12.81)

Відповідно до формули Остроградського-Гаусса перший (об'ємний) інтеграл із (12.81) дорівнює (точне співвідношення) інтегралу по поверхні цього об'єму:

$$\int_{\Omega} \vec{\nabla} \cdot (\boldsymbol{\sigma} \cdot \delta \vec{u}) d\Omega = \int_{S} (\boldsymbol{\sigma} \cdot \vec{v}) \cdot \delta \vec{u} \, dS \,. \tag{12.82}$$

Згідно з природними граничними умовами (12.71), а саме $\vec{P} = \boldsymbol{\sigma} \cdot \vec{v}|_{S_p}$, на поверхні S_p задане силове навантаження \vec{P} . Інша частина поверхні тіла $S \setminus S_p$ є вільною від навантажень, тому $(\boldsymbol{\sigma} \cdot \vec{v})|_{S \setminus S_p} = 0$ та $\int_{S \setminus S_p} (\boldsymbol{\sigma} \cdot \vec{v}) \cdot \delta \vec{u} \, dS = 0$. Тому співвід-

ношення (12.82) можна змінити на

$$\int_{\Omega} \vec{\nabla} \cdot (\boldsymbol{\sigma} \cdot \delta \vec{u}) d\Omega = \int_{S_p} \vec{P} \cdot \delta \vec{u} \, dS \,. \tag{12.83}$$

З використанням (12.83) вираз (12.81) запишеться у вигляді:

$$\int_{S_p} \vec{P} \cdot \delta \vec{u} \, dS - \int_{\Omega} \boldsymbol{\sigma} \cdot (\vec{\nabla} \cdot \delta \vec{u}) \, d\Omega + \int_{\Omega} \vec{O} \cdot \delta \vec{u} \, d\Omega = 0 \,.$$
(12.84)

Оскільки тензор напружень σ є симетричним, то можемо записати тотожність:

$$\boldsymbol{\sigma} \cdot (\vec{\nabla} \cdot \delta \vec{u}) \equiv \boldsymbol{\sigma} \cdot 0.5 (\vec{\nabla} \cdot \delta \vec{u} + \vec{\nabla} \cdot \delta \vec{u}) = \boldsymbol{\sigma} \cdot \delta \left[0.5 (\vec{\nabla} \cdot \vec{u} + \vec{\nabla} \cdot \vec{u}) \right] = \boldsymbol{\sigma} : \delta \boldsymbol{\varepsilon} , \qquad (12.85)$$

де введено позначення (через вектори та компоненти):

$$\delta \boldsymbol{\varepsilon} = 0.5(\vec{\nabla} \cdot \delta \vec{u} + \vec{\nabla} \cdot \delta \vec{u}); \quad \delta \boldsymbol{\varepsilon}_{ij} = 0.5(\nabla_i \delta u_j + \nabla_j \delta u_i). \tag{12.86}$$

Співвідношення (12.86) точно співпадають з варіаціями від виразів (12.51) для нескінченно малих деформацій ε_{ij} . Саме тому тут застосовано таке ж позначення, як і в (12.51).

Остаточно замість (12.84) маємо (із заміною знаків на протилежні):

$$\Psi(\delta) = \int_{\Omega} \boldsymbol{\sigma} : \delta \boldsymbol{\varepsilon} \ d\Omega - \int_{\Omega} \vec{O} \cdot \delta \vec{u} \ d\Omega - \int_{S_p} \vec{P} \cdot \delta \vec{u} \ dS = 0.$$
(12.87)

Позначимо:

$$\Pi(\delta) = \int_{\Omega} \boldsymbol{\sigma} : \delta \boldsymbol{\varepsilon} \ d\Omega; \quad \mathbf{A}(\delta) = \int_{\Omega} \vec{O} \cdot \delta \vec{u} \ d\Omega + \int_{S_p} \vec{P} \cdot \delta \vec{u} \ dS.$$
(12.88)

Тут $\prod(\delta)$ – робота напружень на варіаціях деформацій; $A(\delta)$ – робота масових і поверхневих зусиль на варіаціях переміщень (але не варіації робіт).

Тоді (12.87) можна записати як

Частина III. Розділ 12	175

$$\Psi(\delta) = \Pi(\delta) - A(\delta) = 0 \quad \text{afo} \quad \Pi(\delta) = A(\delta). \tag{12.89}$$

Вираз $\Psi(\delta) = \Pi(\delta) - A(\delta)$ визначає *повну віртуальну енергію* тіла ($\Psi = \Pi - A$ є *повною енергією* тіла). Зауважимо, що при варіюванні переміщень припускається незмінність напружень та зусиль.

Вирази (12.87) та (12.89) відображають принцип можливих переміщень. Його зачатки є у "Фізиці" Аристотеля (384-322 до Р.Х.), а його розвиток пов'язують з Г. Галілеєм (1564-1642), І. Бернуллі (1667-1748), Г. Лейбніцем (1646-1716), Л. Ейлером (1707-1783), Ж.Л. Лагранжем (1736-1813), У. Гамільтоном (1805-1865), іншими. На честь Ж.Л. Лагранжа, який ввів поняття варіації та створив основи варіаційного числення, рівняння (12.87) чи (12.89) часто називають варіаційним функціоналом Лагранжа.

Контрольні питання до підрозділу 12.1

1. Які координатні системи застосовують в кінематиці деформівного тіла?

2. Якій тензор характеризує метрику простору?

Контрольні питання до підрозділу 12.2

1. Для чого вводиться оператор однозначного відображення й які властивості він має?

2. Надайте характеристику формулюванням Лагранжа та Ейлера.

Контрольні питання до підрозділу 12.3

1. Запишіть вирази тензора деформацій Гріна Лагранжа через метричні тензори та через компоненти градієнтів переміщень.

2. Що характеризують символи Крістофеля другого роду?

3. Як змінюються вирази тензора деформацій Гріна Лагранжа у випадку нескінченно малих деформацій?

Контрольні питання до підрозділу 12.4

1. Які компоненти та наслідки має закон збереження?

- 2. Запишіть та поясніть закон збереження маси.
- 3. Запишіть та поясніть закон збереження імпульсу (рівняння руху).

4. Назвіть аксіоми Нолла. Яке значення вони мають в механіці деформівного твердого тіла?

Контрольні питання до підрозділу 12.5

1. Які рівняння механіки деформівного тіла враховує принцип можливих переміщень?

2. Яка тотожність є слідством з симетричності тензора напружень Коші?

3. Яку енергію визначає принцип можливих переміщень?

Розділ 13

ПОСТАНОВКИ КРАЙОВИХ ЗАДАЧ ДИНАМІКИ І МІЦНОСТІ МАШИН

При вирішенні проблем динаміки і міцності машин можуть розглядатися наступні основні крайові задачі:

1. Про тепловий стан (ТС) тіла:

- стаціонарна теплопровідність;
- нестаціонарна теплопровідність.

Результат: температурне поле, кінетика його зміни – використовується як початкові дані для більшості з перерахованих нижче задач, оскільки температура дуже впливає на напружено-деформований стан (НДС) всього тіла.

2. Про напружено-деформований (квазі) стаціонарний стан тіла:

- задача (термо) пружності;
- (термо) пружнопластична задача;
- задача (термо) повзучості.

3. Про напружено-деформований (квазі) динамічний стан тіла:

- задача (термо) пружності;
- (термо) пружнопластична задача;
- задача (термо) повзучості.
- 4. Динаміка тіла (частотний аналіз);
- 5. Контактні крайові задачі про визначення ТС і (або) НДС тіл;
- 6. Про втрату стійкості тіла.

При цьому об'єм тіл і їхні поверхні можуть бути як незмінними, так і змінюватися. Результати: поля температур, переміщень, деформацій, напружень, їхніх швидкостей, власні частоти і форми коливань, форми втрати стійкості – використовуються для прогнозу міцності і деформаційних характеристик тіла (виробу, елемента конструкції, агрегату в цілому).

Постановкою крайових задач займаються відповідні розділи механіки суцільних середовищ. У цьому Розділі сформулюємо постановки задач пунктів 1 ... 4, а для пунктів 5 та 6 – у відповідних Розділах.

Розглядаємо тіло об'ємом Ω з поверхнею S. Застосовуємо, якщо інше не оговорено, Лагранжеві координати. Фізичні величини: напруження, деформації, зусилля, переміщення розглядаємо в декартовій системі координат (ДСК). Матеріал будемо вважати ізотропним, а деформації – нескінченно малими, якщо не оговорено інше.

13.1. Постановка незв'язаної крайової задачі теплопровідності

Використаємо рівняння закону збереження в частинних похідних за змінними Ейлера (12.56):

$$\frac{\partial A}{\partial \tau} + \nabla_i (AV^i + B_i) = W, \qquad (13.1)$$

де: $A = A(\vec{x}, \tau) - \text{об'ємна}$ густина величини, що розглядається, в даному випадку $A(\vec{x}, \tau) = c_p(\vec{x}, \tau) \cdot \overline{\rho}(\vec{x}, \tau) \cdot \theta(\vec{x}, \tau) - \text{об'ємна}$ густина теплоти $(c_p(\vec{x}, \tau) - \text{питомi}$ теплоємність (Дж/(кг · град)), $\overline{\rho}(\vec{x}, \tau) -$ густина речовини (кг/м³), $\theta(\vec{x}, \tau) -$ температура); V^i – швидкість переміщення речовини; $B = B(\vec{x}, \tau, \vec{v})$ – швидкість густини потоку через границю області S, в даному випадку потоку тепла $B_j(\vec{x}, \tau, \vec{v}) = q_j(\vec{x}, \tau)$; $W = W(\vec{x}, \tau)$ – приріст об'ємної густини за одиницю часу, викликаний припливом ззовні (задається), в даному випадку $W(\vec{x}, \tau) = \hat{\omega}(\vec{x}, \tau)$ – потужність внутрішнього джерела (або стоку) тепла; \vec{x}, τ – Ейлерові координати. Просторовий оператор Лапласа ∇_i відповідає формулам (12.46) при n = 1.

Перейдемо до Лагранжевих координат, тобто у просторово-часовий простір \vec{x} (координати точки тіла) та t (час), а також проведемо лінеаризацію задачі. Також використаємо закон теплопровідності Фур'є для ізотропної теплопровідності: тепловий потік у напрямку осі є пропорційним відповідному градієнту температури та має напрямок, протилежний градієнту:

$$q_i = -\lambda \cdot \nabla_i \theta, \qquad (13.2-a)$$

де $\lambda = \lambda(\vec{x}, \theta)$ – коефіцієнт теплопровідності матеріалу (Вт/(м·град)).

Отримаємо, що в кожній елементарній одиниці об'єму середовища баланс потоку теплоти визначається співвідношенням

$$c_{P}\overline{\rho}\frac{\partial\theta}{\partial t} + c_{P}\overline{\rho}\cdot(\nabla_{i}\theta)V^{i} - \nabla_{i}(\lambda\nabla_{i}\theta) = \widehat{\omega}$$
(13.3-a)

за початковою умовою

$$\theta(\vec{x}, t_0) = \hat{\theta}_0(\vec{x}). \tag{13.4}$$

У (13.3-а) $\hat{\omega} = \hat{\omega}(\vec{x}, t)$ – потужність внутрішнього джерела (або стоку) тепла. На поверхні тіла (об'єму, рідини) задають граничні умови (ГУ):

• за температурою поверхні (її частини S_{θ}) – ГУ 1-го роду

$$\left. \mathcal{P}(\vec{x},t) \right|_{S_a} = \hat{\theta}(\vec{x},t); \qquad (13.5)$$

• за тепловим потоком (у напрямку зовнішньої нормалі \vec{v} до поверхні) — природні ГУ

$$\left. \lambda \frac{\partial \theta}{\partial \nu} \right|_{S_G} = \widehat{q} \left|_{S_q} + \widetilde{q} \right|_{S_{\alpha}} + \left. \breve{q} \right|_{S_{\beta}}, \tag{13.6}$$

де $\hat{q} = q(\vec{x}, t)$ – відомий потік тепла через границю S_q . Для конвекційної складової теплового потоку через поверхню використовують лінійну залежність

$$\tilde{q}\big|_{S_{\alpha}} = -\alpha \cdot (\theta - \hat{\theta}_{\infty})\big|_{S_{\alpha}}, \qquad (13.7-a)$$

або одну з нелінійних залежностей

$$\tilde{q}\big|_{S_{\alpha}} = -\alpha \cdot (\theta^{\mu} - \hat{\theta}_{\infty}^{\mu})\big|_{S_{\alpha}} = -\alpha \cdot \tilde{Q}\big|_{S_{\alpha}}, \quad \text{de} \quad \tilde{Q} = (\theta^{\mu} - \hat{\theta}_{\infty}^{\mu}); \quad (13.7-6)$$

$$\tilde{q}\big|_{S_{\alpha}} = -\alpha \cdot (\theta - \hat{\theta}_{\infty})^{\eta + 1}\big|_{S_{\alpha}} = -\alpha \cdot \tilde{\tilde{Q}}\big|_{S_{\alpha}}, \quad \text{de} \quad \tilde{\tilde{Q}} = (\theta - \hat{\theta}_{\infty})^{\eta + 1}; \quad (13.7-B)$$

а для променевої складової теплового потоку – вираз

$$\left. \vec{q} \right|_{S_{\beta}} = -\beta f \left[e(\theta + \theta_{abs})^4 - a \hat{\theta}_a^4 \right] \right|_{S_{\beta}} = -\beta f \left[eW - a \hat{\theta}_a^4 \right] \right|_{S_{\beta}}, \text{ de } W = (\theta + \theta_{abs})^4.$$
(13.8)

У формулах (13.7) та (13.8) позначено: α, β – коефіцієнт конвекційної тепловіддачі (Вт/(м²·град)) і постійна Стефана-Больцмана (Вт/(м²·град⁴)) відповідно; поверхня з ГУ $S_G = S_q \cup S_\alpha \cup S_\beta \cup S_\theta$; $\hat{\theta}_{\infty} = \theta_{\infty}(\vec{x},t)$ – температура середовища біля поверхні S_{α} з конвекційним теплообміном; θ_{abs} – зсув розрахункової температури θ від абсолютного нуля; $\hat{\theta}_a = \hat{\theta}_a(\vec{x},t)$ – абсолютна температура тіла, з яким тіло (об'єм, рідина), що розглядається, має променевий теплообміни через поверхню S_{β} ; $0 \le \mu, \eta \le 1$ – експериментально встановлені степеневі показники; $0 \le e \le 1$ та $0 \le a \le 1$ – функції (коефіцієнти), відповідно, спроможності поверхні тіла до поглинання та випромінювання поверхнею джерела. За допомогою *а* можна моделювати такі явища, як часткове екранування джерела (наприклад, закриття сонця хмарами). Функція (коефіцієнт) *е* може враховувати, наприклад, колір поверхні, що поглинає випромінювання. Значок "^" над змінною вказує на те, що її величина відома та задається.

Фактор освітленості f точки m на поверхні S_{β} тіла променевим джерелом з точки n на поверхні випромінювання S_a обчислюється за формулою

$$(f)_{m \leftarrow n} = \frac{1}{(S_{\beta})_m} \int_{(S_{\beta})_m} \int_{(S_a)_n} \frac{\cos \vartheta_m \cos \vartheta_n}{\pi r^2} (dS_a)_n (dS_{\beta})_m, \qquad (13.9)$$

де r – відстань між двома точками n та m на елементах поверхонь $(dS_a)_n$ та $(dS_\beta)_m$; \mathcal{G}_m та \mathcal{G}_n – кути між лінією, що з'єднує ці точки n та m, та нормалями до цих елементів поверхонь в точках n та m.

Рівняння стаціонарної теплопровідності виводиться безпосередньо з (13.3-а) виключенням компонент, що залежать від часу (тобто при $\partial \theta / \partial t \equiv 0$):

$$c_{P}\overline{\rho} (\nabla_{i}\theta) V^{i} - \nabla_{i} (\lambda \nabla_{i}\theta) = \widehat{\omega}. \qquad (13.10-a)$$

Для твердих тіл характерні відносно малі швидкості V^i переміщення точок тіла, тому зазвичай нехтують конвекційним переносом тепла у тілі, тобто у рівняннях (13.3-а) та (13.10-а) вважають $c_p \overline{\rho} (\nabla_i \theta) V^i \equiv 0$, тому рівняння приймають вигляд:

$$c_{P}\overline{\rho}\frac{\partial\theta}{\partial t} - \nabla_{i}(\lambda\nabla_{i}\theta) = \widehat{\omega}; \qquad (13.3-6)$$

$$\nabla_i (\lambda \nabla_i \theta) = -\hat{\omega}. \tag{13.10-6}$$

У залежності від конкретних умов задачі в рівняннях (13.3), (13.6) і (13.10) можлива відсутність деяких складових. Зокрема, заданих теплового потоку $(S_q = 0 \text{ aбo } \hat{q}|_{S_q} = 0)$, конвекційного теплообміну $(S_{\alpha} = 0 \text{ aбo } \tilde{q}|_{S_{\alpha}} = 0)$, променевого теплообміну $(S_{\beta} = 0 \text{ aбo } \tilde{q}|_{S_{\alpha}} = 0)$, об'ємного теплового джерела ($\hat{\omega} = 0$).

Теплоємність матеріалу зазвичай є помірною функцією температури.

А коефіцієнт теплопровідності – теж помірна функція температури, яка зменшується зі збільшенням температури для твердих матеріалів і збільшується зі збільшенням температури для рідин і газів. Крім того, він може залежати від напрямку (бути тензором другого рангу λ_{ij}) в анізотропному матеріалі. Тоді закон Фур'є (13.2-а) прийме вигляд

$$q_i = -\lambda_{ij} \nabla_j \theta \,. \tag{13.2-6}$$

Взагалі, кожна характеристика матеріалу може залежати від температури. Розглянемо на прикладі коефіцієнта теплопровідності $\lambda(\theta)$. Зазвичай відомо значення $\lambda_{RM} = \lambda(\theta_{RM})$, де θ_{RM} (RM — від Reference Material) є температурою випробування матеріалу (якщо вона не вказана, то дорівнює приблизно 20°*C*). Температурну залежність можна задати у вигляді функції температури $F(\theta)$, щоб поточне значення $\lambda(\theta)$ обчислювалося як $\lambda(\theta) = \lambda_{RM} \cdot F(\theta)$, причому $F(\theta_{RM}) \equiv 1$. Якщо функція $F(\theta)$ буде задаватися у табличному (дискретному) вигляді, то для проміжних значень температури повинна застосовуватися апроксимаційна формула. У випадку лінійної апроксимації поточне значення:

$$F(\theta) \approx F(\theta_{(k)}) + \frac{\theta - \theta_{(k)}}{\theta_{(k+1)} - \theta_{(k)}} \Big[F(\theta_{(k+1)}) - F(\theta_{(k)}) \Big], \qquad (13.11)$$

де k – номер точки на графіку $F(\theta)$.

Отже, для кожної точки тіла записано системи диференційних рівнянь (13.3) ... (13.9) для нестаціонарної, та (13.10), (13.5) ... (13.9) для стаціонарної задачі теплопровідності відносно значень температур $\theta(\vec{x},t)$ або $\theta(\vec{x})$. У загальному вигляді ці системи можна уявити як операторні рівняння:

$$C\frac{\partial \vec{u}}{\partial t} + L\vec{u} = \vec{f} \quad \mathbf{B} \quad \Omega; \qquad (13.12)$$

$$L\vec{u} = \vec{f} \quad \mathbf{B} \quad \Omega, \tag{13.13}$$

з ГУ та НУ вигляду

$$K\vec{u} = \vec{g}$$
 Ha S; $\vec{u}(t_0) = \hat{\vec{u}}_0$. (13.14)

13.2. Постановка крайової задачі термопружності

Зазвичай вважається, що в початковий момент t_0 в тілі, що розглядається, переміщення $(u_i)_0 = u_i(\vec{x}, t_0)$, деформації $(\varepsilon_{ij})_0 = \varepsilon_{ij}(\vec{x}, t_0)$, напруження $(\sigma_{nn})_0 = \sigma_{nn}(\vec{x}, t_0)$ мають відомі (частіше – нульові) значення, відомо початкове поле температур $\hat{\theta}_0(\vec{x}) = \hat{\theta}(\vec{x}, t_0)$. Далі вважається, що навантаження змінюється кроками (n номер кроку); що в об'ємі тіла Ω , а також на частині його поверхні $S_G = S_U \cup S_P$ за деякий проміжок часу $\Delta t = t^{n+1} - t^n$ відбудеться зміна навантажень, тобто на момент часу t^{n+1} були прикладені: $\hat{O}_m(\vec{x}, t^{n+1}) = \bar{\rho} \cdot \hat{F}_m(\vec{x}, t^{n+1}) -$ об'ємні сили $(\vec{F}(\vec{x},t^{n+1}) - \text{вектор масової сили}), \hat{P}_m(\vec{x},t^{n+1}) - \text{поверхневі сили на } S_P; \hat{P}_m(\vec{x},t^{n+1}) -$ зосереджені сили в деяких точках; відбулися переміщення $\hat{u}_i(\vec{x},t^{n+1}) -$ на S_U , а також в Ω змінилася температура з $\hat{\theta}_0(\vec{x})$ на $\hat{\theta}(\vec{x},t^{n+1})$. Тоді для визначення в кожній точці тіла (оточуючому її однорідному просторі) величин: $u_i(\vec{x},t^{n+1}) -$ переміщень, $\varepsilon_{ij}(\vec{x},t^{n+1}) -$ деформацій, $\sigma_{mn}(\vec{x},t^{n+1}) -$ напружень, а також інших похідних від них величин, маємо наступну крайову задачу (в ДСК):

• рівняння рівноваги як окремий випадок рівняння руху:

$$\nabla_n \sigma_{mn} + \hat{O}_m = 0; \qquad (13.15)$$

• "геометричні" (12.51) (для нескінченно малих деформацій і поворотів):

$$\varepsilon_{ij} = 0.5(\nabla_i u_j + \nabla_j u_i), \qquad (13.16)$$

а також рівняння сумісності деформацій (12.49) і принцип суперпозиції нескінченно малих деформацій різної природи:

$$\varepsilon_{ij} = \varepsilon^{e}_{ij} + \varepsilon^{\theta}_{ij} + (\varepsilon_{ij})_{0}; \qquad (13.17)$$

• "фізичні" рівняння. Між напруженнями та пружними деформаціями існує однозначна функціональна залежність. Зазвичай вводять поняття функціонала пружної енергії (або пружного потенціалу) *W*, за допомогою якого закон пружності виражається як

$$\sigma_{mn} = (\partial W / \partial \varepsilon_{mn} + \partial W / \partial \varepsilon_{nm}) / 2.$$
(13.18)

Для лінійно-пружної моделі матеріалу це є закон Гука:

$$\sigma_{mn} = E_{mnij} \varepsilon_{ij}^{e} \quad \text{afo} \quad \varepsilon_{ij}^{e} = C_{ijmn} \sigma_{mn}, \qquad (13.19)$$

де $E_{mnij} = E_{mnij}(\theta)$ та $C_{ijmn} = C_{ijmn}(\theta)$ – матеріальні тензори модулів пружності та піддатливості відповідно, в яких модулі можуть залежати від координат (структурно неоднорідне тіло) та/або температури.

До "фізичних" рівнянь належать і рівняння для температурної деформації. Зміна температури пов'язана зі зміною енергії атомів, яка супроводжується зміною амплітуди їх коливань. Тому в умовах Землі *ніщо не може завадити реалізації температурних деформацій*. За це їх називаються такими, "що не стискуються" (до таких деформацій ще відносять деформації фазових перетворень у матеріалі). Оскільки частота коливань атомів дуже велика, то для процесів деформування тіл можна вважати, що температурна деформація є миттєвою, тобто відповідає *поточній* температурі (не має пам'яті).

Для обчислень $\varepsilon_{ij}^{\theta}$ зазвичай використовується формула Ж. Дюамеля (1797-1872 рр.) і Ф. Неймана (1798-1895 рр.):

$$\varepsilon_{ij}^{\theta} = \delta_{ij} \alpha_{\theta} (\theta - \theta_0). \qquad (13.20\text{-a})$$

Але для багатьох матеріалів виявилося, що в різних температурних діапазонах значення коефіцієнта температурного подовження α_{θ} є різними. Щоб не змінювати привичний запис формули для обчислень $\varepsilon_{ij}^{\theta}$, замість α_{θ} вводять $\overline{\alpha}_{\theta}(\theta)$. Згідно з рис.13.1 для волокна матеріалу $\varepsilon^{\theta} = a - b$. Якщо прийняти, що

$$tg(\gamma_1) = \alpha_{\theta}(\theta_0); \quad tg(\gamma_2) = \alpha_{\theta}(\theta),$$
 (13.21)

то $a = tg(\gamma_2) \cdot (\theta - \theta_{RM}) = \alpha_{\theta}(\theta) \cdot (\theta - \theta_{RM})$ та $b = tg(\gamma_1) \cdot (\theta_0 - \theta_{RM}) = \alpha_{\theta}(\theta_0) \cdot (\theta_0 - \theta_{RM})$. Тоді

$$\varepsilon_{ij}^{\theta} = \delta_{ij} \left[\alpha_{\theta}(\theta) \cdot (\theta - \theta_{RM}) - \alpha_{\theta}(\theta_{0}) \cdot (\theta_{0} - \theta_{RM}) \right], \qquad (13.20-6)$$

де $\theta_{\scriptscriptstyle RM}$ – базова температура випробування матеріалу; або

$$\varepsilon_{ij}^{\theta} = \delta_{ij} \overline{\alpha}_{\theta}(\theta) \cdot (\theta - \theta_0), \qquad (13.20-B)$$

де

$$\overline{\alpha}_{\theta}(\theta) = \left[\alpha_{\theta}(\theta) \cdot (\theta - \theta_{RM}) - \alpha_{\theta}(\theta_{0}) \cdot (\theta_{0} - \theta_{RM})\right] / (\theta - \theta_{0}).$$
(13.22)

Примітка 13.1. Застосування формули (13.22) при $\theta \approx \theta_0$ може викликати (див. п.1.2.1) небажано велику похибку обчислення $\bar{\alpha}_{\theta}(\theta)$, тому доцільно користуватися формулою (13.20-б), а не (13.20-в).

Поточні значення $\alpha_{\theta}(\theta)$ обчислюються як

$$\alpha_{\theta}(\theta) = \alpha_{\theta}(\theta_{RM}) \cdot F(\theta), \qquad (13.23)$$

де $F(\theta)$ є відповідною коректувальною функцією, причому $F(\theta_{RM}) \equiv 1$. Якщо в тілі є декілька матеріалів, буде потрібно й декілька таких функцій.

Якщо $\theta_{\rm RM} = \theta_0$, то $\overline{\alpha}_{\theta} = \alpha_{\theta}(\theta)$ та

$$\varepsilon_{ij}^{\theta} = \delta_{ij} \alpha_{\theta}(\theta) \cdot (\theta - \theta_0) = \delta_{ij} \alpha_{\theta}(\theta_{RM}) \cdot F(\theta) \cdot (\theta - \theta_0).$$
(13.20-г)

Якщо $\theta_{RM} = 0$, то з (13.20-б)

$$\varepsilon_{ij}^{\theta} = \delta_{ij} \left[\alpha_{\theta}(\theta) \cdot \theta - \alpha_{\theta}(\theta_{0}) \cdot \theta_{0} \right] = \delta_{ij} \alpha_{\theta}(0) \cdot \left[F(\theta) \cdot \theta - F(\theta_{0}) \cdot \theta_{0} \right].$$
(13.20-д)

Введений у такий спосіб коефіцієнт $\bar{\alpha}_{\theta}(\theta)$ називають "*січним* для температурного діапазону значенням α_{θ} – **SCTE**". Увага: формули (13.20-б) ... (13.20-д) ϵ вірними за умови, що табличні значення $\alpha_{\theta}(\theta)$ відповідають формулам (13.21). Є й інші варіанти врахування змінного α_{θ} , але цей ϵ найбільш обґрунтованим.

Якщо функції $F(\theta)$ будуть задаватися у табличному (дискретному) вигляді, то для проміжних значень



Рис.13.1 До отримання формули (13.20-б)

температури повинна застосовуватися апроксимаційна формула. У випадку лінійної апроксимації поточне значення $F(\theta)$ буде обчислюватися згідно з формулою (13.11). Інші варіанти апроксимацій можна знайти у Розділі 9.

Для ізотропного пружно-*нелінійного* матеріалу "фізичні" рівняння (13.18) зручно представити не як $\sigma_{mn} = E_{mnii}(\varepsilon^{e}_{ii}, \theta)\varepsilon^{e}_{ii}$, а у вигляді трьох законів:

• закон пружної зміни елементарного об'єму

$$\varepsilon_V^a = \sigma_V / 3k; \qquad (13.24)$$

• тензорне співвідношення (закон зміни форми елементарного об'єму)

$$e_{ij}^{a} = \left(\varphi / 2G(\theta)\right) \cdot S_{ij}; \quad \varphi = 3G(\theta) \cdot \varepsilon_{u}^{a} / \sigma_{u}; \quad (13.25)$$

• рівняння стану, що визначається експериментально та задається у вигляді функціональної залежності (функціонала):

$$\sigma_u = K(\varepsilon_u^a, \theta). \tag{13.26}$$

Позначено: $\varepsilon_{v} = \delta_{ij}\varepsilon_{ij}/3$ й $\sigma_{v} = \delta_{ij}\sigma_{ij}/3 - \text{об'ємна деформація й напруження}$ відповідно; $k = k(\theta) - \text{модуль об'ємної деформації; } e_{ij} = \varepsilon_{ij} - \delta_{ij}\varepsilon_{v}$ і $S_{ij} = \sigma_{ij} - \delta_{ij}\sigma_{v}$ - компоненти девіаторних частин тензорів деформацій і напружень відповідно; ε_{u}^{a} і σ_{u} – інтенсивності "активних" деформацій $\varepsilon_{ij}^{a} = \varepsilon_{ij} - \varepsilon_{ij}^{\theta} - (\varepsilon_{ij})_{0}$ і напружень σ_{ij} відповідно. У пружному матеріалі немає необоротних деформацій, тому "активні" деформації – пружні. З урахуванням (13.17) та (13.20) $\varepsilon_{u}^{a} = \varepsilon_{u}$. Якщо матеріал – лінійно-пружний, то (13.26) замінюється на $\sigma_{u} = 3G\varepsilon_{u}^{a}$, де $G = G(\theta)$ – модуль зсуву. Для ізотропного матеріалу $G = E/(1+2\mu)$, де $E = E(\theta)$ – модуль Юнга; $\mu = \mu(\theta)$ – коефіцієнт Пуассона.

Крім статичних, "геометричних" і "фізичних" рівнянь додатково залучаються ГУ на поверхнях S_U (1-го роду) та S_P (природні):

$$u_i \mid_{S_{ii}} = \widehat{u}_i; \tag{13.27}$$

$$\sigma_{mn} v_n \Big|_{S_P} = \widehat{P}_m, \qquad (13.28)$$

причому наявні зосереджені сили \hat{P}_m в деяких точках (загальною кількістю $N_{\bar{P}}$) враховуються теж через (13.28).

Виключивши з (13.15) і (13.28) напруження за допомогою геометричних і *лінійних* фізичних рівнянь, тобто (13.16), (13.17), (13.19) і одного з (13.20), отримаємо в кожній точці $\vec{x} \in \Omega$:

$$\nabla_{n}\left\{E_{mnij}(\theta)\left[0.5(\nabla_{i}u_{j}+\nabla_{j}u_{i})\right]\right\}=\nabla_{n}E_{mnij}(\theta)\delta_{ij}\overline{\alpha}_{\theta}\Delta\widehat{\theta}-\widehat{O}_{m};$$
(13.29)

$$\left\{ E_{mnij}(\theta) \Big[0.5(\nabla_i u_j + \nabla_j u_i) \Big] \right\} v_n \Big|_{S_p} = E_{mnij}(\theta) \delta_{ij} \overline{\alpha}_{\theta} \Delta \widehat{\theta} v_n + \widehat{P}_m + \widehat{\overline{P}}_m.$$
(13.30)

Сукупність виразів (13.29), (13.30) і (13.27) є системою диференційних рівнянь в точці тіла, записаною відносно переміщень (рівняння Дюамеля-Неймана). Очевидно, що ця система в загальному вигляді може бути записана як (13.13) з ГУ вигляду (13.14).

Для розв'язування крайової задачі часто зручніше мати її варіаційне формулювання. Не завжди це можливо, але крайові задачі для твердого тіла, що деформується, це дозволяють.

Для отримання варіаційного формулювання задачі використовують співвідношення (13.15), (13.16) і (13.28), властивості симетрії тензора напружень $\sigma_{mn} = \sigma_{nm}$ і формулу Остроградського-Гаусса. В підсумку отримують функціо-

нал (принцип можливих переміщень) відносно варіацій переміщень і зв'язаних із ними деформацій (див. підрозділ 12.5, формулу (12.87))

$$\Psi(\delta) = \int_{\Omega} \sigma_{mn} \delta \varepsilon_{mn} d\Omega - \int_{\Omega} \hat{O}_m \delta u_m d\Omega - \int_{S_p} \hat{P}_m \delta u_m dS = 0, \qquad (13.31)$$

який в поєднанні з рівнянням (13.16), кінематичними ГУ (13.27) на поверхні S_U визначає незліченну множину можливих (віртуальних) напружено-деформованих станів. Дійсний НДС є одним з віртуальних, але він додатково задовольняє фізичним рівнянням зв'язків між напруженнями та деформаціями. Оскільки цей функціонал не містить фізичних рівнянь, він є придатним і для нелінійних задач термопружності, а також для задач термопружнопластичності та термоповзучості, які будуть розглядатися в наступних підрозділах. У (13.31) зосереджені сили розглядаються як окремий випадок поверхневих сил.

Додатково треба відзначити, що в цій задачі час не є параметром, тобто явно не входить до рівнянь. Час застосовується лише для того, щоб розрізняти початковий стан з наступними.

13.3. Постановка крайової задачі термопружнопластичності

Відома значна кількість моделей термопружнопластичного матеріалу. Розглянемо лише дві популярні моделі.

13.3.1. Узагальнена на термопластичність деформаційна теорія пластичності

До формули (13.16) включно міркування такі ж, як і при постановці задачі термопружності. Замість формули (13.17) маємо

$$\varepsilon_{ij} = \varepsilon_{ij}^{a} + \varepsilon_{ij}^{\theta} + (\varepsilon_{ij})_{0}, \qquad (13.32)$$

де $\varepsilon_{ij}^{a} = \varepsilon_{ij} - \varepsilon_{ij}^{\theta} - (\varepsilon_{ij})_{0} = \varepsilon_{ij}^{e} + \varepsilon_{ij}^{p}$ – компоненти тензора малих пружнопластичних деформацій ("активних" деформацій). При активному навантаженні "фізичні" рівняння пружного матеріалу замінюються на (ізотропний матеріал):

• закон пружної зміни елементарного об'єму

$$\varepsilon_V^a = \sigma_V / 3k; \quad \varepsilon_V^p = 0; \tag{13.33}$$

• тензорне співвідношення (закон зміни форми елементарного об'єму)

$$\mathcal{E}_{ij}^{a} = \left(\varphi / 2G(\theta) \right) \cdot S_{ij}; \quad \varphi = 3G(\theta) \cdot \mathcal{E}_{u}^{a} / \sigma_{u}; \quad (13.34)$$

• рівняння стану (яке визначається експериментально)

$$\sigma_{u} = K(\varepsilon_{u}^{a}, \theta), \qquad (13.35)$$

де $\varepsilon_{v} = \delta_{ij}\varepsilon_{ij}/3$ та $\sigma_{v} = \delta_{ij}\sigma_{ij}/3$ – об'ємні деформація й напруження відповідно; $k = k(\theta)$ – модуль об'ємної деформації; $e_{ij} = \varepsilon_{ij} - \delta_{ij}\varepsilon_{v}$ та $S_{ij} = \sigma_{ij} - \delta_{ij}\sigma_{v}$ – компоненти девіаторних частин тензорів деформацій і напружень відповідно; ε_{u}^{a} – інтенсивності пружнопластичних деформацій (які залишилися після видалення температурних та початкових деформацій); σ_{u} – інтенсивність напружень; $\varphi \ge 1$, причому $\varphi = 1$ у пружному стані (тоді з (13.34) $S_{ij} = 2G(\theta) \cdot e_{ij}^a$ та $\sigma_u = 3G(\theta) \cdot \varepsilon_u^a$).

Вирази (13.20), (13.27) й (13.28) зберігаються. Розвантаження – пружне.

Існує декілька методів приведення сформульованої задачі до системи рівнянь відносно переміщень. Вони розглядаються в теорії пластичності. Отриману систему рівнянь можна представити операторним рівнянням загального вигляду

$$L(\vec{u}) = \vec{f} \quad \mathbf{B} \ \Omega \tag{13.36}$$

з ГУ (13.14), в яких диференційний оператор *L* є нелінійним.

13.3.2. Узагальнена на термопластичність інкрементальна теорія Прандтля-Рейса

До формули (13.15) включно міркування такі ж, як і при постановці задачі термопружності. Замість рівнянь Коші (13.16) (малі деформації) маємо

$$d\varepsilon_{ij} = 0.5(\nabla_i du_j + \nabla_j du_i), \qquad (13.37)$$

а також рівняння сумісності деформацій (12.49) і принцип суперпозиції нескінченно малих приростів деформацій різної природи у вигляді:

$$d\varepsilon_{ij} = d\varepsilon^{e}_{ij} + d\varepsilon^{\theta}_{ij} + d\varepsilon^{p}_{ij}.$$
(13.38)

"Фізичні" рівняння (13.19) і (13.20-б) замінюються на:

$$d\sigma_{mn} = E_{mnij} d\varepsilon_{ij}^{e} + \frac{\partial E_{mnij}}{\partial \theta} d\hat{\theta} \varepsilon_{ij}^{e}; \quad d\varepsilon_{ij}^{e} = C_{ijmn} d\sigma_{mn} + \frac{\partial C_{ijmn}}{\partial \theta} d\hat{\theta} \sigma_{mn}; \quad d\hat{\theta} = \theta - \theta_{0}; (13.39)$$
$$d\varepsilon_{ij}^{\theta} = \delta_{ij} \Big[\alpha_{\theta}(\theta) \cdot (\theta - \theta_{RM}) - \alpha_{\theta}(\theta_{0}) \cdot (\theta_{0} - \theta_{RM}) \Big], \quad (13.40)$$

а при активному навантаженні за межами зони пружності використовуються рівняння теорії пластичності (ізотропний матеріал, що зміцняється):

• закон пружної зміни об'єму

$$d\varepsilon_{V}^{a} = d\sigma_{V} / 3k; \qquad d\varepsilon_{V}^{p} = 0; \qquad (13.41)$$

де компоненти тензора приросту "активних" деформацій $d\varepsilon_{ij}^a = d\varepsilon_{ij} - d\varepsilon_{ij}^{\theta} = d\varepsilon_{ij}^e + d\varepsilon_{ij}^p;$

• закон зміни форми елементарного об'єму (тензорне співвідношення)

$$de_{ij}^{p} = d\lambda^{p}S_{ij}; \quad d\lambda^{p} = \frac{3d\overline{\varepsilon}_{u}^{p}}{2\sigma_{u}} = \frac{3d\chi}{2\sigma_{u}}; \quad (13.42)$$

• рівняння стану (визначається експериментально):

$$\sigma_u = H(\chi, \theta) \quad \text{afo} \quad \sigma_u = K(\varepsilon_u^a, \theta),$$
(13.43)

де $\chi = \int d\chi$ – параметр Одквіста; $d\chi = d\overline{\varepsilon}_{u}^{p} = (2de_{ij}^{p}de_{ij}^{p}/3)^{1/2}$ – інтенсивність приростів пластичних деформацій.

Примітка 13.2. Часто вводять поняття "миттєвої термомеханічної гіперповерхні" f, яка розмежовує пружний та пластичний стан матеріалу, а напрямок зовнішньої нормалі до неї вказує на напрямок розвитку пластичних деформацій (постулат Друккера): $de_{mn}^{p} = h \partial f / \partial \sigma_{mn}$, де h – деякий функціонал.

Зазвичай цю поверхню пов'язують з виразами для рівняння стану (13.43): $f = \sigma_u^2 - H^2(\chi, \theta) = 0$ або $f = \sigma_u^2 - K^2(\varepsilon_u^a, \theta) = 0$. З урахуванням виразу $\sigma_u = (3S_{mn}S_{mn}/2)^{1/2}$ можна отримати перший вираз (13.42) та зв'язок між функціоналами *h* та $d\lambda^p$.

Додатково залучаються ГУ (13.27) і (13.28) на S_{U} та S_{P} відповідно.

Існує декілька методів приведення сформульованої задачі до системи рівнянь відносно переміщень. Вони розглядаються в теорії пластичності. В загальному вигляді всі вони приводяться до операторного рівняння (13.36) з ГУ вигляду (13.14).

Диференційній постановці крайової задачі термопластичності ставиться у відповідність варіаційна інтегральна постановка. В підсумку можна отримати варіаційний функціонал (13.31) відносно варіацій переміщень і зв'язаних із ними деформацій, який в поєднанні з рівнянням (13.16) і кінематичними ГУ (13.27) на поверхні S_U визначає незліченну множину можливих (віртуальних) НДС. Дійсний НДС є одним з віртуальних, але він додатково задовольняє нелінійним "фізичним" рівнянням.

Як і в пружній, в цій задачі час також не є параметром, тобто явно не входить до рівнянь. Час застосовується лише для того, щоб розрізняти початковий стан з наступними.

13.4. Постановка крайової задачі термоповзучості

Наведемо постановку задачі з використанням технічної теорії повзучості – *meopiï зміцнення* (ізотропний матеріал).

До формули (13.37) включно міркування такі ж, як і при постановці задачі термопластичності Прандтля-Рейса. Замість формули (13.38) маємо

$$d\varepsilon_{ij} = d\varepsilon^{e}_{ij} + d\varepsilon^{\theta}_{ij} + d\varepsilon^{c}_{ij}.$$
(13.44)

Замість формул (13.41) ... (13.43) при будь-якому рівні навантаження застосовуються "фізичні" рівняння вказаної теорії повзучості:

• закон пружної зміни об'єму

$$d\varepsilon_{V}^{a} = d\sigma_{V} / 3k; \quad d\varepsilon_{V}^{c} = 0, \qquad (13.45)$$

де компоненти тензора приростів "активних" деформацій $d\varepsilon_{ij}^{a} = d\varepsilon_{ij} - d\varepsilon_{ij}^{e} = d\varepsilon_{ij}^{e} + d\varepsilon_{ij}^{c};$

• закон зміни форми елементарного об'єму (тензорне співвідношення)

$$de_{ij}^{c} = d\lambda^{c}S_{ij}; \quad d\lambda^{c} = \frac{3d\overline{\varepsilon}_{u}^{c}}{2\sigma_{u}}; \quad (13.46)$$

• рівняння стану теорії зміцнення (визначається експериментально)

$$\dot{\varepsilon}_{u}^{c} = \frac{d\overline{\varepsilon}_{u}^{c}}{dt} = \Psi(\sigma_{u}, \varepsilon_{u}^{c}, \theta, t), \qquad (13.47)$$

де $d\overline{\varepsilon}_{u}^{c}$ – інтенсивність приростів деформацій повзучості, $\dot{\varepsilon}_{u}^{c}$ – швидкість зміни інтенсивності деформацій повзучості ε_{u}^{c} .

Як і в будь-якій еволюційній задачі, потрібно задати початкові умови. В задачі повзучості з "фізичними" рівняннями технічної теорії зміцнення зазвичай задаються незначні початкові деформації повзучості:

$$\varepsilon_u^c(\vec{x},0) = (\varepsilon_u^c)_0. \tag{13.48}$$

Додатково залучаються ГУ (13.27) і (13.28) на S_{U} та S_{P} відповідно.

Існує декілька методів приведення сформульованої задачі до системи рівнянь відносно переміщень. Вони розглядаються в теорії повзучості. В загальному вигляді всі отримані системи рівнянь можна представити у вигляді операторного рівняння

$$\frac{\partial \vec{u}}{\partial t} + L(\vec{u}) = \vec{f} \quad \mathbf{B} \quad \Omega \tag{13.49}$$

з ГУ (13.14), тобто $K\vec{u} = \vec{g}$ на S.

Диференційній постановці крайової задачі термоповзучості ставиться у відповідність варіаційна інтегральна постановка. В підсумку можна отримати варіаційний функціонал (13.31) відносно варіацій переміщень і зв'язаних із ними деформацій, який в поєднанні з рівнянням (13.16) і кінематичними ГУ (13.27) на поверхні S_U визначає незліченну множину можливих (віртуальних) НДС. Дійсний НДС є одним з віртуальних, але він додатково задовольняє нелінійним "фізичним" рівнянням. У цій задачі, на відміну від задач термопружності та термопластичності, час явно входить до рівнянь, але напруження залежать від часу неявно, як це й потрібно у відповідності до аксіом Нолла (див. підрозділ 12.4).

13.5. Постановка крайової динамічної задачі термопружності

Постановка крайової динамічної задачі термопружності в більшості моментів співпадає з постановками крайової (квазі)стаціонарної задачі термопружності, тому звернемо увагу лише на розбіжності. Зазвичай задаються початкові умови для переміщень, швидкостей та прискорень, тобто при t_0 :

$$u_m(\vec{x},t_0) = \hat{u}_m(\vec{x}), \quad m = 1,2,3;$$
 (13.50)

$$\partial u_m(\vec{x}, t_0) / \partial t = \hat{g}_m(\vec{x}); \qquad (13.51)$$

$$\partial^2 u_m(\vec{x}, t_0) / \partial t^2 = \hat{b}_m(\vec{x}).$$
(13.52)

Згідно з принципом Д'Аламбера, за об'ємну силу розглядають силу з компонентами $\hat{\overline{O}}_m = \hat{O}_m - \overline{\rho} \frac{\partial^2 u_m}{\partial t^2} - \alpha \frac{\partial u_m}{\partial t}$, де $\overline{\rho}$ – густина матеріалу, α – коефіцієнт демпфування. Рівняння рівноваги (13.15), заміняється на повне рівняння руху):

$$\nabla_n \sigma_{mn} + \hat{O}_m = \bar{\rho} \frac{\partial^2 u_m}{\partial t^2} + \alpha \frac{\partial u_m}{\partial t}.$$
 (13.53)

Формула (13.29) перетворюється на

$$\nabla_{n}\left\{E_{mnij}(\theta)\left[0.5(\nabla_{i}u_{j}+\nabla_{j}u_{i})\right]\right\}-\overline{\rho}\frac{\partial^{2}u_{m}}{\partial t^{2}}-\alpha\frac{\partial u_{m}}{\partial t}=\nabla_{n}E_{mnij}(\theta)\delta_{ij}\alpha_{\theta}\Delta\widehat{\theta}-\widehat{O}_{m},\quad(13.54)$$
і в сукупності з ГУ (13.27), (13.30) та початковими умовами (13.50) ... (13.52) є системою диференційних рівнянь в точці тіла, записаною відносно переміщень, розв'язок якої опише зміни НДС в часі з урахуванням динамічного навантаження та в'язкого тертя. Очевидно, що ця система у вигляді операторного рівняння може бути записана як:

$$M\frac{\partial^2 \vec{u}}{\partial t^2} + C\frac{\partial \vec{u}}{\partial t} + L\vec{u} = \vec{f} \quad \mathbf{B} \ \Omega \ . \tag{13.55}$$

При варіаційному формулюванні динамічної задачі використовують рівняння руху (13.52). Тому, як і для крайової задачі термопружності, одержується варіаційний функціонал принципу можливих переміщень

$$\int_{\Omega} \overline{\rho} \frac{\partial^2 u_m}{\partial t^2} \delta u_m d\Omega + \int_{\Omega} \alpha \frac{\partial u_m}{\partial t} \delta u_m d\Omega + \int_{\Omega} \sigma_{mn} \delta \varepsilon_{mn} d\Omega - \int_{\Omega} \widehat{O}_m \delta u_m d\Omega - \int_{S_p} \widehat{P}_m \delta u_m dS = 0, \quad (13.56)$$

до якого ще необхідно додати "фізичні" рівняння (закон Гука $\sigma_{mk} = E_{mkij} \varepsilon_{ij}^{e}$), рівняння зв'язку між варіаціями деформацій та переміщень $\delta \varepsilon_{ij} = 0.5(\nabla_i \delta u_j + \nabla_j \delta u_i)$, а також ГУ 1-го роду (13.27), тобто $u_i |_{S_U} = \hat{u}_i$, й початкові умови (13.50) ... (13.52).

Ще можна застосувати рівняння Лагранжа 2-го роду, результат буде той самий.

Можливі такі варіанти динамічних навантажень:

- гармонійне;
- ударне (раптово та швидко прикладене негармонійне);
- випадкове (стохастичне) динамічне;

При всіх цих динамічних навантаженнях можливі обернені задачі про визначення таких характеристик об'єкта розрахунків, як оптимальні значення густини матеріалу або демпфування.

Окремим випадком задачі про гармонійне збудження є задача про визначення власних частот та форм коливань об'єкта розрахунків.

При цьому зазвичай вважають, що температурне поле є незмінним, а інші зовнішні впливи носять періодичний характер:

$$\hat{u}_m(\vec{x},t) = \hat{\tilde{u}}_m(\vec{x})\sin(pt+\beta); \qquad (13.57)$$

$$\widehat{O}_m(\vec{x},t) = \widetilde{O}_m(\vec{x})\sin(pt+\beta); \qquad (13.58)$$

$$\widehat{P}_{m}(\vec{x},t) = \widehat{\widetilde{P}}_{m}(\vec{x})\sin(pt+\beta), \qquad (13.59)$$

де p – задана колова частота зовнішніх впливів, β – фаза коливань (припускається їхня синхронність та синфазність).

Для лінійно-пружного середовища компоненти вектора переміщень, тензорів деформацій і напружень будуть залежати від часу майже так само, як і зовнішні навантаження:

$$u_m(\vec{x},t) = \tilde{u}_m(\vec{x})\sin(pt + \beta_1); \qquad (13.60)$$

$$\varepsilon_{mn}(\vec{x},t) = \tilde{\varepsilon}_{mn}(\vec{x})\sin(pt + \beta_1); \qquad (13.61)$$

$$\sigma_{mn}(\vec{x},t) = \tilde{\sigma}_{mn}(\vec{x})\sin(pt + \beta_1), \qquad (13.62)$$

причому $\beta_1 = \beta$ при $\alpha = 0$ (при відсутності внутрішньої дисипації енергії).

Підставивши (13.60), (13.62) в рівняння руху (13.52), отримаємо:

$$(\nabla_n \tilde{\sigma}_{mn} + \tilde{O}_m + \bar{\rho} p^2 \tilde{u}_m) \sin(pt + \beta_1) - \alpha p \tilde{u}_m \cos(pt + \beta_1) = 0.$$
(13.63)

Виключивши з (13.63) напруження за допомогою "геометричних" і лінійних "фізичних" рівнянь, тобто (13.16) і (13.19), отримаємо в кожній точці $\vec{x} \in \Omega$:

$$\nabla_{n} \left\{ E_{mnij}(\theta) \Big[0.5(\nabla_{i} \tilde{u}_{j} + \nabla_{j} \tilde{u}_{i}) \Big] \right\} + \overline{\rho} p^{2} \tilde{u}_{m} \Big) \sin(pt + \beta_{1}) - \tilde{\rho} \sum_{i=1}^{n} (-i + \beta_{i}) \exp(pt + \beta_{1}) - \tilde{\rho} \sum_{i=1}^{n} (-i + \beta_{i}) \exp(pt + \beta_{1}) - \tilde{\rho} \sum_{i=1}^{n} (-i + \beta_{i}) \exp(pt + \beta_{1}) - \tilde{\rho} \sum_{i=1}^{n} (-i + \beta_{i}) \exp(pt + \beta_{1}) - \tilde{\rho} \sum_{i=1}^{n} (-i + \beta_{i}) \exp(pt + \beta_{1}) - \tilde{\rho} \sum_{i=1}^{n} (-i + \beta_{i}) \exp(pt + \beta_{1}) \exp(pt + \beta_{1}) - \tilde{\rho} \sum_{i=1}^{n} (-i + \beta_{i}) \exp(pt + \beta_{1}) \exp(pt +$$

$$-\alpha p\tilde{u}_m \cos(pt + \beta_1) = -\tilde{O}_m \sin(pt + \beta_1).$$
(13.64)

За відсутності дисипації енергії з (13.64) випливає:

$$\nabla_n \left\{ E_{mnij}(\theta) \Big[0.5 (\nabla_i \tilde{u}_j + \nabla_j \tilde{u}_i) \Big] \right\} + \overline{\rho} p^2 \tilde{u}_m = -\hat{\tilde{O}}_m.$$
(13.65)

3 граничних умов (13.27) і (13.28) можна отримати, що

$$\tilde{u}_m \mid_{S_U} = \tilde{\tilde{u}}_m; \tag{13.66}$$

$$\widetilde{\sigma}_{mn} v_n \Big|_{S_p} = \widehat{\widetilde{P}}_m.$$
(13.67)

Початкові умови в такій задачі не вимагаються.

У рівняннях (13.64) ... (13.67) містяться принаймні дві задачі:

• про розв'язок задачі при заданої частоті збудження *p* (або про *вимушені* гармонійні коливання, чи про гармонійне збудження);

• про власні частоти ω та форми коливань тіла (нетривіальний розв'язок за відсутності зовнішніх термосилових впливів та з урахуванням статичних кінематичних граничних умов, зокрема й при їх відсутності). Без врахування дисипації енергії:

$$\nabla_n \left\{ E_{mnij}(\theta) \Big[0.5(\nabla_i \tilde{u}_j + \nabla_j \tilde{u}_i) \Big] \right\} + \overline{\rho} \omega^2 \tilde{u}_m = 0.$$
(13.68)

Цим значенням відповідають *власні форми коливань*. Вираз (13.68) в операторному вигляді записується як

$$L\tilde{\tilde{u}} + \omega^2 M\tilde{\tilde{u}} = 0 \quad \mathbf{B} \ \Omega \ . \tag{13.69}$$

Чисельні методи й алгоритми розв'язування динамічних задач будуть розглянуті в Розділах 25 і 26.

Контрольні питання до підрозділу 13.1

1. Які закони використовують для формулювання крайової задачі теплопровідності?

2. Пояснить призначення функцій, що входять до граничної умови з променевого теплообміну.

3. Чим відрізняються рівняння стаціонарної теплопровідності від рівнянь нестаціонарної теплопровідності?

Контрольні питання до підрозділу 13.2

1. Яким вважається початковий стан для крайової задачі термопружності?

- 2. Які три групи рівнянь входять у постановку крайової задачі термопружності?
- 3. В який спосіб вводиться "січний" коефіцієнт температурного подовження SCTE?
- 4. Через які три "фізичних" закони описують пружно-нелінійний матеріал?
- 5. Що визначають граничні умови крайової задачі термопружності?

6. Які рівняння враховує принцип можливих переміщень (13.31)?

Контрольні питання до підрозділу 13.3

1. Які три групи рівнянь входять у постановки крайової задачі термопружнопластичності у формулюваннях Іллюшина та Прандтля-Рейса?

2. Що таке "миттєва термомеханічна гіперповерхня"?

Контрольні питання до підрозділу 13.4

1. Які три групи рівнянь входять у постановку крайової задачі термоповзучості за теорією зміцнення?

2. Чому й які вводяться початкові умови?

Контрольні питання до підрозділу 13.5

- 1. Які особливості має постановка крайової динамічної задачі термопружності?
- 2. Назвіть приклади застосування початкової умови для прискорень (13.52).

3. Як отримати задачу про власні форми та частоти коливань пружного тіла?

Розділ 14

ПОНЯТТЯ ПРО АЛГЕБРАЇЗАЦІЮ КРАЙОВИХ ЗАДАЧ

14.1. Поняття про алгебраїзацію крайових задач

У результаті постановки характерних крайових задач (див. Розділ 13) прийшли до висновку, що всі крайові задачі (окрім задачі про власні частоти) можна уявити у вигляді трьох диференційних рівнянь відносно вектора *ū*, які відносяться до рівнянь трьох типів:

• стаціонарні (лінійні задачі еліптичного типу або нелінійні)

$$L\vec{u} = \vec{f} \quad \text{afo} \quad L(\vec{u}) = \vec{f} \quad \mathbf{B} \ \Omega; \tag{14.1}$$

• параболічного (задачі нестаціонарної теплопровідності та повзучості)

$$C\frac{\partial \vec{u}}{\partial t} + L\vec{u} = \vec{f} \quad \mathbf{B} \ \Omega; \tag{14.2}$$

• гіперболічного (задачі динаміки)

$$M\frac{\partial^2 \vec{u}}{\partial t^2} + C\frac{\partial \vec{u}}{\partial t} + L\vec{u} = \vec{f}, \quad \text{abo} \quad M\frac{\partial^2 \vec{u}}{\partial t^2} + L\vec{u} = \vec{f}, \quad \mathbf{B} \ \Omega, \tag{14.3}$$

в яких з трьох операторів *M*, *C*, *L* лише оператор *L* є диференційним.

Граничні умови для них можна уявити у вигляді (13.14), тобто

$$K\vec{u} = \vec{g} \tag{14.4}$$

на відповідній частині поверхні S (інші позначення: Γ або $\partial \Omega$).

Для нелінійних задач, реалізуючи ітераційні чисельні методи розв'язування задачі, в ітераціях застосовують лінійні оператори.

Щоб розглядати далі питання існування розв'язків й їх збіжності, для побудови методів розв'язування необхідно представити задачу у вигляді *операторних* рівнянь в придатних функціональних просторах.

Для задач з диференційними операторами достатньо використати простір Соболева W_2^l , де l – порядок диференційного оператора L. Зовнішній вигляд рівнянь (14.1) ... (14.4) залишиться без змін, але тепер це – операторні рівняння. Основне завдання при такому переході – визначити властивості операторів Lта K: чи будуть вони лінійними та обмеженими, а оператор L – позитивно визначеним, симетричним (або чи можна його зробити симетричним), та інше.

Для нелінійних задач можуть бути додатково залучені різні властивості просторів і вимоги до них.

Наступний крок – *алгебраїзація* (*дискретизація, редукція*) задачі, тобто перехід від неперервних аргументів до дискретних з одночасним позбавленням від операторів диференціювання та/або інтегрування, якщо останні присутні.

Це тому, що цифрова ЕОМ має дискретний числовий простір та "не знає" операцій з диференціювання та інтегрування (тільки +, -, *, :, а також логічні операції та операції порівнювання).

Якщо крайову задачу сформульовано у вигляді рівнянь *параболічного* або *гіперболічного* типу, зазвичай застосовується наступний підхід: проводиться *роздільна* часова і просторова редукція (ще називають *частковою дискретизацією*). Іноді час розглядається як четвертий вимір нарівні з просторовими змінними. При роздільній часовій редукції використовуються так звані *часові шари*, які характеризуються часовими значеннями $t_n, t_{n+1},...,$ де n = 0,1,... – номер часового шару.

Просторова редукція проводиться на основі базисних функцій: точкових або кусково-визначених, а також методом скінченних різниць (МСР).

Всю сукупність чисельних методів розв'язування крайових задач суцільних середовищ, що застосовують *базисні* функції, називають *проєкційно-сітковими* методами.

Основний принцип при проведенні редукції – збереження (не погіршення) властивостей операторів задач. Для широкого класу задач це відбувається автоматично при застосуванні варіаційного підходу. Це не випадково, оскільки варіаційні функціонали адекватно відображають суттєві закономірності фізичних процесів, зокрема в механіці суцільних середовищ.

Для нелінійних задач додатково ще необхідно застосувати той або інший ітераційний метод.

Зрештою в результаті означених перетворень одержують одну або більше систем алгебраїчних рівнянь вигляду Ax = b, де A – матриця системи, x, b – вектори невідомих і правої частини системи відповідно.

Теоретичний аналіз чисельного алгоритму, побудованого для отримання розв'язку крайової задачі, розглядає дві глобальні проблеми:

• визначення похибки наближення при алгебраїзації;

• обгрунтування обчислювальної стійкість розрахункової схеми. При цьому потрібно показати, що при спрямуванні кроку дискретизації до нуля наближений розв'язок безперервно залежить від початкових (вихідних) умов. Тоді він прямує до точного розв'язку.

Приклад. Дослідимо властивості диференційного оператора $L = -d^2 / dx^2$ на відрізку осі $0 \le x \le a$ та на множині функцій, що перетворюються в нуль при x=0 і при x=a.

Нехай u і v – дві такі функції, α та β – дійсні числа. Очевидно, що

$$L(\alpha u + \beta v) = -\frac{d^2}{dx^2}(\alpha u + \beta v) = -\alpha \frac{d^2 u}{dx^2} - \beta \frac{d^2 v}{dx^2} = \alpha L u + \beta L v,$$

© К.М. Рудаков, 2025

тобто оператор – лінійний (див. формулу (2.14)).

Дослідимо оператор на симетричність. Використовуючи правило інтегрування по частинах і те, що u(0) = u(a) = v(0) = v(a) = 0, отримуємо:

$$\int_{0}^{a} u \frac{d^{2}v}{dx^{2}} dx = \left| w = \frac{dv}{dx} \right| = (uw) |_{0}^{a} - \int_{0}^{a} \frac{du}{dx} w dx = \left| z = \frac{du}{dx} \right| =$$
$$= 0 - 0 - \int_{0}^{a} z \frac{dv}{dx} dx = -(zv) |_{0}^{a} + \int_{0}^{a} v \frac{dz}{dx} dx = -0 + 0 + \int_{0}^{a} v \frac{d^{2}u}{dx^{2}} dx.$$

Таким чином, оператор $L = -d^2/dx^2 \in$ симетричним при означених умовах (див. формулу (2.21)).

Досліджуємо оператор на позитивну визначеність (див. підрозділ 2.3). Вважаючи u = v, маємо

$$-\int_{0}^{a} u \frac{d^{2} u}{dx^{2}} dx = \left| w = \frac{du}{dx} \right| = -(uw) \left|_{0}^{a} + \int_{0}^{a} \frac{du}{dx} w dx \right|_{0}^{a} = \int_{0}^{a} \frac{du}{dx} \frac{du}{dx} dx = \int_{0}^{a} \left(\frac{du}{dx} \right)^{2} dx > 0.$$

Враховано, що рівність нулю можлива лише при $du/dx \equiv 0$ на всьому шляху інтегрування, тобто при постійному значенні функції u(x). Оскільки на кінцях u = 0, то й інтеграл може бути рівним нулю лише при $u \equiv 0$, а цей вироджений випадок розглядати не має сенсу.

Отже, виявлені лінійність, симетричність і позитивна визначеність диференційного оператора $L = -d^2/dx^2$ при зазначених вище умовах.

14.2. Про міри похибок дискретизації та алгебраїзації

Природно, що перехід від неперервних аргументів до дискретних, а також заміна диференційних та інтегральних виразів алгебраїчними супроводжується *похибками*. Необхідно вміти їх вірно оцінювати.

Нехай $\vec{u}_0(\vec{x})$ – точний вектор-розв'язок системи рівнянь $L\vec{u} = \vec{f}$ в області Ω з граничними умовами вигляду $K\vec{u} = \vec{g}$ на S, а $\vec{u}_h(\vec{x})$ – наближений вектор. Локальна (у вузлу) абсолютна похибка визначається просто:

$$\vec{\Delta}_h = (\vec{u}_0(\vec{x}))_h - \vec{u}_h(\vec{x}).$$
(14.5)

Тут $(\vec{u}_0(\vec{x}))_h = P_h(\vec{u}_0(\vec{x}))$, де P_h є оператором проєктування вектора на сітку. Проблема у тому, що вектори $\vec{u}_0(\vec{x})$ та $\vec{u}_h(\vec{x})$ належать різним просторам: безперервному та дискретному відповідно. Щоб порівнювати такі вектори, потрібно один з них спроєктувати на інший простір. Простіше це робити з безперервного простору у дискретний. Якщо $\vec{u}_0(\vec{x})$ є безперервною функцією від \vec{x} , то обчислюються $\vec{u}_0(\vec{x}_i)$, де \vec{x}_i відповідають вузлам дискретного простору. Якщо $\vec{u}_0(\vec{x})$ має розриви першого роду, то в точках розривів потрібно розміщувати вузли дискретного простору та якимось чином усереднювати значення $\vec{u}_0(\vec{x}_i)$, наприклад, за методом середнього арифметичного або інтегрально у деякому околі вузла. Оскільки оцінювання в окремому вузлі не дає загальної картини, вводяться узагальнені міри. Однією з загальновизнаних і зручних мір є *енергетична норма* абсолютної похибки:

$$\left\|\vec{\Delta}_{h}\right\|_{h} = \left[\int_{\Omega} \vec{\Delta}_{h} L_{h} \vec{\Delta}_{h} d\Omega + \int_{S} \vec{\Delta}_{h} K_{h} \vec{\Delta}_{h} dS\right]^{1/2}.$$
(14.6)

За швидкістю прямування $\|\vec{\Delta}_h\|_h$ до нуля при зміні кроку дискретизації *h* можна робити висновки про якість наближення. Але потрібно мати на увазі, що така оцінка із застосуванням узагальненої норми ((14.6) та інші) завжди "м'якше", ніж із застосуванням локальної норми.

Пояснимо причину назви норми (14.6) енергетичною на прикладі задачі механіки. За змістом у задачах механіки вектор \vec{f} – сила, а вектор \vec{u} – переміщення. Очевидно, що $\vec{\Delta}L\vec{\Delta} = (\vec{u}_0 - \vec{u})L(\vec{u}_0 - \vec{u}) = (\vec{u}_0 - \vec{u})(\vec{f}_0 - \vec{f})$ має розмірність роботи (або енергії).

Проте застосовується оцінювання не лише за швидкістю прямування $\|\vec{\Delta}_h\|_h$ до нуля. Є й інші критеріальні оцінки. Їх будемо розглядати далі за необхідністю.

Контрольні питання до Розділу 14

- 1. Чому необхідно проводити дискретизацію та алгебраїзацію крайових задач?
- 2. Назвіть та охарактеризуйте етапи проведення алгебраїзації крайових задач.
- 3. Які міри використовують для оцінки похибок дискретизації та алгебраїзації?

Частина IV АЛГОРИТМИ МЕТОДУ СКІНЧЕННИХ РІЗНИЦЬ

Розділ 15

ПРОСТОРОВЕ НАБЛИЖЕННЯ СТАЦІОНАРНИХ КРАЙОВИХ ЗАДАЧ ЗА МЕТОДОМ СКІНЧЕННИХ РІЗНИЦЬ

Для викладення ідеї та алгоритмів методу скінченних різниць спочатку (у цьому Розділі) розглянемо стаціонарні задачі, що приводяться у гільбертовому просторі *H*₀ до операторного рівняння

$$L\vec{u} = \vec{f} \ \mathbf{B} \ \Omega \tag{15.1}$$

з граничними умовами (ГУ)

$$K\vec{u}\big|_{S} = \hat{\vec{g}} \,. \tag{15.2}$$

Тут $\vec{u} = \vec{u}(\vec{x})$ – шуканий вектор (наприклад, переміщень); \vec{f} – вектор навантажень; L – диференційний оператор, а K – диференційний у випадку природних ГУ (2-го роду) та простий при ГУ 1-го роду, наприклад, $\vec{u}|_{S_U} = \hat{\vec{q}}$.

Нагадаємо, що просторова дискретизація проводиться на основі базисних функцій: точкових, безперервних або кусково-визначених, а також методом скінченних різниць (МСР). В цьому Розділі розглянемо просторову дискретизацію на основі МСР.

15.1. Ідея методу скінченних різниць

Ідея МСР полягає у наступному: а) область, в якій шукається розв'язок крайової задачі, заміняється набором вузлів; б) – замість безперервної функції розв'язку, що шукається, вводиться функція, яка визначена лише у вузлах – *сіткова* функція; в) – здійснюється перехід від диференційних (та, якщо вони є, інтегральних) рівнянь до алгебраїчних і, як результат, створюється система алгебраїчних рівнянь відносно невідомих значень сіткової функції.

Розглянемо ці питання докладніше.

У МСР розв'язок розглядається в скінченній кількості точок простору – вузлах. Така множина називається сіткою. Але це характерно для багатьох наближених методів. Головне те, що функція, яка буде наближати майбутній розв'язок задачі, визначається теж у вузлах, і тільки у вузлах сітки (на сітці), тому називається сітковою функцією. Сітка на відрізку лінії – це N вузлів, які розміщено один від одного на відстані, що називається *кроком* сітки. Крок h може бути рівномірним (*рівномірна сітка*) і нерівномірним (*нерівномірна сітка*). Припустимо для визначеності, що використовуємо декартову систему координат. Позначимо як h_x, h_y, h_z кроки вздовж відповідних координатних напрямків x, y, z. Тоді множина вузлів сітки для нерівномірної і рівномірної сітки може бути заданою, відповідно, як

$$\omega_{xyz} = \{ x_i = \phi(h_x, i); \quad y_j = \phi(h_y, j); \quad z_k = \psi(h_z, k); \quad i = 1, ..., N_x; \; j = 1, ..., N_y; \; k = 1, ..., N_z \} ;
 \omega_{xyz} = \{ x_i = x_o \pm ih_x; \quad y_j = y_o \pm jh_y; \; z_k = z_o \pm kh_z; \\
 i = 0, 1, ..., N_x - 1; \; j = 0, 1, ..., N_y - 1; \; k = 0, 1, ..., N_z - 1 \},$$
(15.3)

де ϕ, ϕ, ψ – деякі функції розподілу вузлів вздовж координатних осей; x_o, y_o, z_o – крайні значення координат.

Розрізняють множини вузлів *внутрішніх* $\overline{\omega}_{xyz}$ і *граничних* $\widetilde{\omega}_{xyz}$, причому $\omega_{xyz} = \overline{\omega}_{xyz} + \widetilde{\omega}_{xyz}$; вузли фіктивні *зовнішні*, що вводяться зазвичай для опису граничних умов, а також вузли *шаблону* (з ближнього околу актуального вузла).

Замість функції $u(\vec{x})$ введемо сіткову функцію $u_h(x_i, y_i, z_i)$ з областю визначення D, тому $\omega_{xyz} \in D$. Наприклад, якщо тіло, що розглядається, є паралелепіпедом, то область визначення $D = \{a_x \le x \le b_x; a_y \le y \le b_y; a_z \le z \le b_z\}$.

Перенумерувавши всі вузли в деякому порядку, сіткову функцію можна уявити у вигляді компонент, які створюють вектор

$$\vec{u}_h = \{u_h^{(1)}, u_h^{(2)}, \dots, u_h^{(N)}\}^T,$$
(15.4)

де N – загальна кількість вузлів. Всі ці дії можна представити у вигляді дій оператора проєктування P_h на вектор \vec{u} :

$$\vec{u}_h = P_h \vec{u} \in H_h \,. \tag{15.5}$$

Аналогічно перетворюються й всі інші вектори, зокрема \vec{f} і $\hat{\vec{g}}$ на \vec{f}_h і $\hat{\vec{g}}_h$ відповідно. Залишилося ввести сітковий аналог диференційних операторів *L* та *K*, тобто L_h і K_h . Для цього у МСР використовують *meopiю скінченних різниць*, яка була розроблена значно раніше (див. підрозділ 9.2). А саме, *диференційні* оператори заміняють (наближують) *алгебраїчними* виразами, причому створюють їх з використанням сіткової функції, яка наближує розв'язок, але не на всіх вузлах сітки тіла, а лише на вузлах шаблона (докладніше – у підрозділі 15.2). Ця заміна повинна проводитися з достатньою точністю та зі збереженням основних властивостей оператора, що наближується.

Отже, операторні рівняння (15.1) і (15.2) крайової задачі перетворюються в сітковий аналог

$$L_h \vec{u}_h = \vec{f}_h; \quad K_h \vec{u}_h \big|_S = \hat{\vec{g}}_h, \qquad (15.6)$$

який є системою алгебраїчних рівнянь (САР).

Методи розв'язування САР викладено у Розділах 5, 7, 8. Про існування та одиничність розв'язку сіткових аналогів крайових задач див. підрозділ 2.6.

МСР з'явився задовго до появи ЕОМ. Він дозволяв розв'язувати ті крайові задачі, що ще не мали аналітичного розв'язку. Це була дуже кропітка робота, оскільки будь-яка помилка на проміжних етапах приводила до необхідності виконувати всю або значну частку роботи знову. З появою ЕОМ, яка "не знає" інших математичних операцій, ніж чотири основних арифметичних (+, –, *, /), МСР став основним чисельним методом, яким розв'язувалися крайові задачі.

Зазвичай для оцінювання збіжності розв'язку використовують не одну сітку, а деяку множину сіток $\{\omega_{xyz}\}$ з різноманітними наборами h_{xyz} , тобто множину сіткових функцій, що залежать від параметра h_{xyz} . Ця множина визначена в лінійному просторі H_h , в якому вводиться норма $\|\cdot\|_h$, що враховує крок сітки h_{xyz} і є сітковим аналогом норми $\|\cdot\|$ у безперервному гільбертовому просторі H_0 . Зазвичай приймають норму, яка враховує розмір кроку сітки:

$$\left\|\vec{u}_{h}\right\|_{h} = \left(\sum_{n=1}^{N-1} (u_{h}^{(n)})^{2} h_{xyz}^{(n)}\right)^{1/2}.$$
(15.7)

Основне питання наближення – оцінити близькість двох розв'язків: точного \vec{u}_0 і наближеного \vec{u}_h . Адже це вектори з різних просторів: H_0 і H_h . Необхідно зробити перехід в один з просторів. Частіше здійснюють перехід в H_h . Для цього відображають H_0 на H_h : кожному вектору $\vec{u} \in H_0$ ставиться у відповідність сітковий аналог $\vec{u}_h(x_i, y_j, z_k) \in H_h$, де $x_i, y_j, z_k \in \omega_{xyz}$. Це відображення за формулою (15.5) може бути реалізовано різноманітними методами, наприклад, для *безперервної* функції зазвичай вважають, що $u_h(x_i, y_j, z_k) = u(x_i, y_j, z_k)$, а для функції з *розривами першого роду* – визначають $u_h(x_i, y_j, z_k)$ у вузлі як інтегральне середнє значення *и* по деякому околу, зазвичай діаметром O(h).

Отже, побудувавши відповідно до формули (15.5) відображення для вектора точного розв'язку \vec{u}_{0h} , утворюють векторну різницю $\vec{u}_h(x_i, y_j, z_k) - -\vec{u}_{0h}(x_i, y_j, z_k) \in H_h$. Близькість значень характеризується нормою $\|\vec{u}_h - \vec{u}_{0h}\|_h$, причому доречно вимагати, щоб

$$\lim_{h \to 0} \|\vec{u}_h\|_h = \|\vec{u}_0\|_h \quad \text{afo} \quad \lim_{h \to 0} \|\vec{u}_h - \vec{u}_{0h}\|_h = 0.$$
(15.8)

15.2. Скінченно-різницеве наближення диференційних операторів

Нехай L – лінійний диференційний оператор, який діє на функцію u(x) яку потрібно знайти.

Замінивши похідні, що входять до *Lu*(*x*), різницевими співвідношеннями (див. підрозділ 9.2), отримаємо *алгебраїчне* різницеве наближення:

$$(L_h u_h(x))_i = \sum_{x_j \in III(x_i)} A_h(x_i, x_j) \cdot u_h(x_j), \qquad (15.9)$$

де $A_h(x_i, x_j)$ – коефіцієнти, $III(x_i)$ – шаблон наближення оператора у вузлі x_i , тобто множина вузлів, сусідніх з вузлом x_i і (не обов'язково) самого вузла x_i , в яких значення сіткової функції $u_h(x)$ використані для наближення оператора L. Застосування вузлових шаблонів (основні з них наведені на рис.15.1) вимагає створення не абияких, а впорядкованих сіток.

Отже, скінченно-різницеве наближення заміняє диференційні оператори алгебраїчними виразами (сумами), в яких коефіцієнти – відомі, а невідомими є значення сіткової функції у вузлах. Нагадаємо, що ця заміна повинна проводитися з достатньою точністю та зі збереженням основних властивостей оператора, що наближується.

Говорять, що при достатній гладкості функції u(x) в околі $Ш(x,h_0)$ при $h < h_0$ різницевий оператор L_h наближує диференційний оператор L з поряд-ком m > 0 у вузлі x, якщо

$$\Delta_h = (\Delta(x))_h = L_h u_h(x) - (Lu(x))_h = O(h^m).$$
(15.10)

Оцінюючи похибку наближення на *всій* сітці, вважають, що різницевий оператор L_h наближує на сітці диференційний оператор L з порядком m > 0, якщо

$$\|\Delta_{h}\|_{h} = \|L_{h}u_{h} - (Lu)_{h}\|_{h} = O(|h|^{m}); \qquad (15.11)$$

або

$$\|\Delta_{h}\|_{h} = \|L_{h}u_{h} - (Lu)_{h}\|_{h} \le \alpha |h|^{m}, \qquad (15.12)$$

де $u_h = P_h u$; $(Lu)_h = P_h (Lu)$; P_h – оператор проєктування такий, що $P_h u = u_h \in H_h$ для будь-якого $u \in H_0$; $\alpha > 0$ не залежить від |h|. Під |h| можна розуміти середньоквадратичне або максимальне значення розмірів кроків сітки. Якщо за напрямками (координатами) наближення відрізняються, то потрібно спочатку проводити оцінювання (15.11) або (15.12) за напрямками, після цього вибрати серед отриманих *m* найменше.

Розглянемо основні (класичні) шаблони для диференційних операторів, що використовуються найчастіше.

1. Lu = du/dx. Зафіксуємо вузол x, візьмемо вузли x - h та x + h, де h > 0 (див. рис.15.1-а). Вводяться поняття (застосовано основні *тотожні* варіанти позначень, які використовуються в літературі): *ліва різницева похідна*

$$L_{h}^{-}u_{h} \equiv \frac{u(x) - u(x - h)}{h} \equiv \frac{u(x_{i}) - u(x_{i-1})}{h} \equiv \frac{u_{i} - u_{i-1}}{h} \equiv u_{x}^{-}$$
(15.13)

і права різницева похідна

$$L_{h}^{+}u_{h} \equiv \frac{u(x+h) - u(x)}{h} \equiv \frac{u(x_{i+1}) - u(x_{i})}{h} \equiv \frac{u_{i+1} - u_{i}}{h} \equiv u_{x}^{+}$$
(15.14)

на шаблоні з двох вузлів (просто: ліва і права різниці).

Можна створити лінійну комбінацію

$$L_{h}^{\omega}u_{h} \equiv \omega u_{x}^{+} + (1 - \omega)u_{x}^{-}, \qquad (15.15)$$

де ω – дійсне число, $0 \le \omega \le 1$. При $\omega = 0.5$ отримаємо центральну (двосторонню) різницеву похідну

© К.М. Рудаков, 2025

$$L_{h}^{o}u_{h} \equiv \frac{1}{2}(u_{x}^{+} + u_{x}^{-}) \equiv \frac{u(x+h) - u(x-h)}{2h} \equiv \frac{u(x_{i+1}) - u(x_{i-1})}{2h} \equiv \frac{u_{i+1} - u_{i-1}}{2h} \equiv u_{x}^{o}.$$
 (15.16)

Визначимо похибки різницевого наближення Lu(x) у вузлі x. Для цього припустимо достатню гладкість функції u(x) і розкладемо її в ряд Тейлора:

$$u(x \pm h) = u(x) \pm hu'(x) + \frac{h^2}{2}u''(x) + O(h^3).$$
(15.17)

Підставивши (15.17) у вирази (15.13), (15.14) і (15.16), отримаємо відповідні похибки

$$\Delta_{x}^{-} = u_{x}^{-} - u'(x) = -\frac{h}{2}u''(x) + O(h^{2}) = O(h);$$

$$\Delta_{x}^{+} = u_{x}^{+} - u'(x) = \frac{h}{2}u''(x) + O(h^{2}) = O(h);$$

$$\Delta_{x}^{o} = u_{x}^{o} - u'(x) = O(h^{2}).$$
(15.18)

Рис.15.1 Шаблони різницевого наближення диференційних операторів для: а) – лівої, правої та центральної різниць оператора d/dx; б) – операторів d^2/dx^2 і $d^2/dx^2 + d^2/dy^2$; в) – оператора d^4/dx^4 ; г) – оператора $\partial^2/\partial x \partial y$







2. $Lu = d^2u/dx^2$. Візьмемо шаблон з трьох вузлів (x-h, x, x+h, y)див. рис. 15.1-б). Оскільки $d^2u/dx^2 = d(du/dx)/dx$, то з використанням (15.13) і (15.14):

$$L_{h}u_{h} \equiv \frac{u_{x}^{+} - u_{x}^{-}}{h} \equiv \frac{u(x-h) - 2u(x) + u(x+h)}{h^{2}} \equiv \frac{u(x_{i-1}) - 2u(x_{i}) + u(x_{i+1})}{h^{2}} \equiv \frac{u_{i-1} - 2u_{i} + u_{i+1}}{h^{2}} \equiv u_{xx}^{+-},$$
(15.19)

причому похибка наближення

$$\Delta_{xx}^{+-} = u_{xx}^{+-} - u''(x) = \frac{h^2}{12} u^{(IV)}(x) + O(h^4) = O(h^2).$$
(15.20)

3. $Lu = d^4u/dx^4$. Візьмемо шаблон з п'яти вузлів (x-2h, x-h, x, x+h, x+2h,див. рис.15.1-в). Тоді, оскільки $d^4u/dx^4 = d^2(d^2u/dx^2)/dx^2$, з використанням (15.19)

$$L_{h}u_{h} \equiv \frac{u_{xx}^{+-}(x-h) - 2u_{xx}^{+-}(x) + u_{xx}^{+-}(x+h)}{h^{2}} \equiv \frac{u_{i-2} - 4u_{i-1} + 6u_{i} - 4u_{i+1} + u_{i+2}}{h^{4}} \equiv u_{xxxx}^{+-+-},$$
(15.21)

а похибка наближення

$$\Delta_{xxxx}^{+-+-} = u_{xxxx}^{+-+-} - u^{(IV)}(x) = \frac{h^2}{6} u^{(VI)} + O(h^4) = O(h^2).$$
(15.22)

Примітка 15.1. Виявляється, що розклад похибки типу (15.10) за мірами *h* можна використати для *підвищення порядку наближення*.

Наприклад, з (15.22) одержимо $u^{(IV)}(x) = u_{xxxx}^{+-+-} - O(h^2)$, це підставимо в (15.20), перенесемо $\frac{h^2}{12}u_{xxxx}^{+-+-}$ в ліву частину, отримаємо:

$$\breve{\Delta}_{xx}^{+-} = u_{xx}^{+-} - \frac{h^2}{12} u_{xxxx}^{+-+-} - u''(x) = \breve{u}_{xx}^{+-} - u''(x) = O(h^4), \qquad (15.23)$$

тобто одержимо нове наближення оператора:

$$\vec{u}_{xx}^{+-} = u_{xx}^{+-} - \frac{h^2}{12} u_{xxxx}^{+-+-}.$$
(15.24)

Таким чином, порядок наближення другої похідної можна підвищити до m = 4. Але при цьому шаблон збільшується до п'ятьох вузлів (замість трьох), що призводить при застосуванні (15.24) до зниження інших важливих якостей оператора (збільшується обсяг обчислювальної роботи, ускладняється завдання граничних умов та інше).

Ще один приклад: центрально-різницева апроксимація (15.16) вперше була отримана в результаті підвищення точності апроксимації до m = 2 праворізницевої апроксимації u_x^+ . З другого виразу (15.18): $\breve{\Delta}_x^+ = u_x^+ - \frac{h}{2}u''(x) - u'(x) = O(h^2)$, тому, з урахуванням (15.19): $\breve{u}_x^+ = u_x^+ - \frac{h}{2}u''(x) \approx u_x^+ - \frac{h}{2}u_{xx}^{+-} = \frac{u_{i-1} - u_{i+1}}{2h} = u_x^o$.

4. $Lu = d^2u / dx^2 ha$ нерівномірній сітці. Візьмемо нерегулярний шаблон з трьох вузлів ($x - h^-, x, x + h^+; h^- \neq h^+$). Введемо позначення:

$$\widetilde{u}_{x}^{-} = \frac{u(x) - u(x - h^{-})}{h^{-}}; \quad \widetilde{u}_{x}^{+} = \frac{u(x + h^{+}) - u(x)}{h^{+}}; \quad \hbar = \frac{h^{-} + h^{+}}{2}.$$
(15.25)

Визначимо:

$$L_{h}u_{h} \equiv \frac{\tilde{u}_{x}^{+} - \tilde{u}_{x}^{-}}{\hbar} \equiv \frac{1}{\hbar} \left[\frac{u(x+h^{+}) - u(x)}{h^{+}} - \frac{u(x) - u(x-h^{-})}{h^{-}} \right] \equiv \tilde{u}_{xx}^{+-}.$$
 (15.26)

Можна показати, що похибка наближення

$$\Delta_{xx}^{+-} = \tilde{u}_{xx}^{+-} - u''(x) = \frac{h^+ - h^-}{3} u'''(x) + O(\hbar^2) = O(\hbar), \qquad (15.27)$$

а також вказати, що на практиці вона поступово наближається до $O(\hbar^2)$ при зменшенні різниці між h^+ та h^- .

5. $Lu = \partial^2 u / \partial x^2 + \partial^2 u / \partial y^2 - \partial вовимірний оператор Лапласа. Тут функція <math>u = u(x, y)$. Правила, що використовуються, ідентичні раніше викладеним для одновимірного випадку. Візьмемо шаблони з трьох вузлів $(x - h_x, x, x + h_x)$ і $(y - h_y, y, y + h_y)$ з рівномірними кроками, причому h_x і h_y можуть бути різними (див. рис.15.1-б). Отримаємо:

$$L_{h}u_{h} \equiv \frac{u(x-h_{x}, y) - 2u(x, y) + u(x+h_{x}, y)}{h_{x}^{2}} + \frac{u(x, y-h_{y}) - 2u(x, y) + u(x, y+h_{y})}{h_{y}^{2}} \equiv \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h_{x}^{2}} + \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h_{y}^{2}} \equiv u_{xxyy}^{+-+-} \equiv u_{xx}^{+-} + u_{yy}^{+-}.$$
(15.28)

Справедливі оцінки похибки (15.20) за напрямками х і у.

6. $Lu = \frac{\partial^2 u}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial y} \right)$ – *мішана* похідна. Диференційні оператори з

мішаною похідною теж наближуються за принципом *послідовності дій*. Тому похідна наближується як

$$L_{h}u_{h} \equiv \frac{(u_{y}^{o})_{i+1,j} - (u_{y}^{o})_{i-1,j}}{2h_{x}} \equiv \frac{u_{i-1,j-1} - u_{i-1,j+1} - u_{i+1,j-1} + u_{i+1,j+1}}{4h_{x}h_{y}} \equiv u_{xy}^{oo};$$
(15.29)

має, згідно з (15.18), другі порядки наближення по кожному з напрямків (див. рис.15.1-г).

Примітка 15.2. Для інших ортогональних систем координат застосовують ті ж самі наближення, а у випадках наявності при диференційному операторі множника, в якому фігурує значення координати, застосовують це значення для центрального вузла шаблона. Наприклад, вираз $(1/\rho) \cdot (\partial u / \partial \theta)$ для формул (12.53) апроксимується з другим порядком точності як $(u_{i,j+1,k} - u_{i,j-1,k}) / (\rho_{ijk} 2h_{\theta})$ (тривимірний випадок, циліндричні координати), причому обов'язково $\rho_{ijk} > 0$.

Примітка 15.3. Існують інші варіанти наближень диференційних операторів різницевими виразами. Зокрема, для косокутної сітки, для трикутних сіток з рівносторонніх або нерівносторонніх трикутників, для не ортогональних систем координат, з підвищеними порядками наближення, тощо.

15.3. Скінченно-різницеве наближення крайових задач

Для отримання скінченно-різницевого наближення крайової задачі (сіткового аналога крайової задачі) потрібно зробити такі дії:

1) створити сіткове наближення геометричної моделі тіла;

2) обрати сіткову функцію (або декілька сіткових функцій);

3) провести наближення похідних, що є в рівняннях крайовій задачі (включно з граничними умовами) алгебраїчними виразами достатньої точності;

4) записати скінченно-різницеве наближення рівнянь крайової задачі у вузлі (сітковий аналог рівнянь крайової задачі в елементарному об'ємі);

5) отримані в п. 4) наближення рівнянь записати в кожному *i*-му вузлі тіла (шляхом "накладення шаблонів"). Це буде САР відносно невідомих значень сіткових функцій. Вивчити властивості САР та обрати метод її розв'язування.

Розглянемо різницеве наближення крайових задач на прикладі одновимірних задач й двовимірної задачі.

15.3.1. Одновимірна лінійна крайова задача Діріхле

Припустимо, що необхідно знайти функцію u(x) на відрізку $a \le x \le b$ таку, що задовольнить диференційному рівнянню

$$Lu = f \tag{15.30}$$

з граничними умовами тільки 1-го роду (перша основна крайова задача, або задача Діріхле)

$$u(a) = u_a; \quad u(b) = u_b,$$
 (15.31)

де диференційний оператор L приймемо, наприклад, у вигляді

$$L = d^2 / dx^2. (15.32)$$

Така задача відповідає крайовій задачі про стаціонарну теплопровідність для стрижня з постійним коефіцієнтом теплопровідності λ , розподіленим джерелом тепла $\omega = -f\lambda$ і при підтриманні заданої температури кінців стрижня.

Для алгебраїзації крайової задачі введемо сітку з рівномірним кроком $h_x = h = (b-a)/(N-1)$, де N – кількість вузлів, враховуючи граничні. Після цього перейдемо від диференційного рівняння (15.30) до скінченно-різницевого

$$L_h u_h = f_h; \quad L_h u_h = u_{xx}^{+-}; \quad \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} = f(x_i) = f_i$$
 (15.33)

з граничними умовами 1-го роду

$$u_h(a) = u_a; \quad u_h(b) = u_b.$$
 (15.34)

Очевидно, що рівняння (15.33) може бути застосовано лише до *внутрішніх* вузлів, тобто при i = 2, 3, ..., N - 1. Але в задачі Діріхле значення функції в *граничних* вузлах відомі (задані в ГУ 1-го роду). Тому (15.33) з урахуванням (15.34) являє собою систему лінійних алгебраїчних рівнянь (СЛАР) вигляду

$$(A_{mn})_{h}(u_{n})_{h} = (f_{m})_{h}$$
 abo $A_{h}\vec{u}_{h} = \vec{f}_{h}$ (15.35)

з (N-2) невідомими.

Якщо прийняти N = 6, то відповідно до (15.33) при i = 2, 3, 4, 5:

 $u_1 - 2u_2 + u_3 = f_2h^2$; $u_2 - 2u_3 + u_4 = f_3h^2$; $u_3 - 2u_4 + u_5 = f_4h^2$; $u_4 - 2u_5 + u_6 = f_5h^2$, причому із ГУ (15.34) $u_1 = u_a$ та $u_6 = u_b$. Отже, складові СЛАР, породженої МСР, мають вигляд:

$$A_{h} = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}; \quad \vec{u}_{h} = \begin{cases} u_{2} \\ u_{3} \\ u_{4} \\ u_{5} \end{cases}; \quad \vec{f}_{h} = \begin{cases} -h^{2}f_{2} + u_{a} \\ -h^{2}f_{3} \\ -h^{2}f_{4} \\ -h^{2}f_{5} + u_{b} \end{cases}.$$
 (15.36)

Примітка 15.4. Матриця A_h системи є симетричною, позитивно визначеною та розрідженою, в даному випадку – тридіагональною. Це дозволяє використати для розв'язування СЛАР *метод прогонки* (див. підрозділ 5.4).

15.3.2. Одновимірна змішана лінійна крайова задача

Припустимо, що необхідно знайти функцію u(x) на відрізку $a \le x \le b$, що задовольнить диференційному рівнянню (15.30) з граничними умовами, серед яких є ГУ не тільки 1-го, а й 2-го роду, які ще мають назву природних (третя основна задача, або змішана задача):

$$u(a) = u_a; \quad Ku\Big|_{x=b} = g_b,$$
 (15.37)

де диференційний оператор *L* приймемо у вигляді (15.32), а K = d/dx. Така задача відповідає крайовій задачі про стаціонарну теплопровідність для стрижня з постійним коефіцієнтом теплопровідності λ , розподіленим джерелом тепла $\omega = -f\lambda$, при підтриманні заданої температури одного кінця стрижня і при заданому тепловому потоці через інший кінець стрижня.

Для алгебраїзації крайової задачі знову введемо сітку з рівномірним кроком $h_x = h = (b-a)/(N-1)$, де N – кількість вузлів, враховуючи граничні. Після цього перейдемо від диференційного рівняння (15.30) до скінченно-різницевого (15.33) з граничними умовами 1-го та 2-го роду

$$u_h(a) = u_a; \quad K_h u_h \Big|_{x=b} = g_b.$$
 (15.38)

Наближення (15.33) може бути застосоване лише до *внутрішніх* вузлів. В *граничному* вузлі з x = a значення функції відоме (задане в ГУ 1-го роду). Оператор *К* ми можемо наблизити *лівою різницевою похідною*:

$$K_h^- u_h = u_x^- = (u_N - u_{N-1}) / h.$$
(15.39)

Застосувавши (15.33) з урахуванням (15.38) і (15.39), отримаємо систему рівнянь вигляду (15.35) з (N-1) невідомими. Якщо знову прийняти N = 6, то складові системи мають вигляд:

$$A_{h} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}; \quad \vec{u}_{h} = \begin{cases} u_{2} \\ u_{3} \\ u_{4} \\ u_{5} \\ u_{6} \end{cases}; \quad \vec{f}_{h} = \begin{cases} -h^{2}f_{2} + u_{a} \\ -h^{2}f_{3} \\ -h^{2}f_{4} \\ -h^{2}f_{5} \\ hg_{b} \end{cases}.$$
(15.40)

Врахування ГУ 2-го роду можна провести, зберігши розмірність системи, що дорівнює (N-2). Використовуючи (15.38) і (15.39), отримаємо, що $u_6 = u_5 + hg_b$. Підставивши в (15.33), отримаємо СЛАР

$$A_{h} = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}; \quad \vec{u}_{h} = \begin{cases} u_{2} \\ u_{3} \\ u_{4} \\ u_{5} \end{cases}; \quad \vec{f}_{h} = \begin{cases} -h^{2}f_{2} + u_{a} \\ -h^{2}f_{3} \\ -h^{2}f_{4} \\ -h^{2}f_{5} + hg_{b} \end{cases}.$$
 (15.41)

Недолік наближення задачі у вигляді (15.40) і (15.41) в тому, що низький рівень точності наближення граничної умови 2-го роду (O(h)) призводить до зниження точності наближення всієї задачі з $O(h^2)$ до O(h).

Означений недолік можна усунути, ввівши додатковий, *фіктивний* (N+1)-й вузол з координатою $x_N + h$. Тоді й для граничного вузла N використаємо наближення (15.33), тобто $u_{N-1} - 2u_N + u_{N+1} = f_N h^2$, а для ГУ 2-го роду застосуємо центральне різницеве наближення (15.16):

$$K_{h}^{o}u_{h} = \frac{u_{N+1} - u_{N-1}}{2h} = g_{b}$$
(15.42)

з порядком наближення $O(h^2)$. Для вилучення з розгляду значення u_{N+1} у фіктивному вузлі виразимо з (15.42) $u_{N+1} = u_{N-1} + 2hg_b$ і підставимо в (15.33) при i = N, тобто отримаємо рівняння $u_{N-1} - 2u_N + u_{N-1} + 2hg_b = f_N h^2$ при N = 6. Остаточно отримаємо систему рівнянь (15.35) зі складовими

$$A_{h} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}; \quad \vec{u}_{h} = \begin{cases} u_{2} \\ u_{3} \\ u_{4} \\ u_{5} \\ u_{6} \end{cases}; \quad \vec{f}_{h} = \begin{cases} -h^{2}f_{2} + u_{a} \\ -h^{2}f_{3} \\ -h^{2}f_{4} \\ -h^{2}f_{5} \\ -0.5h^{2}f_{6} + hg_{b} \end{cases}.$$
(15.43)

Ми знов повернулися до збільшеної системи з (N-1) невідомими, але наближення всіх рівнянь крайової задачі має один і той же високий порядок $O(h^2)$, тобто і наближення всієї задачі $O(h^2)$.

15.3.3. Одновимірна нелінійна крайова задача Діріхле

Припустимо, що необхідно знайти функцію u(x) на відрізку $a \le x \le b$, що задовольняє нелінійному диференційному рівнянню

$$Lu = \frac{d}{dx} \left(\lambda(u) \frac{du}{dx} \right) = f$$
(15.44)

з граничними умовами 1-го роду

$$u(a) = u_a; \quad u(b) = u_b.$$
 (15.45)

У (15.44) $\lambda(u)$ – коефіцієнт теплопровідності, що залежить від u (температури), тобто розглядаємо стаціонарну нелінійну задачу теплопровідності. Покажемо, як МСР наближує нелінійну крайову задачу.

Позначимо $v = \lambda(u) \cdot du/dx$. Введемо для Gv = dv/dx центральну різницеву похідну на половині кроку

$$G_{h}^{o}v_{h} = \frac{v(x+h/2) - v(x-h/2)}{h} = \frac{v(x_{i+1/2}) - v(x_{i-1/2})}{h} = \frac{v_{i+1/2} - v_{i-1/2}}{h} = v_{x}^{o}$$
(15.46)

і застосуємо її для основного рівняння (15.44). Отримаємо:

$$L_{h}u_{h} = G_{h}^{o}v_{h} = \frac{\lambda(u_{i+1/2}) \cdot \left(\frac{du}{dx}\right)_{i+1/2} - \lambda(u_{i-1/2}) \cdot \left(\frac{du}{dx}\right)_{i-1/2}}{h} = f_{i}.$$
 (15.47)

Замінимо $(du/dx)_{i\pm 1/2}$ відповідними центрально-різницевими похідними вигляду (15.16), отримаємо:

$$L_{h}u_{h} = \frac{\lambda(u_{i-1/2}) \cdot u_{i-1} - (\lambda(u_{i+1/2}) + \lambda(u_{i-1/2})) \cdot u_{i} + \lambda(u_{i+1/2}) \cdot u_{i+1}}{h^{2}} = f_{i}.$$
 (15.48)

Застосовуючи (15.48) для всіх внутрішніх вузлів і (15.45) – в граничних, отримаємо *нелінійну* систему алгебраїчних рівнянь вигляду

$$A_{h}(\vec{u}_{h})\vec{u}_{h} = \vec{f}_{h}$$
(15.49)

відносно (*N*-2) невідомих. Нелінійна система (15.49) може бути розв'язана стандартними *ітераційними* методами. За методом простих ітерацій

$$A_h(\vec{u}_h^{(k-1)})\vec{u}_h^{(k)} = \vec{f}_h, \qquad (15.50)$$

де k = 1, 2, ... - номер ітерації.

15.3.4. Двовимірна змішана лінійна крайова задача

Розглянемо задачу про кручення пружного стрижня з прямокутним перетином. Оскільки в перетині є дві осі симетрії, розглянемо чверть перетину, тобто областю розгляду візьмемо прямокутник $0 \le x \le a$ й $0 \le y \le b$. В цій області необхідно знайти *функцію кручення Прандтля u*(*x*, *y*), яка задовольняє диференційному рівнянню типу (15.30) у вигляді:

$$Lu = \partial^2 u / \partial x^2 + \partial^2 u / \partial y^2 = -2, \qquad (15.51)$$

з граничними умовами 1-го та 2-го роду (третя основна задача, або змішана задача)

$$u(a, y) = u(x, b) = 0; \quad \partial u / \partial x \mid_{x=0} = \partial u / \partial y \mid_{y=0} = 0; \quad (15.52)$$

причому другий вираз в ГУ випливає з умов симетрії розв'язку відносно осей *х* та *у*. Після цього потрібно знайти відповідний заданому куту скручування стрижня *θ* крутильний момент:

$$M_{KP} = 2G\theta \int_{\Omega} u(x, y) dx dy, \qquad (15.53)$$

де G – модуль зсуву. При необхідності дотичні напруження в перетині:

$$\tau_{zx} = G\theta \frac{\partial u(x, y)}{\partial y}; \quad \tau_{zy} = -G\theta \frac{\partial u(x, y)}{\partial x}.$$
(15.54)

Для алгебраїзації крайової задачі введемо сітку з рівномірними за напрямками кроками $h_x = a/(N_x - 1)$ і $h_y = b/(N_y - 1)$, де N_x , $N_y -$ кількості вузлів уздовж координатних осей, враховуючи граничні вузли. Вузли будемо нумерувати так:

204

 $i = 0,1,...,N_x - 1; j = 0,1,...,N_y - 1.$ Перейдемо від диференційного рівняння (15.51) до скінченно-різницевого:

$$L_{h}u_{h} = u_{xx}^{+-} + u_{yy}^{+-} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h_{x}^{2}} + \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h_{y}^{2}} = -2$$
(15.55)

з граничними умовами 1-го та 2-го роду

$$u_{h}(a, y) = u_{h}(x, b) = 0; \quad (\partial u / \partial x)_{h} |_{x=0} = (\partial u / \partial y)_{h} |_{y=0} = 0.$$
(15.56)

Наближення (15.55) може бути застосоване лише до *внутрішніх* вузлів перетину, тобто при $i=1,...,N_x-2$; $j=1,...,N_y-2$. В *граничному* вузлі з x=a значення функції відоме (задане в ГУ 1-го роду). ГУ 2-го роду для збереження другого порядку наближення ми можемо наблизити центрально-різницевими похідними (з фіктивними вузлами)

$$u_{x}^{-}\Big|_{x=0} = \frac{u_{1,j} - u_{-1,j}}{2h_{x}} = 0; \ j = 0, ..., N_{y} - 1; \quad u_{y}^{-}\Big|_{y=0} = \frac{u_{i,1} - u_{i,-1}}{2h_{y}} = 0, \ i = 0, ..., N_{x} - 1, \ (15.57)$$

з яких є очевидними рівності

$$u_{-1,j} = u_{1,j} \Big|_{x=0}; \quad u_{i,-1} = u_{i,1} \Big|_{y=0},$$
 (15.58)

що дозволять виключити із співвідношень (15.55) значення функції у фіктивних вузлах. Введення фіктивних вузлів дозволяє збільшити кількість вузлів, в яких можна застосувати наближення (15.55): при $i = 0, ..., N_x - 2$; $j = 0, ..., N_y - 2$.

Позначимо: $s = -1/h_x^2$; $t = -1/h_y^2$; $q = 2(1/h_x^2 + 1/h_y^2) = -2(s+t)$. Приймемо $N_x = 4$ та $N_y = 3$. На рис.15.3 для цього випадку показано сітку вузлів задачі, зокрема й фіктивні вузли (помічені як ×).

В підсумку отримаємо систему лінійних алгебраїчних рівнянь

$$\begin{bmatrix} q & 2s & 0 & 2t & 0 & 0 \\ s & q & s & 0 & 2t & 0 \\ 0 & s & q & 0 & 0 & 2t \\ t & 0 & 0 & q & 2s & 0 \\ 0 & t & 0 & s & q & s \\ 0 & 0 & t & 0 & s & q \end{bmatrix} \cdot \begin{bmatrix} u_{0,0} \\ u_{1,0} \\ u_{2,0} \\ \vdots \\ u_{1,1} \\ u_{2,1} \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix}.$$
(15.59)

Очевидно, що матриця системи є несиметричною, що викликане введенням фіктивних вузлів для збереження другого порядку наближення всієї задачі.

Прийнявши $h_x = h_y = 1$, отримаємо вектор-розв'язок

$$(u_{0,0}, u_{1,0}, u_{2,0}, u_{0,1}, u_{1,1}, u_{2,1})^T =$$

= (3.1370, 2.8866, 1.9971, 2.3873, 2.2062, 1.5508). (15.60)

Відомо, що точного аналітичного розв'язку задача не має, але існують розв'язки, наприклад, із застосуванням тригонометричних рядів. Використовуючи формулу трапецій при інтегруванні (15.53), можна отримати при $G\theta = 1$, що $M_{KP} \approx 65.41$ (значення розв'язку із застосуванням тригонометричних рядів дорівнює 76.40). Застосовуючи наближення лівою різницею, можна отримати





 $|\tau_{\text{max}}| \approx 2.39$ (у вузлу 0.2, тобто посередині довгої сторони перетину; розв'язок із застосуванням тригонометричних рядів дає значення 2.96).

15.4. Збіжність розв'язку методу скінченних різниць. Поняття про екстраполяцію за Річардсоном. Переваги та недоліки методу

Очевидно, що для покращення результатів, одержаних за допомогою MCP, необхідно згущати сітку. Для

оцінювання точності розв'язку можна використати розв'язок "в рядах" – вектор значень типу (15.60) у вузлах, відповідних до скінченно-різницевого наближення, а також формули оцінки похибки наближення на всій сітці (15.11) або (15.12).

Маючи не менше двох розв'язків МСР на сітках із різним згущенням (розв'язки \vec{u}_1 і \vec{u}_2), можна провести так звану *екстраполяцію за Річардсоном* (Lewis Fry Richardson, 1881-1953 рр.) для отримання більш точних значень (авторська назва: екстраполяція до межі). В кожному вузлі екстрапольоване значення:

$$u \approx \frac{u_1 h_2^{\ m} - u_2 h_1^{\ m}}{h_2^{\ m} - h_1^{\ m}},$$
(15.61)

де h_1 і h_2 – кроки по сітці розв'язків 1 і 2 (середні або максимальні з h_x і h_y); m – порядок наближення МСР. Ця формула – не точна, не цілком коректна, однак, як правило, дає краще наближення. Спочатку записується, що відношення норм векторів похибок наближення

$$\frac{\|\vec{u} - \vec{u}_1\|}{\|\vec{u} - \vec{u}_2\|} \approx \frac{O(h_1^m)}{O(h_2^m)} = f(h_1, h_2) = g(h_1, h_2) \cdot \left(\frac{h_1}{h_2}\right)^m.$$
(15.62)

3 припущення, що функція $g(h_1, h_2) \approx 1$, вираз (15.62) спрощується:

$$\frac{\|\vec{u} - \vec{u}_1\|}{\|\vec{u} - \vec{u}_2\|} \approx \left(\frac{h_1}{h_2}\right)^m.$$
(15.63)

З (15.63), для кожного вузла, випливає формула (15.61). Якщо $h_2 = h_1 / n$, де n є деяким числом, то з (15.61) легко отримати, що

$$u \approx \frac{n^m u_2 - u_1}{n^m - 1}.$$
 (15.64)

Оскільки вузли в сітках з різними кроками в загальному випадку не співпадають, то для застосування (15.61) додатково ще необхідно визначитися з

206

новою сіткою (для нового розв'язку \vec{u}), перенести на неї розв'язки \vec{u}_1 та \vec{u}_2 , а лише потім застосувати формулу (15.61) в кожному вузлі сітки.

Сучасна теорія екстраполяції за Річардсоном складна і велика (див., наприклад, главу 6 книги [38]). А сама ідея екстраполяції за Річардсоном базується на тому, що розв'язки для деякої функції можна представити як екстрапольовані нескінченними рядами з різними кроками, але з однаковими коефіцієнтами та степенями при цих кроках:

$$u = u(h) + a_1 \cdot h^{m_1} + a_2 \cdot h^{m_2} + \dots; \qquad (15.65)$$

$$u = u(h / n) + a_1 \cdot (h / n)^{m_1} + a_2 \cdot (h / n)^{m_2} + \dots$$
(15.66)

Якщо позначити $h_1 = h$; $h_2 = h_1 / n = h / n$; $u_1 = u(h)$; $u_2 = u(h_2) = u(h / n)$ та прийняти, що всі $a_i = 0$ при i > 1 та $m_1 = m$, то залишаться вирази $u \approx u_1 + a_1 \cdot (h_1)^m$ та $u \approx u_2 + a_1 \cdot (h_2)^m$, з яких $(u - u_1) / (u - u_2) \approx (a_1 / a_1) \cdot (h_1 / h_2)^m = (h_1 / h_2)^m$ (див. формулу (15.63)).

Сформулюємо очевидні переваги і недоліки методу скінченних різниць.

Переваги МСР:

• метод дозволяє фактично не збирати в явному вигляді матрицю САР, не зберігати її і не виділяти для неї операційну пам'ять ЕОМ, оскільки матриця має однотипне заповнення, обумовлене обраним шаблоном, геометрією тіла та характером граничних умов. Це дозволяє розв'язувати величезні за кількістю невідомих системи рівнянь на ЕОМ з малим розміром оперативної пам'яті;

• МСР добре досліджений теоретично;

• діагональний характер заповнення матриці САР дозволяє застосовувати спеціалізовані методи розв'язування систем рівнянь.

Недоліки МСР:

• не має універсального формулювання граничних умов;

• реалізація методу на нерегулярних сітках вимагає введення складних шаблонів і ускладненої логіки при складанні системи рівнянь; точність наближення на сітках зі змінними кроками є меншою, ніж при однакових.

• локальне згущення сітки не є раціональним, тому розв'язування задач з концентраторами напружень є ускладненим.

Завдяки своїм перевагам МСР був практично єдиним чисельним методом, за допомогою якого вдавалося розв'язувати актуальні задачі у ті часи, коли ЕОМ були малопотужними. Для довідки: легендарна ЕОМ другої генерації "БЭСМ-6" мала для користувача всього 32 кілобайта оперативної пам'яті, але досить велику швидкодію (біля 1 млн. машинних операцій за секунду) і швидкі накопичувачі барабанного типу (з непорушними магнітними головками). Використовували блочні методи розв'язування систем рівнянь, коли в ОЗП (оперативному запам'ятовуючому пристрої) формувалася лише необхідна на даному етапі частина (блок) системи рівнянь, проводилися обчислення, результати запам'ятовувалися на магнітних носіях, після цього формувався новий блок та проводилися обчислення. МСР був розвинений на некласичні випадки, зокрема для розв'язування задач з великими деформаціями і переміщеннями, з винесенням матеріалу із зони деформування. Наприклад, задачі про пробивання броні бронебійним або кумулятивним снарядом. При цьому кроки скінченно-різницевої сітки суттєво різні, змінні у часі, а геометрія – достатньо довільна, причому безупинно змінюється. За допомогою МСР було розв'язано величезну кількість задач, пов'язаних з опануванням атомною енергією і космосом. Адже пік зацікавленості цими проблемами прийшовся якраз на той час, коли ЕОМ вже були, але їхні можливості порівняно з сучасними здаються дуже мізерними. Допомогли унікальні переваги МСР. Тому заслуги МСР переоцінити просто неможливо.

Зараз МСР значною мірою витіснено з розрахункової практики внаслідок появи потужної обчислювальної техніки і чисельних методів, які не мають недоліків, характерних для МСР. Насамперед це метод скінченних елементів (МСЕ). Його ми будемо розглядати починаючи з Розділу 19.

Контрольні питання до підрозділу 15.1

1. Сформулюйте ідею методу скінченних різниць.

2. Для чого вводять оператор проєктування?

3. Що таке шаблон диференційного оператора?

4. Які диференційні скінченно-різницеві оператори мають перший порядок наближення, а які – другого?

Контрольні питання до підрозділу 15.2

- 1. Назвіть етапи скінченно-різницевого наближення крайових задач.
- 2. Які особливості має скінченно-різницеве наближення крайової задачі Діріхле?
- 3. Які особливості має скінченно-різницеве наближення змішаної крайової задачі?
- 4. Для чого використовують так звані "фіктивні" вузли?

Контрольні питання до підрозділу 15.3

- 1. Для чого використовують екстраполяцію за Річардсоном?
- 2. Назвіть переваги та недоліки методу скінченних різниць.

ПРОСТОРОВО-ЧАСОВЕ НАБЛИЖЕННЯ НЕСТАЦІОНАРНИХ КРАЙОВИХ ЗАДАЧ ЗА МЕТОДОМ СКІНЧЕННИХ РІЗНИЦЬ

16.1. Загальні зауваження

Просторово-часовий характер мають задачі, в яких на деякий початковий момент часу стан системи тіл (тіла) є відомим, необхідно визначити її стани в наступні моменти часу. У Розділі 13 було поставлено дві характерні крайові задачі, що залежать від часового аргументу:

• задача нестаціонарної теплопровідності, яка описується в об'ємі рівнянням *параболічного*_типу

$$\frac{\partial \vec{u}}{\partial t} + L(\vec{u}) = \hat{\vec{f}}$$
(16.1)

з ГУ вигляду

$$K\vec{u} = \hat{\vec{g}} \tag{16.2}$$

за початкових умов

$$\vec{u}(\vec{x},0) = \hat{\vec{u}}(\vec{x}).$$
 (16.3)

• задача про динамічну поведінку тіла, що деформується, яка описується в об'ємі рівнянням *гіперболічного* типу

$$M\frac{\partial^2 \vec{u}}{\partial t^2} + C\frac{\partial \vec{u}}{\partial t} + L(\vec{u}) = \hat{\vec{f}}$$
(16.4)

з ГУ вигляду (16.2) за початкових умов

$$\vec{u}(\vec{x},0) = \hat{\vec{u}}_0(\vec{x}); \quad \frac{\partial \vec{u}(\vec{x},0)}{\partial t} = \hat{\vec{h}}(\vec{x}).$$
(16.5)

Просторово-часова алгебраїзація задачі може проводитися:

• методом скінченних різниць (МСР);

• методом повної або часткової алгебраїзації з використанням наступних типів базисних функцій: точкових, кусково-визначених або неперервних.

За методом часткової алгебраїзації для часової редукції застосовується МСР або пряме інтегрування у часі, а для просторової – МСР або перераховані вище типи базисних функцій. У методі повної алгебраїзації вся редукція проводиться на основі базисних функцій. Він застосовується порівняно рідко (використаємо його в підрозділі 25.2).

Час позначимо як параметр t, який не зменшується, з початковим значенням $t_0 = 0$. Графічно це часова вісь з часовими точками (шарами), "відстань" між якими визначається часовими кроками $\Delta t_n = t_{n+1} - t_n$. Значення часу шару $t_n = t_0 + \sum_{k=1}^{n} \Delta t_k$. Поточний час можна визначити як

Частина IV. Розділ 16		
	$t = t_n + \omega \Delta t_n = (1 - \omega)t_n + \omega \cdot t_{n+1},$	(16.6)

де 0 ≤ ω ≤1. Вираз (16.6) можна переписати у вигляді

$$\omega = (t - t_n) / (t_{n+1} - t_n) = (t - t_n) / \Delta t_n.$$
(16.7)

Використання вагового коефіцієнта ω дозволяє одержувати в рамках єдиного запису різні за властивостями схеми алгебраїзації (наближення).

16.2. Основні наближення для часових диференційних операторів

Щоб проводити алгебраїзацію крайових задач МСР, вводяться наближення як просторових, так і часових диференційних операторів. Можна навіть не в водити наближення останніх, а застосовувати пряме інтегрування у часі, причому цей підхід є більш універсальним при побудові схем алгебраїзації крайових задач і буде використовуватися надалі. Але все ж основні наближення для часових диференційних операторів потрібно знати.

Вони є подібними різницевим наближенням просторових диференційних операторів, розглянутих у підрозділі 15.2, тому нічого принципово нового тут немає. Нагадаємо (див. підрозділ 15.1), що ідея скінченно-різницевого наближення (МСР) така: диференційні оператори заміняються алгебраїчними виразами, в яких коефіцієнти відомі, а невідомими є значення сіткової функції у вузлах.

Нехай \mathfrak{T} – часовий диференційний оператор, що діє на функцію $u(\vec{x},t)$. Замінивши різницевим співвідношенням похідну, що входить до $\mathfrak{I}u(\vec{x},t)$, отримаємо різницевий вираз (наближення):

$$(\mathfrak{T}_{t}\boldsymbol{u}_{t}(\vec{x},t))_{n} = \sum_{t_{m}\in\mathcal{III}(t_{n})} A_{t}(t_{m},t_{n}) \cdot \boldsymbol{u}_{t}(\vec{x},t_{m}), \qquad (16.8)$$

де $A_t(t_m, t_n)$ – відомі коефіцієнти, III(t) – шаблон у момент часу t.

Перш ніж проводити різницеве наближення, необхідно вибрати шаблон у момент часу t_n , тобто часові шари, сусідні з часом t_n , в яких значення сіткової функції $u_i(\vec{x},t)$ буде використано для наближення оператора \mathfrak{I} .

Розглянемо основні скінченно-різницеві наближення для часових диференційних операторів.

1. $\Im u = \partial u / \partial t$. Зафіксуємо час t_n , візьмемо час $t_n - \Delta t_{n-1}$ і $t_n + \Delta t_n$, де $\Delta t_{n-1}, \Delta t_n > 0$. Вводяться поняття: *ліва* різницева похідна

$$\mathfrak{T}_{t} u_{t} = \frac{u(\vec{x}, t_{n}) - u(\vec{x}, t_{n} - \Delta t_{n-1})}{\Delta t_{n-1}} = \frac{u(\vec{x}, t_{n}) - u(\vec{x}, t_{n-1})}{\Delta t_{n-1}} = \frac{u^{n} - u^{n-1}}{\Delta t_{n-1}} = u_{t}^{-}$$
(16.9)

і права різницева похідна

210

$$\mathfrak{T}_{t}^{+}u_{t} \equiv \frac{u(\vec{x}, t_{n} + \Delta t_{n}) - u(\vec{x}, t_{n})}{\Delta t_{n}} \equiv \frac{u(\vec{x}, t_{n+1}) - u(\vec{x}, t_{n})}{\Delta t_{n}} \equiv \frac{u^{n+1} - u^{n}}{\Delta t_{n}} \equiv u_{t}^{+}$$
(16.10)

на шаблоні з двох точок (просто: ліва і права різниці).

Примітка 16.1. Індекси, що позначають часовий крок, при невідомій функції (або векторі) у МСР зазвичай пишуть зверху, оскільки при цій функції ще є нижні індекси, пов'язані з просторовими вимірами.

Можна розглядати лінійну комбінацію (схему з ваговим коефіцієнтом)

$$\mathfrak{T}_t^{\omega} u_t \equiv \omega u_t^+ + (1 - \omega) u_t^-, \tag{16.11}$$

де $0 \le \omega \le 1 - д$ ійсне число. Якщо прийняти $\Delta t_{n-1} = \Delta t_n = \Delta t$, то

$$\Im_{t}^{\omega}u_{t} = \frac{\omega u(\vec{x}, t_{n} + \Delta t_{n}) + (1 - 2\omega)u(\vec{x}, t_{n}) - (1 - \omega)u(\vec{x}, t_{n} - \Delta t_{n})}{\Delta t} = \frac{\omega u^{n+1} + (1 - 2\omega)u^{n} - (1 - \omega)u^{n-1}}{\Delta t} = u_{t}^{\omega}.$$
(16.12)

При $\omega = 0.5$ з (16.12) отримаємо центральну (двосторонню) різницеву похідну

$$\mathfrak{T}_{t}^{o}u_{t} = \frac{u(\vec{x}, t_{n} + \Delta t) - u(\vec{x}, t_{n} - \Delta t)}{2\Delta t} = \frac{u(\vec{x}, t_{n+1}) - u(\vec{x}, t_{n-1})}{2\Delta t} = \frac{u^{n+1} - u^{n-1}}{2\Delta t} = \frac{1}{2}(u_{t}^{+} + u_{t}^{-}) = u_{t}^{o}.$$
(16.13)

Маємо повну подібність з просторовими різницевими операторами (див. підрозділ 15.2). Тому оцінки похибок запишемо відразу:

$$\Delta_t^- = u_t^- - \dot{u}(\vec{x}, t) = -\frac{\Delta t_n}{2} \ddot{u}(\vec{x}, t) + O(\Delta t^2) = O(\Delta t); \qquad (16.14)$$

$$\Delta_{t}^{+} = u_{t}^{+} - \dot{u}(\vec{x}, t) = \frac{\Delta t}{2} \ddot{u}(\vec{x}, t) + O(\Delta t^{2}) = O(\Delta t); \qquad (16.15)$$

$$\Delta_{t}^{\omega} = u_{t}^{\omega} - \dot{u}(\vec{x}, t) = (\omega - 0.5)\Delta t \ddot{u}(\vec{x}, t) + O(\Delta t^{2}) = O(\Delta t); \quad \omega \neq 0.5;$$
(16.16)

$$\Delta_t^0 = u_t^0 - \dot{u}(\vec{x}, t) = O(\Delta t^2).$$
(16.17)

2. $\Im u = \partial^2 u / \partial t^2$. Візьмемо шаблон з трьох точок $(t - \Delta t, t, t + \Delta t)$. Тоді з використанням (16.9) і (16.10)

$$\Im_{t}u_{t} = \frac{u_{t}^{+} - u_{t}^{-}}{\Delta t} = \frac{u(\vec{x}, t + \Delta t) - 2u(\vec{x}, t) + u(\vec{x}, t - \Delta t)}{\Delta t^{2}} = \frac{u(\vec{x}, t^{n+1}) - 2u(\vec{x}, t^{n}) + u(\vec{x}, t^{n-1})}{\Delta t^{2}} = \frac{u^{n+1} - 2u^{n} + u^{n-1}}{\Delta t^{2}} = u_{tt}^{+-},$$
(16.18)

причому похибка наближення на рівномірному кроці Δt :

$$\Delta_{tt}^{+-} = u_{tt}^{+-} - \ddot{u}(\vec{x}, t) = \frac{\Delta t^2}{12} \ddot{\ddot{u}}(\vec{x}, t) + O(\Delta t^4) = O(\Delta t^2).$$
(16.19)

3. $\Im u = \partial^2 u / \partial t^2$ на нерівномірній сітці. Візьмемо нерегулярний триточковий шаблон $(t - \Delta t_n, t, t + \Delta t_{n+1}; \Delta t_n \neq \Delta t_{n+1})$. Введемо позначення

$$\widetilde{u}_{t}^{-} = \frac{u(\vec{x},t) - u(\vec{x},t - \Delta t_{n})}{\Delta t_{n}}; \quad \widetilde{u}_{t}^{+} = \frac{u(\vec{x},t + \Delta t_{n+1}) - u(\vec{x},t)}{\Delta t_{n+1}}; \quad \Delta \widetilde{t} = \frac{\Delta t_{n} + \Delta t_{n+1}}{2}. \quad (16.20)$$

Визначимо апроксимацію оператора:

$$\Im_{t}u_{t} = \frac{\tilde{u}_{t}^{+} - \tilde{u}_{t}^{-}}{\Delta \tilde{t}} = \frac{1}{\Delta \tilde{t}} \left[\frac{u(\vec{x}, t + \Delta t_{n+1}) - u(\vec{x}, t)}{\Delta t_{n+1}} - \frac{u(\vec{x}, t) - u(\vec{x}, t - \Delta t_{n})}{\Delta t_{n}} \right] = \tilde{u}_{tt}^{+-}.$$
 (16.21)

Можна показати, що похибка такого наближення

$$\tilde{\Delta}_{tt}^{+-} = \tilde{u}_{tt}^{+-} - \ddot{u}(\vec{x}, t) = \frac{\Delta t_{n+1} - \Delta t_n}{3} \ddot{u}(\vec{x}, t) + O(\Delta \tilde{t}^2) = O(\Delta \tilde{t}).$$
(16.22)

© К.М. Рудаков, 2025

16.3. Основні вимоги до просторово-часових схем

Перш ніж перейти до конкретного розгляду чисельних схем просторовочасового наближення крайових задач, сформулюємо основні вимоги до просторово-часових схем.

Поставлена крайова задача повинна мати якісний розв'язок, тобто схема наближення й алгоритм повинні відповідати певним критеріям, а саме:

• точності;

- стійкості просторово-часового оператора при обчисленнях;
- відсутності осциляцій розв'язку;
- ефективності алгоритму.
- Говорять, що задачу поставлено коректно, якщо:

• задача має однозначні розв'язки при будь-яких початкових даних з певного класу;

• розв'язок задачі безперервно залежить від початкових даних (задача А-стійка).

Під початковими даними розуміють початкові та граничні умови, а також праві частини САР, що створюються в наслідок алгебраїзації задачі.

Існують задачі, які в принципі не можна коректно поставити – некоректні задачі, для розв'язку яких розроблено відповідні методи. Але вони не характерні для динаміки і міцності машин, тому некоректні задачі тут не розглядаються.

Наближення крайової задачі теж є коректним за тих же умов, що й вихідне її формулювання (див. вище).

Доведено *теорему*: якщо для крайової задачі схема є коректною і має наближення, то наближений розв'язок задачі збігається до точного при прагненні до нуля кроку сітки, що наближує, причому порядок точності схеми на шаблоні і порядок наближення задачі співпадають.

Ця теорема пов'язує воєдино поняття *стійкості* та збіжності будь-яких схем наближення.

Порядок наближення просторово-часового оператора на шаблоні визначається при його розробці. Введемо оператор \mathfrak{T}_t – часовий сітковий аналог операторів $\mathfrak{T} = \partial/\partial t$ або $\mathfrak{T} = \partial^2/\partial t^2$.

Говорять, що при достатній гладкості функції $\vec{u}(t)$ (векторної або скалярної) в околі часового шаблону $\mathcal{U}(t, \Delta t_*)$ при $\Delta t < \Delta t_*$ часовий різницевий оператор \mathfrak{T}_t наближує диференційний оператор \mathfrak{T} з порядком k > 0 в момент часу t, якщо похибка

$$\Delta(t) = \mathfrak{I}_t \vec{u}_t(t) - (\mathfrak{I} \vec{u}(t))_t = O(\Delta t^k).$$
(16.23)

Оцінюючи похибку часового наближення на всій сітці, вважають, що різницевий оператор \Im_t наближує на сітці диференційний оператор \Im з порядком k > 0, якщо

$$\|\Delta_{t}\|_{t} = \|\Im_{t}\vec{u}_{t} - (\Im\vec{u})_{t}\|_{t} = O(\Delta t^{k}) \quad \text{afo} \quad \|\Delta_{t}\|_{t} = \|\Im_{t}\vec{u}_{t} - (\Im\vec{u})_{t}\|_{t} \le \beta\Delta t^{k}, \quad (16.24)$$

де $\vec{u}_t = P_t \vec{u}$; $(\Im \vec{u})_t = P_t (\Im \vec{u})$; P_t – оператор проєктування такий, що $P_t \vec{u} = \vec{u}_t \in H_t$ для будь-якого $\vec{u} \in H_0$; $\beta > 0$ і не залежить від Δt .

Для того, щоб оцінити похибку наближення на всій сітці, узагальнюють наведені в підрозділі 15.2 оцінювання наближення просторового оператора на просторово-часовий.

Введемо просторово-часовий різницевий оператор $\Re_{th} = \Im_t + L_h$. Вважають, що оператор \Re_{th} наближує на сітці диференційний оператор \Re з порядками m > 0 і k > 0, якщо

$$\|\Delta\|_{th} = \|\Re_{th}\vec{u}_{th} - \vec{f}_{th}\|_{th} = O(\Delta t^{k} + |h|^{m}) \quad \text{afo} \quad \|\Delta\|_{th} = \|\Re_{th}\vec{u}_{th} - \vec{f}_{th}\|_{th} \le \beta\Delta t^{k} + \alpha |h|^{m}, (16.25)$$

де $\vec{u}_{th} = P_{th}\vec{u}$; P_{th} – оператор проєктування на сітку; $\alpha, \beta > 0$ і не залежать від |h| і Δt . Під |h| можна розуміти середньоквадратичне або максимальне значення кроків просторової сітки. Якщо за напрямками (координатами) наближення відрізняються, то потрібно спочатку провести оцінювання (16.25) за напрямками, після цього слід вибрати серед отриманих значень *m* найменше значення.

Ще раз відзначимо, що під А-стійкістю розуміють спроможність просторово-часової схеми не збільшувати на часових кроках похибок (збурень) початкових даних. Поняття А-стійкості (за початковими даними) співпадає з поняттям безперервної залежності розв'язку різницевої задачі від початкових даних.

При дослідженнях стійкості часових схем для задач параболічного типу використовують поняття просторово-часового *оператора переходу* (кроку) $\aleph = \aleph(\Delta \vec{x}, \Delta t)$, який входить у зв'язок (двошарові схеми):

$$\vec{u}^{n+1} = \aleph \vec{u}^n + \Delta t \cdot S \vec{f}^n, \qquad (16.26)$$

де ще один оператор $S = S(\Delta \vec{x}, \Delta t)$ – *оператор джерела*; \vec{u}^{n+1} і \vec{u}^n – вектори наступного і попереднього розв'язків; \vec{f}^n – вектор джерела ("навантаження").

Доведено, що схема наближення (16.26) в лінійному нормованому просторі ϵ стійкою за початковими даними (А-стійкою) для заданого (будь-якого або з обмеженнями) часового кроку Δt за достатніх умов

$$\left\| \aleph(\Delta \vec{x}, \Delta t) \right\| \le 1; \quad \left\| S(\Delta \vec{x}, \Delta t) \right\| \le \beta, \tag{16.27}$$

де $\beta \ge 0$ – будь-яке обмежене число.

Дійсно, при відсутності джерела, тобто при $\vec{f}^n = \vec{0}$ значення $\|\vec{u}^{n+1}\|$ не повинно перевищувати $\|\vec{u}^n\|$ (консервативність системи), а при необмеженому $S(\Delta \vec{x}, \Delta t)$ навіть дуже мале джерело \vec{f}^n буде викликати нескінченне зростання $\|\vec{u}^{n+1}\|$, що не характерно для механічних та теплових процесів.

Достатньою умовою відсутності *осциляцій* розв'язку є умова позитивного знака оператора переходу $\aleph(\Delta \vec{x}, \Delta t)$ *в кожній точці тіла*, тобто

$$\aleph(\Delta \vec{x}, \Delta t) \ge 0. \tag{16.28}$$

Тоді при відсутності джерела компоненти вектора \vec{u}^{n+1} будуть мати той же знак, що й відповідні компоненти \vec{u}^{n} .

© К.М. Рудаков, 2025

Асимптотичною стійкістю називається спроможність схеми одержувати наближення стаціонарного режиму при $t \to \infty$. Вважається, що схема має асимптотичну стійкість, якщо оператор джерела є обмеженим та

$$\lim_{d \to \infty} \left\| \tilde{\aleph}(\Delta \vec{x}, t) \right\| = 0 \quad \text{afo} \quad \lim_{\Delta t \to \infty} \left\| \aleph(\Delta \vec{x}, \Delta t) \right\| = 0, \tag{16.29}$$

де оператор переходу $\tilde{\aleph}(\Delta \vec{x}, t)$ відповідає формулі

$$\vec{u}^n = \tilde{\aleph}(\Delta \vec{x}, t) \vec{u}^0; \qquad (16.30)$$

а \vec{u}^0 – початкове значення; $t = n\Delta t$ при $n \to \infty$.

Якщо схема відповідає умовам асимптотичної стійкості, то вона одночасно є стійкою за початковими даними (А-стійкою).

Вважається, що просторово-часовий оператор є *узгодженим*, якщо величина часового кроку лежить в деяких межах, які залежать від розміру просторового кроку (МСР, МСЕ та інші методи, що використовують дискретні базисні функції). В МСР часовий крок може бути обмеженим тільки згори, в МСЕ таке обмеження буває і згори, і знизу. Ці залежності виникають при дослідженні алгоритму на стійкість, тому умови погодженості та стійкості або розглядають одночасно, або зовсім не розрізняють як окремі.

Під ефективністю алгоритму (схеми) розуміють мінімальність машинного часу для отримання наближеного розв'язку задачі зі заздалегідь заданою точністю. Оскільки час розв'язування задачі залежить від швидкодії пристрою (колись таким пристроєм була людина, зараз – ЕОМ), то ефективність краще оцінювати кількістю арифметичних (елементарних, або визначальних) операцій при достатній якості алгоритму.

Під якістю алгоритму розуміють його відповідність задачі, що розв'язується (точність наближення, стійкість (погодженість), ефективність застосованого методу розв'язування САР та інше).

Далі розглянемо основні скінченно-різницеві схеми для крайових задач параболічного та гіперболічного типу.

16.4. Крайові задачі параболічного типу

Маємо систему рівнянь (16.1) ... (16.3). Введемо сітки:

п

$$\omega_{xyz} = \{x_i = \phi(h_x, i); \quad y_j = \phi(h_y, j); \quad z_k = \psi(h_z, k); \\ i = 0, 1, ..., N_x - 1; \quad j = 0, 1, ..., N_y - 1; \quad k = 0, 1, ..., N_z - 1\};$$
(16.31)
$$\omega_t = \{t_n = t_0 + \sum \Delta t_n, \quad n = 0, 1, ..., N_t - 1\},$$
(16.32)

А також введемо сітку в області $\omega_{xyzt} = \omega_{xyz} \times \omega_t$ з постійними просторовими h_x, h_y, h_z і часовим Δt кроками. Позначимо $u_{i,j,k}^n = u(x_i, y_j, z_k, t_n)$.

Перший крок алгебраїзації — за просторовими змінними. Розглянемо задачу теплопровідності (13.3-б), (13.4) … (13.6), яка в операторному вигляді представляється як (16.1) … (16.3). Замінимо в (16.1) просторовий оператор L просторово-різницевим оператором L_h . Для спрощення припустимо, що опера-

тор L є лінійним. З тією ж метою припустимо, що характеристики матеріалу задачі теплопровідності $c_p, \bar{\rho}, \lambda$ – постійні.

Для задачі теплопровідності (13.3-б), (13.4) ... (13.6) просторовий оператор L – оператор Лапласа. Візьмемо тривузлові шаблони з рівномірними кроками $(x-h_x,x,x+h_x)$, $(y-h_y,y,y+h_y)$ і $(z-h_z,z,z+h_z)$, причому h_x,h_y і h_z можуть бути різними. Отримаємо:

$$(L_{h}u_{h})_{i,j,k} = -\frac{u_{i-1,j,k} - 2u_{i,j,k} + u_{i+1,j,k}}{h_{x}^{2}} - \frac{u_{i,j-1,k} - 2u_{i,j,k} + u_{i,j+1,k}}{h_{y}^{2}} - \frac{u_{i,j,k-1} - 2u_{i,j,k} + u_{i,j,k+1}}{h_{z}^{2}} = -(u_{xxyyzz}^{+-++-})_{i,j,k}.$$
(16.33)

Справедливим є оцінювання похибки типу (15.20) за напрямками.

Рівняння теплопровідності (13.3-б) представимо спочатку у вигляді:

$$c_{P}\overline{\rho}\frac{\partial u_{i,j,k}}{\partial t} = \lambda \cdot (u_{xxyyzz}^{+-+-+})_{i,j,k} + f_{i,j,k}.$$
(16.34)

Простою заміною $\tau = t\lambda/c_P \bar{\rho}$; $q = f/\lambda$; $d\tau = dt\lambda/c_P \bar{\rho}$ можна привести (16.34) до безрозмірного вигляду. Тому далі замість (16.34) використовуємо:

$$\frac{\partial u_{i,j,k}}{\partial \tau} = (u_{xxyyzz}^{+-+-})_{i,j,k} + q_{i,j,k} \,. \tag{16.35}$$

Застосуємо пряме інтегрування у часі, оскільки цей підхід є більш універсальним, ніж використання часових диференційних операторів. Для цього спочатку помножимо (16.35) на $d\tau$ та проставимо інтеграли:

$$\int du_{i,j,k} = \int \left[(u_{xxyyzz}^{+-+-+})_{i,j,k} + q_{i,j,k} \right] d\tau .$$
(16.36)

Якщо інтегрувати в часових межах від τ_n до τ_{n+1} , то отримаємо сім'ю *двошарових* схем, а якщо від τ_{n-1} до τ_{n+1} , то – сім'ю *тришарових* схем.

16.4.1. Двошарові просторово-часові схеми

Проведемо інтегрування (16.36) в межах від τ_n до τ_{n+1} , тобто часового кроку $\Delta \tau$. Інтеграл у лівій частині (16.36) обчислюється точно:

$$\int_{u^n}^{u^{n+1}} du_{i,j,k} = u_{i,j,k}^{n+1} - u_{i,j,k}^n, \qquad (16.37)$$

а в правій — ні, оскільки містить невідомі значення $u = u(\vec{x}, \tau)$ у виразі для $(u_{xyyzz}^{++++-})_{i,i,k}$ (див. формулу (16.33)).

Згідно з однією із теорем про середнє при інтегруванні функцій, в інтервалі інтегрування неперервної функції завжди є таке значення ξ , при якому

$$\int_{\tau_{n}}^{\tau_{n+1}} u(\vec{x},\tau) d\tau = \tilde{u}(\vec{x},\xi)(\tau_{n+1} - \tau_{n}) = \tilde{u} \Delta \tau, \qquad (16.38)$$



де $\tilde{u} = u(\vec{x},\xi)$. Якщо в погодженому діапазоні [0,1] ввести ваговий коефіцієнт $\omega \in [0,1]$, який відповідає $\xi \in [\tau_n, \tau_{n+1}]$ (див. рис.16.1) і теж є невідомим, то цю формулу можна переписати у вигляді

$$\tilde{u} = \frac{1}{\Delta \tau} \int_{\tau_n}^{\tau_{n+1}} u(\vec{x}, \tau) d\tau = \omega u(\vec{x}, \tau_{n+1}) + (1 - \omega) u(\vec{x}, \tau_n) =$$

= $\omega u^{n+1} = (1 - \omega) u^n; \quad \omega \in [0, 1], \quad (16.39)$

Рис.16.1 До відповідності $\omega \in [0,1]$ величині $\xi \in [\tau_n, \tau_{n+1}]$: 1 – реальна функція u; 2 – її лінійна апроксимація тобто виразити \tilde{u} через значення u^n та u^{n+1} . Оскільки вектор "навантаження" може залежати від функції, яка шукається, тобто від u, його теж зазвичай представляють через вираз з ваговим коефіцієнтом:

$$\tilde{q} = \frac{1}{\Delta \tau} \int_{\tau_n}^{\tau_{n+1}} q(\vec{x}, u, \tau) d\tau = \omega q^{n+1} + (1 - \omega) q^n; \quad \omega \in [0, 1].$$
(16.40)

Примітка 16.2. Оскільки функції u та q не є подібними, то вагові коефіцієнти у формулах (16.39) і (16.40) повинні бути різними. Але традиційно при малих $\Delta \tau$ їх приймають однаковими, тому і позначають однаково.

Примітка 16.3. Застосування вагових схем (16.39) і (16.40) зовсім не є обов'язковим або єдино можливим варіантом. Варіантів – безліч.

Для будь-якого просторового оператора L маємо право записати, що

$$L_{h}\tilde{u}_{h} = L_{h}P_{h}\left(\frac{1}{\Delta\tau}\int_{\tau_{n}}^{\tau_{n+1}}ud\tau\right) = \frac{1}{\Delta\tau}\int_{\tau_{n}}^{\tau_{n+1}}L_{h}u_{h}d\tau, \qquad (16.41)$$

оскільки інтегрування проводиться за часом (*P_h* – оператор проєктування). Тому права частина (16.36) запишеться у вигляді

$$\int_{\tau_n}^{n+1} \left[(u_{xxyyzz}^{+-+--})_{i,j,k} + q_{i,j,k} \right] d\tau = \left[(\tilde{u}_{xxyyzz}^{+-+--})_{i,j,k} + \tilde{q}_{i,j,k} \right] \Delta \tau .$$
(16.42)

3 урахуванням (16.39)

$$\int_{\tau_n}^{\tau_{n+1}} \left[(u_{xxyyzz}^{+-+-+-})_{i,j,k} + q_{i,j,k} \right] d\tau = \left[\omega (u_{xxyyzz}^{+-++--})_{i,j,k}^{n+1} + (1-\omega) (u_{xxyyzz}^{+-++--})_{i,j,k}^{n} + \tilde{q}_{i,j,k} \right] \Delta \tau . \quad (16.43)$$

Отже, вирази (16.37) і (16.43) дозволяють записати рівняння теплопровідності (13.3-б) у вигляді наближення

$$\frac{u_{i,j,k}^{n+1} - u_{i,j,k}^{n}}{\Delta \tau} = \omega (u_{xxyyzz}^{+-+-+})_{i,j,k}^{n+1} + (1 - \omega) (u_{xxyyzz}^{+-++-})_{i,j,k}^{n} + \tilde{q}_{i,j,k}; \quad \omega \in [0,1].$$
(16.44)

Проблема лише у тому, що ані ξ , ані ω не є відомими. Якщо часовий інтервал $\Delta \tau$ вважати малим, а функцію $u(\vec{x},\tau)$ такою, яка змінюється у часі не дуже швидко, то з досить значною точністю можна обирати будь-яке значення

 $\omega \in [0,1]$ у всьому діапазоні. Таких варіантів – безліч. Але серед них є *класичні* варіанти:

• *явна* схема *Ейлера* (при ω = 0):

$$\frac{u_{i,j,k}^{n+1} - u_{i,j,k}^{n}}{\Delta \tau} = \frac{u_{i-1,j,k}^{n} - 2u_{i,j,k}^{n} + u_{i+1,j,k}^{n}}{h_{x}^{2}} + \frac{u_{i,j-1,k}^{n} - 2u_{i,j,k}^{n} + u_{i,j+1,k}^{n}}{h_{y}^{2}} + \frac{u_{i,j,k-1}^{n} - 2u_{i,j,k}^{n} + u_{i,j,k+1}^{n}}{h_{z}^{2}} + \tilde{q}_{i,j,k}.$$
(16.45)

У скороченому вигляді

$$(u_{\tau}^{+} - u_{xxyyzz}^{+-++-})^{n} = \tilde{q}.$$
 (16.46)

Відповідно до (16.45) значення $u_{i,j,k}^{n+1}$ знаходиться для точки тіла відразу, без збирання САР для всього тіла (явна схема).

На рис.16.2-а показано одновимірний шаблон явної схеми Ейлера, тобто при u = u(x,t).

• *неявна* схема Ейлера (при ω = 1):

$$\frac{u_{i,j,k}^{n+1} - u_{i,j,k}^{n}}{\Delta \tau} = \frac{u_{i-1,j,k}^{n+1} - 2u_{i,j,k}^{n+1} + u_{i+1,j,k}^{n+1}}{h_{x}^{2}} + \frac{u_{i,j-1,k}^{n+1} - 2u_{i,j,k}^{n+1} + u_{i,j+1,k}^{n+1}}{h_{y}^{2}} + \frac{u_{i,j,k-1}^{n+1} - 2u_{i,j,k}^{n+1} + u_{i,j,k+1}^{n+1}}{h_{z}^{2}} + \tilde{q}_{i,j,k},$$
(16.47)

або в скороченому вигляді

$$(u_{\tau}^{-} - u_{xxyyzz}^{++++-})^{n+1} = \tilde{q} .$$
 (16.48)

Одновимірний шаблон цієї неявної схеми показано на рис.16.2-б.

• схема Гальоркіна (при $\omega = 2/3$). У скороченому вигляді:

$$u_{\tau}^{+} = \frac{2}{3} (u_{xxyyzz}^{+-+-+})^{n+1} + \frac{1}{3} (u_{xxyyzz}^{+-+-+-})^{n} = \tilde{q}.$$
(16.49)

Всі ці схеми мають порядок наближення $O(\Delta \tau + |h|^2)$.

• схема Кранка-Ніколсона (при $\omega = 0.5$). У скороченому вигляді:

$$u_{\tau}^{+} = 0.5 \Big[(u_{xxyyzz}^{+-+-+-})^{n+1} + (u_{xxyyzz}^{+-+-+-})^{n} \Big] = \tilde{q} .$$
(16.50)

Цю схему можна розглядати як результат позмінного застосування явної і неявної схем з кроками $\Delta \tau/2$. Схема Кранка-Ніколсона – єдина з усіх вагових схем із (16.44), яка має порядок наближення $O(\Delta \tau^2 + |h|^2)$. Шаблони всіх схем з $0 < \omega < 1$, показано на рис.16.2-в.





Примітка 16.4. Якщо застосовувати інший підхід для створення розрахункових схем, а саме – використовувати різницеві наближення (16.9) чи (16.10) диференційного оператора $\partial u / \partial t$, тобто ліву чи праву різницю, які мають перший порядок наближення, то стає незрозуміло, чому схема Кранка-Ніколсона є схемою з другим порядком наближення.

Які показники стійкості мають ці схеми?

Для визначення цих умов використовують такі підходи: частотний аналіз, енергетичні нерівності і принцип максимуму.

Явна схема Ейлера виявилася лише *умовно стійкою* за початковими даними. А.А. Самарський застосував принцип максимуму. При цьому спочатку явну схему було записано у вигляді:

$$u^{n+1} = \left(1 - 2\Delta\tau \sum_{\alpha=1}^{p} \frac{1}{h_{\alpha}^{2}}\right) \cdot u^{n} + \left(\sum_{\alpha=1}^{p} \frac{1}{h_{\alpha}^{2}} u_{+h_{\alpha}}^{n} + \sum_{\alpha=1}^{p} \frac{1}{h_{\alpha}^{2}} u_{-h_{\alpha}}^{n} + q^{n}\right) \Delta\tau, \qquad (16.51)$$

де α – номер координатної осі; p – кількість ступенів свободи вузла; $u^n = u_{i,j,k}^n$; $u_{\pm h_1}^n = u_{i\pm 1,j,k}^n$; $u_{\pm h_2}^n = u_{i,j\pm 1,k}^n$ тощо. За принципом максимуму знаки виразів при u^n , $u_{\pm h_2}^n$ та $u_{-h_2}^n$ повинні бути однакові, тому

$$1 - 2\Delta \tau \sum_{\alpha=1}^{p} \frac{1}{h_{\alpha}^{2}} \ge 0.$$
 (16.52)

З (16.52) випливає *умова стійкості схеми* (згадаємо, що тут $\Delta \tau = \Delta t \lambda / c_P \overline{\rho}$):

$$\Delta \tau \leq \frac{1}{2} \left(\sum_{\alpha=1}^{p} \frac{1}{h_{\alpha}^{2}} \right)^{-1} \quad \text{afo} \quad \Delta t \leq \frac{c_{p} \overline{\rho}}{2\lambda} \cdot \left(\sum_{\alpha=1}^{p} \frac{1}{h_{\alpha}^{2}} \right)^{-1}.$$
(16.53)

Якщо позначити $h = \min\{h_{\alpha}\}$, то

$$\Delta \tau \leq \frac{h^2}{2p} \quad \text{afo} \quad \Delta t \leq \frac{c_P \overline{\rho}}{\lambda} \frac{h^2}{2p}.$$
 (16.54)

При p = 1, тобто для одновимірної задачі, необхідно $\Delta t \leq c_p \overline{\rho} h^2 / 2\lambda$.

3 формул (16.53), (16.54) випливає, що для явної двошарової схеми Ейлера:

• часовий крок Δt дуже жорстко обмежено згори;

• припустима величина часового кроку Δt зменшується при зростанні розмірності задачі p та при зменшенні просторового кроку та відношення $c_p \overline{\rho} / \lambda$.

Доведено, що всі двошарові вагові схеми є А-стійкими, якщо

$$\omega \ge \frac{1}{2} - \frac{|h|^2}{4p\Delta\tau}, \text{ afo } \omega \ge \frac{1}{2} - \frac{c_P \overline{\rho}}{\lambda} \cdot \frac{|h|^2}{4p\Delta t}.$$
(16.55)

Оскільки величина других доданків у виразах (16.55) зазвичай мала порівняно з 0.5, то (16.55) зазвичай скорочують до умови $\omega \ge 0.5$.

Тому схеми Кранка-Ніколсона ($\omega = 0.5$), Гальоркіна ($\omega = 2/3$) та неявна Ейлера ($\omega = 1$) є А-стійкими при будь-яких $\Delta \tau$ або Δt , незалежно від величин кроків сітки h_{α} . Але всі схеми, окрім схеми Кранка-Ніколсона, мають порядок наближення $O(\Delta \tau + |h|^2)$, тобто малу точність за часом, що не забезпечує *асимптотичну* збіжність.

На жаль, схема Кранка-Ніколсона піддана осциляціям розв'язку, оскільки не завжди відповідає умові (16.28). Як виявилося, їй також не відповідають й інші розглянуті схеми, окрім неявної з $\omega = 1$.

Розглянемо цю проблему за допомогою частотного аналізу.

Власні значення λ_m , m = 1,...,N матриці [A] САР $\{\dot{u}\} + [A]\{u\} = \{b\}$, яка породжується рівнянням (16.1) після його просторової алгебраїзації, знаходяться з системи $[A]\{w\} + \mu\{w\} = \{0\}$, де $\{w\} \in$ власні вектори, відповідні власним значенням μ_m матриці [A]. Відомо, що із застосуванням власних векторів завжди можна записати, що

$$\{u\}^{n} = \sum_{m=1}^{N} \left(c_{m}^{n} \cdot \{w\}_{m}\right) \quad \text{Ta} \quad \{u\}^{n+1} = \sum_{m=1}^{N} \left(c_{m}^{n+1} \cdot \{w\}_{m}\right), \quad (16.56)$$

де $N \in$ кількістю рівнянь у САР, а $c_m = c_m(\tau)$.

Якщо ці вирази підставити у вагову схему (16.44), то одержимо, що при відсутності вектора навантаження в кожному вузлу тіла

$$c_m^{n+1} = \frac{1 - \mu_m \Delta \tau \cdot (1 - \omega)}{1 + \mu_m \Delta \tau \omega} c_m^n; \quad m = 1, ..., N; \quad \omega \in [0, 1],$$
(16.57)

тобто оператор переходу (див. формулу (16.26))

$$\aleph_m(\Delta \tau) = \frac{1 - \mu_m \Delta \tau \cdot (1 - \omega)}{1 + \mu_m \Delta \tau \omega}; \quad m = 1, \dots, N; \quad \omega \in [0, 1].$$
(16.58)

У задачі нестаціонарної теплопровідності всі μ_m матриці [A] САР $\{\dot{u}\}+[A]\{u\}=\{b\} \in \partial i \ddot{u}$ сними (тобто мають Im $\mu_m=0$) та позитивними (тобто Re $\mu_m > 0$). Тому при фіксованому *m* оператор переходу (16.58) є звичайним алгебраїчним рівнянням з дійсними числами.

Щоб виконувалася умова відсутності осциляцій (16.28), тут $\aleph(\Delta \tau) \ge 0$, із (16.58) отримаємо, що, для кожного μ_m повинна виконуватися умова

$$\frac{1-\mu_m \Delta \tau \cdot (1-\omega)}{1+\mu_m \Delta \tau \omega} \ge 0; \quad m = 1, \dots, N; \quad \omega \in [0,1].$$
(16.59)

Оскільки знаменник $1 + \mu_m \Delta \tau \omega$ завжди більше нуля, то умовою відсутності осциляцій буде вираз $1 - \mu_m \Delta \tau \cdot (1 - \omega) \ge 0$, а для всіх власних векторів матриці

 $1 - \mu_{\max} \Delta \tau \cdot (1 - \omega) \ge 0$ або $\Delta \tau \le 1 / (\mu_{\max} \cdot (1 - \omega))$, (16.60) причому при $\omega = 1$ осциляції відсутні завжди, а при $\omega \to 1$ для дуже значних часових кроків $\Delta \tau$.

Для явної схеми $\omega = 0$, тому для неї

$$\Delta \tau \le 1/\mu_{\rm max} \,. \tag{16.61}$$

Цю умову називають умовою (формулою) Куранта.

© К.М. Рудаков, 2025

На практиці замість значення μ_{max} (яке можна обчислити методами, викладеними у Розділі 6, але це потребує багато часу) зазвичай застосовують оцінку (див. формулу (6.18) у підрозділі 6.3):

$$\mu_{\max} = s(A) = |\overline{\mu}(A)| \le \max_{m} \sum_{n=1}^{N} |a_{mn}|, \qquad (16.62)$$

де a_{mn} є компонентами деякої матриці [A]; N – розмірність матриці [A].

Для ілюстрування обмежень, які мають описані розрахункові схеми щодо осциляцій, зазвичай наводять графіки функції переходу $\aleph(\mu\Delta\tau)$, де μ є одним з власних значень μ_m матриці [A] для однієї з власних форм. Точний вираз дається формулою $\aleph(\mu\Delta\tau) = \exp(-\mu\Delta\tau)$, а загальний для всіх вагових двошарових схем — формулою $\aleph(\mu\Delta\tau) = \frac{1-\mu\Delta\tau\cdot(1-\omega)}{1+\mu\Delta\tau\omega}$ (відповідає формулі (16.58)). На рис.16.3 наведені графіки оператора $\aleph(\mu\Delta\tau)$: точний та у випадках $\omega = 0$ (явна схема Ейлера), $\omega = 0.5$ (схема Кранка-Ніколсона), $\omega = 2/3$ (схема Гальоркіна) та $\omega = 1$ (неявна схема Ейлера). Значення параметра $\mu\Delta\tau$, при яких функція переходу $N(\mu\Delta\tau)$ менше нуля, є неприпустимими з точки зору осциляцій цієї власної форми. Отже, рівність $\aleph(\mu_m\Delta\tau) = 0$ є границею осциляцій m-ої моди, а умова $\aleph(\mu_m\Delta\tau) > 0$ є умовою відсутності осциляції для цих схем, яка відповідає загальній умові (16.28).



Рис.16.3 Графіки функцій оператора переходу для точного розв'язку та класичних вагових двошарових схем

Рис.16.3 та формула (16.60) показують, що схема Гальоркіна допускає без осциляцій більший за розміром часовий крок, ніж схема Кранка-Ніколсона. До того ж на практиці схема Гальоркіна за точністю наближається до схеми Кранка-Ніколсона. Це й є причиною популярності схеми Гальоркіна.

Обмеження часового кроку за умови відсутності осциляцій є настільки жорсткими, що на практиці їх зазвичай не виконують буквально. Тоді в

процесі розв'язування задачі деякі високочастотні моди (з дуже значними μ_m) мають осциляції, з якими або миряться, або згладжують (осереднюють) їх додатковими процедурами.

Вираз (16.58) можна застосувати й для визначення часового кроку із умови стійкості (16.27), тобто || №_m(Δτ) ||≤1, якщо його записати як

$$-1 \leq \frac{1 - \mu_m \Delta \tau \cdot (1 - \omega)}{1 + \mu_m \Delta \tau \omega}; \quad \frac{1 - \mu_m \Delta \tau \cdot (1 - \omega)}{1 + \mu_m \Delta \tau \omega} \leq 1; \quad m = 1, \dots, N; \quad \omega \in [0, 1].$$
(16.63)

Друга умова виконується завжди, оскільки після звільнення від знаменника та приведення подібних залишається $-\mu_m \Delta \tau \leq 0$. З першої умови аналогічно можна отримати, що $\mu_m \cdot (1-2\omega) \cdot \Delta \tau \leq 2$, або для всіх власних форм матриці

$$u_{\max} \cdot (1 - 2\omega) \cdot \Delta \tau \le 2; \quad \omega \in [0, 1].$$
(16.64)

Очевидно, що ця умова виконується для будь-яких часових кроків $\Delta \tau$ при

$$\omega \ge \frac{1}{2} - \frac{1}{\mu_{max} \Delta \tau} \tag{16.65}$$

(порівняйте її з умовою (16.55)) або наближено при $\omega \ge 0.5$, оскільки зазвичай $\mu_{max} \Delta \tau >> 1$. Тобто (16.65) або $\omega \ge 0.5$ є умовами наявності А-стійкості. При $\omega < 0.5$ вагові двошарові схеми є лише *умовно стійкими*, з кроком

$$\Delta \tau \le 2 / \left(\mu_{\max} \cdot (1 - 2\omega) \right). \tag{16.66}$$

16.4.2. Тришарові просторово-часові схеми

Проведемо інтегрування (16.36) в межах від τ_{n-1} до τ_{n+1} , тобто на часовому кроці $2\Delta \tau = \tau_{n+1} - \tau_{n-1}$ аналогічно викладеному у попередньому пункті 16.4.1. Результат напишемо у вигляді:

$$\frac{u_{i,j,k}^{n+1} - u_{i,j,k}^{n-1}}{2\Delta\tau} = (\tilde{\tilde{u}}_{xxyyzz}^{++++-})_{i,j,k} + \tilde{\tilde{q}}_{i,j,k}, \qquad (16.67)$$

де значок "≈" над величиною вказує, що застосовуються величини:

$$\tilde{\tilde{u}} = \frac{1}{2\Delta\tau} \int_{\tau_{n-1}}^{\tau_{n+1}} u(\vec{x},\tau) d\tau ; \quad \tilde{\tilde{q}} = \frac{1}{2\Delta\tau} \int_{\tau_{n-1}}^{\tau_{n+1}} q(\vec{x},\tau) d\tau , \qquad (16.68)$$

а $\tilde{\tilde{u}} = u(\vec{x},\xi)$ для $\xi \in [\tau_{n-1},\tau_{n+1}]$. Якщо ввести ваговий коефіцієнт $\omega \in [0,2]$, який відповідає $\xi \in [\tau_{n-1},\tau_{n+1}]$, то ці формули можна переписати у вигляді

$$\tilde{\tilde{u}} = \frac{1}{2\Delta\tau} \int_{\tau_{n-1}}^{\tau_{n+1}} u(\vec{x},\tau) d\tau = \omega u^{n+1} + (1-2\omega)u^n + \omega u^{n-1};$$
(16.69)

$$\tilde{\tilde{q}} = \frac{1}{2\Delta\tau} \int_{\tau_{n-1}}^{\tau_{n+1}} q(\vec{x},\tau) d\tau = \omega q^{n+1} + (1-2\omega)q^n + \omega q^{n-1}.$$
(16.70)

У залежності від того, яке значення $\omega \in [0,2]$ застосувати, можна отримати різноманітні тришарові схеми. При $\omega > 0.25$ схеми А-стійки. Оскільки в (16.67) ліва частина має порядок наближення $O(\Delta \tau^2)$, а права – $O(|h|^2)$, то всі тришарові схеми, що відповідають (16.67), мають порядок наближення $O(\Delta \tau^2 + |h|^2)$.

Очевидно, що при $\omega = 0$ маємо $\tilde{\tilde{u}} = u^n$ та $\tilde{\tilde{q}} = q^n$. Саме при такому ω утворюється *явна* схема *Річардсона*:

$$\frac{u_{i,j,k}^{n+1} - u_{i,j,k}^{n-1}}{2\Delta\tau} = (u_{xxyyzz}^{+-+-+})_{i,j,k}^{n} + q_{i,j,k}^{n}.$$
(16.71)

Вона була історично першою тришаровою схемою (1910 рік). Але схема (16.71) нестійка, оскільки в ній $\omega < 0.25$.

Виявилося, що А-стійку явну схему можна отримати з (16.71), якщо в правій частині замінити $2u_{i,j,k}^n$ на $u_{i,j,k}^{n-1} + u_{i,j,k}^{n+1}$. Отримуємо *явну* схему Дюфорта-Франкела (схему "ромб"):

$$\frac{u_{i,j,k}^{n+1} - u_{i,j,k}^{n-1}}{2\Delta\tau} = \frac{u_{i-1,j,k}^{n} - u_{i,j,k}^{n-1} - u_{i,j,k}^{n+1} + u_{i+1,j,k}^{n}}{h_{x}^{2}} + \frac{u_{i,j-1,k}^{n} - u_{i,j,k}^{n-1} - u_{i,j,k}^{n-1} - u_{i,j,k}^{n-1} - u_{i,j,k}^{n-1} + u_{i,j,k+1}^{n}}{h_{z}^{2}} + \frac{\tilde{q}_{i,j,k}}{h_{z}^{2}} + \frac{u_{i,j,k-1}^{n} - u_{i,j,k}^{n-1} - u_{i,j,k}^{n-1} - u_{i,j,k}^{n-1} + u_{i,j,k+1}^{n}}{h_{z}^{2}} + \tilde{q}_{i,j,k}.$$
(16.72)

Вона А-стійка, але має наближення порядку $O(\Delta \tau^2 + |h|^2)$ за умовою, що $\lim(\Delta \tau / |h|) \rightarrow 0$ при $\Delta \tau$, $|h| \rightarrow 0$ (умовне наближення).

Існують й інші тришарові схеми. Однак практика показала, що тришарові схеми програють так званим "економічним" схемам в ефективності. Суттєві недоліки тришарових схем: вимога постійності $\Delta \tau$ та необхідність застосовувати іншу схему на початковій стадії (для знаходження значень $u_{i,j,k}^1$). Тому тришарові схеми застосовуються лише іноді.

16.4.3. "Економічні" схеми (схеми розщеплення)

Явні двошарові схеми мають формули рекурентного типу, тобто не вимагають обертання матриці при розв'язуванні СЛАР. Тому вони економічні за кількістю арифметичних дій на часовому кроці: пропорційне N – кількості невідомих у системі. Але через умовну стійкість ці схеми мають надто малий часовий крок, який до того ж дещо зменшується при зростанні розмірності (дво- та тривимірні) задачі та при зростанні відношення $c_p \bar{\rho} / \lambda$. Існують тришарові явні схеми, А-стійкі, але теж з умовним наближенням. Неявні двошарові схеми (при $\omega > 0$) вимагають обертання матриці при розв'язуванні СЛАР, що різко збільшує кількість арифметичних дій до порядку N^3 . Тому загальні кількості дій для розв'язування крайової задачі явними та неявними схемами на однаковому часовому відрізку є близькими, тобто виграшу від застосування явних схем фактично немає.

Швидше за все розв'язуються СЛАР з діагональними матрицями. Саме такі матриці мають явні двошарові схеми, якщо їх збирати. Всі інші схеми породжують більш заповнені матриці. Відомо, що будь-яку позитивно визначену матрицю можна перетворити у діагональну. Але це трудомістка задача, оскільки потребує знаходження власних значень і форм матриці (див. Розділ 6). З якими ще видами
Частина IV. Розділ 16

матриць СЛАР розв'язується швидко? З тридіагональними та з трикутними. Чи можна створити неявні схеми з такими видами матриць? Як з'ясувалося – можна. Пошуки ефективних схем привели до створення так званих "економічних" неявних схем, які А-стійкі та мають зменшену кількість арифметичних дій на часовому кроці.

16.4.3.1. Поздовжньо-поперечна схема Пісмена-Речфорда

Історично перша "економічна" неявна схема – *поздовжньо-поперечна* схема *Пісмена-Речфорда* (D.W. Peaceman, H.H. Rachford, 1955 рік). Ідея поздовжньо-поперечних схем полягає в тому, що задача розв'язується в p етапів (p > 1 – розмірність задачі), на кожному з яких прогноз-розв'язок здійснюється на основі одного з *координатних* напрямків. Введемо оператори Λ_m та Λ такі, що

$$\Lambda_m u = u_{x^m x^m}^{+-}; \quad m = 1, ..., p; \quad \Lambda = \sum_{m=1}^p \Lambda_m; \quad \Lambda u = \sum_{m=1}^p \Lambda_m u = u_{xxyyzz}^{+-+-+-}.$$
(16.73)

Схема Пісмена-Речфорда працює тільки при p = 2 (при p = 3 – нестійка) і складається з двох етапів (або СЛАР):

$$\begin{cases} \frac{u_{i,j}^{n+1/2} - u_{i,j}^{n}}{0.5\Delta\tau} = \Lambda_{1}u_{i,j}^{n+1/2} + \Lambda_{2}u_{i,j}^{n} + \tilde{q}_{i,j}; \\ \frac{u_{i,j}^{n+1} - u_{i,j}^{n+1/2}}{0.5\Delta\tau} = \Lambda_{1}u_{i,j}^{n+1/2} + \Lambda_{2}u_{i,j}^{n+1} + \tilde{q}_{i,j}. \end{cases}$$
(16.74)

Видно, що на першому етапі схема неявна за напрямком m = 1, але явна за напрямком m = 2; на другому етапі — навпаки. Одержуємо *тридіагональні* матриці СЛАР, системи розв'язуються методом прогонки (див. підрозділ 5.4), який має кількість дій, пропорційну кількості вузлів.

Для визначення порядку наближення цієї схеми зведемо дві СЛАР (16.74) в одну шляхом виключення з них $u_{i,j}^{n+1/2}$. Для цього спочатку помножимо їх на $0.5\Delta\tau$ і приведемо подібні (I – одиничний оператор. Для скорочення записів тут і далі опустимо нижні індекси при u):

$$(I - 0.5\Delta\tau\Lambda_1)u^{n+1/2} = (I + 0.5\Delta\tau\Lambda_2)u^n + 0.5\Delta\tau\tilde{q};$$
(16.75)

$$(I - 0.5\Delta\tau\Lambda_2)u^{n+1} = (I + 0.5\Delta\tau\Lambda_1)u^{n+1/2} + 0.5\Delta\tau\tilde{q}.$$
(16.76)

Потім помножимо (16.75) на $(I - 0.5\Delta \tau \Lambda_1)^{-1}$, отриманий відносно $u^{n+1/2}$ вираз підставимо в (16.76), результат помножимо на $(I - 0.5\Delta \tau \Lambda_1)$, щоб позбутися оберненого оператора. Отримаємо:

$$(I - 0.5\Delta\tau\Lambda_1)(I - 0.5\Delta\tau\Lambda_2)u^{n+1} = (I + 0.5\Delta\tau\Lambda_1)(I + 0.5\Delta\tau\Lambda_2)u^n + (I + 0.5\Delta\tau\Lambda_1)0.5\Delta\tau\tilde{q} + 0.5\Delta\tau\tilde{q}.$$
(16.77)

Розкриємо дужки та напишемо результат у вигляді:

$$I [I - 0.5\Delta\tau (\Lambda_1 + \Lambda_2)] u^{n+1} + (0.5\Delta\tau)^2 \Lambda_1 \Lambda_2 u^{n+1} = I [I + 0.5\Delta\tau (\Lambda_1 + \Lambda_2)] u^n + (0.5\Delta\tau)^2 \Lambda_1 \Lambda_2 u^n + \Delta\tau \tilde{q} + (0.5\Delta\tau)^2 \Lambda_1 \tilde{q}.$$
(16.78)

Врахуємо, що $\Lambda_1 + \Lambda_2 = \Lambda$. Всі три члена, що містять $(0.5\Delta \tau)^2$, можна відкинути як складові другого порядку малості порівняно з іншими. Отримаємо:

$$I - 0.5\Delta\tau\Lambda \big) u^{n+1} = \big(I + 0.5\Delta\tau\Lambda \big) u^n + \Delta\tau\tilde{q} \,. \tag{16.79}$$

Залишилося розкрити дужки та переписати його у вигляді

$$\frac{u^{n+1} - u^n}{\Delta \tau} = \Lambda \frac{u^{n+1} + u^n}{2} + \tilde{q} , \qquad (16.80)$$

який точно відповідає схемі Кранка-Ніколсона (див. формули (16.50) і (16.73)), яка має порядок наближення $O(\Delta \tau^2 + |h|^2)$.

Є лема Келога, яка стверджує, що

$$\| (I - \sigma \Lambda)(I + \sigma \Lambda)^{-1} \| \le 1; \quad \| (I - \sigma \Lambda)(I + \sigma \Lambda)^{-1} \| < 1, \tag{16.81}$$

якщо
$$\Lambda \ge 0$$
 та $\Lambda > 0$ відповідно, а також $\sigma > 0$. Вираз (16.79) перепишемо як

$$u^{n+1} = \left(I - 0.5\Delta\tau\Lambda\right)^{-1} \left(I + 0.5\Delta\tau\Lambda\right) u^n + \Delta\tau\left(I - 0.5\Delta\tau\Lambda\right)^{-1} \tilde{q} .$$
(16.82)

Згідно з (16.26) одержали оператори переходу та джерела відповідно

$$\aleph(\Delta \vec{x}, \Delta t) = (I - 0.5\Delta \tau \Lambda)^{-1} (I + 0.5\Delta \tau \Lambda); \quad S(\Delta \vec{x}, \Delta t) = (I - 0.5\Delta \tau \Lambda)^{-1}.$$
(16.83)

У нашому випадку
 $\sigma=0.5\Delta\tau>0$ та $\Lambda<0$ (див. приклад у підрозділі 14.1), тому згідно з
лемою Келога

$$\|(I - 0.5\Delta \tau \Lambda)^{-1}(I + 0.5\Delta \tau \Lambda)\| < 1,$$
 (16.84)

тобто поздовжньо-поперечна схема Пісмена-Речфорда є А-стійкою. Оператор джерела $S(\Delta \vec{x}, \Delta t)$ є обмеженим за модулем, оскільки $\sigma = 0.5\Delta \tau > 0$ та $\Lambda < 0$.

Додатковими дослідженнями доведено, що схема Пісмена-Речфорда є асимптотичне стійкою і має умовний сумарний (адитивний) порядок наближення $O(\Delta \tau^2 + |h|^2)$. Умови стосуються наявності обмежених значень деяких вищих частинних похідних, тому не є жорсткими. Схему узагальнено на випадок нелінійного рівняння (зі змінними коефіцієнтами), при цьому сумарний порядок наближення $O(\Delta \tau^2 + |h|^2)$ збережено при спеціальному способі задавання граничних умов.

Але поздовжньо-поперечна схема Пісмена-Речфорда може працювати лише для двовимірної задачі.

16.4.3.2. Схеми з трикутними факторизованими операторами

Заміна класичного поняття наближення на слабшу умову *сумарного* наближення значно полегшує побудову економічних схем і приводить до поняття *адитивних* схем з факторизованими операторами, в яких похибка на кожному з кроків не визначена, а визначена лише адитивна (сумарна) похибка.

Викладки (16.75) ... (16.80) підказали ідею створення економічних схем з факторизованими операторами.

Спочатку введемо поняття факторизованого оператора. Розглянемо (16.44), тобто групу можливих двошарових схем. Будь-яку двошарову схему завжди можна представити в *канонічному* вигляді

$$Bu_{\tau}^{+} = Ru^{n} + \tilde{q}, \quad \text{тобто} \quad B\frac{u^{n+1} - u^{n}}{\Delta\tau} = Ru^{n} + \tilde{q}. \quad (16.85)$$

Помножимо на $\Delta \tau$ і приведемо подібні. Для u^{n+1} одержуємо систему рівнянь

$$Bu^{n+1} = F^n, (16.86)$$

де відомо праву частину

$$F^{n} = (B + \Delta \tau R)u^{n} + \Delta \tau \tilde{q}. \qquad (16.87)$$

Припустимо, що оператор *В* можна записати у *факторизованому* вигляді $B = B_1 \cdot ... \cdot B_s$, (16.88) де B_m – "економічні" оператори зі зменшеною кількістю арифметичних дій.

Тоді й система (16.86) — "економічна", оскільки її можна розкласти на *s* систем з "економічними" операторами B_m :

$$B_1 y_1 = F^n; \quad B_m y_m = y_{m-1}; \quad m = 2, ..., s; \quad u^{n+1} = y_s,$$
 (16.89)

де y_m – проміжні розв'язки.

Зокрема, розглянуту вище поздовжньо-поперечну схему Пісмена-Речфорда ми фактично представили у факторизованому вигляді формулою (16.77), яка після знехтування $(0.5\Delta \tau)^2 \Lambda_1 \tilde{q}$ отримує вигляд:

$$(I - 0.5\Delta\tau\Lambda_1)(I - 0.5\Delta\tau\Lambda_2)u^{n+1} = (I + 0.5\Delta\tau\Lambda_1)(I + 0.5\Delta\tau\Lambda_2)u^n + \Delta\tau\tilde{q}^n.$$
(16.90)

Як бачимо, схема Пісмена-Речфорда має два факторизованих "економічних" оператора (*s* = *p* = 2):

$$B_m = I - 0.5\Delta \tau \Lambda_m; \quad m = 1, 2.$$
 (16.91)

Можливим є інший, ніж (16.91), варіант факторизованого оператора B_m , що знімає обмеження з розмірності області, для якої розв'язується крайова задача. Замість (16.91) напишемо фактично аналогічний вираз факторизованого оператора

$$B_m = I - 0.5\Delta \tau R_m; \quad m = 1, ..., 2s; \quad s = 1, 2, ...,$$
(16.92)

де 2*s* – парна кількість таких операторів. Для *R_m* повинні виконуватися наступні умови:

$$R = \sum_{m=1}^{2s} R_m; \quad Ru = \sum_{m=1}^{2s} R_m u = u_{xxyyzz}^{+-+-+-}.$$
 (16.93)

Згадаємо, що просторова похідна другого порядку (одновимірний випадок)

$$u_{h}^{+-} = \frac{u_{i-1} - 2u_{i} + u_{i+1}}{h^{2}} = \frac{u_{i}^{+} - u_{i}^{-}}{h} = \frac{1}{h} (u_{i}^{+} - u_{i}^{-}).$$
(16.94)

Приймемо, що 2s = 2, а

$$R_{1}u = -\sum_{\alpha=1}^{p} u_{x_{\alpha}}^{-} / h_{\alpha} ; \quad R_{2}u = \sum_{\alpha=1}^{p} u_{x_{\alpha}}^{+} / h_{\alpha} .$$
 (16.95)

де p – розмірність крайової задачі. Тоді R_m – "трикутні" оператори, тобто такі, що не призводять до заповнення відповідно верхньої та нижньої частини матриць в СЛАР, що будуть породжуватися. Оскільки одиничний оператор I породжує діагональну матрицю, то оператор $B_m = I - 0.5\Delta \tau R_m$ теж трикутний. СЛАР з трикутними матрицями розв'язуються простою підстановкою, що у

методі скінченних різниць вимагає дій порядку $O(N^2)$, де N – кількість рівнянь у системі.

Всі схеми, в яких вихідний просторовий оператор представляється у вигляді суми декількох операторів, називають *схемами розщеплення*.

Одна з популярних адитивних схем розщеплення має вигляд:

$$\begin{cases} \frac{u^{n+1/2} - u^n}{0.5\Delta\tau} - R_1 u^{n+1/2} - R_2 u^n = \tilde{q};\\ \frac{u^{n+1} - u^{n+1/2}}{0.5\Delta\tau} - R_1 u^{n+1/2} - R_2 u^{n+1} = \tilde{q}. \end{cases}$$
(16.96)

який формально співпадає за написом зі схемою Пісмена-Речфорда (див. формули (16.74)), теж має умовний *сумарний (адитивний*) порядок наближення $O(\Delta \tau^2 + |h|^2)$, теж А-стійка та асимптотичне стійка. Але ця "економічна" схема, як й інші схеми з трикутними факторизованими операторами, не має обмежень щодо розмірності крайової задачі.

Якщо у рівнянні (16.85) оператор $B^{-1}R \le 0$ і не залежить від часу, то рекомендують застосовувати наступну адитивну "економічну" схему (схему розщеплення), яку називають схемою методу *стабілізації*. Представимо задачу у вигляді двошарової схеми з ваговим коефіцієнтом

$$\frac{u^{n+1} - u^n}{\Delta \tau} = \omega (u_{xxyyzz}^{+-+--})^{n+1} + (1 - \omega) (u_{xxyyzz}^{+-+-+-})^n + \tilde{q} ; \quad \omega \in [0,1].$$
(16.97)

Використовуючи позначення (16.93) і (16.95), спочатку отримаємо:

$$\frac{u^{n+1} - u^n}{\Delta \tau} = \omega (R_1 + R_2) u^{n+1} + (1 - \omega) R u^n + \tilde{q} .$$
(16.98)

Помножив на $\Delta \tau$, позбудемося знаменника і приведемо подібні:

$$[I - \omega \Delta \tau (R_1 + R_2)]u^{n+1} = [I + (1 - \omega)\Delta \tau R]u^n + \Delta \tau \tilde{q}.$$
(16.99)

Можемо стверджувати, що схема (16.99) (як і вихідна (16.97)) – А-стійка при виконанні умови (16.55) для ω .

Позначимо $u^{n+1} = u^n + \eta$ і підставимо в (16.99). Після скорочення однакових членів при u^n отримаємо:

$$[I - \omega \Delta \tau (R_1 + R_2)]\eta = \Delta \tau \cdot (Ru^n + \tilde{q}).$$
(16.100)

Представимо ліву частину у *факторизованому* вигляді (з точністю до членів порядку $O(\Delta \tau^2)$):

$$(I - \omega \Delta \tau R_1)(I - \omega \Delta \tau R_2)\eta = \Delta \tau \cdot (Ru^n + \tilde{q}).$$
(16.101)

Відповідно до (16.88) і (16.89) замінимо (16.101) на вирази, що виконуються послідовно:

$$\begin{cases} (I - \omega \Delta \tau R_1) y = \Delta \tau \cdot (Ru^n + \tilde{q}); \\ (I - \omega \Delta \tau R_2) \eta = y; \\ u^{n+1} = u^n + \eta, \end{cases}$$
(16.102)

де y – проміжний розв'язок; $\omega \in [0,1]$.

Ми отримали адитивну "економічну" схему. Вона А-стійка при виконанні наступних умов:

- незалежності операторів R_1 і R_2 від часу;
- при $R_1 \le 0$ та $R_2 \le 0$;
- $\omega \ge 0.5$.

Схема має порядок наближення $O(|h|^2 + \Delta \tau^2)$ лише при $\omega = 0.5$.

Слід додати наступне. Якщо оператор $R = R_1 + R_2$ є симетричним, то $R_1^T = R_2$ та $R_2^T = R_1$. Тоді можна збирати, наприклад, верхню частину матриці, яка відповідає оператору R (позначимо її члени як a_{mn}^+), і поділити діагональні члени на двійку. Отримаємо компоненти матриці, відповідні оператору R_2 , а для R_1 компоненти утворюються простим транспонуванням отриманої для R_2 матриці.

Окрім схеми методу стабілізації є ще одна *аналогічна за основними показниками* схема, створена у рамках методу прогноз-коректор (предікторкоректор).

Загальна ідея методу прогноз-коректор полягає у тому, що розв'язок знаходиться у два етапи, причому на першому робиться прогноз розв'язку з меншою точністю, ніж це потрібно, а на другому етапі цей прогнозований розв'язок використовується для отримання (корекції) остаточного розв'язку.

Із застосуванням "трикутних" операторів *R_m* ця схема записується як

$$\begin{cases} \frac{u^{n+1/4} - u^{n}}{0.5\Delta\tau} - R_{1}u^{n+1/4} = \tilde{q};\\ \frac{u^{n+1/2} - u^{n+1/4}}{0.5\Delta\tau} - R_{2}u^{n+1/2} = 0;\\ \frac{u^{n+1} - u^{n}}{\Delta\tau} = Ru^{n+1/2} + q^{n+1/2}. \end{cases}$$
(16.103)

Перші дві СЛАР є неявними "економічними" схемами з кроком у часі $\Delta \tau/2$ для одержання $u^{n+1/2}$, яке у третій СЛАР з наближенням $O(\Delta \tau^2 + |h|^2)$ використовується для отримання u^{n+1} . Оскільки третя СЛАР містить явну схему, то загальна кількість дій у (16.103) приблизно така, як і у (16.102).

Якщо основні оператори залежать від часу, тобто $R_m = R_m(\tau)$, то замість методів стабілізації та прогноз-коректор потрібно застосовувати іншу схему, зокрема *двоциклічну* схему *покомпонентного* розщеплення:

$$\begin{cases} (I - 0.5\Delta\tau R_{1})\eta = (I + 0.5\Delta\tau R_{1})u^{n-1}; \\ (I - 0.5\Delta\tau R_{2})(\pi - \Delta\tau q^{n}) = (I + 0.5\Delta\tau R_{2})\eta; \\ (I - 0.5\Delta\tau R_{2})\chi = (I + 0.5\Delta\tau R_{2})(\pi + \Delta\tau q^{n}); \\ (I - 0.5\Delta\tau R_{1})u^{n+1} = (I + 0.5\Delta\tau R_{1})\chi. \end{cases}$$
(16.104)

або (інша форма її запису)

$$\begin{cases} (I - 0.5\Delta\tau R_{1})\eta = (I + 0.5\Delta\tau R_{1})u^{n-1}; \\ (I - 0.5\Delta\tau R_{2})\pi = (I + 0.5\Delta\tau R_{2})\eta; \\ \mu = \pi + 2\Delta\tau q^{n}; \\ (I - 0.5\Delta\tau R_{2})\chi = (I + 0.5\Delta\tau R_{2})\mu; \\ (I - 0.5\Delta\tau R_{1})u^{n+1} = (I + 0.5\Delta\tau R_{1})\chi. \end{cases}$$
(16.105)

В цих формулах η, π, χ, μ – допоміжні (проміжні) розв'язки, а оператори R_1 та R_2 обчислюються при $\tau = \tau_n$, яке при попередньому застосуванні схеми (16.104) або (16.105) дорівнювало τ_{n+1} .

У підрозділі 16.6 розглянуто питання обирання часового кроку, який забезпечить достатню точність наближення факторизованими операторами.

16.5. Крайові задачі гіперболічного типу

Маємо систему рівнянь (16.4), (16.5), (16.2). Введемо сітки (16.31) і (16.32). Основне рівняння (16.4), приведемо до вигляду

$$\frac{\partial^2 \vec{u}}{\partial t^2} + G \frac{\partial \vec{u}}{\partial t} = R \vec{u} + \vec{q} , \qquad (16.106)$$

де позначено $G = M^{-1}C$, $R\vec{u} = -M^{-1}L(\vec{u})$; $\vec{q} = M^{-1}\vec{f}$.

Скінченно-різницеві наближення операторів $\partial/\partial t$ та $\partial^2/\partial t^2$ розглянуто у підрозділі 16.2. Формула (16.18) для наближення оператора $\partial^2/\partial t^2$ спирається на три часових вузла, тому наближення рівняння (16.106) буде проводитися тришаровими просторово-часовими схемами. Порядок наближення цього оператора – другий. Тому й для оператора $\partial/\partial t$ потрібно обрати скінченно-різницеве наближення другого порядку, а саме центральне-різницеве (16.13). Отже, рівняння (16.106) після скінченно-різницевого наближення буде мати вигляд

$$\frac{u_{i,j,k}^{n-1} - 2u_{i,j,k}^{n} + u_{i,j,k}^{n+1}}{\Delta t^2} + G \frac{u_{i,j,k}^{n+1} - u_{i,j,k}^{n-1}}{2\Delta t} = Ru_{i,j,k} + \tilde{\tilde{q}}_{i,j,k} \,.$$
(16.107)

Доведено, що схема (16.107) буде безумовно А-стійкою, якщо прийняти

$$Ru_{i,j,k} = R \frac{u_{i,j,k}^{n-1} + u_{i,j,k}^{n+1}}{2},$$
(16.108)

де просторовий диференційний оператор *R* не залежить від часу. Отже, отримаємо безумовно А-стійку неявну схему другого порядку наближення

$$\frac{u_{i,j,k}^{n-1} - 2u_{i,j,k}^{n} + u_{i,j,k}^{n+1}}{\Delta t^2} + G \frac{u_{i,j,k}^{n+1} - u_{i,j,k}^{n-1}}{2\Delta t} = R \frac{u_{i,j,k}^{n-1} + u_{i,j,k}^{n+1}}{2} + \tilde{\tilde{q}}_{i,j,k}$$
(16.109-a)

або у скороченому вигляді

$$u_{tt}^{+-} + Gu_{t}^{0} = R \frac{u^{n-1} + u^{n+1}}{2} + \tilde{\tilde{q}}.$$
 (16.109-6)

Це одна з безлічі можливих схем.

Якщо оператор *G* породжує *діагональну* матрицю [\overline{G}], або таку, яку легко привести до діагональної матриці [\overline{G}], часто застосовують схему, в якій замість (16.108) обирають, що $Ru_{i,j,k} = Ru_{i,j,k}^n$. Тоді (16.107) можна записати у вигляді явної (рекурентної) схеми:

$$u_{i,j,k}^{n+1} = \left(\frac{1}{\Delta t^{2}} + \bar{G}\frac{1}{2\Delta t}\right)^{-1} \left[\left(\frac{2}{\Delta t^{2}} + R\right) u_{i,j,k}^{n} - \left(\frac{1}{\Delta t^{2}} - \bar{G}\frac{1}{2\Delta t}\right) u_{i,j,k}^{n-1} + \tilde{\tilde{q}}_{i,j,k} \right].$$
(16.110)

яку називають *схемою центральних різниць*. Ця схема має умовну А-стійкість, тобто часовий крок є обмеженим зверху

$$\Delta t \le T_{\min} / \pi , \qquad (16.111)$$

де T_{min} – найменший з періодів власних коливань тіла. Це дуже жорстке обмеження, але велика кількість кроків компенсується відсутністю необхідності у збиранні та розв'язуванні СЛАР з недіагональною матрицею.

У випадку відсутності дисипативного члена, як і для задач параболічного типу, розроблені "економічні" схеми з факторизованим оператором.

Напишемо вираз (16.109-а) при $G \equiv 0$ і опущених нижніх індексах у канонічному вигляді:

$$B\frac{u^{n-1}-2u^n+u^{n+1}}{\Delta t^2} = R\frac{u^{n-1}+u^{n+1}}{2} + \tilde{\tilde{q}}.$$
 (16.112)

Значення u^{n-1} і u^n вважаються відомими. Для u^{n+1} одержуємо систему рівнянь

$$Bu^{n+1} = F^n, (16.113)$$

де відома права частина

$$F^{n} = (2B + \Delta t^{2}R)u^{n} - Bu^{n-1} + \Delta t^{2}\tilde{\tilde{q}}.$$
 (16.114)

Припустимо, що оператор В можна записати у факторизованому вигляді

$$B = B_1 \cdot \dots \cdot B_s, \qquad (16.115)$$

де B_k — "економічні" оператори зі зменшеною кількістю арифметичних дій. Тоді і система (16.113) є "економічною", оскільки її можна розкласти на *s* систем з операторами B_k :

$$B_1 y^{(1)} = F^n; \quad B_k y^{(k)} = y^{(k-1)}; \quad k = 2, \dots, s; \quad u^{n+1} = y^{(S)}, \quad (16.116)$$

які виконуються послідовно, і де $y^{(k)}$ – проміжні розв'язки.

Оператори *В_k* зручно прийняти у вигляді:

$$B_k = [I - 0.5\Delta t^2 R_k], \qquad (16.117)$$

де необхідно, щоб

$$R = \sum_{k=1}^{S} R_k ; \quad -R_k \ge 0.$$
 (16.118)

Наприклад, при s = 2 маємо

$$B = (I - 0.5\Delta t^2 R_1)(I - 0.5\Delta t^2 R_2) = (I - 0.5\Delta t^2 R) + O(\Delta t^2).$$
(16.119)

Підставивши результат у (16.113), отримаємо після скорочення подібних

$$\frac{u^{n-1} - 2u^n + u^{n+1}}{\Delta t^2} = R \frac{u^{n-1} + u^{n+1}}{2} + \tilde{\tilde{q}}, \qquad (16.120)$$

тобто з точністю $O(\Delta t^2)$ повертаємося до безумовно А-стійкої схеми (16.109) при $G \equiv 0$ (відсутності дисипації).

Схема центральних різниць (16.110) при $\overline{G} = 0$ спрощується до

$$\frac{1}{\Delta t^2} u_{i,j,k}^{n+1} = \left(\frac{2}{\Delta t^2} - R\right) u_{i,j,k}^n - \frac{1}{\Delta t^2} u_{i,j,k}^{n-1} + \tilde{\tilde{q}}_{i,j,k}, \qquad (16.121)$$

причому обмеження (16.111) часового кроку зберігається.

Відзначимо також, що розроблено ще один підхід до створення "економічних" схем розв'язування задач гіперболічного типу – зведення задачі до системи рівнянь параболічного типу.

Старт розрахунків. Для початку розрахунків за тришаровими схемами необхідно мати значення $u_{i,j,k}^0$ та $u_{i,j,k}^1$. Щоб зберегти другий порядок наближення за часом і на першому часовому кроці, обчислимо $u_{i,j,k}^1$ як результат розкладу в ряд Тейлора в околі t = 0 на часовому кроці Δt (надалі нижні індекси для спрощення запису опускаємо):

$$u^{1} = u^{0} + \Delta t \dot{u}^{0} + \Delta t^{2} \ddot{u}^{0} / 2 + O(\Delta t^{3}).$$
(16.122)

Як початкові умови в задачах гіперболічного типу задаються $u^0 = u(0) = \hat{u}_0$; $\dot{u}_0 = \partial u(0) / \partial t = \hat{h}$ (див. рівняння (16.5)). Враховуючи, що відповідно до основного рівняння (16.106)

$$\ddot{u} = \partial^2 u / \partial t^2 = -G\dot{u} + Ru + q, \qquad (16.123)$$

з (16.122) отримаємо:

$$u^{1} = \hat{u}^{0} + \Delta t \hat{h} + \Delta t^{2} [-G^{0} \dot{u}^{0} + R^{0} u^{0} + q^{0}] / 2 + O(\Delta t^{3}) \approx$$

$$\approx (I + 0.5\Delta t^{2} R^{0}) \hat{u}^{0} + \Delta t \cdot (I - 0.5\Delta t G^{0}) \hat{h} + 0.5\Delta t^{2} q^{0}. \qquad (16.124)$$

Отже, застосування формули (16.124) забезпечує другий порядок наближення за часом на першому часовому кроці.

16.6. Про обирання часового кроку, який забезпечить достатню точність наближення факторизованого оператора

У п.п.16.4.3.2 та підрозділі 16.5 введено поняття схем з факторизованими операторами.

Було визначено, що формалізація крайових задач нестаціонарної теплопровідності та динамічного навантаження призводить до систем алгебраїчних рівнянь, які можна записати у канонічному операторному вигляді (16.86) та (16.113), а саме:

$$Bu^{n+1} = F^n, (16.125)$$

де права частина відома, оператор $B = I - \omega \Delta t R$ або $B = I - 0.5 \Delta t^2 R$ породжує квадратну матрицю розмірністю $N \cdot N$; верхній індекс вказує номер часового кроку (шару), а $0 < \omega \le 1$. Проблема полягає у великій кількості арифметичних дій щодо знаходження оберненої матриці: пропорційне N^3 . Вона вирішується шляхом застосування факторизованих операторів B_m , які наближують B за формулою $B = B_1 \cdot ... \cdot B_s$; $s \ge 2$. Ці оператори для всіх неявних схем повинні мати

кількість дій, пропорційну N^2 (явні схеми мають цю кількість, пропорційну N, але дуже жорстке обмеження на часовий крок Δt). Зазвичай, щоб обернення B_m потребувало кількість дій, не більшу, ніж пропорційну N^2 , оператори B_m обирають такими, що породжують трикутні матриці. Факторизовані оператори, відповідно до типів крайових задач, що розглядаються, були розглянуті вище:

$$B_m = I - \omega \Delta t R_m; \qquad (16.126-a)$$

$$B_m = I - 0.5\Delta t^2 R_m, \qquad (16.126-6)$$

де $\sum_{m=1}^{s} R_m = R$ є обов'язковою умовою, причому введено оператори R < 0 та $R_m < 0$. Тоді, наприклад, при s = 2, маємо наближення

$$B = I + \omega \Delta t | R |\approx (I + \omega \Delta t | R_1 |) (I + \omega \Delta t | R_2 |) =$$

= $I + \omega \Delta t (| R_1 | + | R_2 |) + (\omega \Delta t)^2 | R_1 | \cdot | R_2 | = (I + \omega \Delta t | R |) + O(\Delta t^2);$ (16.127-a)
$$B = I + 0.5 \Delta t^2 | R |\approx (I + 0.5 \Delta t^2 | R_1 |) (I + 0.5 \Delta t^2 | R_2 |) =$$

$$= I + 0.5\Delta t^{2}(|R_{1}| + |R_{2}|) + 0.25\Delta t^{4} |R_{1}| \cdot |R_{2}| = (I + 0.5\Delta t^{2} |R|) + O(\Delta t^{4}).$$
(16.127-6)

У відкинутих у виразах (16.127) частинах (підкреслені) присутні результати перемноження двох матриць. Може виявитися, що вони переважають Δt^2 та/або Δt^4 , тому точність наближення буде незначна. Отже, потрібно отримати відповідні вирази, що обмежують Δt , виходячи з необхідної точності розв'язку крайових задач нестаціонарної теплопровідності та динамічного навантаження із застосуванням факторизованих операторів.

Оскільки оператор $|-R| \in$ позитивно визначеним, то $-\det R > 0$ й $-\det B > 0$, як і будь-яка інша їхня норма $|| \bullet ||$. Тепер можна застосувати нерівність Міньковського щодо норм, і записати, що

$$||I + \omega \Delta t | R | || \le ||I|| + \omega \Delta t || |R||;$$
 (16.128-a)

$$||I + 0.5\Delta t^{2} |R||| \le ||I|| + 0.5\Delta t^{2} ||R|||.$$
(16.128-6)

Призначимо $\delta > 0$ – мале число, яке визначає відносну похибку на кожному часовому кроці, викликану застосуванням факторизованих операторів. Наприклад, $\delta = 0.001$, тобто 0.1%. Тоді умова порівняної малості підкреслених частин виразів, якими нехтуємо:

$$\omega^{2} \Delta t^{2} \parallel R_{1} R_{2} \parallel \leq \delta \parallel I - \omega \Delta t R \parallel; \qquad (16.129-a)$$

$$0.25\Delta t^{4} \|R_{1}R_{2}\| \leq \delta \|I - 0.5\Delta t^{2}R\|.$$
(16.129-6)

3 урахуванням || *I* ||=1 й (16.128), із (16.129) маємо

$$\omega^{2} \Delta t^{2} \| R_{1} R_{2} \| \leq \delta + \delta \omega \Delta t \| | R | \|; \qquad (16.130-a)$$

$$0.25\Delta t^{4} \|R_{1}R_{2}\| \le \delta + \delta 0.5\Delta t^{2} \||R|\|.$$
(16.130-6)

Коли є рівність, маємо два відомих квадратичних алгебраїчних рівняння відносно приросту часового кроку Δt чи Δt^2 . Позначимо коефіцієнти:

$$a = \omega^2 ||R_1R_2||; \quad b = -\delta\omega ||R||; \quad c = -\delta;$$
 (16.131-a)

$$\beta = 0.25 || R_1 R_2 ||; \quad \gamma = -\delta 0.5 || |R| ||; \quad c = -\delta.$$
(16.131-a)

Тоді із умови відносної малості відкинутої з (16.127) частини для застосування виразів (16.126) маємо обмеження на часовий крок (за виключенням $\omega = 0$):

$$\Delta t \le \frac{-b + \sqrt{b^2 - 4ac}}{2a}; \tag{16.132-a}$$

$$\Delta t \le \sqrt{\frac{-\gamma + \sqrt{\gamma^2 - 4\beta c}}{2\beta}} \,. \tag{16.132-6}$$

Ці обмеження є додатковими, оскільки на часовий крок є основні обмеженні, отримані за умов абсолютної чи асимптотичної стійкості та відсутності осциляцій розв'язку. Відомі формула Куранта (16.61), а саме

$$\Delta t \le 1/[\mu_{\max}(1-\omega)]$$
 (16.133)

для крайової задачі нестаціонарної теплопровідності (визначає умову відсутності осциляцій розв'язку), а також формула $\Delta t \leq T_{\min} / \pi$ для крайової динамічної задачі. Тут позначені μ_{\max} — максимальне власне значення матриці R; T_{\min} мінімальний період коливань (із всіх ступенів свободи об'єкта розрахунків).

Виникає питання, яку норму застосувати? Для обчислення детермінанта матриці потрібна кількість "довгих" операцій (типу помножити / поділити) порядку N^3 , що неприйнятне. Але детермінант – лише одна із можливих норм матриці. Для нерівності Міньковського можна застосовувати будь-яку норму, але тут тільки ту, що якимсь чином пов'язана з μ_{max} . Відома оцінка μ_{max} через її *l*-норму (октаедричну), див. (16.62):

$$\mu_{\max} \le \max_{j} \sum_{i} |r_{ij}| = ||R||_{l}, \qquad (16.134)$$

яка зовсім не потребує для обчислення "довгих" операцій. Визначення цієї норми можна легко робити одночасно з операцією збирання глобальної матриці із компонент матриць скінченних елементів. Це повністю скасовує операції із вибирання необхідних компонент глобальної матриці із пам'яті комп'ютера.

Крім того, відома така властивість будь-яких норм: $||YZ||=||Y|| \cdot ||Z||$. Тому й $||R_1R_2||=||R_1|| \cdot ||R_2||$. Оскільки оператор R в задачах, що розглядаються, є симетричним, то зазвичай приймають, що $R_2 = R_1^T$ (при цьому, як зазначено вище, обов'язково $\sum_{m=1}^{S} R_m = R$). Тоді $||R_1R_2||=||R_1||^2=||R_2||^2$, що значно прискорює обчислення $||R_1R_2||$, оскільки не потрібно попередньо перемножувати дві матриці.

Відомо, що для крайових задач нестаціонарної теплопровідності зазвичай у програмних комплексах, наприклад Nastran, замість умови Куранта застосовують формулу

$$\Delta t \le \frac{C_p \rho}{\lambda} \frac{h^2}{10},\tag{16.135}$$

де C_p , ρ – питомі теплоємність та густина матеріалу; λ – його коефіцієнт теплопровідності; h – мінімальний діаметр скінченного елемента в моделі

об'єкта. Ця формула гарантує наявність абсолютної стійкості явної схеми, тобто і всіх неявних схем. Але вона не має відношення до умови відсутності осциляцій розв'язку, тому дає значення Δt , інше, ніж за формулою Куранта. Отже, для задачі нестаціонарної теплопровідності можна отримати три оцінки Δt , потім з них призначити найменшу. Для динамічної задачі таких значень Δt буде два.

Контрольні питання до підрозділу 16.2

1. Які є скінченно-різницеві наближення для часових диференційних операторів, їхні шаблони та яка їхня точність?

Контрольні питання до підрозділу 16.3

1. В чому полягає поняття А-стійкості розв'язків, одержаних застосуванням часових схем?

2. Яка теорема пов'язує воєдино поняття стійкості та збіжності будь-яких просторовочасових схем?

3. Якими показниками оцінюється порядок наближення просторово-часового оператора крайової задачі просторово-часовим сітковим оператором $\Re_{th} = \Im_t + L_h$?

4. Що пов'язує просторово-часовий оператор переходу (кроку) $\aleph = \aleph(\Delta \vec{x}, \Delta t)$?

5. Що пов'язує просторовий оператор джерела $S = S(\Delta \vec{x}, \Delta t)$?

6. Яка властивість оператора переходу є достатньою для асимптотичній стійкості просторово-часових схем?

7. Яка властивість оператора переходу є достатньою для відсутності осциляцій розв'язку при застосуванні просторово-часових схем?

8. Що значить: просторово-часовий оператор є узгодженим?

Контрольні питання до підрозділу 16.4

1. Як виглядає загальна двошарова вагова схема скінченно-різницевого наближення крайових задач теплопровідності.

2. Які є класичні двошарові вагові схеми скінченно-різницевого наближення крайових задач теплопровідності, їхні шаблони.

3. Вкажіть властивості класичних двошарових вагових схем скінченно-різницевого наближення крайових задач теплопровідності щодо А-стійкості.

4. Вкажіть властивості класичних двошарових вагових схем скінченно-різницевого наближення крайових задач теплопровідності щодо наявності осциляцій розв'язків.

5. Як виглядають графіки функцій оператора переходу для точного розв'язку та класичних вагових двошарових схем?

6. Які недоліки мають тришарові просторово-часові схеми?

7. Що означає поняття "економічні схеми"?

8. "Економічна" схема Пісмена-Речфорда, її переваги та недоліки.

9. В чому полягає поняття факторизованих операторів?

10. Запишіть інший, ніж у схемі Пісмена-Речфорда, варіант факторизованого оператора B_m , який не має недоліків цієї схеми.

11. Які схеми називають схемами розщеплення?

12. Які схеми називають адитивними схемами?

Контрольні питання до підрозділу 16.5

1. Що потрібно зробити з тришаровою схемою скінченно-різницевого наближення крайових динамічних задач, щоб отримати А-стійку схему?

2. Яку схему називають схемою центральних різниць? Чи є вона А-стійкою?

Контрольні питання до підрозділу 16.6

1. Чому можливо керувати точністю результатів розв'язку крайових задач нестаціонарної теплопровідності та динамічного навантаження у схемах із факторизованими операторами за рахунок обмеження часового кроку?

2. За рахунок чого алгоритм обчислення значення обмеженого часового кроку потрібує відносно невелику кількість математичних операцій?

Частина V МЕТОД ЗВАЖЕНИХ ПОХИБОК НАБЛИЖЕННЯ ТА ІНШІ МЕТОДИ АЛГЕБРАЇЗАЦІЇ КРАЙОВИХ ЗАДАЧ

Розділ 17

ОСНОВНІ МЕТОДИ РОЗВ'ЯЗУВАННЯ КРАЙОВИХ ЗАДАЧ. ПОСЛАБЛЕННЯ МЕТОДУ ЗВАЖЕНИХ ПОХИБОК НАБЛИЖЕННЯ

17.1. Ідеї основних методів алгебраїзації крайових задач за просторовими змінними

Для алгебраїзації крайових задач, окрім методу скінченних різниць (МСР) застосовують і інші. На відміну від МСР ці методи є загальними: дозволяють одержувати і аналітичні, і чисельні розв'язки.

Будемо вважати, що розглядається одна з крайових задач, поставлених у Розділі 13. Шуканий результат – вектор $\vec{u}(\vec{x},t)$. Основний оператор, що діє на $\vec{u}(\vec{x},t)$ у об'ємі Ω – оператор Λ , граничні умови на поверхні *S* містять просторовий оператор *K*. Розглядається крайова задача

$$\Lambda(\vec{u}) = \vec{f} \quad \mathbf{B} \ \Omega \tag{17.1}$$

з граничними (ГУ) та початковими (ПУ) умовами

$$K(\vec{u}) = \vec{g}$$
 на S; $\vec{u}(\vec{x}, t_0) = \vec{u}_0(\vec{x})$ в Ω та на S, (17.2)

де $\vec{u} = \vec{u}(\vec{x},t) \in D(\Lambda) \subset X$; X – частина гільбертового простору H. Для стаціонарної задачі оператор $\Lambda = L$, для нестаціонарної задачі параболічного типу – $\Lambda = C \frac{\partial}{\partial t} + L$, а для нестаціонарної задачі гіперболічного типу $\Lambda = M \frac{\partial^2}{\partial t^2} + C \frac{\partial}{\partial t} + L$, причому в нестаціонарних задачах до початкових умов з (17.2) додаються ще умови, пов'язані з відомими першими та/чи другими похідними у часі. Тут оператори L, M та C – просторові.

17.1.1. Наближення розв'язків крайових задач лінійною комбінацією базисних векторів (метод Фур'є)

Нехай $\vec{u}_0 = \vec{u}_0(\vec{x},t)$ – відомий вектор-розв'язок задачі в області $\Omega \subset X$, яка обмежена поверхнею *S*. Завдання: знайти вектор $\vec{u}(\vec{x},t)$, що його наближує.

У 1807 році Жан Батист Жозеф Фур'є (G.B.G. Fourier, 1768-1830 рр.) у повідомленні "Трактат про поширення тепла у твердих тілах" запропонував знаходити розв'язок задачі стаціонарної теплопровідності у вигляді лінійної суперпозиції (рядом) так званих власних розв'язків з тригонометричними функціями синуса та косинуса. Важливо, що ці функції є ортогональними та нормованими.

Згодом виявилося, що це була геніальна ідея, тому на його честь цей ряд був названий тригонометричним рядом Фур'є, а метод – тригонометричним методом Фур'є. Всі необхідні теореми були доведені згодом П.Г.Л. Дирихле та Г.Ф.Б. Риманом.

Цей метод був узагальнений на дво- та тривимірні статичні, еволюційні та динамічні задачі за методом відокремлення змінних, а також на випадок суперпозиції інших ортогональних (і не зовсім) нормованих функцій (теорема Вейерштрасса "про апроксимування", яка доводить, що такими функціями можуть бути поліноми; теорія подвійності Понтрягина, тощо).

Такий узагальнений метод далі в цієї книзі будемо (для спрощення) називати методом Фур'є.

Відповідно до методу Фур'є вводиться система *базисних* (інші назви: які наближують, пробних) векторів

$$\vec{\psi}(\vec{x},t), \ \vec{\phi}_i(\vec{x}); \ i=1,2,...,$$
 (17.3)

на основі яких точний розв'язок задачі шукається у вигляді розкладу по базису

$$\vec{u}_0(\vec{x},t) = \vec{\psi}(\vec{x},t) + \sum_{i=-\infty}^{\infty} c_i(t)\vec{\varphi}_i(\vec{x}), \qquad (17.4-a)$$

де $c_i(t)$ – функції часу (для стаціонарних задач – коефіцієнти), які потрібно знайти.

Наближений розв'язок створюється з використанням деякої обмеженої кількості (позначимо її як *J*) членів розкладу Фур'є (17.4-а), а саме у вигляді

$$\vec{u}_0(\vec{x},t) \approx \vec{u}(\vec{x},t) = \vec{\psi}(\vec{x},t) + \sum_{i=1}^J c_i(t)\vec{\varphi}_i(\vec{x}).$$
(17.4-6)

Від кількості членів розкладу (17.4-б) залежить точність розв'язку крайової задачі. У загальному випадку збільшення цієї кількості підвищує точність, але одночасно збільшує час обчислень. На практиці застосовують відносно невеликі значення J. Оцінювання отриманої точності розв'язку (за величиною норми похибки наближення $\|\vec{u} - \vec{u}_0\|$) є обов'язковим.

Базисні вектори в розкладі (17.4) повинні мати такі властивості:

- складати повну за енергією систему векторів;
- бути лінійно незалежними;
- бути визначеними в усій області Ω, включаючи її границю S;
- мати всі необхідні для крайової задачі похідні;
- сумарно (через (17.4)) задовольняти всім рівнянням крайової задачі;

• породжувати добре обумовлені (див. Розділ 5.2) системи алгебраїчних рівнянь (САР);

• бути простими у обчисленні (не трудомісткими);

• надавати можливості при відносно малої їхньої кількості (число *J*) отримувати достатньо точні наближення.

Крім того, обиратися вони можуть з того ж простору, що й $\vec{u}(\vec{x},t)$.

Найбільш складне завдання – задовольнити останню умову, оскільки зазвичай точний розв'язок крайової задачі не є відомим. Однак із загальних міркувань було показано, що найкращі показники завжди мають тригонометричні многочлени $\sin kx$, $\cos kx$, дещо гірші – алгебраїчні x^p . Ці многочлени задовольняють й іншим вимогам. Щодо кращої обумовленості, то її мають САР з рідко заповненими матрицями з переважаючими діагональними членами. Такі матриці можуть породжуватися ортогональними базисними векторами на основі тригонометричних многочленів, а також "майже ортогональними", що дорівнюють нулю у всій Ω за винятком незначної її частини (це так звані локальні носії).

Є три можливі варіанти вибору базисних векторів у розкладі (17.4) для створення розв'язку задачі, а саме таких, що тільки за їхній рахунок (при довільних $c_i(t)$ або c_i) розклад Фур'є (17.4):

а/ тотожно задовольняє граничні умови на S, але не задовольняє основне рівняння в Ω . Якщо оператор K є лінійним (поширена ситуація), то приймають $\vec{\psi}$ та $\vec{\varphi}_i$ такими, щоб $K\vec{\psi}|_S = \vec{g}$ та $K\vec{\varphi}_i|_S = \vec{0}$; тоді на S функції $c_i(t)$ (коефіцієнти c_i) не зв'язані граничними значеннями, тобто будь-які. При нелінійному операторі K виникають труднощі при обиранні необхідних $\vec{\psi}$ та $\vec{\varphi}_i$. Цей варіант найчастіше застосовується при канонічних обрисах тіла;

б/ тотожно задовольняє основне рівняння в Ω , але граничні умови на *S* не задовольняє, що вимагає призначити $\vec{\psi} \equiv \vec{0}$. Застосовується відносно рідко, оскільки існують значні проблеми з вибором таких базисних функцій в Ω (див. підрозділ 17.4). Цей варіант називають методом граничного розв'язку;

в/ не задовольняє тотожно рівняння задачі ані в Ω , ані на S. Тоді зазвичай приймають $\vec{\psi} \equiv \vec{0}$.

У методі, запропонованому та розвинутому Володимиром Логвиновичем Рвачёвим (1926-2005 рр.) у 1963 році, розклад (17.4-б) замінюється на

$$\vec{u}_{0}(\vec{x},t) \approx \vec{u}(\vec{x},t) = \vec{\psi}(\vec{x},t) + \omega(\vec{x}) \sum_{i=1}^{J} c_{i}(t) \,\vec{\vartheta}_{i}(\vec{x}), \qquad (17.5)$$

причому

$$\vec{\psi}|_{s} = \vec{u}|_{s}; \quad \omega|_{s} \equiv 0; \quad \omega|_{\Omega \setminus S} > 0, \qquad (17.6)$$

а $\vec{\mathcal{G}}_i(\vec{x})$ – повна, лінійно незалежна система векторів, як і $\vec{\varphi}_i$ у (17.4-б).

Функція $\omega(\vec{x})$ повинна описувати поверхню *S* будь-якої складності. Призначення $\omega(\vec{x})$ у загальному випадку проводиться за допомогою булевої алгебри. Функції $\omega(\vec{x})$ отримали назву R-функцій. З методикою їх призначення можна ознайомитися з фундаментальної монографії [61] та інших публікацій В.Л. Рвачёва та його послідовників.

Метод R-функцій (R-Functions Method, RFM) – один з методів вибору базисних векторів варіанта а) їхнього вибору (див. вище), оскільки очевидно, що завжди можна прийняти $\omega(\vec{x})\vec{\vartheta}_i(\vec{x}) = \vec{\varphi}_i(\vec{x})$.

Примітка 17.1. Окрім методів Фур'є та Рвачёва існують принаймні ще два методи (апроксимації розкладанням у ряд): Трефтца (Е.І. Trefftz, 1888-1970 рр.) та Канторовича (Л.В. Канторович, 1912-1986 рр.). Не розглядаємо, оскільки для розв'язування крайових задач з Розділу 13 достатньо застосовувати узагальнений метод Фур'є.

Але мати апроксимацію (17.4) чи іншу, замало. Ще потрібно отримати систему рівнянь відносно невідомих складових наближення. Для цього розроблено декілька методів.

17.1.2. Ідея методу найменших квадратів

Метод найменших квадратів (OLS – Ordinary Least Squares) був створений Карлом Фрідріхом Гауссом (C.F. Gauss, 1777-1855 рр.) та Адрієн Марі Лежандром (A.M. Legendre, 1752-1833 рр.) в 1795 та 1805 роках, для знаходження розв'язків перевизначених систем рівнянь, що виникали при обробці експериментальних даних. Мінімізується функціонал з суми інтегралів від квадратів векторів-похибок по всій області їхнього визначення:

$$F(\vec{u}) = \int_{\Omega} \vec{R}_{\Omega}^{2} d\Omega + \int_{S} \vec{R}_{S}^{2} dS = \int_{\Omega} (\vec{R}_{\Omega}, \vec{R}_{\Omega}) d\Omega + \int_{S} (\vec{R}_{S}, \vec{R}_{S}) dS , \qquad (17.7)$$

де для крайової задачі вектори похибок наближення

$$\vec{R}_{\Omega} = \Lambda(\vec{u}) - \vec{f}; \quad \vec{R}_{S} = K(\vec{u}) - \hat{\vec{g}}, \qquad (17.8)$$

причому вектор \vec{u} наближується виразом (17.4), часто при $\vec{\psi} \equiv \vec{0}$. Умова мінімізації набуває вигляду $\partial F / \partial c_i = 0$, тому з (17.7)

$$\int_{\Omega} (\vec{R}_{\Omega}, \vec{w}_{j}) d\Omega + \int_{S} (\vec{R}_{S}, \vec{\tilde{w}}_{j}) dS = 0; \quad j = 1, 2, ..., J, \qquad (17.9)$$

де позначені вектори

$$\vec{w}_{j} = \partial \left[\Lambda \left(\vec{\psi} + \sum_{i=1}^{J} c_{i} \vec{\varphi}_{i} \right) \right] / \partial c_{j}; \quad \vec{\tilde{w}}_{j} = \partial \left[K \left(\vec{\psi} + \sum_{i=1}^{J} c_{i} \vec{\varphi}_{i} \right) \right] / \partial c_{j}. \quad (17.10)$$

Якщо оператори Л та К є лінійними, то

$$\vec{w}_j = \Lambda \vec{\varphi}_j; \quad \vec{\tilde{w}}_j = K \vec{\varphi}_j.$$
(17.11)

Вектори \vec{w}_j та \vec{w}_j у (17.9) називають *ваговими* векторами, тобто є такими, що "зважують" вектори похибок наближення. Потрібно мати на увазі, що при цьому вектори похибок наближення не повинні дорівнювати нулю, а будуть мінімізуватися в сенсі деякого середнього зваженого в області визначення.

Якщо в (17.1) і (17.2) оператори Λ і K є симетричними, то вираз (17.9) призводить до САР із симетричною матрицею.

Існує основна лема механіки суцільних середовищ: якщо у цілком щільній області Ω для будь-якого $\Omega_m \subset \Omega$ інтеграл $\int_{\Omega_m} \theta d\Omega = 0$, то функція $\theta \equiv 0$ у всій

Ω. У виразі (17.9) функцією під інтегралами є скалярні добутки двох векторів, а скалярні добутки дорівнюють нулю лише при ортогональності векторів. Тобто вектори похибок наближення *ортогональні* ваговим векторам.

Нехай $\eta = \min F(\vec{u})$. Послідовність $\{\vec{u}_J\}$ апроксимації (17.4-б) називають такою, що мінімізує функціонал $F(\vec{u})$, якщо $\lim_{J\to\infty} F(\vec{u}_J) = F(\vec{u}_0) = \eta$. Доведено, що ця послідовність дійсно є такою, що мінімізує функціонал (17.7), тобто при цьому є розв'язком задачі (17.1): $\lim_{J\to\infty} ||\vec{u}_J - \vec{u}_0||_{\Omega} = 0$.

17.1.3. Ідея методу Релея-Рітца (Рітца)

Варіаційний метод, який зазвичай називають методом Рітца, фактично був створений Релеєм (Sir Jon William Strutt (Джон Уильямс Стретт), Lord Rayleigh, 1842-1919 рр.) у 1870 році, потім був презентований швейцарським фізиком Вальтером Рітцем (W. Ritz) у 1908 році як новий метод. Теоретично обґрунтований у 1918 році М.М. Криловим (умови збіжності).

Метод Релея-Рітца використовує варіаційну теорему 2 з Розділу 2.4. Згідно з нею вектор-розв'язок $\vec{u}(\vec{x}) \rightarrow \vec{u}_0(\vec{x})$ рівняння (17.1) мінімізує функціонал

$$F(\vec{u}) = 0.5(\Lambda \vec{u}, \vec{u}) - (\vec{u}, \vec{f}) \text{ abo } F(\vec{u}) = 0.5 \int_{\Omega} (\Lambda \vec{u}, \vec{u}) d\Omega - \int_{\Omega} (\vec{u}, \vec{f}) d\Omega, \quad (17.12)$$

в якому враховані ГУ (17.2). Зворотне твердження теж є вірним. При цьому оператор Λ в операторному рівнянні (17.1) повинен бути *лінійним, симетричним* та *визначеним як позитивний*. Це характерно, зокрема, для лінійних задач теплопровідності та термопружності.

Якщо вектор $\vec{u}_0(\vec{x})$ наближується виразом (17.4-б) у варіанті а/ (див. п.17.1.1), то умовою мінімуму функціонала (17.12) є вираз

$$\partial F(\vec{u}) / \partial c_j = 0; \quad j = 1, 2, \dots, J.$$
 (17.13)

Як і для методу найменших квадратів доведено, що послідовність $\lim_{J\to\infty} F(\vec{u}_J)$ є такою, що мінімізує функціонал (17.12), тобто при цьому є розв'язком задачі (17.1): $\lim_{J\to\infty} ||\vec{u}_J - \vec{u}_0||_{\Omega} = 0$.

17.1.4. Ідея методу Бубнова-Гальоркіна (Гальоркіна)

Метод був створений у 1911 році російським морським офіцером, проектувальником, будівником та вченим одночасно (він керував проектуванням та будівництвом військових кораблів, підводних човнів) І.Г. Бубновим (1872-1919 pp.), опублікований в 1913 й 1914 роках, зокрема в частині 2 книги "Строительная механика корабля". Б.Г. Гальоркін (1871-1945 рр.) у 1915 році показав, що для цього методу не обов'язкова умова позитивної визначеності оператора L, тобто метод не пов'язаний з методом Релея-Рітца, є самостійним. У подальшому Б.Г. Гальоркін широко застосовував цей метод, тому зазвичай його називають методом Гальоркіна. Цей метод набув теоретичного обгрунтування М.В. Келдишем (1911-1978 pp.) у 1942 р. та С.Г. Міхліним (1908-1990 pp.) у 1948 році.

Нагадаємо, що у п.2.6.2 було вказано, що якщо у нелінійному рівнянні (17.1) крайової задачі, тобто у $\Lambda(\vec{u}) = \vec{f}$, оператор Λ є потенційним, то і у рівнянні $R(\vec{u}) = L(\vec{u}) - \vec{f}$ оператор R теж є потенційним. А потенційні нелінійні оператори мають наступну важливу властивість: якщо для (17.1) існує розв'язок $\vec{u}_0 \in X$, який є критичною точкою для деякого функціонала $P(\vec{u})$, градієнтом якого є $R(\vec{u})$ (див. Розділ 2.5), то

$$(R(\vec{u}_0), \vec{w}) = 0$$
, inakune $((L(\vec{u}_0) - \vec{f}), \vec{w}) = 0$ (17.14-a)

для $\forall \vec{w} \in H$. Тобто існує ортогональність векторів $L(\vec{u}_0) - \vec{f}$ та $\vec{w} \in X$.

Отже, якщо $R_{\Omega}(\vec{u}) = \Lambda(\vec{u}) - \vec{f}$ та $R_S(\vec{u}) = K(\vec{u}) - \vec{g}$ містять тільки потенційні оператори, то властивість (17.14-а) можна застосовувати для отримання розв'язку відповідної нелінійної задачі. Доказано, що достатньою умовою для створення потенційного оператора є *симетричність* всіх операторів, що входять до функціонала, у нашому випадку – операторів Λ та *K* крайової задачі. Саме це характерно для більшості крайових задач механіки деформівного твердого тіла. У параграфі 23 книги [7] розглядаються додаткові умови до операторів крайової задачі, зокрема, що виникають внаслідок дискретності простору.

Відзначимо, що $R_{\Omega}(\vec{u}) = \Lambda(\vec{u}) - \vec{f}$ та $R_{S}(\vec{u}) = K(\vec{u}) - \vec{g}$ є похибками розв'язку крайової задачі $\Lambda(\vec{u}) = \vec{f}$ в Ω з граничною умовою $R_{S}(\vec{u}) = K(\vec{u}) - \vec{g}$.

Отже, якщо простір має визначену границю S, яка обмежує замкнений простір $\Omega \subset H$, а наближення шукається з використанням (17.4-б), то записується необхідна кількість рівнянь типу (17.14-а) — рівняння методу Бубнова-Гальоркіна:

$$F_{j} = \int_{\Omega} ((\Lambda(\vec{u}) - \vec{f}), \vec{w}_{j}) d\Omega + \int_{S} ((K(\vec{u}) - \vec{g}), \vec{\tilde{w}}_{j}) dS = 0, \quad j = 1, 2, \dots$$
(17.14-6)

Для отримання квадратної матриці в загальній САР (відносно $c_i(t)$ або c_i), яка буде породжуватися виразами (17.14-б), потрібна однакова кількість вагових \vec{w}_i , \vec{w}_i та базисних векторів $\vec{\phi}_i$, тобто j = 1, 2, ..., J.

Ваговими векторами \vec{w}_j , \vec{w}_j можна обирати базисні вектори $\vec{\varphi}_j$ (запропоновано ще І.Г. Бубновим). Тоді породжувана матриця САР буде *симетричною*, що полегшить розв'язування САР та покращить точність розв'язку.

Як окремий, але дуже зручний випадок обирання компонент базисних і вагових векторів одночасно, в методі Бубнова-Гальоркіна використовують тригонометричні функції, наприклад (матрична форма запису вектора):

 $\vec{\varphi}_j = \{\sin(j\pi x_1/L_{x_1}), \sin(j\pi x_2/L_{x_2}), \sin(j\pi x_3/L_{x_3})\}^T; j = 1, 2, ..., J,$ (17.15) де L_{x_n} – довжина області в напрямку координати x_n .

Внаслідок ортогональності базисних векторів, яка виражається у тому, що

$$\int_{0}^{L_{x_n}} \sin(i\pi x_n / L_{x_n}) \sin(j\pi x_n / L_{x_n}) dx_n = \begin{cases} L_{x_n} / 2; & i = j; \\ 0; & i \neq j, \end{cases}$$
(17.16)

матриця САР виявляється діагональною, тому відразу ж можна отримати шукані коефіцієнти. Отже, ортогональність базисних функцій є важливою властивістю, що дозволяє різко спростити процес розв'язування САР (частково використовується у методі скінченних елементів).

17.1.5. Ідея методу зважених похибок наближення

Метод зважених похибок наближення розробив наприкінці 1930-х років Г.І. Петров (1912-1987 рр.). Ще метод називають методом моментів при особливому вибору вагових векторів (див. наприкінці цього пункту), а також методом Петрова, методом Гальоркіна-Петрова, оскільки метод Бубнова-Гальоркіна (див.п.17.1.4) можна трактувати як окремий випадок методу Петрова. Далі будемо називати його методом зважених похибок наближення (МЗПН), оскільки саме таку назву найчастіше застосовують у світовій літературі (зокрема, англійською: Weighted Residual Methods – WRM).

Згідно з цим методом наближений розв'язок задачі шукається із умови, що

$$\int_{\Omega} (\vec{R}_{\Omega}(\vec{u}), \vec{w}_{j}) d\Omega = 0 \text{ alo } \int_{\Omega} (\vec{R}_{\Omega}(\vec{u}), \vec{w}_{j}) d\Omega + \int_{S} (\vec{R}_{S}(\vec{u}), \vec{\tilde{w}}_{j}) dS = 0; \quad j = 1, 2, \dots, J, \quad (17.17)$$

де $\vec{w}_j(\vec{x}), \vec{w}_j(\vec{x})$ – повні за енергією системи вагових векторів, причому ці системи в загальному випадку можуть бути незалежними; $\vec{R}_{\Omega}(\vec{u})$ і $\vec{R}_S(\vec{u})$ – вектори похибок наближення (17.8) всередині області і на її поверхні відповідно. Рівність нулю (17.17) відбиває загальну вимогу збіжності $\vec{u} \to \vec{u}_0$ при $J \to \infty$.

Доведено теорему, що при наявності необхідних похідних від $\vec{\psi}$ та $\vec{\varphi}_i$ при умові $J \to \infty$ вектор $\vec{u} \to \vec{u}_0$.

Теоретичною основою методу є фундаментальна теорема про проекції (див. підрозділ 2.1), згідно з якою для кожного вектору $\vec{u} \in \Omega \subset H$, де $\Omega \in$ замкненим простором, існує лише один вектор $\vec{u}^* \in \Omega$ такий, що

$$\|\vec{u} - \vec{u}^*\| < \|\vec{u} - \vec{u}^{\#}\|, \qquad (17.18)$$

де вектор $\vec{u}^{\#}$ – будь-який інший (не \vec{u}^{*}). Необхідною та достатньою умовою виконання цієї нерівності є ортогональність вектора $\vec{u} - \vec{u}^{*}$ будь-якому вектору $\vec{w}(\vec{x}) \in \Omega$. Ця теорема *не накладає ніяких обмежень* на оператори крайової задачі щодо лінійності/нелінійності, потенційності або залежності/незалежності від часу, тому МЗПН є універсальним для задач у *замкненому* просторі Ω .

МЗПН дає велику свободу для вибору вагових векторів $\vec{w}_j, \vec{\tilde{w}}_j$. Обов'язкові умови лише такі: це повинні бути повні за енергією лінійно незалежні системи. Тому метод найменших квадратів (див. п.17.1.2) та метод Бубнова-Гальоркіна (див. п.17.1.4) можна розглядати як окремі випадки МЗПН. Зазначимо, що в останньому в ролі вагових векторів використовуються саме базисні вектори.

Відомі й інші класичні варіанти обирання компонент вагових векторів. У методі поточкових колокацій вони задаються дельта-функціями Дірака

$$w_{j} = \widetilde{w}_{j} = \delta(\vec{x} - \vec{x}_{j}) = \delta(x^{1} - x_{j}^{1}) \cdot \delta(x^{2} - x_{j}^{2}) \cdot \delta(x^{3} - x_{j}^{3})$$
(17.19)

з такими властивостями

$$\delta(\vec{x} - \vec{x}_{j}) = 0, \quad \vec{x} \neq \vec{x}_{j}; \quad \delta(\vec{x} - \vec{x}_{j}) = \infty, \quad \vec{x} = \vec{x}_{j};$$

$$\int_{x^{1} < x_{j}^{1}} \int_{x^{2} < x_{j}^{2}} \int_{x^{3} < x_{j}^{3}} \varphi(\vec{x}) \delta(\vec{x} - \vec{x}_{j}) dx^{1} dx^{2} dx^{3} = \varphi(\vec{x}).$$
(17.20)

Це еквівалентне тому, що похибки наближення вважаються рівними нулю в заданих точках \vec{x}_i .

У методі колокацій по підобластям компоненти вагових векторів:

$$w_{j} = \tilde{w}_{j} = \begin{cases} 1, & x_{j}^{1} < x^{1} < x_{j+1}^{1}, & x_{j}^{2} < x^{2} < x_{j+1}^{2}, & x_{j}^{3} < x^{3} < x_{j+1}^{3}; \\ 0, & x^{1} < x_{j}^{1}, & x^{1} > x_{j+1}^{1}, & x^{2} < x_{j}^{2}, & x^{2} > x_{j+1}^{2}, & x^{3} < x_{j}^{3}, & x^{3} > x_{j+1}^{3}, \end{cases}$$
(17.21)

тому рівняння (17.17) відповідають деяким підобластям основної області.

У методі *моментів* компоненти вагових векторів збираються з елементів $(x)^{j-1}$, де $x \in$ одною з компонент просторового базису. Геометрична інтерпретація: момент площі під кривою, яка відповідає похибці наближення.

17.1.6. Застосування універсальних варіаційних принципів

Якщо крайова задача може бути сформульована у вигляді якогось варіаційного функціонала, то цей функціонал теж можна застосовувати для розв'язування задачі.

Наприклад, варіаційний функціонал (13.31)

$$\Psi(\delta) = \int_{\Omega} \sigma_{mn} \delta \varepsilon_{mn} d\Omega - \int_{\Omega} \hat{O}_m \delta u_m d\Omega - \int_{S_p} \hat{P}_m \delta u_m dS = 0,$$

який є справедливим для всіх (лінійних та нелінійних) крайових задач про напружено-деформований стан, є основою для побудови методів розв'язування відповідних крайових задач і буде використаний у подальших Розділах.

Часто застосування універсальних варіаційних принципів називають енергетичними методами розв'язування крайових задач.

17.1.7. Застосування прямих методів

Якщо вихідну систему рівнянь безпосередньо записати у вигляді, що дозволяє знайти розв'язок крайової задачі, то це є прямий метод розв'язуванні задачі.

Зокрема, метод скінченних різниць (МСР) є одним з таких методів. Іноді застосування варіаційних принципів теж вважають прямим методом.

Важливо, що після отримання прямим методом системи алгебраїчних рівнянь для деякого класу задач потрібно виясняти властивості матриці цієї системи: чи є вона симетричною, позитивно визначеною, добре обумовленою, тощо.

17.2. Послаблена форма методу зважених похибок наближення

Якщо в основній формулі зважених похибок наближення (17.17) в об'ємному інтегралі є просторові похідні високих порядків, то часто можна використати інтегральні формули, що дозволяють знизити порядок похідних – отримати *послаблену форму* методу. Ці формули також можуть допомогти позбутися поверхневих інтегралів по всій поверхні тіла. Крім того, зниження найвищого порядку похідної в крайовій задачі знижує вимоги до вектору (функції або поліному), що апроксимує розв'язок задачі. Це дозволяє з меншими витратами одержувати не менш точні розв'язки.

Нижче отримаємо послаблені форми методу зважених похибок наближення для задачі теплопровідності та для задачі про напружено-деформований стан тіла.

17.2.1. Послаблена форма методу зважених похибок наближення для задачі теплопровідності

Просторову похибку наближення R_{Ω} створимо з основного рівняння крайової задачі, тобто з (13.3-а), яке відображає баланс потоку тепла в кожній точці тіла, а саме:

$$R_{\Omega} = c_{P}\overline{\rho}\frac{\partial\theta}{\partial t} + c_{P}\overline{\rho}(\nabla_{k}\theta)V^{k} - \nabla_{i}(\lambda\nabla_{i}\theta) - \widehat{\omega}. \qquad (17.22)$$

Поверхневу похибку наближення *R_s* створимо з рівняння поверхневих граничних умов (13.6) за тепловим потоком:

$$R_{S} = \lambda \frac{\partial \theta}{\partial v}\Big|_{S_{G}} - \hat{q}\Big|_{S_{q}} - \tilde{q}\Big|_{S_{\alpha}} - \breve{q}\Big|_{S_{\beta}}.$$
(17.23)

Початкові умови (13.4) та ГУ 1-го роду (13.5) враховуються окремо.

Шуканий результат – це єдина скалярна функція: температура. Відповідно до основної формули МЗПН (17.17) складемо *J* функціоналів

$$F_{j} = \int_{\Omega} \left(c_{P} \overline{\rho} \frac{\partial \theta}{\partial t} + c_{P} \overline{\rho} (\nabla_{k} \theta) V^{k} - \nabla_{k} (\lambda \nabla_{k} \theta) - \widehat{\omega} \right) w_{j} d\Omega +$$
$$+ \int_{S} \left(\lambda \frac{\partial \theta}{\partial \nu} \Big|_{S_{q}} - \widehat{q} \Big|_{S_{q}} - \widetilde{q} \Big|_{S_{q}} - \widetilde{q} \Big|_{S_{\beta}} \right) \widetilde{w}_{j} dS = 0; \quad j = 1, 2, ..., J.$$
(17.24)

На вільній від ГУ другого роду поверхні тіла, яку позначимо як $S_0 = S \setminus S_G$, інтеграл $\int_{S_0} \lambda \frac{\partial \theta}{\partial v} \tilde{w}_j dS$ тотожно дорівнює нулю. Тому замість інтеграла $\int_{S_G} \lambda \frac{\partial \theta}{\partial v} \tilde{w}_j dS$ можемо розглядати інтеграл $\int_{S} \lambda \frac{\partial \theta}{\partial v} \tilde{w}_j dS$, тобто на повній поверхні, та навпаки.

Згідно з формулою Гріна-Стокса для скалярних функцій θ та w

$$\int_{\Omega} \frac{\partial \theta}{\partial v} w dS = \int_{\Omega} \nabla \theta \nabla w d\Omega + \int_{\Omega} \nabla (\nabla \theta) w d\Omega, \qquad (17.25)$$

тому інтеграл $\int_{S} \lambda \frac{\partial \theta}{\partial v} \tilde{w}_j dS$ в (17.24) можна замінити сумою інтегралів $\int_{\Omega} \nabla_k \theta \lambda \nabla_k \tilde{w}_j d\Omega + \int_{\Omega} \nabla_k (\lambda \nabla_k \theta) \tilde{w}_j d\Omega$, але при цьому прийняти, що коефіцієнт теплопровідності λ є постійною величиною. Для того, щоб скоротити другий інтеграл цього виразу з інтегралом $\int_{\Omega} \nabla_k (\lambda \nabla_k \theta) w_j d\Omega$ із (17.24), який має протилежний знак, потрібно прийняти, що $\tilde{w}_j = w_j$. Це можна робити, оскільки ці системи – довільні, що повинні задовольняти лише вимогам повноти за енергією. Отже:

$$F_{j} = \int_{\Omega} \left(c_{P} \overline{\rho} \frac{\partial \theta}{\partial t} + c_{P} \overline{\rho} (\nabla_{k} \theta) V^{k} + \nabla_{k} \theta \lambda \nabla_{k} - \widehat{\omega} \right) w_{j} d\Omega - \int_{S_{q}} \widehat{q} w_{j} dS - \int_{S_{\alpha}} \widetilde{q} w_{j} dS - \int_{S_{\beta}} \widetilde{q} w_{j} dS = 0; \quad j = 1, 2, ..., J.$$
(17.26)

Бачимо, що формула Гріна-Стокса (17.25) для скалярних функцій w_i і θ дозволила понизити найвищий порядок похідних до першого (у виразі (17.24) найвищий порядок похідних – другий (член $\nabla_i (\lambda \nabla_i \theta)$)), а також водночас виключити з (17.24) інтеграл по всій поверхні тіла *S*.

Це й є послаблена форма МЗПН (Петрова, Гальоркіна-Петрова) для задачі теплопровідності. Зазначимо, що вимоги до вагових функцій змінилися: окрім повноти вони ще повинні бути один раз диференційованими, що прийнятно. А от базисні функції тепер будуть диференціюватися лише один раз, а не двічі, що знижує вимоги до них. Оскільки формула Гріна-Стокса є точною, то послаблена форма МЗПН не погіршує майбутній розв'язок.

Примітка 17.2. Можна застосовувати й перший варіант основного рівняння МЗПН (17.17) з просторовою похибкою наближення (17.22). Тоді після використання формули Гріна-Стокса (17.25) з'являється можливість врахувати граничні умови (13.6) задачі. Результатом буде той же вираз (17.26).

17.2.2. Послаблена форма методу зважених похибок наближення для задачі про напружено-деформований стан тіла

Використаємо перший варіант основного рівняння МЗПН (17.17):

$$\int_{\Omega} (\vec{R}_{\Omega}, \vec{w}_j) d\Omega = 0; \quad j = 1, 2, ..., J.$$
(17.27)

Із застосуванням рівняння рівноваги (13.15) складемо вектор похибок наближення \vec{R}_{Ω} з компонентами:

$$R_{\Omega})_m = \nabla_n \sigma_{mn} + \hat{O}_m; \quad m, n = 1, 2, 3.$$
 (17.28)

Відповідно до (17.27) складемо Ј виразів (функціоналів)

Частина V. Розділ 17

$$(F_{j})_{m} = \int_{\Omega} (\nabla_{n} \sigma_{mn} + \hat{O}_{m}) w_{j} d\Omega = 0; \quad m, n = 1, 2, 3; \quad j = 1, 2, ..., J.$$
(17.29)

Оскільки $\nabla_n(\sigma_{mn}w_j) = (\nabla_n\sigma_{mn})w_j + \sigma_{mn}(\nabla_nw_j)$, тобто $(\nabla_n\sigma_{mn})w_j = \nabla_n(\sigma_{mn}w_j) - -\sigma_{mn}(\nabla_nw_j)$, то замість (17.29) маємо:

$$(F_j)_m = \int_{\Omega} \nabla_n (\sigma_{mn} w_j) d\Omega - \int_{\Omega} \sigma_{mn} (\nabla_n w_j) d\Omega + \int_{\Omega} \widehat{O}_m w_j d\Omega = 0.$$
(17.30)

Відзначимо, що тим самим на вагові функції *w_j* накладається додаткова умова: наявність просторових похідних першого порядку.

Відповідно до формули Остроградського-Гаусса:

$$\int_{\Omega} \nabla_n (\sigma_{mn} w_j) d\Omega = \int_{S} \sigma_{mn} w_j v_n dS. \qquad (17.31)$$

Але з граничних умов (13.26) вираз $(\sigma_{mn}v_n)|_{S_p} = \hat{P}_m$, а інша частина поверхні (якщо вона є) вільна від навантаження. Тому, підставивши (13.26) у (17.31) і результат – у (17.30), одержимо, помінявши знаки на зворотні:

$$(F_j)_m = -\int_{S_p} \widehat{P}_m w_j dS + \int_{\Omega} \sigma_{mn} (\nabla_n w_j) d\Omega - \int_{\Omega} \widehat{O}_m w_j d\Omega = 0.$$
(17.32)

Оскільки тензор σ_{mn} – симетричний, то справедлива рівність

$$\sigma_{mn}(\nabla_n w_j) = \sigma_{mn} \frac{1}{2} (\nabla_j w_n + \nabla_n w_j). \qquad (17.33)$$

З рівнянь (13.16), (13.19) і принципу суперпозиції малих деформацій різноманітної природи отримаємо:

$$\sigma_{mn} = E_{mnkl} \varepsilon_{kl}^{e} = E_{mnkl} (\varepsilon_{kl} - \varepsilon_{kl}^{\theta} - \varepsilon_{kl}^{pc0}) = E_{mnkl} \Big[0.5 (\nabla_{k} u_{l} + \nabla_{l} u_{k}) - \varepsilon_{kl}^{\theta} - \varepsilon_{kl}^{pc0} \Big], \quad (17.34)$$

де ε_{kl}^{pc0} – компоненти сумарного тензора деформацій пластичності, повзучості та початкових: $\varepsilon_{kl}^{pc0} = \varepsilon_{kl}^{p} + \varepsilon_{kl}^{c} + (\varepsilon_{kl})_{0}$.

Підставляючи (17.33) і (17.34) у (17.32), одержимо, що:

$$(F_{j})_{m} = \int_{\Omega} E_{mnkl} \Big[0.5(\nabla_{k}u_{l} + \nabla_{l}u_{k}) - \varepsilon_{kl}^{\theta} - \varepsilon_{kl}^{pc0} \Big] \Big[0.5(\nabla_{j}w_{n} + \nabla_{n}w_{j}) \Big] d\Omega - \int_{S_{p}} \widehat{P}_{m}w_{j}dS - \int_{\Omega} \widehat{O}_{m}w_{j}d\Omega = 0; \quad m, n = 1, 2, 3; \quad j = 1, 2, ..., J.$$
(17.35)

Це й є послаблена форма МЗПН (Петрова, Гальоркіна-Петрова) для крайової задачі про напружено-деформований стан тіла.

Нагадаємо, що вимоги для вагових функцій змінилися: окрім повноти за енергією та лінійної незалежності, вони ще повинні бути один раз диференційованими, що прийнятно. А от базисні функції тепер будуть диференційованими лише один раз, а не двічі, що знижує вимоги до них.

Оскільки формула Остроградського-Гаусса є точною, то послаблена форма МЗПН не погіршує майбутній розв'язок.

17.3. Приклад застосування методу граничного розв'язку до крайової задачі з двовимірним рівнянням Лапласа

У Розділі 17.1 відмічалося, що при варіанті б) обирання компонент наближення $\vec{u}(\vec{x},t)$ існують значні проблеми з вибором таких базисних функцій в Ω . На прикладі крайової задачі з двовимірним рівнянням Лапласа $\nabla^2 u(x, y) = 0$ покажемо один з варіантів подолання цієї проблеми.

Відомо, що для будь-якої аналітичної функції комплексної змінної f(z) = u(x, y) + iq(x, y) повинна виконуватися умова $\nabla^2 u(x, y) = \nabla^2 q(x, y) = 0$. Тому для вибору функцій, що задовольнять рівнянню Лапласа, можна безпосередньо використати аналітичну функцію $f(z) = z^n = (x + iy)^n$. Тоді маємо при

n = 0:	u = 1;	q=0;	
n = 1:	u = x;	q = y;	
n = 2:	$u=x^2-y^2;$	q = 2xy;	
n = 3:	$u=x^3-3xy^2;$	$q=3x^2y-y^3;$	
n = 4:	$u = x^4 - 6x^2y^2 + y^4;$	$q=4x^3y-4xy^3;$	та ін.

З усього нескінченного набору можливих базисних функцій потрібно виділити деяку кількість допустимих. Для цього необхідно конкретизувати задачу.

Припустимо, що розглядаємо задачу про кручення стрижня прямокутного перетину. Вона описується диференційним рівнянням (15.51), тобто $\partial^2 u / \partial x^2 + \partial^2 u / \partial y^2 = -2$ в прямокутній області $-a \le x \le a$, $-b \le y \le b$, з граничною умовою $u \mid_s = 0$.

Щоб від неоднорідного рівняння Лапласа перейти до однорідного, введемо нову функцію v таку, що $u = v - (x^2 + y^2)/2$. Підставивши цей вираз в (15.51), отримаємо $\partial^2 v / \partial x^2 + \partial^2 v / \partial y^2 = 0$. Гранична умова задачі набуває вигляд $v|_s = v_s = (x^2 + y^2)/2|_s$.

Тепер можна сформулювати вимоги до базисних функцій, що висуваються цією задачею. Очевидно, що її розв'язок повинен бути *симетричним* відносно координатних осей x і y. Тому допустимими є базисні функції з *парними* степенями при x і y. Із виписаних вище можливих базисних функцій це лише такі:

$$\varphi_1 = 1;$$
 $\varphi_2 = x^2 - y^2;$ $\varphi_3 = x^4 - 6x^2y^2 + y^4,$ тобто приймемо $J = 3$. Отже, наближення

$$v \approx c_1 + c_2(x^2 - y^2) + c_3(x^4 - 6x^2y^2 + y^4)$$

точно задовольняє умовам задачі в Ω при будь-яких значеннях c_j , що і вимагається в методі граничного розв'язку. Друге рівняння (17.17) методу зважених похибок наближення, з урахуванням того, що в нашому випадку вектор R_s є скалярною функцією, виглядає як

 $\int_{S} R_{S} \widetilde{w}_{j} dS = 0; \quad j = 1, 2, 3,$

де $R_s = v_s - [(x^2 + y^2)/2] \Big|_s = [c_1 + c_2(x^2 - y^2) + c_3(x^4 - 6x^2y^2 + y^4) - (x^2 + y^2)/2] \Big|_s.$

Якщо застосувати ідею Бубнова для обирання вагових функцій, то $\widetilde{w}_{_j} = \varphi_{_j}$. Після інтегрування виразів (підкреслено вагові функції):

$$2\int_{0}^{a} [c_{1} + c_{2}(x^{2} - b^{2}) + c_{3}(x^{4} - 6x^{2}b^{2} + b^{4}) - (x^{2} + b^{2})/2] \cdot \underline{1} \cdot dx + + 2\int_{0}^{b} [c_{1} + c_{2}(a^{2} - y^{2}) + c_{3}(a^{4} - 6a^{2}y^{2} + y^{4}) - (a^{2} + y^{2})/2] \cdot \underline{1} \cdot dy = 0;$$

$$2\int_{0}^{a} [c_{1} + c_{2}(x^{2} - b^{2}) + c_{3}(x^{4} - 6x^{2}b^{2} + b^{4}) - (x^{2} + b^{2})/2] \cdot \underline{(x^{2} - b^{2})} \cdot dx + + 2\int_{0}^{b} [c_{1} + c_{2}(a^{2} - y^{2}) + c_{3}(a^{4} - 6a^{2}y^{2} + y^{4}) - (a^{2} + y^{2})/2] \cdot \underline{(a^{2} - y^{2})} \cdot dy = 0;$$

$$2\int_{0}^{a} [c_{1} + c_{2}(x^{2} - b^{2}) + c_{3}(x^{4} - 6x^{2}b^{2} + b^{4}) - (x^{2} + b^{2})/2] \cdot \underline{(x^{4} - 6x^{2}b^{2} + b^{4})} \cdot dx + + 2\int_{0}^{b} [c_{1} + c_{2}(a^{2} - y^{2}) + c_{3}(a^{4} - 6a^{2}y^{2} + y^{4}) - (a^{2} + y^{2})/2] \cdot \underline{(a^{4} - 6a^{2}y^{2} + y^{4})} \cdot dx + + 2\int_{0}^{b} [c_{1} + c_{2}(a^{2} - y^{2}) + c_{3}(a^{4} - 6a^{2}y^{2} + y^{4}) - (a^{2} + y^{2})/2] \cdot \underline{(a^{4} - 6a^{2}y^{2} + y^{4})} \cdot dy = 0$$

отримаємо СЛАР $A_{ij}c_j = b_i$; *i*, *j* = 1,2,3, де компоненти

$$\begin{aligned} A_{11} &= a + b \,; \quad A_{12} = A_{21} = (a^3 - 3ab^2 + 3a^2b - b^3)/3 \,; \\ A_{13} &= A_{31} = a^5/5 - 2a^3b^2 + ab^4 + a^4b - 2a^2b^3 + b^5/5 \,; \\ A_{22} &= a^5/5 - 2a^3b^2/3 + ab^4 + a^4b - 2a^2b^3/3 + b^5/5 \,; \\ A_{23} &= A_{32} = a^7/7 - 7a^5b^2/5 + 7a^3b^4/3 - ab^6 + a^6b - 7a^4b^3/3 + 7a^2b^5/5 - b^7/7 \,; \\ A_{33} &= a^9/9 - 12a^7b^2/7 + 38a^5b^4/5 - 4a^3b^6 + ab^8 + a^8b - 4a^6b^3 + 38a^4b^5/5 - 12a^2b^7/7 + b^9/9 \,; \end{aligned}$$

$$b_1 = (a^3 + 3ab^2 + 3a^2b + b^3)/6; \quad b_2 = (a^5 - 5ab^4 + 5a^4b - b^5)/10;$$

$$b_3 = (a^7/7 - a^5b^2 - 5a^3b^4/3 + ab^6 + a^6b - 5a^4b^3/3 - a^2b^5 + b^7/7)/2.$$

Якщо прийняти розміри a = 3; b = 2, то СЛАР отримає компоненти

	5	12.3333	-95	C_1		20.8333	
12	.3333	145	98.543	$\left\{c_2\right\}$	} = <	78.1	}.
	-95	98.543	18170.4	$\left[c_{3}\right]$		- 539.643	

Наближений розв'язок системи такий: $c_1 \approx 3.2154$; $c_2 \approx -0.2749$; $c_3 \approx -0.01438$. Тому

 $u \approx 3.2154 - 0.2749(x^2 - y^2) - 0.001438(x^4 - 6x^2y^2 + y^4) - (x^2 + y^2)/2.$

За умови $G\theta = 1$ можна отримати, що крутильний момент $M_{KP} = 2G\theta \int_{\Omega} u(x, y) dx dy \approx 75.51$ (значення розв'язку в рядах дорівнює 76.40, а за МСР в Розділі 15.3.4 одержали $M_{KP} = 65.41$).

Контрольні питання до підрозділу 17.1

1. Яку ідею закладено в метод Фур'є?

- 2. Які властивості повинні мати базисні функції (вектори) методу Фур'є?
- 3. Вкажіть три можливі варіанти вибору базисних функцій (векторів).
- 4. Яка особливість обирання базисних функцій в методі Рвачева?
- 5. Які особливості має метод найменших квадратів?
- 6. Які особливості має метод Релея-Рітца (Рітца)?
- 7. Які особливості має метод Бубнова-Гальоркіна (Гальоркіна)?

8. Яка сутність рекомендації І.Г. Бубнова з вибору вагових функцій? Що є слідством вибору вагових функцій за рекомендацією І.Г. Бубнова?

9. Які особливості має метод зважених похибок наближення?

10. Чи можливе застосування універсальних варіаційних принципів та прямих методів для отримання розв'язків крайових задач?

Контрольні питання до підрозділу 17.2

1. Яку інтегральну формулу застосовують для отримання послабленої форми методу зважених похибок наближення в задаче про теплопровідність? Як це робиться?

2. Яку інтегральну формулу застосовують для отримання послабленої форми методу зважених похибок наближення в задаче про напружено-деформований стан? Як це робиться?

Контрольні питання до підрозділу 17.3

1. Чому розв'язок крайової задачі у прикладі повинен бути симетричним відносно координатних осей *x* і *y*?

2. Який метод отримання розв'язку крайової задачі виявився точнішим: МКР чи МЗПН? Чому?

Розділ 18

ПРОСТОРОВО-ЧАСОВЕ НАБЛИЖЕННЯ КРАЙОВИХ ЗАДАЧ ЗА МЕТОДОМ ЗВАЖЕНИХ ПОХИБОК НАБЛИЖЕННЯ

18.1. Метод часткової алгебраїзації неперервними базисними функціями в методі зважених похибок наближення

Нагадаємо (див. Розділ 16.1), що просторово-часове наближення крайової задачі може проводитися:

• методом скінченних різниць (МСР, розглянуто у Розділі 16);

• методом повної або часткової алгебраїзації з використанням наступних типів базисних функцій: точкових, кусково-визначених або неперервних.

За методом часткової алгебраїзації для часової редукції застосовується МСР або пряме інтегрування у часі, а для просторової – перераховані вище типи базисних функцій. У методі повної алгебраїзації вся редукція проводиться на основі базисних функцій.

У цьому Розділі розглянемо просторову алгебраїзацію задачі з використанням базисних функцій та методу зважених похибок наближення. Оскільки цей підхід є загальним у тому сенсі, що дозволяє одержувати і аналітичні, і чисельні наближені розв'язки, то в цьому Розділі не розглядаються сіткові аналоги крайових задач та методів (як і у Розділі 17).

Точний вектор-розв'язок $\vec{u}_0(\vec{x},t)$ наблизимо виразом (17.4-б), а саме:

$$\vec{u}_{0}(\vec{x},t) \approx \vec{u}(\vec{x},t) = \vec{\psi}(\vec{x},t) + \sum_{i=1}^{J} c_{i}(t) \vec{\varphi}_{i}(\vec{x}), \qquad (18.1)$$

де $\vec{\psi}(\vec{x},t)$ та $\vec{\varphi}_i(\vec{x})$ є базисними векторами, а $c_i(t)$ – невідомі функції часу.

Застосуємо метод зважених похибок наближення. Випишемо основне рівняння методу – друге рівняння (17.17), а саме:

$$\int_{\Omega} (\vec{R}_{\Omega}, \vec{w}_{j}) d\Omega + \int_{S} (\vec{R}_{S}, \vec{\tilde{w}}_{j}) dS = 0; \quad j = 1, 2, ..., J,$$
(18.2)

де $\vec{w}_j, \vec{\tilde{w}}_j$ – повні системи вагових векторів, причому системи в загальному випадку можуть бути незалежними; \vec{R}_{Ω} і \vec{R}_s – вектори похибок наближення всередині області та на її поверхні відповідно.

Розглянемо крайову задачу *параболічного* типу, що описується узагальненим диференційним рівнянням

$$\frac{\partial \vec{u}}{\partial t} + L(\vec{u}) = \vec{f} \quad \mathbf{B} \quad \Omega \tag{18.3}$$

з ГУ вигляду

$$K(\vec{u}) = \vec{g} \text{ Ha } S \tag{18.4}$$

за початковою умовою

$$\vec{u}(\vec{x},0) = \hat{\vec{u}}_0(\vec{x})$$
. (18.5)

Зберемо вектори похибок наближення

$$\vec{R}_{\Omega} = \frac{\partial \vec{u}}{\partial t} + L(\vec{u}) - \vec{f} \quad \mathbf{B} \ \Omega; \quad \vec{R}_{S} = K(\vec{u}) - \vec{g} \quad \mathbf{Ha} \ S.$$
(18.6)

Підставивши наближення $\vec{u}(\vec{x},t)$ (18.1) у (18.6), результат – у (18.2), отримаємо систему диференційних рівнянь вигляду

$$G\frac{d\vec{c}}{dt} + A(\vec{c}) = \vec{b}, \qquad (18.7)$$

де у випадку *лінійності* операторів L і K матриці [G] і [A] та вектору $\{b\}$ мають компоненти

$$G_{ij} = \int_{\Omega} \vec{\varphi}_i \vec{w}_j d\Omega; \quad i, j = 1, 2, ..., J;$$
(18.8)

$$A_{ij} = \int_{\Omega} (L\vec{\varphi}_i) \vec{w}_j d\Omega + \int_{S} (K\vec{\varphi}_i) \vec{\tilde{w}}_j dS ; \quad i, j = 1, 2, ..., J ;$$
(18.9)

$$b_{j} = \int_{\Omega} (\vec{f} - L\vec{\psi}) \vec{w}_{j} d\Omega + \int_{S} (\vec{g} - K\vec{\psi}) \vec{\widetilde{w}}_{j} dS - \int_{\Omega} \frac{\partial \vec{\psi}}{\partial t} \vec{w}_{j} d\Omega; \quad j = 1, 2, ..., J.$$
(18.10)

Якщо призначити $\vec{\psi} = 0$, то (18.10) значно спрощується:

$$b_{j} = \int_{\Omega} \vec{f} \vec{w}_{j} d\Omega + \int_{S} \vec{g} \vec{\tilde{w}}_{j} dS; \quad j = 1, 2, ..., J.$$
(18.11)

Якщо оператори *L* і *K* – нелінійні, то компоненти матриці [*A*] і вектору {*b*} будуть мати інший вигляд і будуть залежати від \vec{c} , а вираз (18.8) для компонент матриці [*G*] зберігається. При цьому зазвичай приймають $\vec{\psi} = 0$.

Розглянемо також крайову задачу *гіперболічного* типу, що описується узагальненим диференційним (операторним) рівнянням

$$M\frac{\partial^2 \vec{u}}{\partial t^2} + C\frac{\partial \vec{u}}{\partial t} + L(\vec{u}) = \vec{f} \quad \mathbf{B} \quad \Omega$$
(18.12)

з ГУ вигляду (18.4) за початкових умов

$$\vec{u}(\vec{x},0) = \hat{\vec{u}}_0(\vec{x}); \quad \frac{\partial \vec{u}(\vec{x},0)}{\partial t} = \hat{\vec{h}}(\vec{x}); \quad \frac{\partial^2 \vec{u}(\vec{x},0)}{\partial t^2} = \hat{\vec{g}}(\vec{x}).$$
(18.13)

Зберемо вектори похибок наближення

$$\vec{R}_{\Omega} = M \frac{\partial^2 \vec{u}}{\partial t^2} + C \frac{\partial \vec{u}}{\partial t} + L(\vec{u}) - \vec{f} \quad \mathbf{B} \quad \Omega; \quad \vec{R}_S = K(\vec{u}) - \vec{g} \quad \mathbf{Ha} \quad S.$$
(18.14)

Підставивши (18.1) у (18.14), результат – у (18.2), отримаємо систему диференційних рівнянь вигляду

$$H\frac{d^{2}\vec{c}}{dt^{2}} + G\frac{d\vec{c}}{dt} + A(\vec{c}) = \vec{b} , \qquad (18.15)$$

де у випадку *лінійності* операторів L і K матриці [H], [G] і [A] мають компоненти

$$H_{ij} = \int_{\Omega} (M \,\vec{\varphi}_i) \,\vec{w}_j d\Omega \; ; \quad i, j = 1, 2, ..., J \; ; \tag{18.16}$$

$$G_{ij} = \int_{\Omega} (C\vec{\varphi}_i) \vec{w}_j d\Omega; \quad i, j = 1, 2, ..., J;$$
(18.17)

Ч	251		
$A_{ij} = \int (L\vec{\varphi}_i)\vec{w}_j d\Omega +$	$\int (K\vec{\varphi}_i)\vec{\tilde{w}}_j dS;$	i, j = 1, 2,, J;	(18.18)

$$b_{j} = \int_{\Omega} (\vec{f} - L\vec{\psi}) \vec{w}_{j} d\Omega + \int_{S} (\vec{g} - K\vec{\psi}) \vec{\tilde{w}}_{j} dS - \frac{\partial^{2}}{\partial t^{2}} \int_{\Omega} (M\vec{\psi}) \vec{w}_{j} d\Omega - \frac{\partial}{\partial t} \int_{\Omega} (C\vec{\psi}) \vec{w}_{j} d\Omega; \quad j = 1, 2, ..., J.$$
(18.19)

Якщо призначити $\vec{\psi} \equiv 0$, то (18.19) спрощується до виразу (18.11).

Якщо оператори *L* і *K* – нелінійні, то компоненти матриці [*A*] і вектору {*b*} будуть мати інший вигляд і будуть залежати від \vec{c} , а вирази (18.16) і (18.17) зберігаються. При цьому зазвичай приймають $\vec{\psi} \equiv 0$.

Для вибору вагових векторів \vec{w}_j і $\tilde{\vec{w}}_j$, а також базисних векторів $\vec{\psi}(\vec{x},t)$ і $\vec{\phi}_i(\vec{x})$ використовуються ті ж методи, що і в стаціонарних задачах.

Таким чином, отримуємо дві системи звичайних алгебро-диференційних рівнянь (18.7) і (18.15), які необхідно розв'язати. Проблема полягає у позбавленні від операторів d/dt або d^2/dt^2 , а також у алгебраїзації L і K. Тут розглянемо тільки проблему позбавлення від часових операторів.

18.2. Крайові задачі параболічного типу

Розглянемо диференційне рівняння (18.7), зведемо його до вигляду

$$\frac{d\vec{c}}{dt} = R\vec{c} + \vec{q} , \qquad (18.20)$$

де позначено

$$R\vec{c} = -G^{-1}A(\vec{c}); \quad \vec{q} = G^{-1}\vec{b}.$$
 (18.21)

Ця задача допускає пряме інтегрування за часом.

18.2.1. Класичні двошарові схеми

Помножимо (18.20) на dt і проведемо пряме інтегрування у часі на часовому кроці $\Delta t_n = t_{n+1} - t_n$:

$$\int_{\vec{c}^n}^{\vec{c}^{n+1}} d\vec{c} = \int_{t_n}^{t_{n+1}} (R\vec{c} + \vec{q}) dt \,.$$
(18.22)

Інтеграл в лівій частині обчислюється точно, а саме дорівнює $\vec{c}^{n+1} - \vec{c}^n$. Підінтегральний вираз в правій частині (18.22) містить невідомий вектор \vec{c} . Розкладемо його в ряд Тейлора в околі t_n і, використовуючи рівняння (18.20), запишемо, що:

$$\vec{c}^{n+1} \approx \vec{c}^{n} + \int_{t_{n}}^{t_{n+1}} \left[(R^{n}\vec{c}^{n} + \vec{q}^{n}) + \frac{d}{dt}(R\vec{c} + \vec{q}) \Big|_{t_{n}} \Delta t + O(\Delta t^{2}) \right] dt = = \vec{c}^{n} + (R^{n}\vec{c}^{n} + \vec{q}^{n})\Delta t + \int_{t_{n}}^{t_{n+1}} \left[(\dot{R}\vec{c} + R\dot{\vec{c}} + \dot{\vec{q}}) \Big|_{t_{n}} \Delta t \right] dt + O(\Delta t^{3}) =$$

$$= \vec{c}^{n} + (R^{n}\vec{c}^{n} + \vec{q}^{n})\Delta t + \int_{t_{n}}^{t_{n+1}} \left\{ \left[\dot{R}\vec{c} + R(R\vec{c} + \vec{q}) + \dot{\vec{q}} \right] \Big|_{t_{n}} \Delta t \right\} dt + O(\Delta t^{3}) = = \vec{c}^{n} + (R^{n}\vec{c}^{n} + \vec{q}^{n})\Delta t + \left[\dot{R}^{n}\vec{c}^{n} + (R^{n})^{2}\vec{c}^{n} + R^{n}\vec{q}^{n} + \dot{\vec{q}}^{n} \right] \Delta t^{2} / 2 + O(\Delta t^{3}),$$
(18.23)

де $\dot{R} = dR/dt$, $\dot{\vec{q}} = d\vec{q}/dt$, а верхній індекс *n* вказує на приналежність до часу t_n .

Якщо розклад в ряд Тейлора провести в околі t_{n+1} , то:

$$\vec{c}^{n} \approx \vec{c}^{n+1} - (R^{n+1}\vec{c}^{n+1} + q^{n+1})\Delta t + + [\dot{R}^{n+1}\vec{c}^{n+1} + (R^{n+1})^{2}\vec{c}^{n+1} + R^{n+1}\vec{q}^{n+1} + \dot{\vec{q}}^{n+1}]\Delta t^{2} / 2 - O(\Delta t^{3}).$$
(18.24)

Подальші дії можуть бути різноманітними, можуть приводити до різних схем з різною точністю інтегрування.

Помножимо (18.23) на $(1-\omega)$, а (18.24) – на ω , результати складемо. Як і в підрозділі 16.4, ваговий коефіцієнт $\omega \in [0,1]$. Отримаємо:

$$\vec{c}^{n+1} = \vec{c}^{n} + (1-\omega)\{(R^{n}\vec{c}^{n} + \vec{q}^{n})\Delta t + [\dot{R}^{n}\vec{c}^{n} + (R^{n})^{2}\vec{c}^{n} + R^{n}\vec{q}^{n} + \dot{\vec{q}}^{n}]\Delta t^{2}/2\} + \omega\{(R^{n+1}\vec{c}^{n+1} + \vec{q}^{n+1})\Delta t - [\dot{R}^{n+1}\vec{c}^{n+1} + (R^{n+1})^{2}\vec{c}^{n+1} + R^{n+1}\vec{q}^{n+1} + \dot{\vec{q}}^{n+1}]\Delta t^{2}/2\} + O(\Delta t^{3}).$$
(18.25)

Відкинемо в (18.25) члени другого порядку наближення і зведемо подібні:

$$(I - \omega \Delta t R^{n+1}) \vec{c}^{n+1} = [I + (1 - \omega) \Delta t R^n] \vec{c}^n + [(1 - \omega) \vec{q}^n + \omega \vec{q}^{n+1})] \Delta t, \qquad (18.26)$$

де I – одиничний оператор. Якщо оператор ($I - \omega \Delta t R^{n+1}$) породжує не вироджену матрицю (зазвичай), то

$$\vec{c}^{n+1} = (I - \omega \Delta t R^{n+1})^{-1} \cdot \{ [I + (1 - \omega) \Delta t R^n] \vec{c}^n + [(1 - \omega) \vec{q}^n + \omega \vec{q}^{n+1})] \Delta t \}.$$
 (18.27)

При $\omega = 0$ з (18.27) отримаємо *явну схему Ейлера* інтегрування з першим порядком наближення відносно Δt :

$$\vec{c}^{n+1} \approx (I + \Delta t R^n) \vec{c}^n + \vec{q}^n \Delta t .$$
(18.28)

Якщо оператор G породжує діагональну матрицю (діагональною вона автоматично утворюється в МСР і методом конденсації в інших методах), то обертати її не потрібно, основна дія, породжена схемою (18.28), це множення матриці на вектор, що здійснюється значно швидше обертання.

Як і в МСР, стійкість явної схеми умовна, причому величина часового кроку, що дозволяється, є дуже малою та повинна відповідати умові $||I + \Delta t R^n|| \le 1$ (див. умову стійкості (16.27) з підрозділу 16.3).

При $\omega = 1$ з (18.27) отримуємо неявну схему Ейлера:

$$\vec{c}^{n+1} \approx (I - \Delta t R^{n+1})^{-1} (\vec{c}^n + \vec{q}^{n+1} \Delta t).$$
 (18.29)

Ще одна визнана схема цієї групи — схема Гальоркіна. Вона утворюється при $\omega = 2/3$.

Всі ці три схеми мають перший порядок наближення відносно Δt , тому не мають асимптотичної стійкості.

Як і в МСР, стійкість неявних схем є абсолютною, тобто досягається при будь-яких значеннях часового кроку лише при $\omega \ge 0.5$.

При $\omega = 0.5$ з (18.27) отримаємо схему інтегрування *Кранка-Ніколсона* з *другим* порядком наближення відносно Δt :

$$\vec{c}^{n+1} = (I - 0.5\Delta t R^{n+1})^{-1} [(I + 0.5\Delta t R^n)\vec{c}^n + 0.5(\vec{q}^n + \vec{q}^{n+1})\Delta t].$$
(18.30)

Існує лема Келога, за якою $||(I-0.5\Delta tR)^{-1}(I+0.5\Delta tR)|| = ||N(\Delta \vec{x}, \Delta t)|| < 1$, якщо R < 0 та $\Delta t > 0$. В нашому випадку $R = -G^{-1}A < 0$, оскільки оператор (матриця) A в задачі теплопровідності є позитивно визначеним. Отже, схема Кранка-Ніколсона має другий порядок наближення, стійка при будь-яких часових кроках. Але відомо, що вона схильна до осциляцій розв'язків (див. п.16.4.1).

18.2.2. Двошарові схеми Розенброка та Ісполова-Шаброва

Розглянемо тепер ще одну групу двошарових схем. Одна з них має виключні показники точності та стійкості при відсутності осциляцій.

Система (18.20) замінюється на еквівалентну

$$\begin{cases} (I - \alpha \Delta t R^{n+\omega})\vec{\eta} = \Delta t (R^{n+\omega} \vec{c}^n + \vec{q}^{n+\omega}); \\ (I - \alpha \Delta t R^{n+\omega})\vec{\chi} = \Delta t [R^{n+\omega} (\vec{c}^n + \beta \vec{\eta}) + \vec{q}^{n+\omega}]; \\ \vec{c}^{n+1} = \vec{c}^n + \vec{\chi} , \end{cases}$$
(18.31)

де постійні α, β, ω необхідно призначити на основі міркувань про точність і стійкість розв'язку; $R^{n+\omega} = R(\vec{x}, t + \omega \Delta t); \ \vec{q}^{n+\omega} = \vec{q}(\vec{x}, t + \omega \Delta t).$

Підставивши перший та третій вирази з (18.31) у другий, розкладаючи в ряд Тейлора та обмежуючись членами $O(\Delta t^2)$, отримаємо:

$$\vec{c}^{n+1} = \vec{c}^n + (R^n \vec{c}^n + \vec{q}^n) \Delta t + + [2\omega \dot{R}^n \vec{c}^n + 2(\alpha + \beta)(R^n)^2 \vec{c}^n + 2(\alpha + \beta)R^n \vec{q}^n + 2\omega \dot{\vec{q}}^n] \Delta t^2 / 2 + O(\Delta t^3).$$
(18.32)

Просте порівняння (18.32) з (18.23), а саме з $\vec{c}^{n+1} = \vec{c}^n + (R^n \vec{c}^n + \vec{q}^n) \Delta t + + [\dot{R}^n \vec{c}^n + (R^n)^2 \vec{c}^n + R^n \vec{q}^n + \dot{\vec{q}}^n] \Delta t^2 / 2 + O(\Delta t^3)$ приводить до висновку, що

$$\alpha + \beta = 0.5; \quad \omega = 0.5.$$
 (18.33)

Але цього недостатньо для визначення α та β . Додаткове рівняння отримаємо із умови *стійкості* схеми, причому в даному випадку без будь-яких обмежень можна розглянути потрібний окремий випадок.

Власні значення μ_j , j = 1,...,N матриці [R] САР $\{\dot{c}\} = [R]\{c\} + \{b\}$, яка породжується рівнянням (18.20) після його просторово-часової алгебраїзації, знаходяться зі системи $[R]\{w\}_m = \mu_m\{w\}_m$, де $\{w\}_m \in$ власні вектори, які відповідають власним значенням μ_m матриці [R]. Відомо, що із застосуванням власних векторів завжди можна записати, що

$$\{c\}^{n+1} = \sum_{m=1}^{N} s_m^{n+1} \{w\}_m, \quad \{c\}^n = \sum_{m=1}^{N} s_m^n \{w\}_m \quad \text{Ta} \quad \{c\}^0 = \sum_{m=1}^{N} s_m^0 \{w\}_m, \quad (18.34)$$

де $N \in кількістю рівнянь у САР, а <math>s_m = s_m(t)$.

Якщо вирази (18.34) підставити у рівняння (18.31), з яких попередньо виключені вектори $\vec{\eta}$ та $\vec{\chi}$ й прийнято $\vec{q}^{n+\omega} = \vec{0}$, то, з урахуванням $[R]\{w\}_m = \mu_m\{w\}_m$ одержимо, що для кожного власного значення μ_m

$$s_m^{n+1} = \frac{1 + (1 - 2\alpha)\Delta t \mu_m + (\alpha^2 - \alpha + \beta)(\Delta t \mu_m)^2}{(1 - \alpha \Delta t \mu_m)^2} \cdot s_m^n; \quad m = 1, ..., N,$$
(18.35)

тобто оператор переходу (див. формулу (16.26))

Частина V. Розділ 18

$$\aleph_{m}(\Delta t) = \frac{1 + (1 - 2\alpha)\Delta t\mu_{m} + (\alpha^{2} - \alpha + \beta)(\Delta t\mu_{m})^{2}}{(1 - \alpha\Delta t\mu_{m})^{2}}; \quad m = 1, ..., N.$$
(18.36)

У задачі теплопровідності всі власні числа μ_m матриці [R] САР $\{\dot{c}\} = [R]\{c\} + \{b\} \in \partial i \ddot{u} c h u m u$ та ві $\partial' \epsilon m h u m u$. Тому при фіксованому *m* оператор переходу (18.36) є звичайним алгебраїчним рівнянням з дійсними числами.

Для асимптотичної стійкості схеми потрібно, щоб виконувалася друга умова (16.29), тут $\lim_{\Delta t \to \infty} \aleph(\Delta t) = 0$. Очевидно, що для виконання цієї умови достатньо покласти, щоб у чисельнику (18.36)

$$\alpha^2 - \alpha + \beta = 0. \tag{18.37}$$

Отримали систему з двох рівнянь: перше з (18.33) та (18.37). Вона має два варіанти розв'язку:

$$\alpha_1 = 0.5(2 - \sqrt{2}); \quad \beta_1 = 0.5(\sqrt{2} - 1);$$
 (18.38)

$$\alpha_2 = 0.5(2 + \sqrt{2}); \quad \beta_2 = -0.5(\sqrt{2} + 1).$$
 (18.39)

Підставивши перші значення α, β й ω у систему (18.31), отримаємо:

$$\begin{cases} (I - 0.5(2 - \sqrt{2})\Delta t R^{n+1/2})\vec{\eta} = \Delta t (R^{n+1/2}\vec{c}^n + \vec{q}^{n+1/2}); \\ (I - 0.5(2 - \sqrt{2})\Delta t R^{n+1/2})\vec{\chi} = \Delta t \{R^{n+1/2}[\vec{c}^n + 0.5(\sqrt{2} - 1)\vec{\eta}] + \vec{q}^{n+1/2}\}; \\ \vec{c}^{n+1} = \vec{c}^n + \vec{\chi}. \end{cases}$$
(18.40)

Це є схема, яка входить до сімейства схем Розенброка.

Підставивши другі значення α, β й ω у систему (18.31), отримаємо:

$$\begin{cases} (I - 0.5(2 + \sqrt{2})\Delta t R^{n+1/2})\vec{\eta} = \Delta t (R^{n+1/2}\vec{c}^n + \vec{q}^{n+1/2}); \\ (I - 0.5(2 + \sqrt{2})\Delta t R^{n+1/2})\vec{\chi} = \Delta t \{R^{n+1/2}[\vec{c}^n - 0.5(\sqrt{2} + 1)\vec{\eta}] + \vec{q}^{n+1/2}\}; \\ \vec{c}^{n+1} = \vec{c}^n + \vec{\chi}. \end{cases}$$
(18.41)

Це є схема, яку опублікували Ю.Г. Ісполов і М.М. Шабров у 1989 році.

Крім того, щоб виконувалася умова (16.28) відсутності осциляцій розв'язків, тут $\aleph(\Delta t) \ge 0$, із (18.36) з урахуванням (18.37) отримаємо, що для кожного $\mu_m < 0$ повинна виконуватися умова

$$\frac{1 + (1 - 2\alpha)\Delta t \mu_m}{(1 - \alpha \Delta t \mu_m)^2} \ge 0; \quad m = 1, ..., N.$$
(18.42)

Оскільки знаменник $(1 + \alpha \Delta t | \mu_m |)^2$ завжди більше нуля, то саме із чисельника $1 - (1 - 2\alpha) \Delta t | \mu_m | \ge 0$ і утворюється обмеження часового кроку для *відсутності осциляцій* розв'язків. Для схеми Розенброка

$$\Delta t \le 1 / \left[(1 - 2\alpha_1) \mid \mu_m \mid \right] = (1 + \sqrt{2}) / \mid \mu_m \mid, \qquad (18.43)$$

або єдина для всіх власних значень матриці $\Delta t \leq (1 + \sqrt{2})/|\mu_{\max}|$. Це жорстке обмеження. Для схеми Ісполова-Шаброва обмежень на Δt немає.

Для ілюстрування наведемо графіки функції переходу $N(\mu\Delta t)$, де $\mu > 0$ є модулем одного з власних значень μ_m матриці [R] для однієї з власних форм, тобто позначимо $\mu = |\mu_m| = -\mu_m$. Точний вираз дається формулою

 $N(\mu\Delta t) = \exp(-\mu\Delta t)$, а для схем, що розглядаються, – формулою $N(\mu\Delta t) = [1-(1-2\alpha)\Delta t\mu]/(1+\alpha\Delta t\mu)^2$, яка відповідає формулі (18.42).

На рис.18.1 наведені графіки оператора $N(\mu\Delta t)$: точний та у випадках схем з α із (18.38), що відповідає схемі Розенброка, та (18.39) для схеми Ісполова-Шаброва. Значення параметра $\mu\Delta t$, при якому функція переходу $N(\mu\Delta t)$ менше нуля, є неприпустимими з точки зору осциляцій цієї власної форми, тому значення $N(\mu_m\Delta t) = 0$ є границею осциляцій, а умова $N(\mu_m\Delta t) > 0$ є умовою відсутності осциляцій розв'язків для цих схем, яка відповідає загальній умові (16.28).

Рис.18.1 показує, що схема Розенброка при перевищенні кроком деякого значення, яке обчислюється за формулою (18.43), буде мати осциляції відповідних складових. А схема Ісполова-Шаброва не має ніяких обмежень шоло часового кроку, асимптотичне та абсолютно стійка та позбавлена осциляцій, порядок має другий наближення часовим за Крім аргументом. того.



Рис.18.1 Графіки функцій оператора переходу для точного розв'язку, схем Розенброка та Ісполова-Шаброва

автори в чисельному експерименті показали її високу стійкість при різких змінах теплового навантаження. Якщо оператор R та часовий крок Δt є незмінними, то із застосуванням методу квадратних коренів (див. п.5.3.5) або схем Холецького (див. п.5.3.6 та п.5.3.7) для розв'язування СЛАР схема (18.41) є ефективною. Якщо оператор (матриця) R залежить від часу (зазвичай – через вектор $\vec{u}(\vec{x},t)$), то в ітераціях для обчислення компонент поточного $R^{n+1/2}$

додатково обчислюється вектор

$$\vec{u}^{n+1/2} = \vec{\psi}(\vec{x}, t + \Delta t/2) + \sum_{i=1}^{J} c_i(t + \Delta t/2) \vec{\varphi}_i(\vec{x}),$$

причому можна наблизити значення $c_i(t + \Delta t/2) \approx (c^n + c^{n+1})/2$, де c^{n+1} відповідає поточній ітерації.

18.2.3. Двошарові "економічні" факторизовані схеми

Бажання якнайшвидше знаходити обернену матрицю в схемах (18.29), (18.30), (18.41) та інших привело до поняття "економічних" схем з факторизованими операторами. Про це вже йшлося у п.16.4.3, але для методу скінченних різниць, де є своя специфіка. Тому знов розглянемо це питання.

Нагадаємо, що будь-яку двошарову схему завжди можна представити в канонічному вигляді Частина V. Розділ 18

$$B\frac{\vec{u}^{n+1} - \vec{u}^n}{\Delta t} = R\vec{u}^n + Q.$$
 (18.44)

(18.45)

Помножимо на Δt і приведемо подібні. Для \vec{u}^{n+1} одержали систему $B\vec{u}^{n+1} = \vec{F}^n$,

де відомо праву частину

$$\vec{F}^n = (B + \Delta t R)\vec{u}^n + \Delta t Q. \qquad (18.46)$$

Відомо, що для отримання з не виродженої квадратної матриці її оберненої матриці потрібна кількість арифметичних операцій, пропорційна N^3 , де N – кількість рядків/стовпців у матриці.

Припустимо, що оператор *В* можна записати у *факторизованому* вигляді (факторизація – це розкладення на множники)

$$B = B_1 \cdot \dots \cdot B_s, \qquad (18.47)$$

де B_m – "економічні" оператори з кількістю арифметичних дій для знаходження обернених B_m^{-1} , пропорційною кількості N, на худий кінець – пропорційною N^2 . Тоді й САР (18.45) стане економічною, оскільки її можна розкласти на S систем з "економічними" операторами B_m :

$$B_1 \vec{y}_1 = F^n; \quad B_m \vec{y}_m = \vec{y}_{m-1}; \quad m = 2, \dots, s; \quad \vec{u}^{n+1} = \vec{y}_s, \tag{18.48}$$

де \vec{y}_m – проміжні вектори-розв'язки.

У схемі (18.26), а саме $(I - \omega \Delta t R^{n+1}) \vec{c}^{n+1} = [I + (1 - \omega) \Delta t R^n] \vec{c}^n + [(1 - \omega) \vec{q}^n + \omega \vec{q}^{n+1})] \Delta t$ маємо такий вираз для оператора *B*:

$$B = I - \omega \Delta t R^{n+1}. \tag{18.49}$$

Виникає питання: чи можна знайти розклад цього оператора при $\omega > 0$ на "економічні" оператори B_m ? Відповідь така: точний швидкий розклад — не можливий (не відомий), а наближений — можливий, і наближених варіантів може бути безліч.

Популярним є такий варіант отримання "економічних" операторів.

Для оператора $B = I - \omega \Delta t R$ з (18.49) оператор R визначають як суму двох операторів (це завжди можна зробити):

$$R = R_1 + R_2. (18.50)$$

Далі розглядають добуток двох операторів:

$$(I - \omega \Delta t R_1)(I - \omega \Delta t R_2) = I - \omega \Delta t R_1 - \omega \Delta t R_2 + \omega^2 \Delta t^2 R_1 R_2.$$
(18.51)

Якщо підкреслений у (18.51) доданок $\omega^2 \Delta t^2 R_1 R_2$ є малим порівняно з іншими членами цього виразу, то цім доданком можна знехтувати та отримати таке наближення оператора *B*:

$$B = I - \omega \Delta t R \approx (I - \omega \Delta t R_1) (I - \omega \Delta t R_2).$$
(18.52)

Отже, маємо два факторизованих оператора: $(I - \omega \Delta t R_1)$ та $(I - \omega \Delta t R_2)$. Такий спосіб їх отримання називають розщепленням.

Така похибка наближення нерівномірно розподіляється по ступеням свободи СЛАР, порушуючи однорідність розв'язку там, де цього не повинне бути. Це явище можна згладити, зменшуючи Δt . Зазначимо, що в крайовій задаче, що розглядається, оператори *A*, *G*, *R*, *I*, *B* породжують квадратні матриці [*A*], [*G*], [*R*], [*I*] та [*B*] відповідно.

Про обирання часового кроку, який забезпечить достатню точність наближення факторизованого оператора, достатньо повно викладено у підрозділі 16.6.

Залишилося питання про те, як зробити виявлені факторизовані матриці (оператори) ($[I] - \omega \Delta t[R_1]$) та ($[I] - \omega \Delta t[R_2]$) "економічними".

Відомо, що будь-яку позитивно визначену квадратну матрицю можна привести до діагональної. Але це є задачею про отримання всіх власних значень матриці, яка має кількість арифметичних операцій, пропорційну N^3 . Тобто цей варіант відпадає. Ще відомо, що обернення трикутної матриці потребує кількість арифметичних операцій, пропорційну N^2 . Саме такий варіант і використовують.

Для прикладу розглянемо найпростішу схему з факторизованими операторами, яку вже розглядали в методі скінченних різниць (див. формули (16.101) з п.16.4.3.2). яка має назву схеми *стабілізації*. Важливо, що *ця схема справедлива* при незалежності оператора R від часу.

Отже, систему (18.26) замінимо на близьку їй

$$(I - \omega \Delta t R_2) \vec{z} = \Delta t [R \vec{c}^n + (1 - \omega) \vec{q}^n + \omega \vec{q}^{n+1}];$$

$$(I - \omega \Delta t R_1) \vec{\eta} = \vec{z}; \qquad \vec{c}^{n+1} = \vec{c}^n + \vec{\eta},$$
(18.53)

де \vec{z} і $\vec{\eta}$ – проміжні вектори, а сума операторів $R_1 + R_2 = R$, причому обов'язкова вимога: $R_1 \le 0$ і $R_2 \le 0$ (тобто вихідні оператори $A_1 \ge 0$ і $A_2 \ge 0$, оскільки G > 0 завжди, див. формули (18.8) і (18.21)).

Матриці системи (18.53) можна прийняти *трикутними*. Тоді система розв'язується швидко, оскільки кількість дій пропорційна N^2 . Цей метод спочатку було розроблено для розв'язування еволюційних задач методом скінченних різниць, де матриця [G] завжди діагональна, і обчислення оберненої матриці [G]⁻¹ для (18.21) елементарне. Але розроблено методи діагоналізації матриці [G] і в інших методах, що дозволяє використовувати цю схему.

Оскільки матриця [R] – симетрична, то зручно прийняти, що

$$(R_{ij})_{1} = \begin{cases} R_{ij}; & i < j; \\ 0.5 \cdot R_{ij}; & i = j; \\ 0 & ; & i > j; \end{cases} \quad (R_{ij})_{2} = (R_{ij})_{1}^{T}.$$
(18.54)

Очевидно, що матриці з такими компонентами є трикутними.

18.2.4. Схема покомпонентного розщеплення

Для випадку залежності оператора R від часу, тобто при R = R(t) або $R = R(\vec{c}(t))$, розроблена схема *покомпонентного розщеплення*, при цьому доведено всі необхідні теореми.

У цій схемі вихідне рівняння (18.20) замінюється на систему (R^n – оператор *R* на *n*-му часовому шарі):

$$\begin{cases} (I - \omega \Delta t R_1^n) \vec{\eta} = (I + \omega \Delta t R_1^n) \vec{c}^{n-1}; \\ (I - \omega \Delta t R_2^n) (\vec{c}^n - \Delta t \vec{q}^n) = (I + \omega \Delta t R_2^n) \vec{\eta}; \\ (I - \omega \Delta t R_2^n) \vec{\chi} = (I + \omega \Delta t R_2^n) (\vec{c}^n + \Delta t \vec{\tilde{q}}^n); \\ (I - \omega \Delta t R_1^n) \vec{c}^{n+1} = (I + \omega \Delta t R_1^n) \vec{\chi}; \end{cases}$$
(18.55)

або (інша форма запису):

$$\begin{cases} (I - \omega \Delta t R_1^n) \vec{\eta} = (I + \omega \Delta t R_1^n) \vec{c}^{n-1}; \\ (I - \omega \Delta t R_2^n) \vec{\xi} = (I + \omega \Delta t R_2^n) \vec{\eta}; \\ \vec{\mu} = \vec{\xi} + 2 \Delta t \vec{q}^n; \\ (I - \omega \Delta t R 2_2^n) \vec{\chi} = (I + \omega \Delta t R_2^n) \vec{\mu}; \\ (I - \omega \Delta t R_1^n) \vec{c}^{n+1} = (I + \omega \Delta t R_1^n) \vec{\chi}. \end{cases}$$
(18.56)

Значення $\vec{c}^n = \vec{c}(t_n)$, необхідне для обчислення R^n , маємо з розв'язку на попередньому часовому шарі (тоді це було \vec{c}^{n+1}).

Необхідні властивості цієї схеми доведено лише при $\omega = 0.5$.

Використання допоміжних векторів, які введено до системи, не означає введення додаткових часових шарів, хоча в літературі часто такі допоміжні вектори позначають як \vec{c} з дробовим номером часового шару.

У методах розщеплення крім схем стабілізації та покомпонентного розщеплення ще розроблено схему *предиктор-коректор*. Усі вони багато у чому схожі, мають однаковий (другий) порядок наближення за часовим аргументом. Тому немає особливих причин віддавати перевагу будь-якій з цих схем (за точністю), але схема стабілізації простіша в реалізації.

18.2.5. Тришарові схеми

Всі наведені вище схеми (18.26) ... (18.31) є *двошаровими*, оскільки шуканий вектор \vec{c} розглядається на двох часових шарах. Якщо ж (18.20) помножити на dt і провести інтегрування на часовому кроці $2\Delta t = t_{n+1} - t_{n-1}$, то отримаємо:

$$\int_{\vec{c}^{n+1}}^{\vec{c}^{n+1}} d\vec{c} = \int_{t_{n-1}}^{t_{n+1}} (R\vec{c} + \vec{q}) dt \,.$$
(18.57)

Припустимо, що оператор *R* – незмінний у часі. Підінтегральну функцію в правій частині (18.57) часто представляють у вигляді:

$$R\vec{c} + \vec{q} = \left[\omega\left(R\vec{c}^{n+1} + \vec{q}^{n+1}\right) + (1 - 2\omega)\left(R\vec{c}^{n} + \vec{q}^{n}\right) + \omega\left(R\vec{c}^{n-1} + \vec{q}^{n-1}\right)\right] + O(\Delta t), \quad (18.58)$$

тоді при $\omega > 0.25$ всі схеми будуть А-стійкими. Після інтегрування і зведення подібних членів отримаємо вагову тришарову схему з наближенням $O(\Delta t^2)$ у часі (схему Dupont):

$$(I - 2\omega\Delta tR)\vec{c}^{n+1} =$$

= $2\Delta t(1 - 2\omega)R\vec{c}^n + (I + 2\omega\Delta tR)\vec{c}^{n-1} + 2\Delta t[\omega\vec{q}^{n-1} + (1 - 2\omega)\vec{q}^n + \omega\vec{q}^{n+1}].$ (18.59)
Якщо прийняти $\omega = 0$, то $R\vec{c} + \vec{q} \approx R\vec{c}^n + \vec{q}^n + O(\Delta t)$, схема стає абсолютно нестійкою: це історично перша тришарова схема – *схема Річардсона*.

Розглянемо питання про можливість збереження другого порядку наближення за часом на першому часовому кроці при застосуванні тришарової схеми. Проблема вирішується досить просто.

Обчислимо \vec{c}^1 як результат розкладу $\vec{c}(t)$ в ряд Тейлора в околі t = 0 на кроці Δt :

$$\vec{c}^{1} = \vec{c}^{0} + \Delta t \dot{\vec{c}}^{0} + O(\Delta t^{2}).$$
(18.60)

Як початкова умова в задачах параболічного типу задається $\vec{c}^0 = \vec{c}(0) = \hat{\vec{c}}^0$. Враховуючи, що відповідно до основного рівняння (18.20)

$$\dot{\vec{c}} = d\vec{c} / dt = R\vec{c} + \vec{q}$$
, (18.61)

з (18.60) отримаємо

$$\vec{c}^{1} = \vec{c}^{0} + \Delta t (R\vec{c}^{0} + \vec{q}^{0}) + O(\Delta t^{2}) \approx (I + \Delta t R)\vec{c}^{0} + \Delta t \vec{q}^{0}.$$
(18.62)

Таким чином, перш ніж застосовувати тришарову схему, необхідно спочатку застосувати вираз (18.62). Тоді другий порядок наближення за часом на першому часовому кроці буде забезпечено.

Одна з негативних властивостей тришарових схем: часовий крок Δt повинен бути незмінним.

18.3. Крайові задачі гіперболічного типу

Розглянемо рівняння (18.12), приведемо його до вигляду

$$\frac{d^2\vec{c}}{dt^2} + G\frac{d\vec{c}}{dt} = R\vec{c} + \vec{q} , \qquad (18.63)$$

де позначено $G = M^{-1}C$, $R\vec{c} = -M^{-1}L(\vec{c})$; $\vec{q} = M^{-1}\vec{b}$. Початкові умови (18.13) запишемо у вигляді $\vec{c}(0) = \hat{\vec{c}}^0$; $d\vec{c}(0)/dt = \hat{\vec{h}}$.

Застосуємо сітковий шаблон для часового диференційного оператора $d^2 \vec{c} / dt^2$ у вигляді (16.18), тобто

$$\Im_t \vec{c}_t = \frac{\vec{c}_t^+ - \vec{c}_t^-}{\Delta t} = \frac{\vec{c}^{n+1} - 2\vec{c}^n + \vec{c}^{n-1}}{\Delta t^2} = \vec{c}_{tt}^{+-}$$
(18.64)

з похибкою наближення $O(\Delta t^2)$. Для часового диференційного оператора $d\vec{c}/dt$ доцільно також використовувати наближення з похибкою наближення $O(\Delta t^2)$, тобто у вигляді (16.13):

$$\mathfrak{F}_{t}\vec{c}_{t} = \frac{\vec{c}^{n+1} - \vec{c}^{n-1}}{2\Delta t} = \vec{c}_{t}^{0}.$$
(18.65)

Застосуємо (18.64) та (18.65) до (18.63):

$$\vec{c}_{tt}^{+-} + G\vec{c}_{t}^{0} = R\vec{c} + \vec{q} .$$
(18.66)

Праву частину виразу (18.66) $A\vec{c} + \vec{q}$ можна записати у вигляді трьох розкладів у ряди Тейлора:

$$(R\vec{c}+\vec{q})_{-}^{n}=R^{n-1}\vec{c}^{n-1}+\vec{q}^{n-1}+(\dot{R}^{n-1}\vec{c}^{n-1}+R^{n-1}\dot{\vec{c}}^{n-1}+\dot{\vec{q}}^{n-1})\Delta t+O(\Delta t^{2}); \qquad (18.67)$$

$$(R\vec{c} + \vec{q})^n = R^n \vec{c}^n + \vec{q}^n;$$
(18.68)

Частина V. Розділ 18

$$(R\vec{c}+\vec{q})_{+}^{n} = R^{n+1}\vec{c}^{n+1} + \vec{q}^{n+1} - (\dot{R}^{n+1}\vec{c}^{n+1} + R^{n+1}\dot{\vec{c}}^{n+1} + \dot{\vec{q}}^{n+1})\Delta t + O(\Delta t^{2}); \qquad (18.69)$$

або як симетричну їхню комбінацію

$$R\vec{c} + \vec{q} \approx \omega (R\vec{c} + \vec{q})_{+}^{n} + (1 - 2\omega)(R\vec{c} + \vec{q})^{n} + \omega (R\vec{c} + \vec{q})_{-}^{n}, \qquad (18.70)$$

де $0 \le \omega \le 1$ – ваговий коефіцієнт.

Як і для задач параболічного типу, можна скомбінувати з виразів (18.67) ... (18.70) велику кількість схем: явних і неявних, умовно і безумовно стійких.

Зокрема доведено, що безумовно стійкою є схема

$$\vec{c}_{tt}^{+-} + G\vec{c}_{t}^{0} = R \frac{\vec{c}^{n+1} + \vec{c}^{n-1}}{2} + \vec{q}^{n}, \qquad (18.71)$$

в якій оператор *R* не залежить від часу. Вона має похибку наближення $O(\Delta t^2)$ за часовим аргументом.

Якщо оператор *G* породжує *діагональну* матрицю [\overline{G}], або таку, яку легко привести до діагональної матриці [\overline{G}], часто застосовують схему, в якій у (18.70) призначають $\omega = 0$, тобто наближують $R\vec{c} + \vec{q} \approx (R\vec{c} + \vec{q})^n$. Тоді (18.66) можна записати у вигляді явної (рекурентної) схеми:

$$\vec{c}^{n+1} = \left(\frac{1}{\Delta t^2} + \vec{G}\frac{1}{2\Delta t}\right)^{-1} \left[\left(\frac{2}{\Delta t^2} + R\right)\vec{c}^n - \left(\frac{1}{\Delta t^2} - \vec{G}\frac{1}{2\Delta t}\right)\vec{c}^{n-1} + q^n \right], \quad (18.72)$$

яку називають *схемою центральних різниць* (див. підрозділ 16.5). Ця схема має умовну стійкість, тобто часовий крок є обмеженим зверху: $\Delta t \leq T_{min} / \pi$ (див. формулу (16.111)), де T_{min} – найменший з періодів власних коливань тіла. Це дуже жорстке обмеження, але велика кількість кроків компенсується відсутністю необхідності у збиранні та розв'язуванні СЛАР з недіагональною матрицею.

У випадку відсутності дисипативного члена, як і для задач параболічного типу, розроблені "економічні" схеми з факторизованим оператором.

Перетворимо вираз (18.66) на

$$B\vec{c}_{tt}^{+-} = R\vec{c}^{n} + \vec{q}^{n} \quad \text{afo} \quad B\frac{\vec{c}^{n+1} - 2\vec{c}^{n} + \vec{c}^{n-1}}{\Delta t^{2}} = R\vec{c}^{n} + \vec{q}^{n}.$$
(18.73)

Нехай значення \vec{c}^{n-1} і \vec{c}^n відомі. Для \vec{c}^{n+1} одержуємо вираз $B\vec{c}^{n+1} = \vec{F}^n$, (18.74)

де відома права частина

$$\vec{F}^{n} = (2B + \Delta t^{2}R)\vec{c}^{n} - B\vec{c}^{n-1} + \Delta t^{2}\vec{q}^{n}.$$
(18.75)

Припустимо, що оператор *В* можна записати у *факторизованому* вигляді $B = B_1 \cdot ... \cdot B_s$, (18.76)

де B_k – "економічні" оператори. Тоді й система (18.74) – "економічна", оскільки її можна розкласти на *s* систем з операторами B_k :

$$B_{1}\vec{y}_{(1)} = \vec{F}^{n}; \quad B_{k}\vec{y}_{(k)} = \vec{y}_{(k-1)}; \quad k = 2,...,s; \quad \vec{c}^{n+1} = \vec{y}_{(S)}, \quad (18.77)$$

де $\vec{y}_{(k)}$ – проміжні розв'язки.

Оператори В_к зручно прийняти у вигляді:

Частина V. Роз	діл 18 26 ⁻

$$B_k = (I - 0.5\Delta t^2 R_k), \tag{18.78}$$

причому необхідно, щоб виконувалися умови

$$R = \sum_{k=1}^{S} R_k ; \quad -R_k \ge 0.$$
 (18.79)

Наприклад, при s = 2 маємо $B = (I - 0.5\Delta t^2 R_1)(I - 0.5\Delta t^2 R_2) = (I - 0.5\Delta t^2 R) + O(\Delta t^2)$. Підставивши це в (18.73), отримаємо "економічну" схему

$$\begin{cases} (I - 0.5\Delta t^2 R_1) \vec{y} = \vec{F}^n; \\ (I - 0.5\Delta t^2 R_2) \vec{c}^{n+1} = \vec{y}, \end{cases}$$
(18.80)

де вектор $\vec{F}^n = (I - 0.5\Delta t^2 R)(2\vec{c}^n - \vec{c}^{n-1}) + (R\vec{c}^n + \vec{q}^n)\Delta t^2$ є відомим. Вона з точністю $O(\Delta t^2)$ наближує безумовно стійку схему (18.71) при $G \equiv 0$, тобто при відсутності дисипації.

Явна схема центральних різниць (18.72) при $\overline{G} = 0$ спрощується до

$$\vec{c}^{n+1} = \Delta t^2 \left[\left(\frac{2}{\Delta t^2} + R \right) \vec{c}^n + q^n \right] - \vec{c}^{n-1}, \qquad (18.81)$$

причому обмеження часового кроку $\Delta t \leq T_{min} / \pi$ зберігається.

Відзначимо також, що розроблено ще один підхід до створення "економічних" схем розв'язування задач гіперболічного типу: зведення задачі до системи рівнянь параболічного типу.

Для початку розрахунків за тришаровими схемами необхідно мати значення \vec{c} як на нульовому, так і на першому часовому шарі. Щоб зберегти другий порядок наближення за часом на першому часовому кроці, обчислюють \vec{c}^1 як результат розкладу в ряд Тейлора в околі t = 0 на часовому кроці Δt :

$$\vec{c}^{1} = \vec{c}^{0} + \Delta t \dot{\vec{c}}^{0} + \Delta t^{2} \ddot{\vec{c}}^{0} / 2 + O(\Delta t^{3}).$$
(18.82)

Як початкові умови в задачах гіперболічного типу задаються $\vec{c}^0 = \vec{c}(0) = \hat{\vec{c}}^0$ та $\dot{\vec{c}}^0 = d\vec{c}(0)/dt = \hat{\vec{h}}$. Враховуючи, що відповідно до основного рівняння (18.63) $\ddot{\vec{c}} = d^2\vec{c}/dt^2 = -G\dot{\vec{c}} + R\vec{c} + \vec{q}$, (18.83)

з (18.82) остаточно маємо наближення:

$$\vec{c}^{1} = \hat{\vec{c}}^{0} + \Delta t \hat{\vec{h}} + \Delta t^{2} [-G\hat{\vec{h}} + R^{0}\hat{\vec{c}}^{0} + \vec{q}^{0}]/2 + O(\Delta t^{3}) \approx (I + 0.5\Delta t^{2}R^{0})\hat{\vec{c}}^{0} + \Delta t (I - 0.5\Delta tG)\hat{\vec{h}} + 0.5\Delta t^{2}\vec{q}^{0}.$$
(18.84)

Отже, перш ніж застосовувати тришарову схему, необхідно спочатку застосувати формулу (18.84). Тоді другий порядок наближення крайової задачі за часом буде забезпечено.

Контрольні питання до Розділу 18.1

1. Як проводиться часткова алгебраїзація крайової задачі параболічного типу неперервними базисними функціями в методі зважених похибок наближення?

2. Як проводиться часткова алгебраїзація крайової задачі гіперболічного типу неперервними базисними функціями в методі зважених похибок наближення?

Контрольні питання до Розділу 18.2

- 1. Як отримують класичні двошарові схеми для крайової задачі параболічного типу?
- 2. У чому особливості двошарових схем Розенброка та Ісполова-Шаброва?
- 3. Як будуються двошарові економічні факторизовані схеми?
- 4. У чому особливості побудови тришарових вагових схем?

Контрольні питання до Розділу 18.3

- 1. Як отримують класичні тришарові схеми для крайової задачі гіперболічного типу?
- 2. Опишіть, як отримують схему центральних різниць.

Частина VI МЕТОД СКІНЧЕННИХ ЕЛЕМЕНТІВ

Розділ 19

ІДЕЯ МЕТОДУ СКІНЧЕННИХ ЕЛЕМЕНТІВ

У 1943 році Р. Курантом (R. Courant) було опубліковано статтю з назвою "Variational methods for the solution of problems of equilibrium and vibrations" ("Варіаційні методи для розв'язування проблем рівноваги і коливань"), в якій він, зокрема, розв'язував крайову задачу про кручення стрижня варіаційним методом Релея-Рітца, причому перетин стрижня представив у вигляді сукупності трикутників з кусково-лінійним наближенням функції кручення у кожному з трикутників. Саме тоді Р. Курант уперше виписав всі основні ідеї методу, який лише у 1960 році Р. Клафом (R.W. Clough) був названий методом скінченних елементів (МСЕ).

Ідеї представлення складних конструкцій у вигляді сукупності простіших об'єктів виникали ще з початку XX сторіччя. Наприклад, при розрахунках кораблів та літаків такими об'єктами були стрингери, стійки, панелі, частини обшивки, листи корпусу корабля тощо. А у 1941 році А. Хренніковим (А. Hrennikoff) було навити сформульовано "метод каркасів", за яким двовимірна пружна крайова задача розв'язувалась за допомогою набору брусів і балок.

Приблизно у 1920 р. данський інженер А. Ostenfeld (1866-1931 р.р.) остаточно сформулював та дав назву "метод переміщень" методу розв'язування крайової задачі про напружено-деформований стан стрижневої системи, в якому розглядається не рівновага вузлів, як це робиться у методі сил, а сумісність деформацій окремих стрижнів у цих вузлах. Для ферм його дещо раніше розвивав Є.О. Патон (1870-1953 р.р.). Як результат утворювалася система лінійних алгебраїчних рівнянь (СЛАР) відносно переміщень та кутів поворотів кінців стрижнів (у методі сил, як відомо, невідомими СЛАР є вузлові сили). Але такий метод виявився не дуже зручним у застосуванні (ЕОМ ще не було), тому не набув значного поширення у практиці розрахунків. Найчастіше застосовується у будівельній механіці.

Є принаймні *три* трактування МСЕ: *інженерне* (метод переміщень), *варіаційне* (запропоновано Р. Курантом), *апроксимаційне*. Яке з них більш відповідає сутності методу? Інженерне – безумовно вузьке. Не всі крайові задачі мають свій варіаційний варіант (функціонал), а МСЕ посильні всі крайові задачі. І лише *апроксимаційне трактування є універсальним*, оскільки МСЕ дозволяє використовувати для отримання САР або варіаційний функціонал (застосовано ще Р. Курантом), або інший метод розв'язування крайових задач, зокрема, методи Бубнова-Гальоркіна, зважених похибок наближення, принцип можливих переміщень, тощо. Тому доцільно *MCE* розглядати як *метод апроксимації простору, часу та розв'язку, що шукається*. МСЕ вважається проєкційно-сітковим методом.

Ще не маючи сучасної назви, МСЕ спочатку був застосований для розв'язування одновимірних та двовимірних крайових задач, згодом, з появою потужних ЕОМ, – тривимірних.

19.1. Ідея методу скінченних елементів

Розглядаючи в попередніх Розділах методи розв'язування крайових задач, які спиралися на наближення типу (17.4), ми обов'язково припускали, що кожна з базисних функцій визначена *єдиним виразом у всій області* Ω.

У MCE:

• вся область Ω ділиться на підобласті Ω^e (скінченні елементи, СЕ), що *не перекривають одна одну та щільно заповнюють область* Ω . Позначивши як *Е* загальну кількість СЕ, а як E^* – кількість СЕ, що примикають до *S* та мають поверхні $S^e \subset S$, можемо записати, що

$$\Omega \approx \Omega_{h} = \bigcup_{e=1}^{E} \Omega^{e}; \quad \prod_{e=1}^{E} \Omega^{e} = 0; \qquad S \approx S_{h} = \bigcup_{e=1}^{E^{*}} S^{e}; \quad \prod_{e=1}^{E^{*}} S^{e} = 0, \quad \text{de } S^{e} \subset S.$$
(19.1)

Якщо поверхня *S* області Ω апроксимована точно, то у (19.1) рівності стають точними;

• визначаються базисні функції в кожному СЕ окремо, причому базисні функції СЕ за його межами теж визначені, але *тотожно дорівнюють нулю*, тобто базисні функції визначаються у кусковий спосіб;

• для визначення базисних функцій кожний СЕ має декілька вузлів;

• глобальні нумерації СЕ (e = 1,...,E) та вузлів (n = 1,...,N) є довільними. В кожному СЕ для визначеності вузли мають локальну нумерацію ($m = 1,...,M_e$). Відповідність локальних вузлових номерів глобальним номерам цих вузлів встановлюється оператором відповідності (інцидентності) (Λ_m^n)^{*e*}:

 $(\Lambda_n^m)^e = \begin{cases} 1; \; якщо вузлу \; з \; глобальним номером \; n \; відповідає \; вузел \; з \; локальним номером \; m \; y \; CE \; з \; номером \; e; \; (19.2) \; 0; \; інакше. \end{cases}$

Цей оператор (масив) фактично описує операцію *розбивки* області Ω на СЕ (підобласті Ω^e). Можна ввести й обернений за смислом масив інцидентності $(\Xi_m^n)^e$, який буде описувати операцію *зв'язування* скінченно-елементної моделі.

19.1.1. Скінченне-елементне наближення базисними функціями

Розглянемо найпростіші локально визначені базисні функції. Для цього використаємо задачу наближення одновимірної функції $u_0(x)$, тобто будемо будувати функцію $u(x) \approx u_0(x)$, коли $u_0(x)$ є відомою.

На рис.19.1-а функція $u_0(x)$ наближується базисними функціями, які визначаються у *центрах* СЕ і можуть бути лише *постійними*. Наближення є розривним на границях СЕ. На границях діапазону визначення функція $u_0(x)$ теж наближується з похибками. Тому потрібно шукати інші варіанти наближення.





На рис.19.1-б та же функція наближується у N вузлах, розташованих по *межах* СЕ, тобто кусково-лінійними базисними функціями. Бачимо значне покращення наближення функції $u_0(x)$. Рис.19.1-б підказує, що для ще більшої точності наближення функції в кожному СЕ можна ввести *проміжний (внутрі-иній)* вузол СЕ, після цього через три вузла провести криву, наприклад, степеневу другого порядку. Таких проміжних вузлів може бути й більше.

Доведено (є така теорема), що при $E \to \infty$ (тут E – кількість CE) $u(x) \to u_0(x)$, тобто наближення дійсно існує, тому відповідні їм базисні функції створюють *повну* систему.

Розглянемо формалізацію скінченно-елементного наближення.

У кожному СЕ вводяться локальні апроксимаційні функції $\tilde{u}^{e}(x)$ (або $\tilde{u}^{e}(r)$, якщо введені локальні координати r, але останнє – не обов'язково).

Між локальними та глобальними координатами є взаємно-однозначні зв'язки, наприклад, у другому випадку (див. рис.19.1-б) $(r)^e = f(x) = (x - x_m)^e$, де x_m визначає початок СЕ, а m – номер вузла. З використанням масивів (операторів) інцидентності $(\Lambda_n^m)^e$ та $(\Xi_m^n)^e$ ці зв'язки у вузлах можна формалізувати як

$$(r_m)^e = f((x_m)^e); \quad x_n = (\Lambda_n^m)^e (x_m)^e; \quad (x_m)^e = (\Xi_m^n)^e x_n; \quad n = 1, ..., N; \quad m = 1, ..., M_e, \quad (19.3)$$

де: M_e – кількість вузлів у СЕ з номером e; N – загальна кількість вузлів у Ω . Пояснення: перебираються всі вузли СЕ, для кожного з них (як би) збирається сума добутків елементів масиву $(\Lambda_n^m)^e$ з координатами $(x_m)^e$, але реально у сумі будуть всі нулі, окрім *одного* члена суми.

З використанням введених у СЕ функцій наближення $\tilde{u}^e(x)$ встановлюється відповідність вузлових значень $\tilde{u}^e_m = \tilde{u}^e(x_m)$ вузловим значенням $u_n = u(x_n)$. Це робиться аналогічними взаємно-однозначними зв'язками (відношеннями інцидентності), які прямо випливають з (19.3):

$$\tilde{u}_m^e = \tilde{u}^e(x_m) = (\Xi_m^n)^e u(x_n) = (\Xi_m^n)^e u_n; \quad n = 1, ..., N; \ m = 1, ..., M_e.$$
(19.4)

Оскільки функції $\tilde{u}^{e}(x)$ визначені і за границями СЕ, де тотожно дорівнюють нулю, а СЕ не перетинаються та щільно заповнюють простір (тут – вздовж осі x), можна записати апроксимацію функції $u_0(x)$ в Ω виразом

$$u_0(x) \approx u(x) = \sum_{e=1}^{E} \tilde{u}^e(x)$$
. (19.5)

Щоб абстрагуватися від конкретних функцій, що апроксимуються, вводяться *нормовані* локальні базисні функції, а саме $\varphi_m^e(x)$ за таким правилом:

$$\varphi_{m}^{e}(x_{k}) = \delta_{mk}; \quad m, k = 1, ..., M_{e},$$
(19.6)

де δ_{mk} – компоненти одиничної матриці; *m* та *k* – номери вузлів, що належать СЕ. Тобто в МСЕ у "своєму" *m*-му вузлі базисна функція $\varphi_m^e(x_m) \equiv 1$, а у всіх інших ("чужих") дорівнює нулю. Саме такі базисні функції зображені на рис.19.2.

Для (19.5) у межах СЕ функції наближення $\tilde{u}^{e}(x)$ описуються формулою апроксимації на базисі, тобто збирається сума результатів перемноження базисних функцій $\varphi_{m}^{e}(x)$ на деякі вузлові значення c_{m} , які шукаються:

$$\tilde{u}^{e}(x) = \sum_{m=1}^{M_{e}} (\varphi_{m}^{e}(x) \cdot c_{m}).$$
(19.7-a)



Рис.19.2 Наближення функції $u_0(x)$ нормованими базисними функціями СЕ: а) – кусково-постійними; б) – кусково-лінійними

Примітка 19.1. Якщо бути точними, то маємо визначити, що за допомогою $\varphi_m^e(x)$ функції наближення $\tilde{u}^e(x)$ повинні описуватися як

$$\tilde{u}^{e}(x) = \sum_{m=1}^{M_{e}} (s_{m} \cdot \varphi_{m}^{e}(x) \cdot c_{m}), \qquad (19.7-6)$$

де c_m є невідомими, а множник s_m визначається кратністю k_m вузла т СЕ:

$$s_m = \begin{cases} 1; \ k_m = 0; \\ 1/k_m; \ k_m > 0. \end{cases}$$
(19.8)

Пояснення: якщо у вузлі *m* стикуються *k* CE, то такий вузол (на границі CE) має кратність k_m (всередині CE вважають k = 0).

В більшості публікацій у формулах типу (19.7) множник s_m опускають. Це можна робити з математичних причин, які тут немає сенсу обговорювати. Тому надалі множник s_m теж будемо опускати як такий, що не впливає на остаточний результат.

Якщо підставити (19.7-а) у (19.5), то наближення функції $u_0(x)$ в Ω (всі суми запишемо в явному вигляді):

$$u_{0}(x) \approx u_{h}(x) = \sum_{e=1}^{E} \left(\sum_{m=1}^{M_{e}} \left(\varphi_{m}^{e}(x) \cdot c_{m} \right) \right),$$
(19.9)

тобто містить суму обмеженої кількості базисних функцій, помножених на невідомі коефіцієнти, тим самим відповідає формулі (17.4-б) методу Фур'є з одним уточненням: (19.9) не має окремої функції для описування граничних умов, тобто прийнято, що $\vec{\psi} \equiv 0$. Оскільки вибір базисних функцій $\varphi_m^e(x)$ ніяк не пов'язаний з об'єктом наближення, то наближення МСЕ відноситься до варіанта в) вибору базисних функцій (див. п.17.1.1).

З огляду на властивість нормованих базисних функцій (19.6) та на вираз (19.7-а), приходимо до важливого висновку, що значення

$$c_m = \tilde{u}^e(x_m) = \tilde{u}^e_m, \qquad (19.10)$$

тобто невідомими коефіцієнтами *c_m* є *вузлові значення функції*, що наближується. Тому вирази (19.7-а) і (19.9) замінимо на

$$\tilde{u}^{e}(x) = \sum_{m=1}^{M_{e}} \left(\varphi_{m}^{e}(x) \cdot \tilde{u}_{m}^{e} \right); \quad u_{0}(x) \approx u_{h}(x) = \sum_{e=1}^{E} \left(\sum_{m=1}^{M_{e}} \left(\varphi_{m}^{e}(x) \cdot \tilde{u}_{m}^{e} \right) \right), \quad (19.11)$$

які визначають наближення функції $u_0(x)$ у будь-якій точці СЕ та тіла.

Щодо наближення в СЕ, *тільки вузлові значення функції* $\tilde{u}^e(x_m) = \tilde{u}_m^e$ *та безперервні значення базисних функцій* $\varphi_m^e(x)$, $m = 1, ..., M_e$ актуального *СЕ* з номером *е* визначають наближені значення функції $u_0(x)$ у цьому СЕ.

Примітка 19.2. Є й дещо інший варіант формалізації наближення у МСЕ. Вводиться функція приналежності до СЕ з номером *e*:

$$\chi^{e}(x) = \begin{cases} 1; & x \in \Omega^{e}; \\ 0; & x \notin \Omega^{e}. \end{cases}$$
(19.12)

Позначимо:

$$\tilde{\varphi}_m(x) = \sum_{\Omega^e \subset \Upsilon_m} \left(\chi^e(x) \cdot \varphi^e_m(x) \right), \tag{19.13}$$

де Υ_m – множина СЕ, які містять вузол номер *m*. Тоді, з урахуванням (19.2), замість (19.9) маємо такий запис наближення:

$$u_{0}(x) \approx u(x) = \sum_{n=1}^{N} (\Xi_{n}^{m})^{e} \left(c_{m} \cdot \sum_{\Omega^{e} \subset \Upsilon_{m}} \left(\chi^{e}(x) \cdot \varphi_{m}^{e}(x) \right) \right); \quad m = 1, ..., M_{e},$$
(19.14)

де *N* – загальна кількість вузлів у Ω.

Надалі будемо застосовувати варіант (19.9) як зручніший.

Кусково-постійні і кусково-лінійні базисні функції, що природно, можна визначити для одновимірного, двовимірного й тривимірного простору, тобто вектор координат \vec{x} може мати необхідну розмірність. Для двовимірного випадку застосовують СЕ у формі трикутника або чотирикутника, також і криволінійні. В тривимірному випадку *поверхні* СЕ мають форму і трикутника, і чотирикутника (тетра-, пента-, гексагональні або інші СЕ). Докладніше – в Розділі 20.

19.1.2. Про побудову системи алгебраїчних рівнянь для розв'язування крайових задач за методом скінченних елементів

Зрозуміло, в який спосіб МСЕ застосовується в крайових задачах.

Розглянемо стаціонарну крайову задачу, яка в загальному вигляді представлена як $L(\vec{u}) = \vec{f}$ в Ω , з граничними умовами $K(\vec{u})|_{s} = \vec{g}$.

Тоді для формул (17.14-б) методу Бубнова-Галеркіна з п.17.1.4 або (17.17) методу зважених похибок наближення з п.17.1.5 вектори

$$\vec{R}_{\Omega} = L(\vec{u}) - \vec{f} \ y \ \Omega; \quad \vec{R}_s = K(\vec{u}) - \vec{g} \ \text{Ha} \ S.$$
 (19.15)

Тут $\vec{u} = (\vec{u}(\vec{x}))_h$ – шуканий вектор-розв'язок, який наближується подібно до (19.9); L і К – оператори, які для вказаних методів можуть бути як лінійними так і нелінійними. Для лінійних крайових задач скінченно-елементну модель задачі можна отримувати за допомогою методу Релея-Рітца (див. п.17.1.3).

Багато крайових задач припускають варіаційну постановку. Зокрема, це стосується задач механіки твердого тіла, що деформується. І в цьому випадку МСЕ дозволяє одержувати розв'язки, зокрема і в нелінійних задачах.

У випадку використання функціонала (13.31) інтеграли відображають роботи внутрішніх зусиль та зовнішніх сил. Роботу можна додавати в алгебраїчний спосіб лише тоді, коли кожна складова роботи обчислена для окремих об'єктів або на взаємно ортогональних напрямках. Оскільки СЕ не перетинаються, то можемо кожний інтеграл у (13.31) записати у вигляді алгебраїчної суми з окремих інтегралів, які відповідають окремим СЕ:

$$\Psi(\delta) = \sum_{e=1}^{E} \int_{\Omega^{e}} \sigma_{mn} \delta \varepsilon_{mn} d\Omega - \sum_{e=1}^{E} \int_{\Omega^{e}} \widehat{O}_{m} \delta U_{m} d\Omega - \sum_{e=1}^{E^{*}} \int_{S_{p}^{e}} \widehat{P}_{m} \delta U_{m} dS = 0.$$
(19.16)

Ще виникає питання про базисні функції: диференційні оператори L і K висувають певні вимоги до базисних функцій, а саме: необхідні похідні принаймні не повинні приймати нульові та нескінченні значення. Це означає, що якщо диференційний оператор має найвищий порядок похідних *p*, то і базисна функція повинна мати обмежені та ненульові похідні того ж порядку. Якщо базисна функція є поліномом, його порядок повинен бути не нижче *p*-го.

Якщо розв'язок крайової задачі залежить від часу, то, згідно з методом Фур'є в наближенні (19.9), як і у (17.4-б), маємо $c_m = c_m(t)$, тому:

$$u_{0}(\vec{x},t) \approx u_{h}(\vec{x},t) = \sum_{e=1}^{E} \left(\sum_{m=1}^{M_{e}} \left(\varphi_{m}^{e}(\vec{x}) \cdot c_{m}(t) \right) \right),$$
(19.17)

або в межах CE з номером e:

$$u_0(\vec{x},t) \approx u_h(\vec{x},t) = \sum_{m=1}^{M_e} \left(\varphi_m^e(\vec{x}) \cdot c_m(t) \right).$$
(19.18)

Вираз (19.10) замінюється на

269

Частина	VI.	Розділ	19
			-

$$c_m(t) = \tilde{u}^e(x_m, t) = \tilde{u}^e_m(t).$$
 (19.19)

Більш докладно МСЕ розглянемо в наступних Розділах, оскільки тут йдеться лише про ідею МСЕ.

19.2. Переваги та недоліки методу скінченних елементів

МСЕ має ряд очевидних переваг:

• приймаючи доволі просту *однотипну* форму підобластей (СЕ), можна і базисні функції вибирати однотипними, тобто створювати бібліотеку СЕ;

• чим відносно дрібніші СЕ, тим більш докладно і точно можна наближати *поверхню* області Ω. Це особливо важливо при розрахунках реальних тіл, які зазвичай мають складну форму.

• наявність СЕ різноманітної геометричної форми та розмірів дозволяє відносно легко згущувати сітку СЕ в місцях очікуваних значних градієнтів розв'язку;

• базисні функції в підобластях (в СЕ) значно легше підібрати такими, що задовольняють локальним умовам крайової задачі, зокрема і граничним;

• за межами своєї підобласті (свого СЕ) базисні функції дорівнюють нулю. Це призводить до розрідженості матриці СЛАР, породженої методом, що прискорює розв'язок СЛАР порівняно зі СЛАР, яка має щільно заповнену матрицю;

• оскільки підобласті (СЕ) не перекривають одна одну, можна інтеграли по всій області обчислювати як *суму* інтегралів по підобластях (по СЕ). При цьому частина інтегралів зникає, оскільки в них підінтегральна функція тотожно дорівнює нулю. Це стосується інтегралів по поверхні області, тобто до опису граничних умов: зазвичай ГУ задаються лише на *частині* поверхні тіла;

• із застосуванням МСЕ можна розв'язувати будь-які нелінійні задачі з будь-якими властивостями середовища;

• МСЕ при однакових розмірах сітки є точнішим, ніж метод скінченних різниць (МСР), оскільки у МСР наближення проводиться за *напрямками*, а у МСЕ – в усьому *об'ємі* Ω;

• алгоритм МСЕ легко програмується, його можна створювати в модульному вигляді, що дозволяє легко розширювати клас розглядуваних задач.

Природно, метод має і певні недоліки:

• значення компонент матриці САР не є заздалегідь відомими, як це відбувається у методі скінченних різниць;

• матриця САР, що породжується МСЕ, більш заповнена, ніж в методі скінченних різниць, що призводить до величезних обсягів даних, вимагає ЕОМ з великими об'ємами пам'яті;

• необхідно створювати (і зберігати) додаткову інформацію про скінченно-елементну сітку;

• кускове завдання базисних функцій не гарантує гладкості наближення по всій області: якщо сама функція (див. формули (19.11)) і може бути набли-

жена безперервно, то похідні будуть (або можуть) мати розриви на границях CE. Необхідно вирішувати питання про *припустимість* таких розривів.

Перших три недоліки тривалий час стримували розвиток МСЕ для тривимірних крайових задач, оскільки потужності ЕОМ були недостатніми.

Контрольні питання до Розділу 19

1. У чому полягає ідея методу скінченних елементів?

- 2. Як зветься масив (оператор) $(\Lambda_n^m)^e$, й які властивості він має?
- 3. Які висновки отримали з наближення функції кусковими базисними функціями?
- 4. Які властивості повинні мати базисні функції скінченного елемента?
- 5. Які переваги та недоліки має метод скінченних елементів?

Розділ 20

СКІНЧЕННІ ЕЛЕМЕНТИ, БАЗИСНІ ФУНКЦІЇ

20.1. Основні поняття, визначення

Скінченно-елементна модель тіла будується з урахуванням таких особливостей і правил:

а) область об'ємом Ω з поверхнею *S* представляється у вигляді сукупності кінцевої кількості підобластей Ω^e – скінченних елементів (CE), які *безперервно* заповнюють всю область, *не перетинаючись*. СЕ стикуються між собою по *міжелементних гранях* (*межах*). Дві грані, які стикуються (двох CE), повинні бути *ідентичними*, причому зв'язок між гранями встановлюється у *вузлах* цих граней. Тобто не дозволяється мати не задіяні у зв'язках ("висячі") вузли на таких гранях. Кожному CE надається оригінальний номер *e* (не обов'язково за порядком). Множину CE з об'ємами Ω^e позначимо як $\Re^e = {\Omega^1, \Omega^2, ..., \Omega^E}$, а поверхонь CE *S*^{*e*}, що виходять на загальну поверхню *S*, як $\mathbb{R}^e = {S^1, S^2, ..., S^{E^*}}$, причому:

$$\bigcup_{e=1}^{E} \Omega^{e} = \Omega^{*} = \Omega_{h}; \quad \prod_{e=1}^{E} \Omega^{e} = 0; \quad \bigcup_{e=1}^{E^{*}} S^{e} = S_{h}; \quad \prod_{e=1}^{E^{*}} S^{e} = 0, \quad (20.1)$$

де E – загальна кількість CE; Ω^* – незв'язана (незібрана) область із "розсипаних" CE; Ω_h – зв'язана (зібрана) область із CE, що наближує Ω . Різниця областей $\Omega \setminus \Omega_h = \Omega_{\varepsilon}$ називається *областю відхилення*. Якщо множина Ω_{ε} порожня, або $S_h = S$, то область Ω є наближеною геометрично точно.

б) кожен СЕ має деяку кількість *вузлів*: *зовнішніх* і *внутрішніх* (останні є не завжди). Перелік вузлів і граней в СЕ здійснюється у певному порядку: застосовується авторський або загально прийнятий *шаблон нумерації вузлів і граней*.

в) вводиться глобальна (загальна) нумерація вузлів для всієї області (але не обов'язково за порядком). Загальну кількість вузлів позначимо як N. Перехід від локальної нумерації вузлів СЕ до глобальної називають зв'язуванням, а у зворотному напрямку – розбивкою. Перелік глобальних номерів вузлів, що входять до кожного СЕ, називається monoлогією СЕ, а відповідну таблицю для всіх СЕ, що складають актуальну область – monoлогічною (аналог оператора інцидентності (19.2)). Укладення топологічної таблиці і є по суті операцією poзбивки. Ще розбивкою називають операцію відображення незв'язаної області СЕ Ω^* на зв'язану область Ω_h . Зазвичай топологічну таблицю доповнюють іншою інформацією про СЕ: тип СЕ, кількість вузлів у СЕ, номер матеріалу в СЕ, кількість точок чисельного інтегрування вздовж локальних координат в СЕ та інше. Якщо у вузлі *m* стикуються *k* CE, то такий граничний вузол має кратність k_m ; $k_m = 0$ для вузла, що не належить міжелементній грані (див. Примітку 19.1).

г) розташування кожного вузла в Ω задається його глобальними координатами, для чого також складається відповідна таблиця глобальних координат вузлів. Зазвичай її доповнюють іншою інформацією, безпосередньо пов'язаною з вузлами, наприклад: про систему координат, в якій задані координати; "прапорцями" дозволу або заборони ступенів свободи вузла (можливості зміщення вузла вздовж кожної з осей глобальної координатної системи або повороту навколо названих осей); про значення початкової температури у вузлах, інше.

Примітка 20.1. Для скорочення об'єму пам'яті при зберіганні таблиць топології і глобальних координат вузлів раніше широко застосовували спеціальні розрахункові алгоритми, які генерували повну інформацію з обмеженого об'єму даних. Використовувалися рівномірність і циклічність змін номерів вузлів і координат, а також однорідність початкових температур, умов закріплення вузлів, тощо.

д) кількість вузлів у СЕ визначається *типом* СЕ і порядком наближення в СЕ. Кожний СЕ тіла може мати свій тип і свою кількість вузлів. Типізація дозволяє мати *каталог* СЕ, які розрізняються за ознаками, наприклад:

• за *розмірністю*: одновимірний (лінійний), двовимірний (плоский, оболонковий), тривимірний (оболонковий, осесиметричний, об'ємний), інші;

• за *геометричним обрисом*: для одновимірних – лінійний (прямо або криволінійний); для двовимірних плоских – трикутний в плані, чотирикутний, секторальний; для осесиметричних – тороїд з перетином трикутного, чотирикутного (або іншого) вигляду; для тривимірних – піраміда або призма з трикутною, чотирикутною (або іншою) основою (тетраедр, пентаедр, гексаедр);

• за порядком наближення: перший порядок наближення, другий, третій, інші, а також змішаний;

• за асимптотикою наближення: звичайна або спеціальна (наприклад, для наближення у вершини тріщин).

е) для кожного СЕ задаються нормовані, однозначні та безперервні функції форми $\psi_m^e(\vec{x})$ та базисні функції $\varphi_m^e(\vec{x})$. Функції форми $\psi_m^e(\vec{x})$ призначені для опису *геометрії* СЕ, базисні функції $\varphi_m^e(\vec{x}) - для$ наближення функцій та (можливо) їхніх похідних у скінченному елементі. Ці функції є інтерполяційними (тобто їхні вузлові значення є регламентованими). Кожна з цих функцій відповідає одному з вузлів, причому не обов'язково, щоб всі вузли мали "свої" базисні функції та функції та функції та функції та функції адного СЕ з прямими ребрами достатньо *трьох* вузлів, але він може мати базисні функції другого порядку наближення, які будуть "спиратися" на *шість* вузлів даного СЕ. Або навпаки, двовимірний криволінійний трикутний СЕ має *шість* вузлів для описування геометрії, а будуть застосовуватися базисні функції першого порядку наближення, для чого достатньо використати *три* кутових вузла даного СЕ. За межами СЕ базисні функції кожного

СЕ тотожно дорівнюють *нулю*. В конкретному вузлі СЕ його "рідна" функція дорівнює *одиниці*, а інших вузлів СЕ – *нулю*. Ще одна обов'язкова вимога: сума базисних функцій в будь-якій точці СЕ повинна дорівнювати одиниці;

ж) для описування функцій форми й базисних функцій в СЕ може вводитися локальна координатна система \vec{r} , тоді $\psi_m^e(\vec{r})$ та $\varphi_m^e(\vec{r})$.

Деякі визначення, що застосовуються для оцінювання якості скінченноелементного наближення:

• діаметром δ^e CE з номером *е* називають максимально можливу в CE відстань між двома точками *і та j*, що належать CE, тобто

$$\delta^{e} = \max \left\| \vec{x}_{i} - \vec{x}_{j} \right\|, \quad \text{de } \vec{x}_{i}, \quad \vec{x}_{j} \in \Omega^{e}; \quad (20.2)$$

• мірою дрібноти δ скінченно-елементної моделі всього тіла називають максимальний діаметр її CE, тобто

$$\delta = \max\left\{\delta^1, \delta^2, \dots, \delta^E\right\}; \tag{20.3}$$

• подрібнення (зменшення розмірів CE) називають *регулярним*, якщо кожний вузол і кожна міжелементна межа у вихідній моделі є вузлом і міжелементною межею і після подрібнення скінченно-елементної сітки (інакше – нерегулярним). Тобто при регулярному подрібненні дрібняться вже існуючі CE. Подрібнення називають таким, що *рівномірно тоншає*, якщо зменшуються як максимальний, так і мінімальний діаметри всій сукупності CE.

20.2. Симплексні моделі СЕ в евклідовому просторі

20.2.1. Поняття симплекса

Симплексом у k-вимірному евклідовому просторі називають опуклу множину Θ , яка визначається сукупністю $M_e = k + 1$ вершин (вузлів) з координатами $(x^i)_m$, де i = 1,...,k є координатним напрямом, а $m = 1,...,M_e$ – номером вузла, причому вузли не належать до однієї (k-1)-вимірної гіперплощині. Якщо k = 1, то це є лінією з 2-ма вузлами на кінцях (див. рис.20.1-а,г), при k = 2маємо трикутник з 3-ма вузлами у кутах (див. рис.20.1-б,д), при k = 3 – тетраедр з 4-ма вузлами у кутах (див. рис.20.1-в,е). На погляд із декартової системи, якщо k -вимірний простір має лінійний базис, то й симплекс є лінійним, а якщо нелінійний (наприклад, циліндрична система координат), то й симплекс є нелінійним.



Рис.20.1 Симплексні моделі СЕ: а), б), в) – у лінійному базисі; г), д), е) – у нелінійному базисі

Координати будь-якої точки симплекса (множини $\Theta \equiv \Omega^e$, CE) можуть бути обчислені за формулою

$$x^{i} = c_{1} \cdot (x^{i})_{1} + \dots + c_{M_{e}} \cdot (x^{i})_{M_{e}} = \sum_{m=1}^{M_{e}} c_{m} \cdot (x^{i})_{m}$$
, $\exists e \quad c_{m} \ge 0$ $\exists x \quad \sum_{m=1}^{M_{e}} c_{m} = 1$. (20.4)

Позначимо $x^1 = x$, $x^2 = y$, а $x^3 = z$. Наближення функції $u = u(\vec{x}) = u(x^i)$ при $x^i \in \Theta$ відбувається за формулою симплекса, тобто *лінійного* наближення

$$u(\vec{x}) = a_0 + \sum_{i=1}^k a_i x^i = a_0 + a_1 x + a_2 y + a_3 z; \quad x^i \in \Theta; \quad k = 3.$$
 (20.5)

Для визначення $k+1 = M_e$ коефіцієнтів a_0, a_i використовується M_e відомих вузлових значень $u(\vec{x}_m) = u_m$, $m = 1, ..., M_e$. Тоді M_e рівнянь (20.5) створюють СЛАР з позитивно визначеною матрицею. У тривимірному випадку це СЛАР

$$\begin{bmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ 1 & x_3 & y_3 & z_3 \\ 1 & x_4 & y_4 & z_4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}.$$
 (20.6)

Після визначення коефіцієнтів a_0, a_i ; i = 1, ..., k можна виразити локальне лінійне наближення (20.5) у вигляді

$$u(\vec{x}) = \sum_{m=1}^{M_e} \varphi_m^e(\vec{x}) \cdot u_m; \quad \vec{x} \in \Theta, \qquad (20.7)$$

де лінійні нормовані інтерполяційні функції $\varphi_m^e(\vec{x})$ іноді називають *барицентричними* координатами, причому після групування їх можна записати як

$$\varphi_m^e(\vec{x}) = (b_0)_m + \sum_{i=1}^k (b_i)_m x^i = (b_0)_m + (b_1)_m x + (b_2)_m y + (b_3)_m z \,; \quad x^i \in \Theta \,, \qquad (20.8)$$

вони мають властивості повноти і безперервності, а також:

$$\varphi_m^e(\vec{x}_n) = \delta_{mn}; \quad \varphi_m^e(\vec{x}) \ge 0; \quad n, m = 1, ..., M_e; \quad \sum_{m=1}^{M_e} \varphi_m^e(\vec{x}) = 1; \quad \vec{x} \in \Theta.$$
(20.9)

Перший вираз (20.9) вказує, що базисні функції є нормованими до одиниці. Він вже фігурував у Розділі 19 (див. формулу (19.6)) як обов'язкова властивість базисних функцій СЕ у МСЕ.

Примітка 20.2. Оскільки за допомогою формули (20.7) можна наблизити будь-яку функцію $u(\vec{x})$, то такими функціями можуть бути функції форми (описують форму CE) та базисні функції (описують наближений розв'язок задачі).

20.2.2. Одновимірна симплексна модель СЕ

При *k* = 1 СЕ має 2 вузла на кінцях відрізка лінії (див. рис.20.1-а,г).

Розглянемо одновимірний простір з координатою $x^1 = x$. Тоді для формули (20.8)

 $(b_0)_1 = x_2/s$; $(b_0)_2 = -x_1/s$; $(b_1)_1 = -1/s$; $(b_1)_2 = 1/s$; $s = x_2 - x_1$ (20.10) і лінійні інтерполяційні функції $\varphi_n(x)$ отримують вирази

$$\varphi_1^e(x) = (x_2 - x) / s; \quad \varphi_2^e(x) = -(x_1 - x) / s.$$
 (20.11)

Формули (20.11) відповідають інтерполяційній формулі Лагранжа (9.18), тому їхні похибки наближення відповідають оцінці (9.22), а саме

$$|R_1(x)| \le \frac{K_2}{2} \cdot |\Pi_2(x)|,$$
 (20.12)

де $K_2 = \max |u''(\xi)|$ при $\xi \in [x_1, x_2]$, а $\Pi_2(x) = (x - x_1)(x - x_2)$.

Для одновимірної симплексної моделі СЕ у випадках дво- та тривимірного простору, коли СЕ не лежить вздовж однієї з осей x^i , вводять локальну систему координат (див. п. 24.6.6).

20.2.3. Двовимірна симплексна модель СЕ

Має 3 вузла у кутах трикутника (див. рис.20.1-б,д). Розглянемо двовимірний простір з координатами $x^1 = x$ та $x^2 = y$. СЛАР (20.6) набуває вигляду

$$\begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}.$$
 (20.12)

Розв'язавши систему рівнянь (20.12), отримаємо замість (20.8):

$$\varphi_m^e(\vec{x}) = \varphi_m^e(x, y) = (\alpha_m + \beta_m x + \gamma_m y) / 2\Delta^e; \quad m = 1, 2, 3, \quad (20.13)$$

де

$$\alpha_1 = x_2 y_3 - x_3 y_2; \quad \alpha_2 = x_3 y_1 - x_1 y_3; \quad \alpha_3 = x_1 y_2 - x_2 y_1; \quad (20.14)$$

$$\beta_1 = y_2 - y_3; \quad \beta_2 = y_3 - y_1; \quad \beta_3 = y_1 - y_2;$$
 (20.15)

$$\gamma_1 = x_2 - x_3; \quad \gamma_2 = x_3 - x_1; \quad \gamma_3 = x_1 - x_2;$$
 (20.16)

$$2\Delta^{e} = \det \begin{vmatrix} 1 & x_{1} & y_{1} \\ 1 & x_{2} & y_{2} \\ 1 & x_{3} & y_{3} \end{vmatrix}.$$
 (20.17)

Інтерполяційні функції (20.13) відповідають властивостям повноти і безперервності, а також умовам (20.9). Їх ще називають трикутними координатами, барицентричними координатами, *L*-координатами, природними координатами, а також координатами площини.

Геометрична інтерпретація Δ^e – площа трикутного СЕ. Оскільки площа не може бути від'ємною, то потрібно мати такий порядок нумерації вузлів: проти годинникової стрілки в правій системі координат. На рис.20.2 зображено геометричні інтерпретації *L*-координат. Рис.20.2-а показує ізолінії *L*-координат (ізолінії паралельні ребру, що лежить проти вузла, для якого визначається значення інтерполяційної функції); а рис.20.2-б – площі, відношення яких дає значення *L*-координати (заштриховано площу (Δ_o)₁):

$$L_{m}(\vec{x}_{Q}) = \varphi_{m}^{e}(\vec{x}_{Q}) = (\Delta_{Q})_{m} / \Delta^{e}.$$
(20.18)

Як виявилося, похибки симплексних наближень при *k* ≥ 2 залежать від взаємних положень вузлів симплекса.



Рис.20.2 Геометричні інтерпретації L-координат двовимірного симплексного СЕ

При k = 2 (трикутний CE) Сінг (J.L. Synge, 1957 р.) отримав оцінку похибки для першої похідної

$$\left| \partial u(\vec{x}) / \partial x^{i} - \partial u_{0}(\vec{x}) / \partial x^{i} \right| \leq \left| \frac{K_{2} \delta_{\Theta}}{2 \sin \theta} \right|; \quad i = 1, 2,$$
 (20.19)

де K_2 є максимальним значенням для всіх $|\partial^2 u_0 / \partial x^i \partial x^j|$, i, j = 1, 2; δ_{Θ} – діаметр CE, а \mathscr{G} є найбільшим з кутів CE. При цьому вважалося, що наближені значення $u(\vec{x})$ у вузлах CE точно дорівнюють значенням функції $u_0(\vec{x})$, яку наближують. Формула показує, що найменша похибка буде у випадку застосування рівностороннього трикутника. Вона швидко збільшується при значному витягуванні CE в будь-якому напрямку (голкоподібні CE), а також при наявності в CE тупого кута, наближеного до 180 градусів.

20.2.4. Тривимірна симплексна модель СЕ

Має 4 вузла у кутах тетраедра (див. рис.20.1-в,е). Позначимо $x^1 = x$, $x^2 = y$ та $x^3 = z$. Вигляд СЛАР (20.6) не змінюється, а лінійні нормовані інтерполяційні функції $\varphi_m(\vec{x})$, введені у формулі (20.8), записуються як

$$\varphi_m^e(\vec{x}) = \varphi_m^e(x, y, z) = (\alpha_m + \beta_m x + \gamma_m y + \lambda_m z) / 6\Delta^e; \quad m = 1, 2, 3, 4, \quad (20.20)$$

де

$$6\Delta^{e} = \det \begin{vmatrix} 1 & x_{1} & y_{1} & z_{1} \\ 1 & x_{2} & y_{2} & z_{2} \\ 1 & x_{3} & y_{3} & z_{3} \\ 1 & x_{4} & y_{4} & z_{4} \end{vmatrix};$$
(20.21)

$$\alpha_{1} = \det \begin{vmatrix} x_{2} & y_{2} & z_{2} \\ x_{3} & y_{3} & z_{3} \\ x_{4} & y_{4} & z_{4} \end{vmatrix}; \quad \beta_{1} = -\det \begin{vmatrix} 1 & y_{2} & z_{2} \\ 1 & y_{3} & z_{3} \\ 1 & y_{4} & z_{4} \end{vmatrix}; \quad (20.22)$$

$$\gamma_{1} = \det \begin{vmatrix} x_{2} & 1 & z_{2} \\ x_{3} & 1 & z_{3} \\ x_{4} & 1 & z_{4} \end{vmatrix}; \quad \lambda_{1} = -\det \begin{vmatrix} x_{2} & y_{2} & 1 \\ x_{3} & y_{3} & 1 \\ x_{4} & y_{4} & 1 \end{vmatrix},$$
(20.23)

а вирази для коефіцієнтів $\alpha_2, \alpha_3, \alpha_4, \beta_2, \beta_3, \beta_4, \gamma_2, \gamma_3, \gamma_4, \lambda_2, \lambda_3, \lambda_4$ утворюються з відповідних формул (20.21) і (20.22) циклічною перестановкою індексів 1,2,3,4.

Інтерполяційні функції (20.20) мають властивості повноти і безперервності, а також властивості, наведені у (20.9). Їх ще називають об'ємними трикутними координатами або *L*-координатами.

Геометрична інтерпретація Δ^e — об'єм тетраедра. Оскільки об'єм не може бути від'ємним, то потрібно мати такий порядок нумерації вузлів: проти годинникової стрілки в правій системі координат.

При k = 3 (СЕ у вигляді тетраедра) Кі (S.W. Key, 1966 р.) отримав оцінку похибки симплексного наближення

$$\left| \partial u(\vec{x}) / \partial x^{i} - \partial u_{0}(\vec{x}) / \partial x^{i} \right| \leq \left| \frac{K_{2} \delta_{\Theta}}{18 \cos(\theta / 2) \cdot \sin \theta} \right|; \quad i = 1, 2, 3,$$

$$(20.24)$$

де кут θ є найбільшим з кутів між ребрами будь-якої трикутної грані тетраедра, а кут θ є найбільшим з кутів між гранню з кутом θ та будь-якою іншою гранню тетраедра. Вважалося, що наближені значення $u(\vec{x})$ у вузлах симплекса точно дорівнюють значенням функції $u_0(\vec{x})$, яку наближують. Формула показує, що найменша похибка буде у випадку застосування рівностороннього тетраедра. Вона швидко збільшується при значному витягуванні СЕ в будь-якому напрямку (голкоподібні СЕ), а також при наявності в СЕ дуже короткого ребра або тупого кута, наближеного до 180 градусів (СЕ наближений до плоского).

20.3. Комплексні та мультиплексні моделі СЕ в евклідовому просторі



Рис.20.3. Комплексні криволінійні моделі

Комплексними моделями називають СЕ, в яких кількість вузлів перевищує значення k+1 (k є розмірністю евклідового простору). Додаткові (порівняно з симплексними моделями) вузли в цих моделях розміщуються на ребрах СЕ та призначаються для забезпечення умови безперервності між сусідніми СЕ у випадках,

коли ребра СЕ не є лініями координат. У зв'язку з цим ребра СЕ можуть бути як лінійними, так й нелінійними.

Форми комплексних СЕ можна застосовувати таки ж самі, як й у симплексних. А саме: при k = 1 (одновимірний СЕ) – лінія з $M_e > k + 1 = 2$ вузлами (2 – граничних, інші – внутрішні, див. рис.20.3-а); при k = 2 (двовимірний СЕ) – трикутник з $M_e > 3$ вузлами (3 – кутові, інші – на ребрах, див. рис.20.3-б); при k = 3 – тетраедр з $M_e > 4$ вузлами (4 – кутові, інші – на ребрах, див. рис.20.3-в).

Зазвичай вузли на ребрах розміщують з рівномірним кроком, але це не обов'язково. Головне, щоб матриця СЛАР, яка збирається для визначення коєфіцієнтів, була позитивно визначеною.

У випадку застосування квадратичного наближення маємо, що

$$u(\vec{x}) = a_0 + \sum_{i=1}^k a_i x^i + \sum_{i=1}^k \sum_{j=1}^k a_{ij} x^i x^j , \qquad (20.25)$$

де $a_{ij} = a_{ji}$, i, j = 1,...,k. Загальна кількість коефіцієнтів дорівнює кількості вузлів CE, а саме $M_e = (k+1)(k+2)/2$. Для їхнього знаходження як і для симплексних моделей записується система алгебраїчних рівнянь, яка визначає відомі значення $u(\vec{x}_m) = u_m$, $m = 1,...,M_e$. Отримана СЛАР буде мати позитивно визначену матрицю, якщо CE не буде виродженим.

Якщо для визначення коефіцієнтів (20.25) використати *вузлові* значення $u(\vec{x})$, а саме $u(\vec{x}_m) = u_m$; $m = 1, ..., M_e$, то отримане СЛАР дозволяє виразити локальне наближення (20.25) у вигляді (20.7). Відповідні інтерполяційні функції $\varphi_m^e(\vec{x})$ можна записати у вигляді

$$\varphi_m^e(\vec{x}) = (b_0)_m + \sum_{i=1}^k (b_i)_m x^i + \sum_{i=1}^k \sum_{j=1}^k (b_{ij})_m x^i x^j , \qquad (20.26)$$

де $(b_{ij})_m = (b_{ji})_m$, i, j = 1, ..., k.

М. Зламал (М. Zlamal, 1968 р.) довів, що при k = 2 та $M_e = 6$, тобто для криволінійного трикутника з квадратичним наближенням (20.25), коли проміжні вузли розташовані посередині ребер, оцінка похибки має вигляд

$$|\partial u(\vec{x}) / \partial x^{i} - \partial u_{0}(\vec{x}) / \partial x^{i}| \leq \frac{2K_{2}\delta_{\Theta}^{2}}{\sin \theta}; \quad i = 1, 2, \qquad (20.27)$$

де $K_2 \epsilon$ граничним значенням для всіх $|\partial^2 u(\vec{x})/\partial x^i \partial x^j|$, i = 1,2; δ_{Θ} – діаметр СЕ, а $\vartheta \epsilon$ найбільшим з кутів СЕ. При цьому вважалося, що наближені значення $\partial u(\vec{x})/\partial x^i$ у вузлах СЕ точно дорівнюють значенням відповідних похідних $\partial u_0(\vec{x})/\partial x^i$ функції, яку наближують.

Мультиплексними моделями називають СЕ, в яких кількість вузлів перевищує значення k+1($k \in$ розмірністю евклідового простору), додаткові (порівняно з симплексними моделями) вузли поміщуються на ребрах



Рис.20.4 Мультиплексні криволінійні моделі

СЕ, а для забезпечення умови безперервності між сусідніми СЕ ребра СЕ є лініями координат. У зв'язку з цим ребра СЕ можуть бути як лінійними, так й нелінійними.

Для двовимірного мультиплексного CE (4 вузла в кутах, рис.20.4-а) застосовується *білінійне* наближення

$$u(\vec{x}) = a_0 + \sum_{i=1}^{2} a_i x^i + a_{12} x^1 x^2.$$
(20.28)

Для тривимірного мультиплексного CE (8 вузлів у кутах, рис.20.4-б) застосовується *трилінійне* наближення

$$u(\vec{x}) = a_0 + \sum_{i=1}^3 a_i x^i + \sum_{i=1}^3 \sum_{j=1}^3 a_{ij} x^i x^j + a_{123} x^1 x^2 x^3.$$
(20.29)

Для тривимірного мультиплексного CE у вигляді призми з трикутною основою (6 вузлів у кутах, рис.20.4-в) застосовується наближення

$$u(\vec{x}) = a_0 + \sum_{i=1}^3 a_i x^i + x^3 \cdot (c \cdot x^1 + d \cdot x^2).$$
 (20.30)

Два варіанта CE у вигляді криволінійного паралелепіпеда з проміжними вузлами на ребрах CE:

• квадратичне наближення (20 вузлів, з них 8 – кутових, по 1 проміжному на ребрах):

$$u(\vec{x}) = a_0 + \sum_{i=1}^3 a_i x^i + \sum_{i=1}^3 \sum_{j=1}^3 a_{ij} x^i x^j + (x^1)^2 (b_2 x^2 + b_3 x^3) + (x^2)^2 (c_1 x^1 + c_3 x^3) + (x^3)^2 (d_1 x^1 + d_2 x^2) + x^1 x^2 x^3 \left(e_0 + \sum_{i=1}^3 e_i x^i \right);$$
(20.31)

• кубічне наближення (32 вузла, з них 8 – кутових, по 2 проміжних на ребрах), в якому до виразу (20.31) додається член:

$$f_{1} \cdot (x^{1})^{3} + f_{2} \cdot (x^{2})^{3} + f_{3} \cdot (x^{3})^{3} + x^{1}x^{2}x^{3} \sum_{i=1}^{3} \sum_{i=1}^{3} g_{ij}x^{i}x^{j} + (x^{1})^{3}(h_{2}x^{2} + h_{3}x^{3}) + (x^{2})^{3}(p_{3}x^{3} + p_{1}x^{1}) + (x^{3})^{3}(q_{1}x^{1} + q_{2}x^{2}), \qquad (20.32)$$

де $a_i, b_i, c_i, d_i, e_i, f_i, g_{ij}, h_i, p_i, q_i$ є коефіцієнтами, а $g_{ij} = 0$ при $i \neq j$.

20.4. Параметричні моделі СЕ в евклідовому просторі

Всі розглянуті вище моделі СЕ потребують обчислювання коефіцієнтів наближення для кожного із СЕ, з яких зібрана скінченно-елементна модель тіла. У випадку симплексних СЕ таких дій відносно небагато, а для інших варіантів ситуація ускладнюється. Є й така проблема: реальні конструкції зазвичай мають складні обриси, зокрема й нелінійні, для наближення яких скінченними елементами бажано мати СЕ, геометрія яких мінімально залежить від координатної системи та може бути без проблем пристосовуватися до реальних потреб. Тому було розроблено інший підхід: застосування для інтерполяції у СЕ локальної системи координат.

20.4.1. Параметричні інтерполяційні функції одновимірних СЕ

Для побудови полінома степеня p потрібно p+1 вузлів СЕ. Наприклад, при p = 2 поліном має вигляд $\varphi(x) = a_0 + a_1 x + a_2 x^2$. Для визначення його коефіцієнтів слід записати систему рівнянь для трьох вузлів. Наприклад, для інтерполяційної функції першого вузла (нижній індекс при координаті x вказує на номер вузла):

Таких систем необхідно розв'язати стільки, скільки вузлів у СЕ. Однак розв'язувати їх кожний раз не має потреби, оскільки відомо, що розв'язком таких систем є фундаментальний многочлен Лагранжа (9.18) необхідного степеня p, а саме:

$$\varphi_m^{(p)} = \Lambda_m^p(x) = \frac{(x - x_1)(x - x_2) \cdot \dots \cdot (x - x_{m-1})(x - x_{m+1}) \cdot \dots \cdot (x - x_{p+1})}{(x_m - x_1)(x_m - x_2) \cdot \dots \cdot (x_m - x_{m-1})(x_m - x_{m+1}) \cdot \dots \cdot (x_m - x_{p+1})}.$$
 (20.34)

Про многочлен Лагранжа відомо, що його параметр можна замінити на інший за умовою лінійної залежності між параметрами, і при цьому зовнішній вигляд многочлена (20.34) не зміниться, тільки поміняється позначення параметра. Введемо локальні для СЕ координати, в нашому одновимірному випадку – координату ξ . Зазвичай вимагають, щоб в межах СЕ було $\xi \in [-1,1]$, тобто локальну систему координат поміщують у геометричний центр СЕ. Тоді лінійний зв'язок між ξ та x буде мати вигляд

$$\xi = 2(x - x_c^e) / s^e; \quad x = x_c^e + \xi(s^e / 2), \qquad (20.35)$$

де x_c^e – координата центру СЕ; s^e – довжина ребра СЕ, на якому розміщено вузли СЕ, що розглядаються.

Тоді з (20.34) отримаємо інтерполяційні поліноми:

• першого порядку (лінійне) наближення (див. рис.20.5-а) $\varphi_1^{(1)}(\xi) = -(\xi - 1)/2; \quad \varphi_2^{(1)}(\xi) = (\xi + 1)/2$ (20.36)

з локальними координатами вузлів $\xi_1 = -1; \xi_2 = 1;$

• другого порядку (квадратичне) наближення (див. рис.20.5-б) $\varphi_1^{(2)}(\xi) = \xi(\xi - 1)/2; \quad \varphi_2^{(2)}(\xi) = \xi(\xi + 1)/2; \quad \varphi_3^{(2)}(\xi) = -(\xi - 1)(\xi + 1) = 1 - \xi^2 (20.37)$ з локальними координатами вузлів $\xi_1 = -1; \quad \xi_2 = 1; \quad \xi_3 = 0;$

• третього порядку (кубічне) наближення (див. рис.20.5-в)

Кількість вузлів і степінь *p* інтерполяційних поліномів можна і далі збільшувати, але на практиці зазвичай обмежуються варіантами (20.36) та (20.37).

Як зазначалося раніше, інтерполяційна функція СЕ повинна дорівнювати одиниці в "своєму" вузлі, в інших вузлах СЕ – нулю, а сума всіх інтерполяційних функцій СЕ в будь-якій точці СЕ повинна дорівнювати одиниці. Неважко перевірити, що вирази (20.36) ... (20.38) відповідають цим умовам.





Ще вузлові інтерполяційні функції СЕ будують з використанням так званого *принципу ієрархії*. Згідно з цим принципом *підвищення* порядку наближення інтерполяційних функцій (а також і СЕ), при якому запроваджуються нові вузли, проводиться за декілька етапів.

Наприклад, при запровадженні проміжного вузла 3 (див. рис.20.6) виконуються такі дії:



Рис.20.6 До застосування принципу ієрархії в одновимірному СЕ

• виписуються інтерполяційні функції для наявних вузлів 1 та 2, тобто у випадку *p* = 1. Це є вирази (20.36);

• вводиться новий вузол 3 та виписується його інтерполяційна функція, це є третій вираз (20.37), тобто $\varphi_3^{(2)}(\xi) = (1 - \xi^2);$

• проводиться корекція інтерполяційної функції для кутового вузла 1 так, щоб її значення у вузлі 3 дорівнювало нулю: $\varphi_1^{(2)}(\xi) = \varphi_1^{(1)}(\xi) - \varphi_3^{(2)}(\xi)/2$. Конкретно це буде такий вираз: $\varphi_1^{(2)}(\xi) = -(\xi-1)/2 - (1-\xi^2)/2 = [(1-\xi)-(1-\xi)(1+\xi)]/2 = (1-\xi)(1-1-\xi)/2 = \xi(\xi-1)/2$, тобто перша формула (20.37);

• проводиться корекція інтерполяційної функції для кутового вузла 2 так, щоб її значення у вузлі 3 дорівнювало нулю: $\varphi_2^{(2)}(\xi) = \varphi_2^{(1)}(\xi) - \varphi_3^{(2)}(\xi)/2$.

Конкретно це буде такий вираз: $\phi_2^{(2)}(\xi) = (\xi+1)/2 - (1-\xi^2)/2 = [(1+\xi)-(1-\xi)(1+\xi)]/2 = (1+\xi)(1-1+\xi)/2 = \xi(\xi+1)/2$, тобто друга формула (20.37).

Чому віднімається саме половина від $\varphi_3^{(2)}(\xi)$? Тому, що при $\xi = 0$ (це координата нового вузла 3) обидві функції $\varphi_1^{(1)}(\xi)$ та $\varphi_2^{(1)}(\xi)$ мають значення 1/2, яке потрібно "прибрати", тобто зробити так, щоб було $\varphi_1^{(2)}(0) = \varphi_2^{(2)}(0) = 0$. Оскільки $\varphi_3^{(2)}(0) = 1$, то $\varphi_3^{(2)}(0)/2 = 1/2$. Саме таке значення нам потрібно "прибрати". Оскільки $\varphi_3^{(2)}(-1) = \varphi_3^{(2)}(1) = 0$, то у вузлах 1 та 2 значення функцій $\varphi_1^{(2)}(\xi)$ та $\varphi_1^{(2)}(\xi)$ не зміняться.

Отже, отримано ті ж самі вирази. Так навіщо потрібен цей принцип? Справа у тому, що застосування принципу ієрархії буває зручним при програмуванні, в результаті чого спочатку програмуються (обчислюються) інтерполяційні функції при p=1, а потім, якщо це потрібно (для отримання p=2), вони корегуються згідно з принципом ієрархії після введення додаткового вузла на ребрі СЕ та його інтерполяційної функції. Ще це зручно при створенні базисних функцій у СЕ серендипового сімейства (див. п.20.4.3).

Формули типу (20.7), а саме

$$u(\vec{x}) = \sum_{m=1}^{M_e} \varphi_m^e(\xi) \cdot u_m \; ; \quad \vec{x} \in \Theta \; , \qquad (20.39)$$

є справедливими для будь-якої функції, зокрема й для координат будь-якої точки CE, тому:

$$x^{j} = \sum_{m=1}^{M_{e}^{*}} \varphi_{m}^{e}(\xi) \cdot (x^{j})_{m}; \quad j = 1, 2, 3.$$
(20.40)

Ці формули вказують, що одновимірні СЕ з параметричною інтерполяцією може бути розташованим у просторі будь-яким чином (як і симплексні СЕ), а не тільки вздовж однієї з осей координат (як мультиплексні моделі).

20.4.2. Параметричні інтерполяційні функції для дво- та тривимірних СЕ лагранжевого сімейства

Для двовимірних і тривимірних СЕ можна застосувати многочлени Лагранжа для кожного з координатних напрямків. Якщо використати добуток многочленів Лагранжа у двовимірному та тривимірному випадках:

$$\varphi_m^{(p)}(\xi,\eta) = \Lambda_m^p(\xi) \cdot \Lambda_m^p(\eta); \qquad (20.41)$$

$$\varphi_m^{(p)}(\xi,\eta,\mu) = \Lambda_m^p(\xi) \cdot \Lambda_m^p(\eta) \cdot \Lambda_m^p(\mu), \qquad (20.42)$$

то отримаємо інтерполяційні функції для двовимірних та тривимірних CE так званого *лагранжевого сімейства* (див. рис.20.7 і 20.8). Тут *ξ*, η, μ є локальними координатами, причому *ξ*, η, μ ∈ [-1,1].

Загальна кількість вузлів отримується теж операцією перемноження кількостей вузлів у кожному напрямку. Наприклад, при p = 2 маємо $M_e = 2 * 2 = 4$ вузла у двовимірному CE, а при p = 3 маємо $M_e = 3 * 3 = 9$ вузла у двовимірному CE та $M_e = 3 * 3 * 3 = 27$ у тривимірному.



Рис.20.7 Двовимірні СЕ лагранжевого сімейства першого (p = 1; $M_e = 2 \cdot 2 = 4$), другого (p = 2; $M_e = 3 \cdot 3 = 9$) та третього (p = 3; $M_e = 4 \cdot 4 = 16$) порядку наближення



Рис.20.8 Тривимірні гексагональні СЕ лагранжевого сімейства першого ($p = 1; M_e = 2 \cdot 2 \cdot 2 = 8$) та другого ($p = 2; M_e = 3 \cdot 3 \cdot 3 = 27$) порядку наближення

20.4.3. Параметричні інтерполяційні функції для дво- та тривимірних СЕ серендіпового сімейства. Ієрархічний підхід

Описаний у п. 20.4.2. варіант створення $\varphi_m^{(p)}$ для дво- та тривимірних СЕ має певні недоліки.

По-перше, при цьому для трикутних СЕ необхідно застосовувати трикутні глобальні координати. А по-друге, в (20.41) і (20.42) при розмірності $k \ge 2$ та порядку наближення $p \ge 2$, як виявилося, є *надмірні* члени, що спричиняє невиправдані додаткові обчислення. На рис.20.9 для прикладу зображений двовимірний СЕ лагранжевого сімейства з інтерполяційними функціями другого порядку наближення. Видно, що СЕ має *внутрішній* вузол за номером 9, наявність якого, як з'ясувалося, зовсім не є обов'язковою.



Рис.20.9 Двовимірний СЕ лагранжевого сімейства другого порядку наближення (а) і дві його інтерполяційні функції (б, в) за двома напрямками

Усунувши надмірність, розробили сімейство СЕ *серендіпового* типу. В них вузли розташовуються на ребрах СЕ, а інтерполяційні функції отримують як

результат перемноження членів необхідного степеня *p* по *основній* локальній координаті (вздовж ребра) на *лінійні* члени по *іншій* локальній координаті. В результаті по ребрах СЕ наближення залишається таким же, як і в СЕ лагранжевого сімейства, тобто степінь гладкості та сумісність сусідніх СЕ зберігаються.

Вузлові інтерполяційні функції СЕ серендіпового сімейства легко будувати з використанням *принципу ієрархії* (див. п. 20.4.1). Згідно з цим принципом модифікація СЕ (запровадження нових вузлів) проводиться алгебраїчним додаванням вже існуючого виразу для вузлових інтерполяційних функцій з поправкою, що приводить значення вузлової інтерполяційної функції до необхідного значення, а саме до нуля; за величиною поправка дорівнює *частині* від значення вузлової функції "нового" сусіднього (на ребрі) вузла.

На рис.20.10 наведено приклад для СЕ серендіпового сімейства другого порядку наближення. При появі проміжного вузла 7 виконуються три корегування:

• вводиться інтерполяційна функція нового вузла 7 за зазначеним вище правилом: перемножуються вираз необхідного степеня (тут p = 2) по основній локальній координаті ξ на лінійний вираз по іншій локальній координаті η : $\varphi_7^{(2)}(\xi,\eta) = (1-\xi^2)(1+\eta)/2$;

• проводиться корекція інтерполяційної функції вузла 3 (для кутового вузла, сусіднього з вузлом 7) так, щоб її значення у вузлі 7 дорівнювало нулю: $\varphi_3^{(2)}(\xi,\eta) = \varphi_3^{(1)}(\xi,\eta) - \varphi_7^{(2)}(\xi,\eta)/2 = (1+\xi)(1+\eta)/4 - (1-\xi^2)(1+\eta)/4 = \xi(1+\xi)(1+\eta)/4;$

• проводиться корекція інтерполяційної функції вузла 4 (для кутового вузла, сусіднього з вузлом 7) так, щоб її значення у вузлі 7 дорівнювало нулю: $\varphi_4^{(2)}(\xi,\eta) = \varphi_4^{(1)}(\xi,\eta) - \varphi_7^{(2)}(\xi,\eta)/2 = (1-\xi)(1+\eta)/4 - (1-\xi^2)(1+\eta)/4 = -\xi(1-\xi)(1+\eta)/4.$



Рис.20.10 Двовимірний СЕ серендіпового сімейства другого порядку наближення (а) і дві його інтерполяційні функції (б, в) за двома напрямками

Щоб не виписувати всі вирази інтерполяційних функцій, вводять такі позначення (всі три локальні координати):

$$\tilde{\xi}_m = \xi_m \xi; \quad \tilde{\eta}_m = \eta_m \eta; \quad \tilde{\mu}_m = \mu_m \mu.$$
(20.43)

З використанням цих позначень запишемо вирази для інтерполяційних функцій для двовимірного CE (шаблон нумерації вузлів відповідає рис.20.10)

• першого порядку (лінійне) наближення:

$$\varphi_m^{(1)}(\xi,\eta) = (1+\tilde{\xi}_m)(1+\tilde{\eta}_m)/4; \quad m = 1,2,3,4;$$
(20.44)

• другого порядку (квадратичне) наближення:

$$\begin{cases} \varphi_m^{(2)}(\xi,\eta) = (1+\tilde{\xi}_m)(1+\tilde{\eta}_m)(\tilde{\xi}_m+\tilde{\eta}_m-1)/4; & m = 1,2,3,4; \\ \varphi_m^{(2)}(\xi,\eta) = (1-\xi^2)(1+\tilde{\eta}_m)/2; & m = 5,7; \\ \varphi_m^{(2)}(\xi,\eta) = (1+\tilde{\xi}_m)(1-\eta^2)/2; & m = 6,8; \end{cases}$$
(20.45)

• третього порядку (кубічне) наближення:

$$\begin{cases} \varphi_m^{(3)}(\xi,\eta) = (1+\tilde{\xi}_m)(1+\tilde{\eta}_m)[-10+9(\xi^2+\eta^2)]/32; & m=1,2,3,4; \\ \varphi_m^{(3)}(\xi,\eta) = (9/32)(1+\tilde{\xi}_m)(1-\eta^2)(1+9\tilde{\eta}_m); & m=7,8,11,12; \\ \varphi_m^{(3)}(\xi,\eta) = (9/32)(1+\tilde{\eta}_m)(1-\xi^2)(1+9\tilde{\xi}_m); & m=5,6,9,10. \end{cases}$$
(20.46)

Останні дві формули (20.46) відрізняються перестановкою координат.

Але при $p \ge 4$ вже недостатньо вузлів тільки на ребрах СЕ, щоб отримати повний многочлен. Тому СЕ серендіпового сімейства при $p \ge 4$ містять і внутрішні вузли, але в значно меншій кількості, ніж СЕ лагранжевого сімейства.

Для тривимірного CE серендіпового сімейства з чотирикутними сторонами інтерполяційні функції мають такий запис для CE (шаблон нумерації вузлів відповідає рис.20.11)

• першого порядку (лінійне) наближення:

$$p_m^{(1)}(\xi,\eta,\mu) = (1+\tilde{\xi}_m)(1+\tilde{\eta}_m)(1+\tilde{\mu}_m)/8; \quad m = 1,2,...,8;$$
(20.47)

- другого порядку (квадратичне) наближення:
- а) для проміжних вузлів

$$\begin{cases} \varphi_m^{(2)}(\xi,\eta,\mu) = (1-\xi^2)(1+\tilde{\eta}_m)(1+\tilde{\mu}_m)/4; & m = 9,11,13,15; \\ \varphi_m^{(2)}(\xi,\eta,\mu) = (1+\tilde{\xi}_m)(1-\eta^2)(1+\tilde{\mu}_m)/4; & m = 10,12,14,16; \\ \varphi_m^{(2)}(\xi,\eta,\mu) = (1+\tilde{\xi}_m)(1+\tilde{\eta}_m)(1-\mu^2)/4; & m = 17,18,19,20. \end{cases}$$
(20.48)

б) для кутових вузлів (номер m) при появі проміжного вузла (номер k) інтерполяційна функція модифікується згідно з принципом ієрархії за формулою

$$\varphi_m^{(2)}(\xi,\eta,\mu) = \varphi_m^{(1)}(\xi,\eta,\mu) - \varphi_k^{(2)}(\xi,\eta,\mu)/2; \quad m = 1,2,...,8.$$
(20.49)

Вище розглянуто тільки основні варіанти інтерполяційних функцій, причому в СЕ з чотирикутними сторонами. Оскільки після пом'якшення формулювання крайових задач механіки твердого тіла, що деформується, як правило, найвища похідна – перша, то для розв'язування таких задач формально достатньо застосовувати в СЕ базисні функції першого порядку наближення, часто застосовують СЕ другого порядку наближення і лише зрідка – третього.

Ще раз нагадаємо про обов'язкову вимогу до інтерполяційних функцій CE: в будь-якій точці CE необхідно, щоб

$$\sum_{m=1}^{M_e} \varphi_m^e(\xi, \eta, \mu) = 1; \quad \xi, \eta, \mu \in [-1, 1].$$
(20.50)

Властивість (20.50) зазвичай використовують для перевірки правильності призначення та програмування інтерполяційних функцій.

Також нагадаємо, що в МСЕ інтерполяційні функції використовуються для описання:

• геометрії СЕ. Їх називають функціями *форми*;

 функцій, що наближують розв'язок крайової задачі. Їх називають базисними функціями.

Якщо у СЕ його геометрія описується з вищим порядком наближення, ніж розв'язок крайової задачі (функції форми мають вищий порядок, ніж базисні),



ють вищий порядок, ніж базисні), то такі СЕ називають *суперпара-* серендіпового сімейства, чотирикутний в плані *метричними*. Навпаки, якщо у СЕ його функції форми мають нижчий порядок, ніж базисні, то такі СЕ називають *субпараметричними*.

Параметричні СЕ, в яких базисні функції та функції форми співпадають, називають *ізопараметричними*. Тобто будь-яка функція f в межах ізопараметричного СЕ описується формулою

$$f(\xi,\eta,\mu) = \sum_{m=1}^{M_e} \varphi_m^e(\xi,\eta,\mu) \cdot f_m , \qquad (20.51)$$

де f_m – значення функції у вузлах СЕ; M_e – кількість вузлів у СЕ. Такими функціями будуть і глобальні координати:

$$x^{j}(\xi,\eta,\mu) = \sum_{m=1}^{M_{e}} \varphi_{m}^{e}(\xi,\eta,\mu) \cdot (x^{j})_{m}; \quad j = 1,2,3$$
(20.52)

для описування геометрії СЕ.

20.5. Про похідні моделі СЕ

Нові моделі СЕ, побудовані на основі існуючих, називають похідними моделями СЕ. Зокрема, якщо у k-вимірному просторі переміщувати (k-1)вимірний об'єкт вздовж твірної, яка описується деякою функцією f(s), де $s \in$ параметром довжини твірної, то можна одержати різноманітні СЕ. Вузли описують лінії (ребра) в k-вимірному просторі, які теж описуються функцією f(s). Є й інші варіанти трансформації СЕ.

Наведемо декілька прикладів похідних моделей СЕ:

• осесиметричні тривимірні СЕ породжуються шляхом повороту двовимірних СЕ навколо осі Z на кут у 2π радіан. Вузли описують окружності (див. рис.20.12-а,б), тобто параметром s є кут $\theta \in [0, 2\pi]$. Такий СЕ застосовують тоді, коли одночасно і тіло, і зовнішні впливи (навантаження, закріплення) є осесиметричними. Оскільки розв'язок крайової задачі не залежить від кута θ , то осесиметричні СЕ зображують у вигляді двовимірних перерізів (на рис.20.12 – заштриховані), для яких застосовують ті ж інтерполяційні функції, що й у двовимірних СЕ. Природно, що значення радіальної координати $\rho = x^1$ будьякого з вузлів тіла не можуть бути від'ємними;



Рис.20.12 Тривимірні СЕ, похідні від двовимірних СЕ: осесиметричні трикутного (а) та чотирьохкутного (б) перерізу; в) – трикутна призма

• з двовимірного трикутного СЕ "витягують" стовпчик потрібної довжини h, тобто параметр $s \in [0, h]$. Створюється трикутна призма (пентагональний СЕ), один з варіантів якої зображено на рис.20.12-в. Для таких, навіть однакових, трикутних призм розроблено різні варіанти інтерполяційних функцій, однак досі немає варіанта, який би не викликав підвищені похибки наближення похідних від функції, що апроксимується, при всіх випадках розташування вузлів;

• двовимірний чотирикутний СЕ трансформують в трикутний шляхом стягування вузлів, що належать одному з ребер (на рис.20.13 – верхньому), в одну точку. При цьому певним чином необхідно модифікувати й інтерполяційні функції. Очевидно, що для двовимірного СЕ таких варіантів – чотири.



Рис.20.13 Параболічний двовимірний трикутний СЕ як вироджений чотирикутний

Замість (історично перших) L-координат (їх три) застосовують локальну ортогональну систему координат ξ , η . Тоді базисні функції мають вирази:

$$\varphi_3^e(\xi,\mu) = (1-\xi)(1-\mu)/4 - (1-\xi)(1-\mu^2)/4 - (1-\xi^2)(1-\mu)/4 + \Delta;$$

$$\varphi_4^e(\xi,\mu) = (1+\xi)(1-\mu)/4 - (1-\xi^2)(1-\mu)/4 - (1+\xi)(1-\mu^2)/4 + \Delta;$$

	Частина	VI.	Розділ	20
--	---------	-----	--------	----

$\varphi_1^e(\xi,\mu) = (\mu+1)/2 - (1-\mu^2)/2;$	$\varphi_6^e(\xi,\mu) = (1-\xi)(1-\mu^2)/2;$	(20.53)
$\varphi_7^e(\xi,\mu) = (1-\mu^2)(1-\mu)/2-2\Delta;$	$\varphi_8^e(\xi,\mu) = (1+\xi)(1-\mu^2)/2,$	

де позначено $\Delta = (1 - \xi^2)(1 - \mu^2) / 8$.

Для тривимірного гексагонального CE варіантів його модифікації шліхом стягуванням ребер в точку значно більше;

• зміщують проміжні вузли у бік кутових, причому інтерполяційні функції не корегують. Ці СЕ застосовуються при розрахунках тіл з тріщинами для поліпшення наближення напружено-деформованого стану у вершин тріщин відповідно до оцінок аналітичних розв'язків. Для збереження властивості неперервності функції, що апроксимується, не тільки у вузлах, а й вздовж ребер, ще й стягують одне ребро у точку: на СЕ з рис.20.13 (розташованому правіше) ще потрібно перемістити вузли з номерами 6 і 8 вверх на чверть висоти. Виявляється, що при цьому змінюються асимптоти розв'язку, що наближуються такими СЕ – на $1/\sqrt{r}$, де r – відстань від вершини тріщини.

Додаткову інформацію у спеціальних випадках модифікацій СЕ можна знайти в довідковій літературі.

20.6. СЕ з ермітовими поліномами для базисних функцій

Деякі моделі крайових задач потребують побудови наближень у СЕ, які окрім самих функцій наближують і їхні перші або навіть вищі похідні. Зокрема, це крайові задачі про згин стержнів, пластин та оболонок. При згині стержнів основна шукана функція – прогин w(x), а в граничних умовах часто фігурує кут повороту площини перетину v(x), причому вважається, що $v(x) = \partial w(x) / \partial x = w'(x)$.

Розглянемо одновимірний СЕ з двома вузлами. Будемо вважати, що відомі вузлові значення не тільки функції, але й її перших похідних, тобто $u_m = u(\xi_m)$; $u'_m = u'(\xi_m)$; m = 1,2. Для врахування 4-х відомих значень потрібно застосувати поліном з p = 3, а саме $u(\xi) = a_0 + \sum_{i=1}^{3} a_i \xi^i$. Будемо вважати, що $\xi \in [-1,1]$. Вираз для першої похідної: $u'(\xi) = a_1 + 2a_2\xi + 3a_3\xi^2$. Складемо лінійну систему рівнянь

$$\begin{cases} u_1 = u(-1) = a_0 - a_1 + a_2 - a_3; \\ u_2 = u(1) = a_0 + a_1 + a_2 + a_3; \\ u'_1 = u'(-1) = a_1 - 2a_2 + 3a_3; \\ u'_2 = u'(1) = a_1 + 2a_2 + 3a_3; \end{cases}$$
(20.53)

яка має розв'язок $a_0 = [2(u_1 + u_2) + (u'_1 - u'_2)]/4$; $a_1 = -[3(u_1 - u_2) + (u'_1 + u'_2)]/4$; $a_2 = -(u'_1 - u'_2)/4$ та $a_3 = [(u_1 - u_2) + (u'_1 + u'_2)]/4$. Якщо згрупувати коефіцієнти відносно u_1, u_2, u'_1, u'_2 , то можемо записати, що

289

Частина VI. Розділ 20

$$u(\xi) = \sum_{m=1}^{2} \varphi_m(\xi) \cdot u_m + \sum_{m=1}^{2} \vartheta_m(\xi) \cdot u'_m, \qquad (20.54)$$

де інтерполяційні функції

 $\varphi_m(\xi) = (2+3\tilde{\xi}_m - \tilde{\xi}_m \xi^2)/4;$ $\vartheta_m(\xi) = -\xi_m (1-\tilde{\xi}_m)(1+\tilde{\xi}_m)^2/4;$ m = 1,2 (20.55) є поліномами Ерміта, записаними для діапазону $\xi \in [-1,1]$ (у довідниках зазвичай їх записують для діапазону $\xi \in [0,1]$). Як і раніше $\tilde{\xi}_m = \xi_m \xi$. Інтерполяційні функції (20.55) зображені на рис.20.13. Легко перевірити, що

$$\sum_{m=1}^{2} \varphi_m(\xi) = 1; \quad \xi \in [-1,1].$$
(20.56)

При необхідності врахувати і другі похідні, тобто $u_m'' = u''(\xi_m)$; m = 1, 2, застосовується многочлен з p = 4, а саме $u(\xi) = a_0 + \sum_{n=1}^4 a_n \xi^n$. І взагалі, щоб в одновимірному СЕ з двома вузлами врахувати похідні q-го порядку, потрібно обирати поліном степені p = q + 2.



Рис.20.13 Інтерполяційні функції $\varphi_m(\xi)$ та $\mathcal{G}_m(\xi)$, m = 1, 2

Осесиметричний оболонковий СЕ, який моделює згин, можна розглядати як похідний від одновимірного (див. підрозділ 20.6), якщо прийняти, що останній належить площині x^1x^3 або ρz , де $\rho = x^1$ та $z = x^3$. Розв'язок не залежить від координати $\theta = x^2$. Для двовимірного СЕ при відомих перших похідних побудова СЕ значно ускладнюється (тут не розглядається).

20.7. Додаткові критерії щодо вибору базисних функцій при розв'язуванні МСЕ крайових задач з диференційними операторами

Очевидно, що кускова інтерполяція простору та розв'язку не дозволяє отримати точний розв'язок задачі, окрім тих випадків, коли ця інтерполяція співпадає з точним розв'язком. Доведена теорема, що у загальному випадку точний розв'язок можна отримати лише при наближенні діаметрів СЕ до нуля, але така дуже щільна вузлова сітка неможлива в ЕОМ. Тобто, практично завжди є похибка розв'язку крайової задачі, пов'язана зі "скінченними" розмірами елементів та обмеженою кількістю вузлів.

Але це не все. Можна помилково призначити такі інтерполяційні функції, що збіжності розв'язків не буде ніколи.

Основне призначення МСЕ – розв'язування крайових задач. Для них характерна наявність похідних першого і вищих порядків за просторовими змінними. В зв'язку з цим потрібно відповісти на питання про необхідний степінь гладкості базисних функцій для розв'язування крайових задач, що містять диференційні похідні, та про повноту цих функцій.

Для найпростішого аналізу ситуації достатньо розглянути одновимірний випадок, для якого потрібно взяти три різноманітні степеневі базисні функції: постійну, лінійну і квадратичну:

$$\phi = a; \quad \psi = a + bx; \quad \theta = a + bx + cx^2,$$
 (20.57)

де а, b, с – відомі константи. Очевидно, що

$$\frac{\partial \varphi}{\partial x} = 0; \quad \frac{\partial \psi}{\partial x} = b; \quad \frac{\partial \theta}{\partial x} = b + 2cx; \\ - \qquad \frac{\partial^2 \psi}{\partial x^2} = 0; \quad \frac{\partial^2 \theta}{\partial x^2} = 2c; \quad (20.58) \\ - \qquad - \qquad \frac{\partial^3 \theta}{\partial x^3} = 0.$$

Хоча в розв'язках задачі похідні інколи й можуть приймати нульові значення, але це не правило, а виняток. Тому базисну функцію φ (див. (20.57)) не можна застосовувати для розв'язування задач з похідними вже першого порядку, ψ – другого, θ – третього. Інакше кажучи, порядок наближення базисної функції p (степеневої) повинен бути не нижчим, ніж найвища похідна q в рівняннях крайової задачі, тобто повинно бути $p \ge q$.

СЕ стикуються по міжелементних поверхнях (гранях, ребрах, вузлах), і саме там відбувається перехід від однієї базисної функції до іншої. Самі базисні функції стикуються без розривів, а от про їхні похідні це можна сказати не завжди, оскільки така вимога до базисних функцій може й не висуватися. Зокрема, у всіх СЕ, розглянутих у підрозділах 20.2 ... 20.5, реалізовувалася C^0 -гладкість. І лише у підрозділі 20.6 була розглянута ідея ермітового наближення, яке дозволяє отримати інтерполяційні функції C^q -гладкості при q > 0.

Нагадаємо, що при стикуванні двох функцій С^{*q*}-гладкість означає гарантію відсутності розривів першого роду для всіх похідних до порядку *q* включно.

Тепер розглянемо проблему повноти. В літературі сформульовано умови, за яких інтерполяційні функції мають властивість повноти. Ця проблема нерозривно пов'язана з проблемою порядку похибки розв'язку.

Припустимо, що при розв'язуванні одновимірної крайової задачі застосовуються СЕ з інтерполяційними функціями у вигляді многочленів порядку p. СЕ мають характерне значення діаметрів δ . Нехай точний розв'язок даної задачі *не має особливих точок*, де хоча б деякі з розглядуваних похідних є нескінченними. Тоді в межах СЕ цей розв'язок можна розкласти в ряд Тейлора в околі актуальної точки C, що входить до СЕ:

$$u(x) \approx u_{c} + u_{c}' \Delta x + u_{c}'' \Delta x^{2} / 2 + \dots + u_{c}^{(p)} \Delta x^{p} / p! + O(\Delta x^{p+1}), \qquad (20.59)$$

де Δx – відстань від точки *С*. Очевидно, що многочлен порядку *р* інтерполяційної функції точно відтворюється відповідним відрізком ряду Тейлора. Тоді

всередині СЕ з діаметром δ максимальна похибка наближення буде $O(\delta^{p+1})$, оскільки max $\{\Delta x\} = \delta$. Аналогічно похибка наближення першої похідної розв'язку u' буде мати порядок $O(\delta^p)$, другої похідної – $O(\delta^{p-1})$, похідної порядку q – порядок $O(\delta^{p+1-q})$, причому найгіршим варіантом, який можна дозволити, буде $O(\delta)$. Отже, для виконання вимог повноти необхідно, щоб

$$p+1-q \ge 1 \quad \text{afo} \quad p \ge q \,. \tag{20.60}$$

Саме таке обмеження вже було виявлено на початку цього Розділу.

Отже, якщо сформульована крайова задача має похідні порядку q, то для забезпечення повноти вимагається застосування базисної функції у вигляді многочлена порядку $p \ge q$, не нижче. Ось чому *послаблене* формулювання задачі в методі зважених похибок наближення є таким важливим в МСЕ: воно дозволяє знизити потрібний порядок многочлена базисної функції.

Зазначене про гладкість і повноту базисних функцій можна повністю перенести і на вагові функції, особливо враховуючи, що в результаті послаблення формулювання задачі від цих функцій, як правило, буде вимагатися однаковий з базисними функціями степінь гладкості і обов'язково – повнота.

Отриману вище для МСЕ умову повноти використовують при *тестуванні* програмних реалізацій алгоритмів. Для такого тестування підбирають крайову задачу з точним розв'язком у вигляді полінома степеня q = p, де p – степінь полінома інтерполяційних функцій СЕ. Тоді результати скінченно-елементного розв'язку повинні *точно* відповідати точному розв'язку в будь-якій точці СЕ. Відхилення вказують на наявність помилки (помилок) в алгоритмі або комп'ютерній програмі.

Окрім розглянутих вище вимог є й інші. Їх сформулювали у вигляді трьох критеріїв відносно базисних функцій, що наближують *переміщення*.

Критерій 1. Базисні функції повинні бути такими, щоб не моделювали паразитні деформації у СЕ при вузлових переміщеннях, викликаних переміщенням СЕ як жорсткого цілого. Доведено, що всі СЕ, в яких виконується умова (20.50), а саме $\sum_{m=1}^{M_e} \varphi_m^e = 1$, задовольняють Критерію 1, що, у свою чергу є умовою збіжності наближення до точного.

Критерій 2. Базисні функції повинні бути такими, щоб вони описували однорідні деформації незалежно від розмірів СЕ. Дійсно, при прямуванні розмірів СЕ до нуля розв'язок прямує до точного, а деформації у СЕ – до однорідної деформації. Але у випадку виникнення у всьому тілі однорідної деформації необхідно, щоб розв'язок не залежав від розмірів СЕ.

Критерій 3. Базисні функції повинні бути такими, щоб *деформації* на границях між суміжними СЕ були обмеженими (навіть тоді, коли деформації на них не є визначеними). Інакше кажучи, цей критерій вимагає безперервність

переміщень на границях між СЕ, а математично постулює вимогу повноти базисних функцій, про яку постійно йшла мова в попередніх Розділах.

Примітка 20.3. Іноді застосовують базисні функції, що не відповідають третьому критерію (*несумісні* СЕ). Але при цьому не повинні використовуватися енергетичні функціонали (наближення за енергією).

Примітка 20.4. Якщо стикуються СЕ, які відповідають третьому критерію, але при цьому їхні *грані на стиках не співпадають повністю*, то розв'язок задачі буде невірним, оскільки при цьому зазвичай не забезпечується безперервність переміщень на границях між СЕ. Це рівнозначно введенню в модель тріщини.

Примітка 20.5. Паразитні деформації у СЕ при його жорсткому повороті на деякий кут (це є одним з варіантів вузлових переміщень при переміщенні СЕ як жорсткого цілого), можуть бути викликані не поганими якостями базисних функцій, а наближеним характером застосованих формул для обчислення деформацій. Зокрема, відомо, що при застосуванні рівнянь для малих деформацій (12.51), тобто $\varepsilon_{mn} = 0.5(\nabla_m u^n + \nabla_n u^m)$, жорсткий поворот елементарного об'єму матеріалу на кут у 90 градусів викличе паразитні деформації, що дорівнюють одиниці, тобто 100 відсотків деформацій та нескінченне значення відносної похибки!

20.8. Оцінка енергетичної норми похибок розв'язків крайових задач у МСЕ. Швидкості *n* – та *p* – збіжності розв'язків

Доведено, що енергетична норма похибок (див. формулу (14.6)) розв'язків крайової задачі пропорційна величині δ (яку визначено згідно з формулою (20.3) як міру дрібноти скінченно-елементної моделі тіла), у деякій степені, а саме

$$\|\vec{\Delta}_h\|_h \le \alpha \delta^{p+1-q} \,. \tag{20.61}$$

Наприклад, при q = 1 (найвища похідна – перша, що характерно для більшості крайових задач) та p = 1 (перший порядок наближення у CE), маємо $\|\vec{\Delta}_h\|_h \le \alpha \delta$, а при p = 2 (другий порядок наближення у CE) $\|\vec{\Delta}_h\|_h \le \alpha \delta^2$.

З (20.61) випливає, що p-збіжність (за рахунок підвищення порядку наближення) швидше, ніж n-збіжність (за рахунок збільшення кількості вузлів). Дійсно, n-збіжність відбувається за рахунок зменшення δ у k разів. При p+1-q=1 величина $\|\vec{\Delta}_h\|_h \le \alpha \delta/k$. Якщо мати p+1-q=2, то $\|\vec{\Delta}_h\|_h \le \alpha (\delta/k)^2$, тобто у другому варіанті швидкість збіжності вища, ніж у першому.

Однак у першому варіанті при тій же кількості вузлів час отримання розв'язку є більшим, оскільки матриця СЛАР, що породжується МСЕ, має більшу щільність заповнення, ніж у другому варіанті. Тому вибір порядку

наближення (перший, другий або ще вищий) та кількості вузлів є індивідуальним для кожної крайової задачі, ЕОМ та обчислювача.

20.9. Завершення

Розроблено стільки варіантів інтерполяційних функцій для СЕ, що для їхнього викладення потрібно багато місця. Тому обмежилися лише основними їх варіантами. Зокрема, є СЕ, що моделюють вершину тріщини у вигляді часткового розрізу СЕ. Особливий випадок – "нескінченні" СЕ, тобто СЕ, що моделюють нескінченне подовження області в деякому напрямку, реально не маючи при цьому значних розмірів.

Нагадаємо (див. підрозділ 20.1), що кожний СЕ, окрім набору функцій форми для описування його геометрії та базисних функцій для наближення розв'язку крайової задачі має ще декілька додаткових характеристик, зокрема:

• індивідуальний номер СЕ у моделі тіла;

• тип CE, який однозначно визначає, зокрема, шаблон нумерації вузлів та граней CE, а також застосовані функцій форми та базисні функції;

• кількість вузлів та набір глобальних номерів вузлів, що входять у склад СЕ (топологічна таблиця);

• номер (назва) матеріалу, з якого "зроблений" СЕ;

• орієнтація осей анізотропії матеріалу у СЕ, якщо він не ізотропний;

• для одновимірного СЕ: геометрія і геометричні характеристики перерізу, орієнтація осей перерізу у глобальних координатах;

• для двовимірного СЕ: геометричні характеристики пластини (оболонки);

• у багатошарових СЕ: кількість шарів та їх орієнтації відносно СЕ або глобальної системи координат;

• інформацію про метод і характеристики наближеного інтегрування в СЕ, якщо таке застосовується;

• інше.

Оскільки всі характеристики, окрім трьох перших, можуть бути однаковими у всіх СЕ або значної їх частині, зазвичай ці характеристики групують у декілька наборів даних, які називають властивостями (англ.: **Property**) CE.

Додамо, що для одновимірних та двовимірних СЕ, окрім наближення базисними функціями, ще використовуються додаткові *моделі деформування*. Зокрема, СЕ типу БАЛКА повинен моделювати такі види деформування: розтяг/стиск, скручування та згин. Якщо при розтягу/стиску та скручуванні зазвичай використовуються прості моделі механіки матеріалів та конструкцій, то при згині, окрім простої моделі (Ейлера-Бернуллі), ще використовують модель "балка Тимошенка". Про це докладніше йдеться в Розділі 24. А в СЕ типу ПЛАСТИНА можуть використовувати як модель Кірхгофа (деформація в напрямку товщини лінійно залежить від відстані від середньої поверхні), так і більш складну модель Рейснера-Міндліна. А також декілька інших.
Контрольні питання до Розділу 20

1. Що означає порожня область відхилення?

- 2. Що таке та для чого існує топологічна таблиця скінченних елементів?
- 3. Наведіть типізацію СЕ.
- 4. Які величини звуться діаметром СЕ та мірою дрібноти скінченно-елементної моделі?

5. Поняття симплекса. Яку систему рівнянь потрібно розв'язувати для визначення коєфіцієнтів симплексного наближення в СЕ?

6. Які похибки наближення мають симплексні СЕ?

7. Що визначають трикутні, барицентричні та *L*-координати?

8. Чим відрізняються комплексні та мультиплексні моделі СЕ від симплексних СЕ?

9. Які характерні ознаки параметричних СЕ?

- 10. Чим відрізняються СЕ лагранжевого та серендіпового сімейств?
- 11. Що таке ієрархічний підхід і для чого він створений?
- 12. Що таке похідні моделі СЕ?
- 13. В яких випадках потрібно за базисні функції застосовувати ермітові поліноми?
- 14. Якім критеріям повинні відповідати базисні функції у СЕ?

15. До яких висновків призводить енергетична оцінка норм похибок розв'язків крайових задач МСЕ щодо *n*- та *p*- збіжності розв'язків?