УКРАЇНСЬКА АКАДЕМІЯ НАУК
ІНСТИТУТ МАТЕМАТИКИ

ACADÉMIE DES SCIENCES D'UKRAINE
INSTITUT DES SCIENCES MATHÉMATIQUES

ЄВГЕН РЕМЕЗ

ПРО МЕТОДИ НАЙКРАЩОГО, В РОЗУМІННІ ЧЕБИШОВА, НАБЛИЖЕНОГО ПРЕДСТАВЛЕННЯ ФУНКЦІЙ

EUGÈNE REMES

SUR LES MÉTHODES POUR RÉALISER LA MEILLEURE APPROXIMATION DES FONCTIONS D'APRÈS LE PRINCIPE DE TCHEBYCHEF

ВИДАВНИЦТВО УКРАЇНСЬКОЇ АКАДЕМІЇ НАУК
КИЇВ — 1935 — KYIV

УКРАЇНСЬКА АКАДЕМІЯ НАУК
ІНСТИТУТ МАТЕМАТИКИ

ACADÉMIE DES SCIENCES D'UKRAINE
INSTITUT DES SCIENCES MATHÉMATIQUES

ЄВГЕН РЕМЕЗ

ПРО МЕТОДИ НАЙКРАЩОГО, В РОЗУМІННІ ЧЕБИШОВА, НАБЛИЖЕНОГО ПРЕДСТАВЛЕННЯ ФУНКЦІЙ

EUGÈNE REMES

SUR LES MÉTHODES POUR RÉALISER LA MEILLEURE APPROXIMATION DES FONCTIONS D'APRÈS LE PRINCIPE DE TCHEBYCHEF

ВИДАВНИЦТВО УКРАЇНСЬКОЇ АКАДЕМІЇ НАУК
КИЇВ — 1935 — KYIV

Бібліографічний опис цього видання вміщено в „Літопису українського друку", „Картковому репертуарі" та інших покажчиках Української книжкової палати.

Відповід. редактор акад. *Д. О. Граве*
Літредактор і учений коректор
М. В. Качеровський

Друкується з розпорядження Української Академії Наук.
Неодмінний секретар УАН акад. *О. В. Палладін.*

З друкарні-літографії Української Академії Наук у Києві.

ВСТУП

Задача Чебишова щодо найкращого наближеного представлення даної функції в певній області за допомогою полінома даного степеня або іншого типу виразу, що залежить від певної кількості параметрів, поруч свого глибокого значення для теорії функцій, яке було виявлено дослідами ряду видатних математиків цього століття, головним же чином основоположними працями С. Н. Бернштейна, має, з другого боку, незаперечне й першорядне практичне значення: досить згадати хоч би застосування до механізмів П. Л. Чебишова або (при трохи поширеному трактуванні проблеми) зв'язаний з іменами П. Л. Чебишова та Д. О. Граве принцип побудування географічних карт, або різноманітні застосування формул Poncelet, або певні питання щодо визначення параметрів емпіричних формул [1] — питання, на які натрапив був ще Laplace і які новішого часу знову висунув бельгієць Ed. Goedseels в сфері практичної геодезії, або нарешті недавні досліди Wilhelm Cauer-а щодо електричних фільтрів. Для практичних застосувань цього роду, що їх обсяг без сумніву є незрівняно ширший, ніж ті спорадичні випадки, до яких апроксимаційний принцип Чебишова досі застосовувався, питання першорядної ваги полягає насамперед в умінні фактично визначати шукані найкращі наближені представлення функцій. Але до останнього часу не розроблено для цієї мети ні в нашій, ні в світовій літературі скількинебудь загальних і водночас практично застосованих методів, які дозволяли б розв'язувати цього роду питання з бажаним ступенем точності. В даній праці дослідження цього питання розгорнуто в двох (точніше в трьох) основних напрямках.

1°. В напрямку розроблення методів розв'язання у скінченнім виді задачі Чебишова в деяких важливих для практики випадках. Ми обрали тут за об'єкт дослідження насамперед поліноміальні наближені представлення лінійного або узагальнено-лінійного типу для функцій від кількох змінних за деяких обмежних умов. Цьому питанню присвячено першу частину праці (розд. I—V). Особливе значення проблеми лінійної апроксимації функцій між іншим в аспекті практики наближеного розв'язання рівнянь алгебричних (а також і трансцендентних), диференціальних та різних функціональних майже не потребує пояснень: через вдалу лінеаризацію рівнянь приходимо в певних випадках до найбільш вивчених типів відповідних функціональних рівнянь, до найпростіших

[1]) Факт незаперечної принципової ваги полягає між іншим у тому, що лише апроксимація за принципом Чебишова допускає суворе поставлення питання про те, чи вкладаються дані досвіду в емпіричну формулу певного типу.

алгебричних рівнянь тощо, і це дає змогу добути наближене розв'язання, що його в деяких випадках можна прийняти навіть за остаточне, а в інших випадках — за попереднє наближення для наступного застосування відповідного методу послідовного наближення. Додамо, що, застосовуючи різні „анаморфози", можемо з формул лінійних або узагальнено-лінійних безпосередньо виводити формули апроксимації інших типів. Приміром, у випадку двох незалежних змінних через перетворення $x=\xi$, $y=\frac{1}{\eta}$ приходимо від „білінійного" виразу $a+bx+cy+\partial xy$ до виразу типу $a+b\xi+\frac{c}{\eta}+\partial\frac{\xi}{\eta}$. Можливість різноманітних загальніших перетворень зрозуміла сама собою.

До апроксимаційних формул розглядуваного типу належать відомі елементарні формули Понселе, які дають для $\sqrt{x^2+y^2}$ та $\sqrt{x^2-y^3}$ наближені вирази виду $ax+by$. Це по суті, зважаючи на однорідність, є формули лінійного наближеного представлення для функції одного змінного, що їх великий геометр визначає на підставі елементарної інтуїції. Спробам аналогічного наближеного представлення для $\sqrt{x^2+y^2+z^2}$ були присвячені роботи Horvath-a, Résal-я та А. А. Маркова [1]). Не кажучи про роботу Horvath-a, яка виконана просто неохайно (саме від неї беруть початок зовсім неприпустимі найгрубіші помилки, які досі маємо в формулах Понселе, вміщених у найпоширеніших технічних довідниках), названі досліди, стосуючись найпростішого дуже частинного випадку, потребують ще багато так в розумінні точності або викінченості результатів, як і в розумінні досконалості застосованого апарату. Horvath та Résal, широко посилаючись на геометричну „очевидність" („il est visible...") приходять до результатів, які виявляються в цілому просто неправдиві (факт досі, здається, непомічений!) — принаймні в даному авторами загальному сформулюванні, як ми це покажемо в розділі III. Щождо роботи нашого славетного покійного академіка, яка трактує власне задачу Понселе в трохи іншому поставленні, то хоч у ній ми й маємо ряд з майстерністю проведених викладок, проте ефективного доведення вона все ж не містить [2]).

2°. Другу частину праці присвячено, поперше, розробленню загального методу для розв'язання задачі Чебишова шляхом послідовного наближення. Ми тут розглядаємо насамперед випадок поліноміального наближеного представлення для функції одного змінного при досить загальних умовах: припускається лише дану функцію $f(x)$ та область, якої стосується апроксимація, обмеженими, не впроваджуючи ніяких інших умов, навіть не вимагаючи цілковитої (однозначної) означеності даної функції. Ідея застосувати нескінченний процес, щоб розв'язати задачу, яка не піддається розв'язанню у скінченному виді, є, звичайно, цілком природна і, мабуть, неминуча і, як така, вона не нова. Але ми не знаємо

[1]) До згаданих тут робіт ми ще повернемося в розділі III цієї роботи.

[2]) Докладніше див. розділ III цієї роботи.

в попередній літературі цього питання такого методу, що базується на нескінченному процесі, який був би практично здійсненний і разом з тим досить загальний. Якщо залишити осторонь ідею граничного переходу від області, що складається із скінченного числа точок, до області неперервної — ідею, зв'язану з іменами P. Kirchberger-a [1]) та de la Vallée-Poussin-a [2]), яка навіть у випадку неперервної функції та неперервної області являє лише теоретичний інтерес [3]), ми маємо два методи: поперше, метод П. Л. Чебишова, вказаний в його мемуарі „Théorie des mécanismes connus sous le nom de parallélogrammes" (§§ 6—8) [4]), і подруге — визначний метод, вказаний С. Н. Бернштейном в його докторській дисертації [5]) у зв'язку з задачею вивчення поліноміальних наближених представлень функції $|x|$. Але методи обох славетних математиків хоч і виправдали себе щодо тих спеціальних задач, які розглядались цими авторами, проте в загальному випадку ці методи, крім інших умов, що дуже істотно обмежують їхнє прикладання, припускають насамперед, що дана функція $f(x)$ є не лише аналітична, але ще й регулярна на всьому розглядуваному інтервалі (a, b). Методи ці зовсім не є застосовні у випадку функції від двох або кількох змінних. Метод послідовного наближення, який ми тут пропонуємо, пов'язуючись у вихідних своїх точках із деякими загальними теоретичними уставленнями Чебишова, Borel-я та de la Vallée-Poussin-a, має характер відмінний від попередніх, який уможливлює прикладання нашого методу до функцій з різноманітними розривами неперервності (і самої функції, і її похідних), також до функцій, визначених на розривній або дискретній множині точок, і нарешті — до певних випадків апроксимації функцій від кількох незалежних змінних. Виявляється при цьому, що задача лінійної апроксимації функції від кількох змінних становить в загальному випадку більш високий ступінь трудності, ніж задача апроксимації нелінійної для функції одного змінного.

3°. В тій самій другій частині праці ми розвиваємо ряд теорем існування, що стосуються розривних або неоднозначно-означених функцій одного змінного (розд. VI) і далі — функцій від кількох змінних (розд. VIII), розглядаючи останнє питання в щільному зв'язку з задачею „мінімального наближення" (термін Goedseels-a — de la Vallée-Poussin-a) системи суперечних лінійних рівнянь. Останні питання, крім свого безпосереднього значення для розроблення сформульованих вище практичних проблем, мають, як нам здається, й ширше самостійне значення. Наскільки

[1]) Mathematische Annalen, Bd. 57, 1903.

[2]) Bull. de l'Académie Roy. de Belgique, 1910.

[3]) Ми не можемо поки сказати більшого і про практичне значення ідеї граничного переходу від наближення за принципом „найменших m-их степенів" до найкращого наближення в розумінні Чебишова (Polya, Comptes Rendus, t. 157; D. Jackson. Transact. of the Amer. Math. Soc., vol. 22; 1921).

[4]) Tchebycheff. Oeuvres, t. I, pp. 111—143.

[5]) С. Бернштейн. О наилучшем приближении непрерывных функций посредств многочленов данной степени. Харьков, 1912. Маємо тут на увазі параметричний С. Н. Бернштейна.

одержані нами в цьому напрямку результати є по суті нові — про це нехай судять ті, хто більш повно, ніж я, знайомі з літературою з цих питань за останні роки.

Переходячи до більш деталізованого огляду по окремих розділах, зауважмо наперед, що прикладання поданих тут і теоретично обгрунтованих методів до розв'язання деяких конкретних проблем становлять істотну складову частину даної праці. Крім того, ми не уникали в окремих випадках і прикладів характеру ілюстративного, пам'ятаючи, що іноді exempla non minus docunt, quam praecepta.

В розділі I трактується задача найкращого наближеного представлення функції $f(x_1, x_2, \ldots, x_n)$, підпорядкованої деяким умовам, у певній відповідній області за допомогою або лінійного виразу $a_0 + \sum_{i=1}^{n} a_i x_i$, або загальнішого виразу типу

$$\begin{aligned}
&\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \ldots\ldots\ldots\ldots\ldots + \alpha_n x_n \\
&+ \alpha_{12} x_1 x_2 + \alpha_{13} x_1 x_3 + \ldots\ldots\ldots\ldots + \alpha_{n-1,n} x_{n-1} x_n \\
&+ \alpha_{123} x_1 x_2 x_3 + \alpha_{124} x_1 x_2 x_4 + \ldots + \alpha_{n-2,\, n-1,\, n} x_{n-2} x_{n-1} x_n \\
&+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&+ \alpha_{123\ldots n} x_1 x_2 x_3 \ldots x_n
\end{aligned}$$

лінійного щодо кожного з n аргументів $x_1, x_2, \ldots, x_n$ окремо. Виясняється та істотна користь, яку можна добути в певних випадках, при розв'язанні цієї задачі, з теорії так званих субгармонічних та супергармонічних функцій. Основна теорема, на якій ми базуємо поданий спосіб розв'язання задачі, зв'язана з деяким узагальненням поняття барицентричних координат для певних типів лінійчастих множин у многомірних просторах. Разом із цим ми відмічаємо й інший можливий підхід, що має за єдиний вихідний пункт теорію потенціалу у многомірних просторах, і цей шлях дозволяє виявити інтимний зв'язок між розглядуваною зачачею лінійної апроксимації та визначною задачею Чебишова — Граве щодо побудування найкращої конформної карти для даної країни.

В розділі II поданий загальний спосіб прикладено до практично-важливої конкретної задачі, яка стосується наближеного лінійного представлення функції, заданої рівнянням виду $w^\rho = x^\rho + y^\rho + \cdots + v^\rho$, де ρ позначає якийсь даний дійсний показник. Ця задача містить як частинний випадок класичну задачу Понселе. Тут таки розглядається приклад застосування при розв'язанні диференціального рівняння.

В розділі III дається цілком викінчене трактування випадків Horvath-а та А. Маркова задачі щодо $\sqrt{x^2 + y^2 + z^2}$.

Розділ IV присвячено задачі найкращої лінійної апроксимації деяких явних алгебричних функцій від кількох аргументів, що мають найближче значення для теорії і практики розв'язування алгебричних рівнянь (метод Греффе, метод Бернуллі), зокрема у випадку дуже близьких коренів.

Розділ V присвячено застосуванню загального методу до кількох різнорідних груп питань. Поперше, тут розглянуто питання щодо найкращої апроксимації суб- або супергармонічної на даному інтервалі функції одного аргумента (тобто функції „конвексної“ або „конкавної“) за допомогою ординати ламаної з наперед даним числом боків. Подруге, застосовано загальний спосіб до деяких елементарних трансцендентних виразів з одним змінним, для яких аналогічного виду формули, але за принципом мінімальної середньої квадратичної похибки, побудував свого часу Шлемільх. Одержані тут нові формули можуть бути корисні при розв'язуванні відповідних трансцендентних рівнянь. Потретє, ми тут зупиняємось на виводі лінійних наближених виразів для x^n, xy, xyz і т. д. За допомогою добутих формул розв'язується деяке цікаве з практичного погляду питання, що має певний зв'язок із задачею апроксимації за принципом Чебишова, але зовсім з останньою задачею не збігається.

У розділах VI і VII після ряду основних означень та теорем подається метод послідовного наближення, придатний для розв'язання задачі найкращої апроксимації при умовах досить загальних. Міркування щодо задачі поліноміального наближеного представлення функцій одного змінного допускають більш або менш безпосередні узагальнення і для інших аналогічних задач, що стосуються, наприклад, тригонометричної апроксимації або апроксимації за допомогою деяких дробових раціональних виразів тощо. Викладений тут метод ми прикладаємо до деяких конкретних задач, що стосуються наближеного представлення функцій $|x|$, $|x|+k\,\mathrm{sg}\,x$. Зокрема відзначмо одержані наближені вирази поліномів найкращого наближення функції $|x|$ на інтервалі $(-1,1)$ до 10-го степеня включно. Разом з деякими результатами, відзначеними на початку розділу V, ці формули мають очевидне значення для загальної задачі поліноміальної апроксимації різноманітних неперервних функцій.

Великий розділ VIII переважно присвячено деяким питанням основного значення дла задачі апроксимації функцій від кількох дійсних змінних. „Les recherches sur l'approximation des fonctions de plusieurs variables sont à peine ébauchées“. Ці слова С. Н. Бернштейна були сказані, в розумінні більш загальному, 22 роки тому в його оглядовій доповіді на Кембріджському конгресі[1]). Навряд чи ми помилимось, коли скажемо, що з того часу стан справи в цікавому для нас питанні не змінився. Ми тут розвиваємо ряд основних теорем існування, що мають на меті поповнити трактування відповідних питань у Kirchberger-a, Tonelli, de la Vallée-Poussin-a, Haar-a. Відзначмо зокрема теореми § 23 n° 8 та § 24 n° 3. В § 25 виясняється можливість поширити викладений вище метод послідовних наближень на деякі випадки апроксимації функцій від кількох змінних та визначення параметрів емпіричних формул.

Деякі з питань, які трактуються в розділах VI та VII, ми частково

[1]) S. Bernstein. Approximation des fonctions continues par des polynômes. Proceedings of the fifth international congress of Mathematicians, Vol. I.

розглянули в двох своїх повідомленнях в Comptes Rendus (Paris), séances 11.VI, 30.VII 1934, та в роботі, що незабаром має вийти друком в Communications de la Soc. Math. de Kharkow, t. X.

Користуюсь з нагоди висловити свою вдячність науковому колективові київських математиків за виявлену увагу під час підготування цієї праці до друку. Окрему подяку я повинен скласти проф. Н. І. Ахієзеру за деякі цінні товариські вказівки, зокрема щодо бібліографії.

ЧАСТИНА ПЕРША

ПРО ДЕЯКІ ПРАКТИЧНОГО ЗНАЧЕННЯ ЗАДАЧІ, ЩО ДОПУСКАЮТЬ РОЗВ'ЯЗАННЯ В СКІНЧЕННОМУ ВИДІ

РОЗДІЛ I

ДЕЯКІ ЗАДАЧІ ЩОДО НАБЛИЖЕНОГО ПРЕДСТАВЛЕННЯ ФУНКЦІЙ ВІД КІЛЬКОХ ЗМІННИХ ТА ОСНОВНІ ЗАСАДИ ДО ЇХ РОЗВ'ЯЗАННЯ

§ 1. Попередня задача. Основна теорема

Функцію $f(x, y, z, \ldots, v)$, що підлягає апроксимації в певній області, яка далі точніше зазначається, ми припустимо покищо однозначно-заданою та неперервною, а в тих випадках, коли за критерій точності береться відносна похибка, ми ще припускатимемо взагалі дану функцію сталого знаку, для конкретності — додатного, в розглядуваній області [1]), відкидаючи останнє припущення при критерії похибки абсолютної. Ми розглядаємо в цьому розділі наближені представлення двох типів:

$$\alpha_0 + \alpha_1 x + \alpha_2 y + \cdots + \alpha_n v = \varphi_1(x, y, \ldots, v; \alpha_0, \alpha_1, \ldots, \alpha_n) \tag{1}$$

та

$$\left.\begin{array}{l} \alpha_0 + \alpha_1 x + \alpha_2 y + \cdots\cdots + \alpha_n v \\ + \alpha_{12} xy + \alpha_{13} xz + \cdots + \alpha_{n-1,\,n} uv \\ + \alpha_{123} xyz + \cdots + \alpha_{n-2,\,n-1,\,n} tuv \\ + \cdots\cdots\cdots\cdots\cdots\cdots \\ + \alpha_{123\ldots n-1,\,n} xyz\ldots uv \end{array}\right\} = \varphi_2(x, y, \ldots, v; \alpha_0, \alpha_1, \ldots, \alpha_{12\ldots n}) \tag{2}$$

В тих випадках, коли ми говоритимемо заразом про наближені представлення типу (1) або (2), не відрізняючи, якого саме з них, писатимемо:

$$\varphi(x, y, \ldots, v, \alpha_0 \alpha_1, \ldots)$$

За область апроксимації в разі наближених представлень типу (1) візьмемо, для конкретності міркувань, область, що виражається системою нерівностей

$$c \geqslant x \geqslant \frac{y}{\lambda} \geqslant \frac{z}{\lambda\mu} \geqslant \ldots \geqslant \frac{u}{\lambda\mu\nu\ldots\pi} \geqslant \frac{v}{\lambda\mu\nu\ldots\pi\rho} \geqslant 0 \tag{T}$$

[1]) В розділах II, IV буде з'ясовано у зв'язку з розв'язанням відповідних конкретних задач, як пристосувати поданий спосіб трактування до таких випадків, коли при відносної похибки функція f у деяких точках розглядуваної області перетворюється

Викладений нижче спосіб розв'язання зберігає силу для загального випадку замкненої області, що виражається системою лінійних нерівностей, які можна звести до канонічного виду (T) за допомогою відповідного (неособливого) лінійного перетворення незалежних змінних. Цим ми хочемо сказати, що в кожному подібному випадку зазначений спосіб застосовується безпосередньо, не вимагаючи справжнього проведення згаданого лінійного перетворення. Числа $\lambda, \mu, \nu, \ldots, \rho$ ми можемо припустити, не порушуючи загальності, додатними.

В разі ж наближеного представлення типу (2) область апроксимації виражатиметься системою нерівностей

$$0 \leqslant x \leqslant a,\ 0 \leqslant y \leqslant b,\ 0 \leqslant z \leqslant c, \ldots, 0 \leqslant v \leqslant h \qquad (P)$$

Тут міркування безпосередньо поширюються на випадок замкненої області, яку дістаємо, застосовуючи до змінних $x, y, \ldots, v$ в системі нерівностей (P) лінійне перетворення

$$x = X + x_0,\ y = Y + y_0, \ldots, v = V + v_0,$$

що переводить вираз (2) на вираз однакового типу.

Позначмо через R відношення наближеного представлення функції $f(x, y, \ldots, v)$ до самої функції:

$$\frac{\varphi(x, y, \ldots, v;\ \alpha_0, \alpha_1, \alpha_2, \ldots)}{f(x, y, \ldots, v)} = R(x, y, \ldots, v;\ \alpha_0, \alpha_1, \alpha_2, \ldots) \qquad (3)$$

Відношення R припускатимемо додатним. Нехай будуть $R_{\max}$ та $R_{\min}$ відповідно найбільше та найменше значення відношення R при зафіксованих параметрах $\alpha_0, \alpha_1, \alpha_2, \ldots$ та при змінних $x, y, \ldots, v$ в області (T) або (P) відповідно до типу наближеного представлення [1]).

Відношення

$$\frac{R_{\max}}{R_{\min}} = \Omega(\alpha_0, \alpha_1, \alpha_2, \ldots) \qquad (4)$$

ми позначатимемо як показник коливання відносної точності наближеного представлення функції f за допомогою виразу φ_1 або φ_2 в області (T) або (P) відповідно.

Попередня задача, яку ми тут розглянемо, полягає якраз в наступній вимозі: дібрати параметри $\alpha_0, \alpha_1, \alpha_2, \ldots$ у виразі φ_1 чи φ_2 (залежно від типу шуканого наближеного представлення) так, щоб показник Ω був якнайменший.

Розв'язок, очевидно, може бути визначений лише до довільного сталого множника.

Ми тут маємо на увазі довести, що за певних умов [2]) шуканий розв'язок одержується, коли застосувати в деякий спосіб шлях інтерполяції, як це зараз буде з'ясовано.

[1]) Далі ми натрапимо на деякі випадки, коли розглядатимемо функцію $f(x, y, \ldots, v)$ не неперервну (але обмежену) або область апроксимації незамкнену. Тоді під $R_{\max}$, $R_{\min}$ розумітимемо відповідні точні границі відношення R.

[2]) Умова (A) нижче.

Визначмо параметри $\alpha_0, \alpha_1, \alpha_2, \ldots$ у виразі φ_1 чи φ_2 так, щоб він точно дорівнював самій функції f у всіх „вершках [1])“ відповідної області апроксимації (T) або (P). „Вершками“ „понадтетраедра“ (T) [відп. „понадпаралелепіпеда“ (P)] ми тут звемо ті системи значень $x, y, \ldots, v$, які одержуємо, замінюючи в системі нерівностей (T) [відп. (P)] n знаків нерівності на знаки рівності. Поставлена вимога приводить негайно до системи лінійних рівнянь, з яких треба визначити шукані параметри. У випадку апроксимації за допомогою виразу φ_1 в області (T) система рівнянь буде, очевидно, така:

$$\left.\begin{array}{ll}
\alpha_0 & = f(0,0,0,\ldots,0) \\
\alpha_0 + \alpha_1 c & = f(c,0,0,\ldots,0) \\
\alpha_0 + \alpha_1 c + \alpha_2 c\lambda & = f(c, c\lambda, 0,\ldots,0) \\
\alpha_0 + \alpha_1 c + \alpha_2 c\lambda + \alpha_3 c\lambda\mu & = f(c, c\lambda, c\lambda\mu, \ldots, 0) \\
\cdots\cdots & \cdots\cdots\cdots\cdots \\
\alpha_0 + \alpha_1 c + \alpha_2 c\lambda + \alpha_3 c\lambda\mu + \cdots \alpha_n c\lambda\mu\nu\ldots\rho & = f(c, c\lambda, c\lambda\mu, \ldots, c\lambda\mu\nu\ldots\rho)
\end{array}\right\} \quad (5)$$

В разі ж апроксимації за допомогою виразу φ_2 в області (P) система рівнянь буде така:

$$\begin{array}{ll}
\alpha_0 & = f(0,0,0,\ldots,0,0,0) \\
\alpha_0 + \alpha_1 a & = f(a,0,0,\ldots,0,0,0) \\
\alpha_0 + \alpha_2 b & = f(0,b,0,\ldots,0,0,0) \\
\alpha_0 + \alpha_3 c & = f(0,0,c,\ldots,0,0,0) \\
\cdots\cdots & \cdots\cdots\cdots\cdots \\
\alpha_0 + \alpha_n h & = f(0,0,0,\ldots,0,0,h) \\
\alpha_0 + \alpha_1 a + \alpha_2 b + \alpha_{12} ab & = f(a,b,0,\ldots,0,0,0) \\
\alpha_0 + \alpha_1 a + \alpha_3 c + \alpha_{13} ac & = f(a,0,c,\ldots,0,0,0) \\
\cdots\cdots & \cdots\cdots\cdots\cdots \\
\alpha_0 + \alpha_{n-1} g + \alpha_n h + \alpha_{n-1,n} gh & = f(0,0,0,\ldots,0,g,h) \\
\alpha_0 + \alpha_1 a + \alpha_2 b + \alpha_3 c + \alpha_{12} ab + & \\
\quad + \alpha_{13} ac + \alpha_{23} bc + \alpha_{123} abc & = f(a,b,c,\ldots,0,0,0) \\
\cdots\cdots & \cdots\cdots\cdots\cdots \\
\alpha_0 + \alpha_{n-2} e + \alpha_{n-1} g + \alpha_n h + & \\
\quad + \alpha_{n-2,\,n-1} eg + \alpha_{n-2,n} eh + & \\
\quad + \alpha_{n-1,\,n} gh + \alpha_{n-2,n-1,n} egh & = f(0,0,0,\ldots,e,g,h) \\
\cdots\cdots & \cdots\cdots\cdots\cdots \\
\cdots\cdots & \cdots\cdots\cdots\cdots\ [2])
\end{array} \quad (6)$$

[1]) Ми тут і в дальшому вживаємо звичайної геометричної мови, говорячи про „точки“ $(x, y, \ldots, v)$, „вершки“, „простір“ (евклідів) тощо виключно для більш стислого вислову понять арифметичних.

[2]) Тут усіх рівнянь $1 + C_n^1 + C_n^2 + C_n^3 + \cdots + C_n^n = 2^n$ і таке саме є число параметрів.

Із цих систем шукані значення параметрів, очевидно, безпосередньо визначаються один по одному.

Знайденим так значенням параметрів дамо позначення

$$\alpha_0 = k_0,\ \alpha_1 = k_1,\ \alpha_2 = k_2, \ldots, \alpha_n = k_n \tag{7}$$

абож

$$\alpha_0 = k_0,\ \alpha_1 = k_1,\ \alpha_2 = k_2, \ldots, \alpha_{12\ldots n} = k_{12\ldots n} \tag{8}$$

залежно від типу наближеного представлення.

Покажімо тепер, що визначені так значення параметрів дають розв'язок поставленої нами попередньої задачі, якщо справджується така умова:

Різниця $f(x, y, \ldots, v) - \varphi(x, y, \ldots, v; k_0, k_1 k_2, \ldots)$ має сталий знак у всій області апроксимації, тобто в області (T) або (P) відповідно	**Умова *A***

Інакше кажучи, треба довести таку теорему:

Теорема. *За умови (A) наближене представлення типу* (1) *або* (2), *при значеннях параметрів* $\alpha_0 = k_0,\ \alpha_1 = k_1,\ \alpha_2 = k_2, \ldots,$ *має екстремальну властивість, яка виражається нерівністю*

$$\Omega(\alpha_0 \alpha_1, \alpha_2, \ldots) \geqslant \Omega(k_0, k_1, k_2, \ldots), \tag{9}$$

де $k_0, k_1, k_2, \ldots$ позначають визначені з рівнянь (5) або (6) (відповідно до типу набл. предст.) значення параметрів, а $\alpha_0, \alpha_1, \alpha_2, \ldots$ позначають довільні значення тих самих параметрів.

Доводячи теорему, розгляньмо спершу випадок наближеного представлення типу (1).

Впровадьмо в розгляд для простору $(x, y, z, \ldots, v)$ певне узагальнення барицентричних координат Мебіуса відносно понадтетраедра (T), підпорядкувавши ще ці узагальнені координати тій умові, щоб сума їх дорівнювала одиниці. Занумеруймо вершки понадтетраедра (T) числами від 1 до $n+1$ в тому порядку, в якому ці вершки з'являються в рівняннях системи (5). Нехай будуть $\xi, \eta, \zeta, \ldots, \upsilon, \omega$ звичайні („декартові") координати якоїсь точки простору $(x, y, z, \ldots, u, v)$. Позначмо через $m_1, m_2, \ldots, m_{n+1}$ узагальнені барицентричні координати тієї самої точки відносно понадтетраедра (T). Згідно з вищесказаним, ці координати повинні справджувати такі умови:

$$\begin{cases} \sum\limits_{i=1}^{n+1} m_i = 1 & (10) \\ \sum\limits_{i=1}^{n+1} m_i x_i = \xi,\ \sum\limits_{i=1}^{n+1} m_i y_i = \eta, \ldots,\ \sum\limits_{i=1}^{n+1} m_i v_i = \omega, & (11) \end{cases}$$

де індекс i при координатах x, y і т. д. позначає нумер вершка з координатами $(x_i, y_i, \ldots, v_i)$.

Напишімо систему (10)—(11) у розгорненому виді:

$$\left.\begin{aligned} m_1+m_2+m_3+m_4+\cdots+m_{n+1} &= 1\\ m_2c+m_3c+m_4c+\cdots+m_{n+1}c &= \xi\\ m_3c\lambda+m_4c\lambda+\cdots+m_{n+1}c\lambda &= \eta\\ m_4c\lambda\mu+\cdots+m_{n+1}c\lambda\mu &= \zeta\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ m_{n+1}c\lambda\mu\nu\ldots\pi\rho &= \omega \end{aligned}\right\} \qquad (12)$$

Варто зауважити, що детермінант цієї системи лінійних рівнянь такий самий, як у системі (5), з заміною лише рядків на стовпці. Це могло б a priori забезпечити існування та єдиність розв'язку. Проте невідомі визначаються тут безпосередньо.

Одержуємо:

$$\left.\begin{aligned} &m_1=1-\frac{\xi}{c},\quad m_2=\frac{\xi}{c}-\frac{\eta}{c\lambda},\quad m_3=\frac{\eta}{c\lambda}-\frac{\zeta}{c\lambda\mu},\ldots\\ &\ldots,\ m_n=\frac{\upsilon}{c\lambda\mu\ldots\pi}-\frac{\omega}{c\lambda\mu\ldots\pi\rho},\quad m_{n+1}=\frac{\omega}{c\lambda\mu\ldots\pi\rho} \end{aligned}\right\} \qquad (13)$$

Бачимо, що всі узагальнені барицентричні координати є невід'ємні, коли точка $(\xi, \eta, \ldots, \omega)$ належить до області (T). Вони додатні, коли точка $(\xi, \eta, \ldots, \omega)$ лежить всередині області (T) у вузькому розумінні, тобто коли всі співвідношення системи нерівностей (T) справджуються без знака $=$.

Як істотно важливий для дальшого наслідок з рівнянь (10) та (11) відзначмо таке співвідношення:

$$\sum_{i=1}^{n+1} m_i\varphi_1(x_i, y_i, \ldots, v_i;\ \alpha_0, \alpha_1, \ldots, \alpha_n) = \varphi_1(\xi, \eta, \ldots, \omega;\ \alpha_0, \alpha_1, \ldots, \alpha_n) \qquad (14)$$

Щоб довести нашу теорему для розглядуваного випадку наближеного представлення типу (1), припустімо для конкретності, що сталий знак різниці $f(x, y, \ldots, v) - \varphi_1(x, y, \ldots, v; k_0, k_1, \ldots, k_{n+1})$ в області (T) є, наприклад, від'ємний, і нехай буде $(\xi, \eta, \ldots, \omega)$ та точка (або одна з тих точок) області (T), де відношення $\varphi_1 : f = R$ досягає свойого maximum-у [1]). Запроваджуючи скорочені позначення

$$\left.\begin{aligned} &R(x_i, y_i, \ldots, v_i;\ \alpha_0, \alpha_1, \ldots, \alpha_n) = R_i\\ &\qquad (i = 1, 2, \ldots, n+1)\\ &R(\xi, \eta, \ldots, \omega;\ \alpha_0, \alpha_1, \ldots, \alpha_n) = R_0 \end{aligned}\right\} \qquad (15)$$

[1]) Нагадаймо, що обидві функції f та φ ми тут припустили для конкретності додатними. В тому разі, коли сталий знак $f-\varphi$ в області (T) був би додатний, довелось би, очевидно, замість maximum-у розглянути minimum з відповідною модифікацією дальших міркувань.

де $R = \varphi_1 : f$, а $\alpha_0, \alpha_1, \ldots, \alpha_n$ позначають будьяку допущенну систему значень параметрів, утворімо далі таку лінійну форму, з невід'ємними коефіцієнтами, від $n+1$ відношень $\frac{R_1}{R_0}, \frac{R_2}{R_0}, \ldots, \frac{R_{n+1}}{R_0}$, що їх, при даних $\xi, \eta, \ldots, \omega$, можна розглядати як функції від $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_n$:

$$\left.\begin{aligned} \Psi(\alpha_0, \alpha_1, \ldots, \alpha_n) = m_1 f(x_1, y_1, \ldots, v_1) \frac{R_1}{R_0} \\ + m_2 f(x_2, y_2, \ldots, v_2) \frac{R_2}{R_0} \\ + \cdots\cdots\cdots \\ + m_{n+1} f(x_{n+1}, \ldots, v_{n+1}) \frac{R_{n+1}}{R_0} \end{aligned}\right\} \tag{16}$$

Беручи на увагу спершу (15) та (3), а потім (14), легко бачити, що

$$\begin{aligned} \Psi(\alpha_0, \alpha_1, \ldots, \alpha_n) = f(\xi, \eta, \ldots, \omega) \cdot \frac{\sum_{i=1}^{n+1} m_i \varphi_1(x_i, y_i, \ldots, v_i; \alpha_0, \alpha_1, \ldots, \alpha_n)}{\varphi_1(\xi, \eta, \ldots, \omega; \alpha_0, \alpha_1, \ldots, \alpha_n)} = \\ = f(\xi, \eta, \ldots, \omega) = \text{const}, \end{aligned} \tag{17}$$

тобто величина форми Ψ не міняється при зміні параметрів $\alpha_0, \alpha_1, \ldots, \alpha_n$.

Але цей висновок негайно переконує нас у правдивості доводжуваної теореми.

Справді бо, з одного боку, якщо для якоїнебудь системи значень параметрів $\alpha_0, \alpha_1, \ldots, \alpha_n$ ті відношення $\frac{R_i}{R_0}$, які мають при собі коефіцієнти $m_i f(x_i, y_i, \ldots, v_i)$ відмінні від нуля, не всі є рівні між собою, то, позначаючи через i_0 нумер того вершка, якому відповідає найменше з цих відношень, матимемо, очевидно, зважаючи на (17) та (16):

$$f(\xi, \eta, \ldots, \omega) > \frac{R_{i0}}{R_0} \cdot \sum_{i=1}^{n+1} m_i f(x_i, y_i, \ldots, v_i) \tag{18}$$

Звідси ж випливає для такої системи значень параметрів $\alpha_0\ \alpha_1, \ldots, \alpha_n$ нерівність

$$\Omega(\alpha_0, \alpha_1, \ldots, \alpha_n) \geqslant \frac{R_0}{R_{i0}} > \frac{\sum_{i=1}^{n+1} m_i f(x_i y_i, \ldots, v_i)}{f(\xi, \eta, \ldots, \omega)} \tag{19}$$

З другого ж боку, для системи значень $\alpha_0 = k_0, \ldots, \alpha_n = k_n$ маємо:

$$\frac{R_1}{R_0} = \frac{R_2}{R_0} = \cdots = \frac{R_{n+1}}{R_0} = \frac{1}{R_0} = \frac{1}{\Omega(k_0, k_1, \ldots, k_n)} \tag{20}$$

і, значить, знов же, на підставі (17) та (16),

$$f(\xi, \eta, \ldots, \omega) = \frac{1}{\Omega(k_0, k_1, \ldots, k_n)} \sum_{i=1}^{n+1} m_i f(x_i y_i, \ldots, v_i)$$

звідки

$$\Omega(k_0, k_1, \ldots, k_n) = \frac{\sum_{i=1}^{n+1} m_i f(x_i y_i, \ldots, v_i)}{f(\xi, \eta, \ldots, \omega)} \qquad 21$$

Звідси безпосередньо маємо такий висновок.

Показник коливання відносної точності $\Omega(\alpha_0, \alpha_1, \ldots, \alpha_n)$ не може бути менший від $\Omega(k_0, k_1, \ldots, k_n)$ ні при яких значеннях параметрів; він лише в тому разі може не перевищувати величину $\Omega(k_0, k_1, \ldots, k_n)$, коли при даній системі значень $\alpha_0, \alpha_1, \ldots, \alpha_n$ відношення R_i має однакові значення для всіх значень індекса i, для яких $m_i \neq 0$, при чому, нагадаємо це, $m_1, m_2, \ldots, m_{n+1}$ є узагальнені барицентричні координати тієї точки $(\xi, \eta, \ldots, \omega)$ області (T), в якій відношення $\frac{\varphi_1(x, y, \ldots, v; k_0, \ldots, k_n)}{f(x, y, \ldots, v)}$ досягає maximum-у (при $f - \varphi < 0$).

Тут можливі два випадки:

1. Точка $(\xi, \eta, \ldots, \omega)$ області (T), в якій відношення $R(x, y, \ldots, v; k_0, \ldots, k_n)$ досягає свойого maximum-у (при $f - \varphi < 0$), знаходиться всередині (у вузькому розумінні) області (T). Тоді узагальнені барицентричні координати $m_1, m_2, \ldots, m_{n+1}$ усі відмінні від нуля (додатні). В цьому разі, базуючись на останньому висновку, легко бачити, що система значень параметрів $\alpha_0 = k_0$, $\alpha_1 = k_1, \ldots, \alpha_n = k_n$ не тільки дає розв'язок поставленої задачі, але цей розв'язок є й єдиний — до довільного сталого множника. Справді, вимога $R_1 = R_2 = \cdots = R_{n+1} = \sigma$, де σ означає довільну сталу, дає для визначення параметрів $\alpha_0, \alpha_1, \ldots, \alpha_n$ систему лінійних рівнянь, що відрізняється від системи (5) лише наявністю довільного сталого множника σ перед правими сторонами, отже для параметрів $\alpha_0, \alpha_1, \ldots, \alpha_n$ одержуємо значення, які відрізняються від $k_0, k_1, \ldots, k_n$ лише тим самим довільним сталим множником σ.

2. Точка $(\xi, \eta, \ldots, \omega)$ знаходиться на границі області (T), тобто при $x = \xi$, $y = \eta, \ldots$, $v = \omega$ в системі нерівностей (T) має місце принаймні один знак $=$. Тоді попередній висновок знову показує, що система значень $\alpha_0 = k_0, \ldots, \alpha_n = k_n$ і тут дає розв'язок поставленої задачі, але цей розв'язок уже взагалі не буде єдиний, бо з системи рівнянь $R_1 = R_2 = \cdots = R_{n+1} = \sigma$ в цьому разі випадають деякі рівняння (випадають ті R_i, що відповідають $m_i = 0$), і ця система вже не визначає параметрів $\alpha_0, \alpha_1, \ldots, \alpha_n$ однозначно. Справжні межі, в яких можна брати значення тих параметрів, які в розглядуваній неповній системі рівнянь залишаються довільні, встановлюється в кожному окремому випадку відповідним дослідженням [1]). Ці межі напевне включають ту систему значень параметрів, яка визначається з рівнянь (5).

Проте, коли множина точок $(\xi, \eta, \ldots, \omega)$, в яких відношення R набирає свого значення, містить більше ніж одну точку, єдиність розв'язку є, на підставі

Переходячи до доводження нашої теореми для випадку наближеного представлення типу (2), треба насамперед умовитись про належне узагальнення барицентричних координат при трактуванні цього випадку. Щоб ясніше виявити суть справи, візьмімо спершу окремий випадок $n=2$, $\varphi_2(x, y; \alpha_0, \ldots, \alpha_{12}) = \alpha_0 + \alpha_1 x + \alpha_2 y + \alpha_{12} xy$. Область (P) незалежних змінних тут є $o \leqslant x \leqslant a$, $o \leqslant y \leqslant b$. Розглядаючи добуток xy як вираз третьої координати $z = xy$, запровадьмо для точок (x, y, xy) відповідної поверхні гіперболічного параболоїда узагальнені барицентричні координати m_1, m_2, m_3, m_4, зосереджуючи чотири маси у вершках (o, o, o), (a, o, o), (o, b, o), (a, b, ab) того просторового чотирикутника на поверхні гіперболічного параболоїда, що проектується (за напрямком OZ) на область (P) площі xoy, та вимагаючи при цьому, щоб сума чотирьох мас дорівнювала одиниці. Ми легко знайдемо значення

$$m_1 = \left(1 - \frac{x}{a}\right)\left(1 - \frac{y}{b}\right); \quad m_2 = \frac{x}{a}\left(1 - \frac{y}{b}\right)$$

$$m_3 = \left(1 - \frac{x}{a}\right)\frac{y}{b}; \qquad m_4 = \frac{x}{a} \cdot \frac{y}{b}$$

Усі чотири координати є невід'ємні, якщо точка (x, y) належить до області (P). Зауважмо, що можливість здійснити таку систему узагальнених барицентричних координат у даному разі була зумовлена a priori відповідною лінійчастою структурою поверхні гіперболічного параболоїда; при чому істотне є й те, що розглянутий просторовий чотирикутник є обмежений чотирма відрізками прямолінійних твірних.

Повертаючись до загального випадку наближеного представлення типу (2) при довільному числі змінних $x, y, \ldots, v$, нам доведеться впровадити в розгляд n-мірну точкову множину в просторі $2^n - 1$ вимірів, розглядаючи не лише $x, y, z, \ldots, t, u, v$, але й добутки $xy, xz, \ldots, uv; xyz, \ldots, tuv$; .....; $xyz \ldots tuv$ як вирази окремих координат точки на зазначеній n-мірній точковій множині. Існує, очевидно, одно-однозначна відповідність між точками цієї точкової множини та точками n-мірного простору $(x, y, \ldots, v)$. На цю відповідність нам доведеться не раз посилатись на найближчих сторінках.

Занумерувавши числами від 1 до 2^n „вершки" області (P) в тому порядку, в якому вони з'являються в системі рівнянь (6), впровадьмо на нашій n-мірній точковій множині $2^n - 1$-мірного простору узагальнені барицентричні координати $m_1, m_2, \ldots, m_{2^n}$, зосереджуючи маси в точках згаданої точкової множини, що відповідають „вершкам" області (P), та вимагаючи при цьому, щоб сума усіх 2^n мас дорівнювала одиниці. Якщо $(\xi, \eta, \zeta, \ldots, \upsilon, \omega)$

міркування, напевне забезпечена, якщо хоч одна з цих точок лежить в
вузькому розумінні) області (T)
далі задач (§ 10; § 14, n° n° 2,3) буде видко, що єдиність розв'язку може
коли жодна точка множини $(\xi, \eta, \ldots, \omega)$ не лежить всередині
(у вузькому розумінні) області апроксимації (див. § 10).

означає будьяку точку простору $(x, y, z, \ldots, u, v)$, то узагальнені барицентричні координати відповідної точки розглядуваної точкової множини визначатимуться із системи лінійних рівнянь:

$$\left.\begin{array}{l} \sum m_i = 1 \\ \sum m_i x_i = \xi, \ldots\ldots\ldots, \sum m_i v_i = \omega \\ \sum m_i x_i y_i = \xi\eta, \ldots\ldots, \sum m_i u_i v_i = \upsilon\omega \\ \sum m_i x_i y_i z_i = \xi\eta\zeta, \ldots, \sum m_i t_i u_i v_i = \tau\upsilon\omega \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum m_i x_i y_i z_i \ldots t_i u_i v_i = \xi\eta\zeta \ldots \tau\upsilon\omega \end{array}\right\} \quad 22)$$

Знак $\sum$ тут скрізь поширюється на значення індекса від $i = 1$ до $i = 2^n$

Важливо зауважити, що детермінант системи (22) є тотожний з детермінантом системи (6), з заміною лише рядків на стовпці. Це забезпечує a priori і існування розв'язку для системи (22), і його єдиність.

Як істотно важливий для дальшого наслідок з рівнянь (22) відмітьмо таке співвідношення, аналогічне до одержаного вище (14):

$$\sum_{i=1}^{2^n} m_i \varphi_2(x_i, y_i, \ldots, v_i; \alpha_0, \alpha_1, \ldots, \alpha_{123\ldots n} = \varphi_2(\xi, \eta, \ldots, \omega; \alpha_0, \alpha_1, \ldots, \alpha_{123\ldots n}) \quad (23)$$

Щодо явних виразів узагальнених барицентричних координат, то вони складаються за таким законом: „вершкові“ № i області (P), що має, скажімо, координати $(a, o, o, \ldots, e, g, o)$, відповідає вміщена у відповідному „вершку“ n-мірної точкової множини $2^n - 1$-мірного простору маса

$$m_i = \frac{\xi}{a}\left(1 - \frac{\eta}{b}\right)\left(1 - \frac{\zeta}{c}\right) \cdots \frac{\tau}{e}\frac{\upsilon}{g}\left(1 - \frac{\omega}{h}\right), \quad (24)$$

якщо точка згаданої точкової множини, для якої визначаються узагальнені барицентричні координати, відповідає, як і раніше, точці $(\xi, \eta, \ldots, \omega)$ простору $(x, y, \ldots, v)$. Тут кожній координаті [„декартовій“, у просторі $(x, y, \ldots, v)$], що дорівнює нулеві в даному вершку, відповідає множник двочленний, а кожній координаті відмінній від нуля — множник одночленний. Вираз (24) показує, що всі узагальнені барицентричні координати є невід'ємні для кожної точки розглядуваної точкової множини, яка відповідає будьякій точці області (P); вони всі додатні, коли точка $(\xi, \eta, \ldots, \omega)$ знаходиться всередині (у вузькому розумінні) області (P), але деякі з них дорівнюють нулеві, коли ця точка знаходиться на границі області (P).

Правдивість виразів (24) ми можемо встановити шляхом повної індукції, перевіривши її спершу для випадку $n = 1$. Позначмо символом

$$S(\xi, \eta, \zeta, \ldots, \tau, \upsilon, \omega; a, b, c, \ldots, e, g, h)$$

суму всіх 2^n виразів типу (24), де n означає, звичайно, і кількість $\xi, \eta, \ldots, \omega$, і кількість величин $a, b, \ldots, h$. Насамперед

вити, переходячи від n до $n+1$, чинність при довільному n першого рівняння системи (22), що його можна було б записати у виді

$$S(\xi, \eta, \zeta, \ldots, \tau, \upsilon, \omega; a, b, c, \ldots, e, g, h) = 1$$

Справді, коли це рівняння вже встановлено для $n = 2$, матимемо негайно

$$S(\xi, \eta, \zeta; a, b, c) = \frac{\zeta}{c} S(\xi, \eta; a, b) + \left(1 - \frac{\zeta}{c}\right) S(\xi, \eta; a, b) = S(\xi, \eta; a, b) = 1$$

і т. д. Після ж того, коли це вже встановлено для довільного n, дуже легко встановити й чинність будьякого іншого з рівнянь (22), бо наприклад,

$$\sum_{i=1}^{2} m_i x_i y_i = ab \cdot \frac{\xi}{a} \cdot \frac{\eta}{b} \cdot S(\zeta, \ldots, \tau, \upsilon, \omega; c, \ldots, e, g, h) = \xi\eta$$

і т. д.

Але, коли все це встановлено, повне доведення теореми для випадку наближеного представлення типу (2) завершується зовсім так само, як у випадку наближеного представлення типу (1): тільки співвідношення (23) в цьому міркуванні заступить місце співвідношення (14) попереднього міркування. Висновки щодо єдиності чи то многозначності розв'язку в даному випадку цілком аналогічні до попереднього.

Увага. В конкретних застосуваннях функція $f(x, y, \ldots, v)$, що підлягає апроксимації в області (T) чи то (P), може допускати певне продовження за межі даної області. Знайдене нами, за умови (A), розв'язання попередньої задачі зберігає ще силу взагалі й при деякому поширенні області апроксимації. Цим ми хочемо сказати, що екстремальна властивість (9) при знайдених значеннях параметрів $k_0, k_1, \ldots$, зберігає ще силу і в тому разі, коли Ω означатиме показник коливання відносної точності знайденого наближеного представлення в області трохи поширеній: треба лише, щоб при цьому поширенні було враховано умови сталості знаків величин $\frac{\varphi}{f} > 0$ та $|f| - |\varphi|$ при $\varphi \equiv \varphi(x, y, \ldots, v; k_0, k_1, \ldots)$ і, крім того, щоб максимум виразу $|\lg R(x, y, \ldots, v; k_0, k_1, k_2, \ldots)|$ в поширеній області не перевищував величину $|\lg R(\xi, \eta, \ldots, w; k_0\ k_1, \ldots)$ Останні умови визначають межі можливого поширення області апроксимації без порушення екстремальної властивості (9) та при незмінному значенні $\Omega(k_0, k_1, \ldots, k_n)$.

§ 2. Остаточне сформулювання задач апроксимації, що відповідають критеріям відносної та абсолютної похибок. Їхнє роз'вязання за умови (A)

Задача 1. Для даної функції $f(x, y, \ldots, v)$ визначити найкраще, за критерієм відносної похибки, наближене представлення типу (1) в області (T). Інакше кажучи, треба визначити параметри $\alpha_0, \alpha_1, \ldots, \alpha_n$ за умовою

$$\max_{(x, y, \ldots, v) \subset (T)} \frac{f(x, y, \ldots, v) - \varphi_1(x, y, \ldots, v; \alpha_0, \alpha_1, \ldots, \alpha_n)}{f(x, y, \ldots, v)} = \min, \qquad (25)$$

де $\subset$ є вживаний в теорії множин знак включення. Щодо природи функції f ми залишаємо покищо в силі, при сформулюванні цієї і дальших задач, ті самі припущення, які було сформульовано на початку § 1.

Задача II. Цілком аналогічна до першої, але стосується наближеного представлення типу (2) в області (P).

Задача III. Визначити для даної функції $f(x, y, \ldots, v)$ найкраще, за критерієм абсолютної похибки, наближене представлення типу (1) в області (T). Тут параметри $\alpha_0, \alpha_1, \ldots, \alpha_n$ треба, зрозуміла річ, визначити за умовою

$$\max_{(x,y,\ldots,v)\subset(T)} f(x,y,\ldots,v) - \varphi_1(x,y,\ldots,v;\ \alpha_0,\alpha_1,\ldots,\alpha_n) = \min \tag{26}$$

Задача IV. Цілком аналогічна до останньої, але стосується наближеного представлення типу (2) в області (P).

Ми зараз покажемо, що значення параметрів $\alpha_0 = k_0$, $\alpha_1 = k_1, \ldots$, визначені з системи рівнянь (5) або (6) відповідно до типу наближеного представлення, дозволяють негайно знайти розв'язання сформульованих тут чотирьох задач за умови (A); для цього залишається лише у випадку задачі I або II помножити $\varphi(x, y, \ldots, v; k_0, k_1, k_2, \ldots)$ на належного сталого множника, а у випадку задачі III або IV — додати до $\varphi(x, y, \ldots, v; k_0, k_1, k_2, \ldots)$ належну аддитивну сталу.

Почнімо з розгляду задач I та II. Відносна похибка при заміні функції f на її наближене представлення φ_1 або φ_2 має взагалі своїм виразом

$$\frac{f-\varphi}{f} = 1 - R(x, y, \ldots, v; \alpha_0, \alpha_1, \ldots) \tag{27}$$

Позначаючи через $\varepsilon = \varepsilon(\alpha_0, \alpha_1, \ldots)$ maximum абсолютного значення виразу (27) при зафіксованих $\alpha_0, \alpha_1, \ldots$ та при змінних $x, y, \ldots v$ у відповідній області апроксимації, маємо, очевидно, визначити параметри α_0, α_1, за умовою

$$\varepsilon = \varepsilon(\alpha_0, \alpha_1, \alpha_2, \ldots) = \min \tag{25'}$$

Але ε дорівнює, очевидно, більшому з двох абсолютних значень $|1 - R_{\max}|$ та $|1 - R_{\min}|$, зберігаючи позначення попереднього параграфа.

Доведімо насамперед нерівність

$$\varepsilon \geqslant 1 - \frac{2}{1+\Omega} \tag{28}$$

яка є справедлива за будьяких значень параметрів $\alpha_0, \alpha_1, \alpha_2, \ldots$, при чому Ω означає, як і досі, відношення $R_{\max} : R_{\min}$ за даної системи значень $\alpha_0, \alpha_1, \alpha_2, \ldots$ Справді, при порівненні величин $R_{\max} - 1$ та $1 - R_{\min}$, більша з яких дає значення ε, можливі лише три випадки:

1 $R_{\max} - 1 = 1 - R_{\min} = \varepsilon$. Замінюючи $R_{\max}$ на $\Omega\, R_{\min}$, дістанемо $R_{\min} = \frac{2}{1+\Omega}$ і водночас $\varepsilon = 1 - R_{\min} = 1 - \frac{2}{1+\Omega}$

2°. $R_{max}-1>1-R_{min}$; $\varepsilon=R_{max}-1$ Тоді $\Omega . R_{min}-1>1-R_{min}$, $R_{min}>\frac{2}{1+\Omega}$; $\varepsilon=\Omega . R_{min}-1>\frac{2\Omega}{1+\Omega}-1=\frac{\Omega-1}{1+\Omega}$, тобто $\varepsilon>1-\frac{2}{1+\Omega}$

3°. $R_{max}-1<1-R_{min}$; $\varepsilon=1-R_{min}$. Тут $\Omega . R_{min}-1<1-R_{min}$, $R_{min}<\frac{2}{1+\Omega}$; $\varepsilon>1-\frac{2}{1+\Omega}$.

Бачимо, що співвідношення (28) має місце у всіх трьох випадках, причому знак = маємо лише тоді, коли задовольняється рівність

$$R_{max}-1=1-R_{min} \tag{29}$$

Зробімо тепер два зауваження. Поперше, $\Omega(\alpha_0, \alpha_1, \ldots)$ не міняється, очевидно, коли всі параметри помножити на одного і того самого множника σ. Згідно з (9), найменше значення Ω є

$$\Omega(k_0\sigma, k_1\sigma, k_2\sigma, \ldots)=\Omega(k_0, k_1, k_2, \ldots) \tag{30}$$

Подруге, якщо для будьякої системи значень параметрів $\alpha_0, \alpha_1, \alpha_2$, рівність (29) не має місця, можна знизити максимальну відносну похибку $\varepsilon(\alpha_0, \alpha_1, \alpha_2 \ldots)$, множачи всі параметри $\alpha_0, \alpha_1, \alpha_2, \ldots$ на спільного множника σ, дібравши цього множника за умовою $\sigma R_{max}-1=1-\sigma R_{min}$, звідки

$$\sigma=\frac{2}{R_{max}+R_{min}} \tag{31}$$

де R_{max} та R_{min} означають, звичайно, найбільше та найменше значення відношення (3) в області (T) чи (P), за даних значень параметрів $\alpha_0, \alpha_1, \alpha_2, \ldots$

Після всього сказаного вже безпосередньо ясно, як добути розв'язок задачі I або II за умови (A) попереднього параграфа:

Визначивши $k_0, k_1, k_2, \ldots$ як значення параметрів, що задовольняють систему лінійних рівнянь (5) або (6) відповідно до типу наближеного представлення, маємо шуканий розв'язок у виді

$$\varphi(x, y, \ldots, v; k_0\sigma, k_1\sigma, k_2\sigma, \ldots)=\sigma . \varphi(x, y, \ldots, v; k_0, k_1, k_2, \ldots), \tag{32}$$

де множник σ має значення

$$\sigma=\frac{2}{1+R(\xi, \eta, \ldots, \omega; k_0, k_1, k_2,} \tag{33}$$

Максимальна відносна похибка знайденого найкращого наближеного представлення, що її ми позначимо через ε_0, матиме, очевидно, своїм виразом

$$\varepsilon_0=1-\frac{2}{1+\Omega(k_0, k_1, \ldots)}=\left|1-\frac{2}{1+R(\xi, \eta, \ldots, \omega; k_0, k_1}\right. \tag{34}$$

Нагадаймо, що $(\xi, \eta, \ldots, \omega)$ означає ту точку області (T) чи (P), в якій $R(x, y, \ldots, v; k_0, k_1, k_2, \ldots)$ досягає свого екстремального значення (максимального при $f-\varphi\leqslant 0$ і мінімального — в противному випадку). Єди-

ність розв'язку є забезпечена лише в тому разі, коли точка $(\xi, \eta, \ldots, \omega)$ лежить всередині області апроксимації (T) чи (P) у вузькому розумінні (тобто виключаючи границю області).

У попередньому ми скрізь дотримувалися звичайно вживаної форми критерія відносної похибки, беручи за вираз відносної похибки $\frac{f-\varphi}{f}$. Але в дійсності має істотні переваги, надто в обчисленнях із значним наближенням, інше означення відносної похибки, як відношення похибки абсолютної до меншої з двох величин — точної або наближеної. Якщо змінимо відповідно сформулювання задач I — II, шукаючи найкращого наближеного представлення в розумінні мінімізації максимума відносної похибки в останньому її означенні, множника σ' доведеться уже обчисляти за новою формулою:

$$\sigma' = \frac{1}{\sqrt{R(\xi, \eta, \ldots, \omega; k_0, k_1, k_2, \ldots)}} \tag{35}$$

А максимальна відносна похибка найкращого наближення, яку ми тут позначимо через ε_0', буде

$$\varepsilon_0' = \sqrt{\Omega(k_0, k_1, k_2, \ldots)} - 1, \tag{36}$$

при чому $\Omega(k_0, k_1, k_2, \ldots)$ дорівнює $R(\xi, \eta, \ldots, \omega; k_0, k_1, \ldots)$, якщо $R > 1$, і оберненому числу в противному разі.

Переходячи далі до розв'язання задач III — IV, тобто до розшукання найкращих наближених представлень за критерієм абсолютної похибки, нам доведеться mutatis mutandis використати ряд попередніх міркувань, починаючи з § 1, при чому, замість величин R, Ω, σ, ε, ε_0, тут ми запровадимо відповідні деякі величини D, A, τ, δ, δ_0. Ми покладемо насамперед

$$\varphi(x, y, \ldots, v; \alpha_0, \alpha_1, \ldots) - f(x, y, \ldots, v) = D(x, y, \ldots, v; \alpha_0, \alpha_1, \ldots) \tag{37}$$

Позначаючи далі через $D_{\max}$ та $D_{\min}$ відповідно найбільше та найменше значення величини (37) при зафіксованих параметрах $\alpha_0, \alpha_1, \ldots$ та при змінних $x, y, \ldots, v$ в області апроксимації (T) чи (P), покладімо

$$D_{\max} - D_{\min} = A(\alpha_0, \alpha_1, \alpha_2, \ldots) \tag{38}$$

Цю величину A, за аналогією з величиною Ω, позначатимемо як показник коливання абсолютної точності наближеного представлення функції f за допомогою виразу φ_1 або φ_2 в області (T) або (P) відповідно.

Нехай будуть, як і раніше, $k_0, k_1, k_2, \ldots$ значення параметрів $\alpha_0, \alpha_1, \alpha_2$, визначені з системи (5) або (6), залежно від типу шуканого наближеного представлення. Ми насамперед доведемо, що значення $k_0, k_1, k_2, \ldots$ при умові (A) попереднього параграфа, поруч з екстремальною властивістю (9), мають ще таку екстремальну властивість:

$$A(\alpha_0, \alpha_1, \alpha_2, \ldots) \geqslant A(k_0, k_1, k_2, \ldots) \tag{39}$$

Припустімо для конкретності, що розглядається наближене представлення типу (1), і нехай буде, наприклад, знов $f-\varphi \leqslant 0$ [1]) в області (T). Якщо $D(x, y, \ldots, v; k_0, k_1, k_2, \ldots,)$ набирає свого максимального значення в точці $(\xi', \eta', \ldots, \omega')$ області (T), введімо, аналогічно до (15), такі позначення:

$$\left.\begin{aligned} D(x_i, y_i, \ldots, v_i;\ \alpha_0, \alpha_1, \ldots, \alpha_n) &= D_i \\ (i = 1, 2, \ldots, n+1) & \\ D(\xi', \eta', \ldots, \omega';\ \alpha_0, \alpha_1, \ldots, \alpha_n) &= D_0 \end{aligned}\right| \qquad (40)$$

і, користуючись узагальненими барицентричними координатами $m'_1, m'_2, \ldots, m'_{n+1}$ точки $(\xi', \eta', \ldots, \omega')$, утворімо в даному разі таку лінійну форму, з невід'ємними коефіцієнтами, від різниць $D_1 - D_0$, $D_2 - D_0, \ldots, D_{n+1} - D_0$, що їх, за даних $\xi', \eta', \ldots, \omega'$, можна розглядати як функції від $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_n$:

$$\chi(\alpha_0, \alpha_1, \ldots, \alpha_n) = m'_1(D_1 - D_0) + m'_2(D_2 - D_0) + \cdots + m'_{n+1}(D_{n+1} - D_0) \qquad (41)$$

Взявши до уваги (40), (37), (14) та (10), одержимо:

$$\left.\begin{aligned} \chi(\alpha_0, \alpha_1, \ldots, \alpha_n) &= \sum m'_i \varphi_1(x_i, y_i, \ldots, v_i;\ \alpha_0, \alpha_1, \ldots, \alpha_n) - \sum m'_i f(x_i y_i, \ldots, v_i) \\ &- \sum m'_i \varphi_1(\xi', \eta', \ldots, \omega';\ \alpha_0, \alpha_1, \ldots, \alpha_n) + \sum m'_i f(\xi', \eta', \ldots, \omega') = \\ &= f(\xi', \eta', \ldots, \omega') - \sum m'_i f(x_i, y_i, \ldots, v_i) = \text{const} \end{aligned}\right| \qquad (42)$$

(де знак $\sum$ скрізь поширюється на значення індекса від $i = 1$ до $i = n+1$), тобто величина форми χ залишається незмінна при зміні параметрів a_0, $\alpha_1, \ldots, \alpha_n$.

Звідси прийдемо до висновку, що, з одного боку,

$$A(k_0, k_1, \ldots, k_n) = \sum_{i=1}^{n+1} m'_i f(x_i, y_i, \ldots, v_i) - f(\xi', \eta', \ldots, \omega'), \qquad (43)$$

а з другого боку

$$A(\alpha_0, \alpha_1, \ldots, \alpha_n) \geqslant \sum_{i=1}^{n+1} m'_i f(x_i, y_i, \ldots, v_i) - f(\xi', \eta', \ldots, \omega') \qquad (44)$$

за будьякої іншої системи значень параметрів. Це завершує доведення екстремальної властивості (39) системи значень $k_0, k_1, \ldots, k_n$. Якщо точка $(\xi', \eta', \ldots, \omega')$ лежить всередині області (T) у вузькому розумінні (тобто виключаючи границю), система значень параметрів $\alpha_0 = k_0$, $\alpha_1 = k_1, \ldots, \alpha_n = k_n$ дає, за умови (A), єдиний — до аддитивної сталої — розв'язок задачі: знайти для даної функції $f(x, y, \ldots, v)$ таке наближене представлення типу (1), щоб показник $A(\alpha_0, \alpha_1, \ldots, \alpha_n)$ коливання абсолютної точності в області (T) був щонайменший.

Коли ж точка $(\xi', \eta', \ldots, \omega')$ знаходиться на границі області (T), знайдений розв'язок зберігає силу, але взагалі він уже не буде єдиний. Зауважмо, що точка $(\xi', \eta', \ldots, \omega')$, яка заступає в цих висновках точку $(\xi, \eta, \ldots, \omega)$ аналогічної задачі попереднього параграфа, що відповідала критерієві відносної похибки, є взагалі відмінна від тієї точки.

Щоб дістати розв'язання поставленої на початку цього параграфа задачі III, нам залишається скористатися ще з аддитивної сталої

$$\tau = -\frac{1}{2} D(\xi', \eta', \ldots, \omega'; k_0, k_1, \ldots, k_n) \tag{45}$$

і ми можемо самий розв'язок подати у виді

$$\varphi_1(x, y, \ldots, v; k_0, k_1, \ldots, k_n) + \tau = (\tau + k_0) + k_1 x + k_2 y + \cdots + k_n v, \tag{46}$$

при чому максимальна абсолютна похибка δ_0 цього найкращого наближеного представлення матиме своїм виразом

$$\delta_0 = \frac{1}{2} A(k_0, k_1, \ldots, k_n) = \frac{1}{2} \left| D(\xi', \eta', \ldots, \omega'; k_0, k_1, \ldots, k_n) \right| \tag{47}$$

В тому випадку, коли точка $(\xi', \eta', \ldots, \omega')$ знаходиться на границі області (T), розв'язок (46), як було уже сказано, взагалі не є єдиний. Відповідна неповна система лінійних рівнянь дає і тут засіб, щоб визначити в с і розв'язки, при чому межі для значень параметрів, які в згаданій неповній системі рівнянь залишаються довільні, визначаються в кожному окремому випадку відповідним дослідженням [1]).

Все сказане про розв'язання задачі III допускає цілком очевидне (на підставі попереднього викладу) узагальнення для задачі IV, отже зупинятись окремо на цій задачі немає потреби.

В світлі всіх попередніх міркувань з'ясовується умовне значення деяких припущень щодо природи функції $f(x, y, z, \ldots, v)$, зроблених на початку цього розділу (§ 1). В той час як умова (A) параграфа 1 мала і с т о т н е значення для наших міркувань [2]), цього не можна сказати про деякі припущення початку § 1, які було запроваджено власне для конкретності та для деякого спрощення міркувань. Зовсім не є необхідна насамперед умова н е п е р е р в н о с т і функції $f(x, y, \ldots, v)$ в області апроксимації. Якби ця функція не була неперервна, але була обмежена в області апроксимації, міркування наші залишилися б в силі; ми лише заступили б функцію f у наших міркуваннях деякими — п і в н е п е р е р в н и м и функціями, як ми це робитимемо в розділах VI — VII при трактуванні інших задач. Далі, для функції f можна було б шукати наближене представлення типу (1) або (2) не у всій області (T) чи (P), а в довільній замкненій або незамкненій, можливо — дискретній, частині згаданої області, аби тільки вона включала всі „вершки" області. Нарешті, для наших міркувань по суті не вимагалося, щоб функція f неодмінно була цілком (однозначно) означена в розглядуваній області: в розділах VI — VIII будуть з'ясовані, в аспекті інших задач, ті модифікації поставлення проблеми, які дозволяють трактувати випадки, коли цієї умови не додержано.

[1]) Пор. відповідну виноску в § 1, а також задачі § 14, n° n° 2, 3.

[2]) Проте, в розділі IV, § 11, побачимо, як за допомогою певної модифікації попередніх міркувань одержані результати можуть бути застосовані і до певних випадків, коли умова (A) не має місця.

§ 3. Деякі прикладання теорії потенціалу. Зв'язок із задачею Чебишова — Граве

При розв'язанні основної екстремум-задачі параграфа 1 або аналогічної до неї, яка відповідає критерієві абсолютної похибки, вирішальне значення має взагалі встановлення наявності умови (A), і це водночас є момент найважчий. Зведення цього моменту на дослідження цілої сукупності екстремальних значень величини $R(x, y, \ldots, v; k_0, k_1, \ldots)$ чи $D(x, y, \ldots, v; k_0, k_1, \ldots)$ в області апроксимації та на її границі звичайними засобами аналізу та алгебри взагалі не можна розглядати як практичне розв'язання питання, бо якраз це дослідження може являти досить серйозні труднощі, і навіть у тих випадках, коли воно є фактично здійсненне, має велику вагу, з різних поглядів, можливість встановити наявність чи відсутність умови (A) незалежно і раніше від нього. Найбільш істотну користь тут можна часто здобути, як ми це зараз з'ясуємо, з належного застосування властивостей функцій субгармонічних та супергармонічних. Вчення про ці функції, як спеціальна галузь теорії потенціалу, систематично почало розвиватись після праць Frederic Riesz-a [1]); розвивалося воно переважно з погляду застосувань до теорії функцій комплексного змінного. Тут ми матимемо приклад застосування іншого роду, яке стосується певних питань апроксимації функцій від кількох дійсних змінних.

Нехай буде $F(x, y, \ldots, v)$ функція однозначна та неперервна [2]) в певній неперервній області $\mathfrak{A}$, що її границю ми позначимо через $\mathfrak{B}$. Через $\mathfrak{A}'$ та $\mathfrak{B}'$ позначатимемо довільну неперервну частину області $\mathfrak{A}$ та її границю, при чому нам досить буде мати на увазі такі частини області $\mathfrak{A}$, до яких застосованість принципу Dirichlet є безсумнівна. Через $\Phi(x, y, \ldots, v)$ позначатимемо гармонічну функцію від $x, y, \ldots, v$, тобто таку, регулярну в певній області, функцію, що справджує в кожній точці цієї області узагальнене Лапласове рівняння

$$\nabla^2\Phi = \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \cdots + \frac{\partial^2 \Phi}{\partial v^2} = 0 \tag{48}$$

Ми скажемо, що дана функція $F(x, y, \ldots, v)$ є субгармонічна в області $\mathfrak{A}$, якщо для кожної області $\mathfrak{A}'$, внутрішньої щодо $\mathfrak{A}$ разом із своєю границею $\mathfrak{B}'$, та для кожної функції $\Phi(x, y, \ldots, v)$ гармонічної в $\mathfrak{A}'$, неперервної на її границі $\mathfrak{B}'$ такої, що

$$F - \Phi \leqslant 0 \tag{49}$$

цій границі, така сама нерівність задовольняється всередині області $\mathfrak{A}'$.

[1]) Riesz. „Sur les fonct. subharmoniques et leur rapport à la théorie du potentiel (Acta t. 48, p. 329 — 343; t. 54, p. 321 — 360). Див. також P. Montel. Sur les fonct. convexes et les fonct. sousharmoniques (Journal de Mathém. pures et appliquées, 1928.

[2]) На загальнішому випадку субгармонічної функції — півнеперервної зверху функції — півнеперервної знизу ми тут не спинимось.

Міняючи в нерівності (49) знак $\leqslant$ на $\geqslant$, ми встановлюємо аналогічно поняття функції супергармонічної в області $\mathfrak{A}$.

У випадку одного незалежного змінного назви субгармонічної та супергармонічної функцій замінюються звичайно назвами конвексної (convexe) та конкавної (concave) функцій. Конвексна функція має графік вгнутий угору, конкавна — вгнутий униз.

При довільній кількості незалежних змінних є відома проста ознака для функції субгармонічної, якщо ця функція має частинні похідні другого порядку в кожній точці області. А саме, за цієї умови функція $F(x, y, \ldots, v)$ є субгармонічна в області $\mathfrak{A}$, якщо

$$\nabla^2 F = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + \cdots + \frac{\partial^2 F}{\partial v^2} \geqslant 0 \tag{50}$$

в кожній точці області $\mathfrak{A}$, і супергармонічна в разі нерівності протилежного знака $\leqslant$.

В наближених представленнях типів (1) та (2) вираз φ_1 або φ_2 є, очевидно, функція гармонічна. Більш того, вона не тільки є гармонічна в самій області (T) чи (P), але й на границі розглядуваної області в тому спеціальному розумінні, що в кожній з тих областей [1]) зменшеного числа вимірів $(n-1, n-2, \ldots, 2, 1)$, які одержуємо, замінюючи в системах нерівностей (T) чи (P) деяку кількість знаків нерівності на знак $=$, лапласіан функції φ відносно тих змінних, які залишаються незалежними, тотожно дорівнює нулеві. Сукупність оцих властивостей функції φ_1 чи φ_2 ми зазначатимемо, кажучи, що функція є кратно-гармонічна в даній області (T) чи (P) та на її границі.

Відповідно до останнього означення ми впровадимо також поняття функції кратно-субгармонічної та функції кратно-супергармонічної в області (T) чи (P) та на її границі. В цікавому для нас випадку, коли функція $F(x, y, \ldots, v)$ має частинні похідні другого порядку, вона буде, наприклад, кратно-субгармонічна, якщо її лапласіан $\nabla^2 F$ є невід'ємний як у самій області (T) чи (P), так і в усіх окремих областях на її границі, розуміючи кожного разу лапласіан як суму відповідних частинних похідних, взятих по кожному з незалежних змінних даної області.

Повертаючись до нашої задачі апроксимації, із самих понять функцій субгармонічної (супергармонічної) та кратно-субгармонічної (кратно-супергармонічної) випливає така достатня ознака щодо існування умови (A) параграфа 1.

Умова (A) напевне має місце, якщо дана функція $f(x, y, \ldots, v)$ є кратно-субгармонічна абож кратно-супергармонічна в області (T) чи (P) та на її границі.

[1]) Не важко підлічити кількість цих областей: на границі області (T) їх буде $C_{n+1}^1 + C_{n+1}^2 + \cdots + C_{n+1}^{n-1} = 2^{n+1} - n - 3$ (напр., при $n = 3$ маємо 4 грані тетраедра та 6 рубів); на границі області (P) їх буде $C_n^1 \cdot 2 + C_n^2 \cdot 2^2 + \cdots + C_n^{n-1} \cdot 2^{n-1} = 3^n - 2^n - 1$ (напр., при $n = 3$ матимемо 6 граней та 12 рубів паралелепіпеда).

Наприклад, узагальнена задача Понселе у випадку чотирьох квадратів сходить до найкращого лінійного наближеного представлення функції $f(y,z,t)=\sqrt{1+y^2+z^2}+t^2$ в області $\lambda \geqslant y \geqslant \frac{z}{\mu} \geqslant \frac{t}{\mu\nu} \geqslant 0$, що її ми позначимо (T_3). Тут ми маємо, поперше, в самій області (T_3):

$$\nabla^2 f=\frac{3+2(y^2+z^2+t^2)}{(\sqrt{1+y^2+z^2+t^2})^3}>0$$

Далі на границі (T_3) маємо чотири області двох вимірів та шість областей одномірних. Відповідні вирази, на які перетворюється f в цих областях, такі:

$$f_1=\sqrt{1+\lambda^2+z^2+t^2};\quad f_2=\sqrt{1+y^2(1+\mu^2)+t^2};$$

$$f_3=\sqrt{1+y^2+z^2(1+\nu^2)};\ f_4=\sqrt{1+y^2+z^2};$$

$$f_5=\sqrt{1+\lambda^2(1+\mu^2)+t^2};\quad f_6=\sqrt{1+\lambda^2+z^2(1+\nu^2)};\ f_7=\sqrt{1+\lambda^2+z^2};$$

$$=\sqrt{1+y^2(1+\mu^2+\mu^2\nu^2)};\ f_9=\sqrt{1+y^2(1+\mu^2)};\quad f_{10}=\sqrt{1+y^2}$$

Дуже легко пересвідчитись, що лапласіан кожного з цих десятьох виразів є додатний у відповідній області, звідки маємо висновок, що умова (A) в даній задачі має місце [1]).

[1]) Додамо деякі міркування, які можуть висвітлити до певної міри питання щодо ступеня загальності запровадженого тут поняття кратно-субгармонічної та кратно-супергармонічної функцій. Досить буде зупинитись, для конкретності, на випадку трьох незалежних змінних x, y, z.

Насамперед зауважмо, що в тому разі, коли всі шість других частинних похідних від $f(x,y,z)$ є одного знака, напр., додатні в точці (x_0,y_0,z_0) та неперервні поблизу неї, цю точку можна обійняти (до того ж на безліч способів) на бажання тетраедром або паралелепіпедом, типу подібного до (T) чи (P), але паралельно зміщеного, — так щоб дана функція $f(x,y,z)$ була кратно-субгармонічна в розглядуваному тетраедрі чи паралелепіпеді та на його границі. Справді, можна буде взяти перший-ліпший тетраедр типу $\varepsilon \geqslant (x-a) \geqslant \frac{y-b}{\lambda} \geqslant \frac{z-c}{\lambda\mu} \geqslant 0$ або паралелепіпед типу $0 \leqslant x-a \leqslant \varepsilon$, $0 \leqslant y-b \leqslant \eta$, $0 \leqslant z-c \leqslant \delta$ за такими умовами, щоб вони, поперше, обіймали точку $(x_0, y_0\ z_0)$ і, подруге, містилися цілком усередині тієї області, де зазначені 6 частинних похідних залишаються додатні. Легко бачити, що лапласіан $\nabla^2 f$ буде додатний не лише в самій області розглядуваного тетраедра чи паралелепіпеда, але й на всіх окремих гранях та рубах границі — в тому розумінні, яке ми з'ясували вище. Досить для цього пересвідчитись, що будуть додатні всі частинні похідні другого порядку, взяті від функції f по тих змінних, що їх розглядаємо як незалежні на даній грані чи рубі. Але це справді є так, бо кожна з цих других частинних похідних виражається, очевидно, лінійною формою від $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2}, \frac{\partial^2 f}{\partial z^2}, \frac{\partial^2 f}{\partial y \partial z}, \frac{\partial^2 f}{\partial z \partial x}$, $\frac{\partial^2}{\partial x \partial y}$, що її коефіцієнти є додатні сталі або нулі (але не всі одночасно нулі).

Розгляньмо тепер, зберігаючи умову щодо неперервності тих самих частинних похідних, загальніший випадок, коли 6 других частинних похідних від $f(x,y,z)$ не всі мають однаковий знак, але лапласіан $\nabla^2 f=\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2}$ є відмінний від нуля в точці

2°. Як дальше застосування теорії потенціалу до розглядуваного кола питань, ми зараз покажемо, що екстремальні властивості (9) та (39) системи значень параметрів $\alpha_0 = k_0$, $\alpha_1, = k_1$, $\alpha_2 = k_2, \ldots$ за умови (A) можна вивести, виходячи лише з того факту, що вирази (1) та (2) є кратно-гармонічні в означеному вище розумінні.

Розгляньмо для конкретності випадок апроксимації за критерієм абсолютної похибки, і нехай іде мова про наближене представлення типу (1), при чому різниця $f - \varphi$ є, напр., від'ємна в (T). За самим способом визначення параметрів $k_0, k_1, \ldots, k_n$ вираз (46) має такі властивості:

а) у всіх „вершках" області (T) абсолютна похибка є однакова й дорівнює $-\tau$;

б) у точці $(\xi', \eta, \ldots, \omega')$, абсолютна похибка дорівнює $+\tau$.

Як би існувала інша система значень параметрів $(\alpha'_0, \alpha'_1, \ldots, \alpha'_n)$, що дає краще наближення, ніж система значень $(\tau + k_0), k_1, \ldots, k_n$, ми повинні були б мати водночас:

$$\varphi_1(x_i, y_i, \ldots, v_i; \alpha'_0, \alpha'_1, \ldots, \alpha'_n) - \varphi_1(x, y_i, \ldots, v_i; \tau + k_0, k_1, \ldots, k_n) > 0 \quad (51)$$
$$(i = 1, 2, \ldots, n+1)$$

$$\varphi_1(\xi', \eta', \ldots, \omega'; \alpha'_0, \alpha'_1, \ldots, \alpha'_n) - \varphi_1(\xi', \eta', \ldots, \omega'; \tau + k_0, k_1, \ldots, k_n) < 0 \quad (52)$$

Але легко бачити, що нерівність (52) є несумісна з (51) через ту добре відому властивість гармонічних функцій, через яку така функція не може бути додатна скрізь на границі області і водночас від'ємна бодай в одній точці всередині області. За нерівностями (51) кратно-гармонічна функція

$$\varphi_1(x, y, \ldots, v; \alpha'_0, \alpha'_1, \ldots, \alpha'_n) - \varphi_1(x, y, \ldots, v; \tau + k_0, k_1, \ldots, k_n)$$

є додатна у всіх „вершках" області (T). Отже вона є додатна насам-

(x_0, y_0, z_0). Тоді бодай один із трьох доданків, напр., $\frac{\partial^2 f}{\partial x^2} \neq 0$ в цій точці. А коли це так, то, як легко бачити, досить буде застосувати перетворення до косокутних координат за формулами

$$\begin{cases} x - x_0 = x' \cos\alpha_1 + y' \cos\alpha_2 + z' \cos\alpha_3 \\ y - y_0 = x' \cos\beta_1 + y' \cos\beta_2 + z' \cos\beta_3 \\ z - z_0 = x' \cos\gamma_1 + y' \cos\gamma_2 + z' \cos\gamma_3 \end{cases}$$

взявши напрямки осей косокутної системи координат так, щоб $\cos\alpha_1$, $\cos\alpha_2$, $\cos\alpha_3$ були досить близькі до $+1$, щоб напевне забезпечити однаковість знаків шістьох других частинних похідних від даної функції по нових аргументах x', y', z', при $x' = y' = z' = 0$. Отже, приходимо до випадку вже розглянутого, лише в цьому разі паралелепіпед чи тетраедр уже не будуть взагалі прямокутні.

Якщо побажаємо поставити ці міркування у зв'язок із задачею наближеного представлення функції $f(x, y, z)$ за допомогою виразу виду (1) або (2), слід зауважити, що коли повернутися від x', y', z' до первісних змінних, то тип наближеного представлення при зворотному перетворенні змінних зберігається лише у випадку виразу (1). Проте, у випадку наближеного представлення типу (2), як легко бачити, можна обійтися без зміни напрямків координатних осей, якщо бодай три частинні похідні $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2}, \frac{\partial^2 f}{\partial z^2}$ мають однаковий знак у точці (x_0, y_0, z_0).

перед у всіх одномірних областях („рубах" понадтетраедра), що добуваються заміною в системі нерівностей (T) $n-1$ знаків $\geqslant$ на знак $=$ Звідси ж матимемо наступний висновок, що та сама функція є додатна також у всіх двомірних областях, що їх дістаємо, замінюючи в системі нерівностей (T) $n-2$ знаки $\geqslant$ на знак $=$ і т. д., поки не вичерпаємо всі області трьох, чотирьох..., $n-1$ вимірів і нарешті й саму область T Це ж і призведе до явної суперечності з нерівністю (52). Отже є доведене, що вираз (46) дає справді найкраще наближення. Певна річ, що заразом ми довели й екстремальну властивість (39). При деякій модифікації міркування тим самим шляхом можемо дослідити й питання щодо єдиності розв'язку.

3°. З'ясований новий шлях дослідження розглядуваних питань відкриває можливість іншого ряду узагальнень, відмінних від тих, які можуть бути зв'язані із застосуванням узагальнених барицентричних координат Наприклад, у випадку двох незалежних змінних аналогічне трактування допускає така задача: неперервну функцію $f(x, y)$ задано в області замкненого многокутника; треба визначити для неї найкраще наближене представлення в тій самій області за допомогою кратно-гармонічної (в області многокутника та на всіх окремих лінійних частинах його границі) функції.

Умова, аналогічна до умови (A), тут напевне має місце, якщо сама дана функція $f(x, y)$ є кратно-субгармонічна або кратно-супергармонічна в тій самій області та на її границі.

Тут шукане наближене представлення все ще залежить від скінченного числа параметрів. Але ми вже безпосередньо підійшли до цікавого факту, що полягає в існуванні інтимного зв'язку поміж останніми нашими міркуваннями та визначною задачею Чебишова — Граве щодо побудови найкращої конформної географічної карти для даної країни — задачі, де шукана функція залежить від добору нескінченної множини значень на контурі області.

В короткому повідомленні „Sur la construction des cartes géographiques" зачитаному 18 січня 1856 року в засіданні Петербурзької Академії Наук, П. Л. Чебишов подав без доведення таку теорему:

„За Лягранжовими позначеннями (Nouveaux Mémoires de l'Acad. de Berlin. Année 1779) масштаб карти виражається так:

$$m = \frac{\sqrt{F'(u+ti) \cdot F'_1(u-ti)}}{\frac{2}{e^u + e^{-u}}}$$

$$\lg m = \frac{1}{2} \lg F'(u+ti) + \frac{1}{2} \lg F'_1(u-ti) - \lg \frac{2}{e^u + e^{-u}}$$

де додатна частина, складена з довільних функцій, є не що інше як інтеграл рівняння

$$\frac{\partial^2 U}{\partial u^2} + \frac{\partial^2 U}{\partial t^2} = 0$$

„Отже відхили масштабу залежать від змін функції $\lg \frac{2}{e^u + e^{-u}}$ та інтегралу цього рівняння.

„Але, зважаючи на визначні властивості цього рівняння, можна показати, що мінімум відхилу його інтегралу від функції $\lg \frac{2}{e^u + e^{-u}}$ в області, обмеженій будьякою кривою, може мати місце лише в тому разі, коли різниця

$$U - \lg \frac{2}{e^u + e^{-u}}$$

має стале значення для точок цієї кривої"

Близько 40 років ця важлива теорема залищалася недоведена. Лише 1894 року Д. О. Граве знайшов доведення цієї теореми, вміщене в його докторській дисертації „Об основных задачах математической теории построения географических карт" (СПБ, 1896). Пізніше він спростив своє перше доведення, надавши йому неперевершеної елегантності, і водночас поширив його на випадок конформної карти будьякої поверхні із сталим знаком кривини [1]).

Суть питання П. Л. Чебишова полягала в доведенні такого положення: гармонічна функція $\Phi(x, y)$, яка найменше відхиляється (за критерієм абсолютного відхилу) від

$$f(x, y) = \lg \frac{2}{e^x + e^{-x}} = -\lg \operatorname{ch} x$$

в якійсь області, обмеженій даним неперервним контуром, відрізняється лише на аддитивну сталу від тієї гармонічної функції, яка на контурі дорівнює даній функції $f(x, y) = -\lg \operatorname{ch} x$.

Коли зауважимо, що $\nabla^2 f(x, y) = \frac{d^2}{dx^2}(-\lg \operatorname{ch} x) = -\frac{1}{\operatorname{ch}^2 x} < 0$, то повне доведення дістанемо негайно, якщо в наших останніх міркуваннях, спеціалізованих щодо випадку двох незалежних змінних, замінити трикутник (T) на довільний неперервний контур, розглядаючи водночас, замість попередніх кратно-гармонічної функції $\varphi_1(x, y)$ та кратно-супергармонічної функції $f(x, y)$, функції гармонічну та супергармонічну в звичайному розумінні. Інтерполяційний вираз $\varphi_1(x, y; k_0, k_1, k_2)$ тут заступає гармонічна функція, що відповідає задачі Dirichlet для даного неперервного контура за даних на ньому значень, що збігаються із значеннями функції $-\lg \operatorname{ch} x$ на цьому контурі.

Повинно бути відзначено, що Д. О. Граве знайшов повне розв'язання питання П. Л. Чебишова задовго до того, як розвинулася сучасна теорія суб- та супергармонічних функцій.

Трактуючи аналогічне питання щодо апроксимації функції суб- або супергармонічної за допомогою функції гармонічної в розрізі теорії

потенціалу, ми приходимо до відповідних екстремальних властивостей еквівалентної Грінової верстви у випадку розподілених всередині області мас сталого знака або, при іншому сформулюванні, до екстремальних властивостей balayage — операції H. Poincaré.

До цих останніх результатів я прийшов цілком незалежно від задачі Чебишова — Граве і раніше того, як я мав щастя ознайомитися з цитованими працями мойого глибокошановного вчителя Д. О. Граве.

РОЗДІЛ II

ЗАДАЧА ЛІНІЙНОГО НАБЛИЖЕНОГО ПРЕДСТАВЛЕННЯ ФУНКЦІЙ, ЩО ВИЗНАЧАЮТЬСЯ РІВНЯННЯМИ ВИДУ $w^\rho = x^\rho + y^\rho + z^\rho + \cdots + v^\rho$. ПРИКЛАД ЗАСТОСУВАННЯ ДО НАБЛИЖЕНОГО ВИЗНАЧЕННЯ РОЗВ'ЯЗКУ ДИФЕРЕНЦІАЛЬНОГО РІВНЯННЯ

В застосуваннях часто трапляються випадки змінних величин, зв'язаних співвідношеннями виду

$$w^\rho = x^\rho + y^\rho + z^\rho + \cdots + v^\rho \tag{53}$$

де сталий показник ρ може мати різноманітні дійсні значення. За основний об'єкт даного розділу править застосування результатів розділу I до задачі визначення найкращого наближеного представлення типу (1) для величини w як функції від $x, y, \ldots, v$.

§ 4. Випадок $\rho = 2$ та його розв'язання за критерієм відносної похибки

В цьому частинному випадку, що є найближче поширення задачі Понселе, ми шукаємо наближену формулу виду

$$\sqrt{x^2 + y^2 + z^2 + \cdots + v^2} \quad \alpha + \beta y + \gamma z + \cdots + \varkappa v \tag{54}$$

в області

$$x \geqslant \frac{y}{\lambda} \geqslant \frac{z}{\lambda\mu} \geqslant \frac{t}{\lambda\mu\nu} \geqslant \ldots \geqslant \frac{u}{\lambda\mu\nu\ldots\pi} \geqslant \frac{v}{\lambda\mu\nu\ldots\pi\rho} \geqslant 0, \tag{55}$$

яку, зважаючи на однорідність шуканої формули (54), можна замінити, не порушуючи загальності, на скінченну область:

$$1 = x \geqslant \frac{y}{\lambda} \geqslant \frac{z}{\lambda\mu} \geqslant \frac{t}{\lambda\mu\nu} \geqslant \ldots \geqslant \frac{v}{\lambda\mu\nu\ldots\pi\rho} \geqslant 0 \tag{56}$$

Інакше кажучи, маємо тут визначити для функції

$$f(y, z, t, \ldots, u, \quad = \sqrt{1 + y^2 + z^2 + t^2 + \cdots + u^2 + v^2} \tag{57}$$

наближене представлення виду

$$\varphi_1(y, z, t, \ldots, u, v; \quad , \beta, \gamma, \ldots, \varkappa) = \alpha + \beta y + \gamma z + \cdots + \varkappa v \tag{58}$$

в області типу (T):

$$\lambda \geqslant y \geqslant \frac{z}{\mu} \geqslant \frac{t}{\mu\nu} \geqslant \ldots \geqslant \frac{v}{\mu\nu\ldots\pi\rho} \geqslant 0 \tag{59}$$

Функція (57) є кратно-субгармонічна в області (59) та на її границі, як це вже було відзначено в попередньому розділі (§ 3, n° 1) [1]). Отже спосіб попереднього розділу тут напевне можна застосувати. Система (5) тут призводить до таких значень параметрів:

$$\left.\begin{aligned}
&\alpha = 1 = k\\
&\beta = \frac{\sqrt{1+\lambda^2}-1}{\lambda} = \frac{\lambda}{\sqrt{1+\lambda^2}+1} = l\\
&\gamma = \frac{\sqrt{1+\lambda^2+\lambda^2\mu^2}-\sqrt{1+\lambda^2}}{\lambda\mu} = \frac{\lambda\mu}{\sqrt{1+\lambda^2+\lambda^2\mu^2}+\sqrt{1+\lambda^2}} = m\\
&\varkappa = \frac{\sqrt{1+\lambda^2+\lambda^2\mu^2+\cdots+\lambda^2\mu^2\nu^2\cdots\pi^2\rho^2}-\sqrt{1+\lambda^2+\lambda^2\mu^2+\cdots+\lambda^2\mu^2\nu^2\ldots\pi^2}}{\lambda\mu\nu\ldots\pi\rho} =\\
&= \frac{\lambda\mu\nu\ldots\pi\rho}{\sqrt{1+\lambda^2+\lambda^2\mu^2+\cdots+\lambda^2\mu^2\nu^2\cdots\pi^2\rho^2}+\sqrt{1+\lambda^2+\lambda^2\mu^2+\cdots+\lambda^2\mu^2\nu^2\ldots\pi^2}} = r
\end{aligned}\right\} \quad 60$$

В даному разі

$$R = \frac{k+ly+mz+\cdots+rv}{\sqrt{1+y^2+z^2+\cdots+v^2}} \quad 61$$

Перетворивши цей вираз до виду

$$R = \frac{k\cdot k+ly+mz+\cdots+rv}{\sqrt{k^2+l^2+m^2+\cdots+r^2}\sqrt{k^2+y^2+z^2+\cdots+v^2}}\cdot\sqrt{k^2+l^2+\cdots+r^2}, \quad 62$$

помічаємо, що перший множник на правій стороні має такі властивості:

1°. Він дорівнює 1 при $y=l, z=m, \ldots, v=r$.

2°. При кожній іншій системі значень $y, z, \ldots, v$ він є менший від 1; в цьому легко пересвідчитись, піднісши цей чинник до квадрата та застосувавши до чисельника узагальнену тотожність Ейлера — Лягранжа.

Звідси робимо висновок, що абсолютно-найбільше значення виразу (62), або (61) при зафіксованих $k, l, m, \ldots, r$, під якими розуміємо значення (60) та при змінних $y, z, \ldots, v$, має місце при $y=l$, $z=m, \ldots, v=r$. Інакше кажучи, точка, що відповідає $(\xi, \eta, \ldots, \omega)$ попереднього розділу, тут, як легко бачити, і є $(\eta, \zeta, \ldots, \omega) \equiv (l, m, \ldots, r)$. Справді, формули (60), коли взяти в них вирази з радикалами в знаменниках, безпосередньо показують, що задовольняються нерівності

$$\lambda > l > \frac{m}{\mu} > \ldots > \frac{r}{\mu\nu\ldots\pi\rho} > 0 \quad 63$$

точка $(l, m, \ldots, r)$, в якій $R = R_{\max} = \sqrt{k^2+l^2+m^2+\cdots+r^2}$ всередині, у вузькому розумінні, області

є забезпечена єдиність розв'язку. Самий розв'язок, тобто шукане найкраще наближене представлення виду (58) для функції (57) в області (59), можемо подати[1]) у виді

$$\sigma\varphi_1(y, z, \ldots, v;\ k, l, m, \ldots, r) = \sigma(k + ly + mz + \cdots + rv) \tag{64}$$

при

$$\sigma = \frac{2}{1 + \sqrt{k^2 + l^2 + m^2 + \cdots + r^2}}; \tag{65}$$

при цьому максимальна відносна похибка є

$$\varepsilon_0 = 1 - \sigma = \frac{\sqrt{k^2 + l^2 + \cdots + r^2} - 1}{\sqrt{k^2 + l^2 + \cdots + r^2} + 1}, \tag{66}$$

де $k, l, m, \ldots, r$ мають значення (60).

Повертаючись до області (55), матимемо остаточно шукану формулу найкращого наближення:

$$\sqrt{x^2 + y^2 + z^2 + \cdots + v^2} = \sigma(kx + ly + mz + \cdots + rv) \tag{67}$$

з максимальною відносною похибкою (66) в області (55), при чому параметри $k, l, m, \ldots, r$ та множник σ визначаються формулами (60) та (65).

В окремому випадку, коли $\lambda = \mu = \nu = \cdots = \pi = \rho = 1$, матимемо область

$$x \geqslant y \geqslant z \geqslant \ldots \geqslant v \geqslant 0 \tag{68}$$

Тоді, позначаючи через N число аргументів, матимемо такі значення параметрів:

$$k = 1, l = \sqrt{2} - 1, m = \sqrt{3} - \sqrt{2}, \ldots, r = \sqrt{N} - \sqrt{N-1} \tag{69}$$

Ми тут випишемо відповідні формули найкращого наближення для $N = 2, 3, 4, \ldots, 10$. Замість $x, y, z, \ldots$ пишемо в цих формулах $x_1, x_2, x_3, \ldots$, при чому область апроксимації скрізь є

$$x_1 \geqslant x_2 \geqslant x_3 \geqslant \ldots \geqslant x_{10} \geqslant 0 \tag{70}$$

$$\left.\begin{aligned}
&\sqrt{x_1^2 + x_2^2} && \approx 0{,}960x_1 + 0{,}398x_2 \text{ з точн. до } 4\% \\
&\sqrt{x_1^2 + x_2^2 + x_3^2} && \approx 0{,}940x_1 + 0{,}389x_2 + 0{,}299x_3 \text{ з точн. до } 6\% \\
&\sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2} && \approx 0{,}926x_1 + 0{,}384x_2 + 0{,}294x_3 + 0{,}248x_4 \\
& && \qquad \text{з точн. до } 7\tfrac{1}{2}\% \\
&\sqrt{x_1^2 + x_2^2 + \cdots + x_5^2} && \approx 0{,}916x_1 + 0{,}379x_2 + 0{,}291x_3 + 0{,}245x_4 \\
& && \qquad + 0{,}216x_5 \text{ з точн. до } 8\tfrac{1}{2}\% \\
&\sqrt{x_1^2 + x_2^2 + \cdots + x_6^2} && \approx 0{,}908x_1 + 0{,}376x_2 + 0{,}289x_3 + 0{,}243x_4 \\
& && \qquad + 0{,}214x_5 + 0{,}194x_6 \text{ з точн. до } 9\tfrac{1}{2}\%
\end{aligned}\right\} \tag{71}$$

[1] Згідно з (32), (33), (34).

$$\sqrt{x_1^2+x_2^2+\cdots+x_7^2} \approx 0{,}902x_1+0{,}373x_2+0{,}287x_3+0{,}242x_4 + 0{,}213x_5+0{,}192x_6+0{,}177x_7 \text{ з точн. до } 10\%$$

$$\sqrt{x_1^2+x_2^2+\cdots+x_8^2} \approx 0{,}896x_1+0{,}371x_2+0{,}285x_3+0{,}240x_4 + 0{,}212x_5+0{,}191x_6+0{,}176x_7+0{,}164x_8 \text{ з точн. до } 10\tfrac{1}{2}\%$$

$$\sqrt{x_1^2+x_2^2+\cdots+x_9^2} \approx 0{,}891x_1+0{,}369x_2+0{,}283x_3+0{,}239x_4 + 0.210x_5+0{,}190x_6+0{,}175x_7+0{,}163x_8 + 0{,}153x_9 \text{ з точн. до } 11\% \qquad 71)$$

$$\sqrt{x_1^2+x_2^2+\cdots+x_{10}^2} \approx 0{,}887x_1+0{,}367x_2+0{,}282x_3+0{,}238x_4 + 0{,}209x_5+0{,}189x_6+0{,}174x_7+0{,}162x_8 + 0{,}152x_9+0{,}144x_{10} \text{ з точн. до } 11\tfrac{1}{2}\%$$

При збільшенні числа квадратів до ∞ максимум похибки йде до 100%, проте щодалі, то повільніше наростає гранична похибка: вона досягає 20% лише за числа членів щось із 45. Справжні ж похибки, зрозуміла річ, лише в рідких випадках досягатимуть зазначених граничних значень, залишаючись взагалі меншими.

При $\lambda, \mu, \nu, \ldots < 1$ граничні похибки будуть, звичайно, менші. Напр., при $\lambda=\mu=\nu=\cdots=\frac{1}{2}$ гранична похибка залишається $<2\%$, навіть коли число квадратів росте безмежно. Навпаки, при $\lambda, \mu, \nu, \ldots > 1$ похибки будуть більші. Напр., при $\lambda=\mu=\infty$ дістаємо формули найкращого наближення

$$\left.\begin{aligned} &\sqrt{x^2+y^2} \approx 0{,}828\,(x+y) \text{ з точн. до } 17{,}2\% \\ &\sqrt{x^2+y^2+z^2} \approx 0{,}732\,(x+y+z) \text{ з точн. до } 26{,}8\% \end{aligned}\right\} \qquad 72)$$

для цілком довільних додатних значень x, y, z.

Подамо ще, нарешті, для випадків $N=2, 3$ при $x_1 \geqslant x_2 \geqslant x_3 \geqslant 0$ формули найкращого наближення, складені теж за критерієм відносної похибки, але при іншому її означенні, якому відповідають формули (35) та (36):

$$\left.\begin{aligned} &\sqrt{x_1^2+x_2^2} \approx 0{,}961x_1+0{,}398x_2 \text{ з точн. до } 4{,}1\% \\ &\sqrt{x_1^2+x_2^2+x_3^2} \approx 0{,}942x_1+0{,}390x_2+0{,}299x_3 \text{ з точн. до } 6{,}2\% \end{aligned}\right\} \qquad 73)$$

[пор. з першими двома формулами (71)].

§ 5. Випадок довільного дійсного показника $\rho \neq 0$

Тут наше завдання насамперед — визначити для відповідної області формули найкращого наближення, за критерієм відносної похибки, виду

$$(x^\rho+y^\rho+z^\rho+\cdots+v^\rho)^{\frac{1}{\rho}} = \alpha x+\beta y+\gamma z+\cdots+\varkappa v \qquad 74)$$

Для можливого спрощення запису ми тут викладемо розв'язання задачі відносно випадку трьох змінних x, y, z; отже шукатимемо формули найкращого наближення виду

$$(x^\rho + y^\rho + z^\rho)^{\frac{1}{\rho}} = \alpha x + \beta y + \gamma z \qquad 75$$

До того ж за область апроксимації візьмімо

$$x \geqslant y \geqslant z \geqslant 0 \qquad 76$$

Аналогічно до того, як це було зроблено в попередньому параграфі, ми й тут, скориставшись з однорідності шуканої формули, прийдемо до задачі — визначити для функції

$$f(y, z) = (1 + y^\rho + z^\rho)^{\frac{1}{\rho}} \qquad 77$$

найкраще наближене представлення виду

$$\varphi_1(y, z; \alpha, \beta, \gamma) = \alpha + \beta y + \gamma z \qquad 78$$

в області типу (T)

$$1 \geqslant y \geqslant z \geqslant 0 \qquad 79$$

Нехай буде спершу $\rho > 0$.

Попереднє дослідження $\nabla^2 f$ в самій області (79) та на її границі, яка складається з трьох одномірних областей, а саме

$$^{a)}\, z = 0;\ 1 \geqslant y \geqslant 0;\ ^{b)} z = y;\ 1 \geqslant y \geqslant 0 \text{ та } ^{c)} y = 1;\ 1 \geqslant z \geqslant 0,$$

дає:

$$\nabla^2 f(y,z) = (\rho - 1)(1 + y^\rho + z^\rho)^{\frac{1}{\rho} - 2}\left[y^{\rho-2} + z^{\rho-2} + y^{\rho-2} z^{\rho-2}(y^2 + z^2)\right]$$

$$\nabla^2 f(y,0) = \frac{d^2}{dy^2}(1 + y^\rho)^{\frac{1}{\rho}} = (\rho - 1) y^{\rho-2}(1 + y^\rho)^{\frac{1}{\rho} - 2}$$

$$\nabla^2 f(y,y) = \frac{d^2}{dy^2}(1 + 2y^\rho)^{\frac{1}{\rho}} = 2(\rho - 1) y^{\rho-2}(1 + 2y^\rho)^{\frac{1}{\rho} - 2} \qquad 80$$

$$\nabla^2 f(1,z) = \frac{d^2}{dz^2}(2 + z^\rho)^{\frac{1}{\rho}} = 2(\rho - 1) z^{\rho-2}(2 + z^\rho)^{\frac{1}{\rho} - 2}$$

Звідси висновок [1]):

при $\rho > 1$ функція $f(y, z)$ є кратно-субгармонічна

при $\rho < 1$ функція $f(y, z)$ є кратно-супергармонічна

в області (79) та на її границі. Отже спосіб розділу I тут напевне можна застосувати.

З системи (5) одержуємо значення параметрів:

$$\alpha = 1 = k,\ \beta = 2^{\frac{1}{\rho}} - 1 = l,\ \gamma = 3^{\frac{1}{\rho}} - 2^{\frac{1}{\rho}} = m$$

Тепер маємо визначити екстремальне значення відношення

$$R(y, z; k, l, m) = \frac{k + ly + mz}{f(y, z)}$$

в області (79), а саме треба знайти:

$$R_{\max} \text{ при } \rho > 1; \quad R_{\min} \text{ при } \rho < 1$$

Для цього доведеться дібрати екстремальне (найбільше при $\rho > 1$, але найменше при $\rho < 1$) із значень, що їх функція $R(y, z; k, l, m)$ набирає в стаціонарних[1]) точках, які є в двомірній області (79) та на окремих одномірних складових частинах її границі.

Стаціонарна точка всередині області (79) визначається з системи двох рівнянь:

$$\frac{y^{\rho-1}}{1+y^{\rho}+z^{\rho}} = \frac{l}{k+ly+mz}, \quad \frac{z^{\rho-1}}{1+y^{\rho}+z^{\rho}} = \frac{m}{k+ly+mz} \tag{83}$$

звідси визначається спершу відношення $y:z$, а потім і самі невідомі[2])

$$y = \left(\frac{l}{k}\right)^{\frac{1}{\rho-1}}, \quad z = \left(\frac{m}{k}\right)^{\frac{1}{\rho-1}} \tag{84}$$

і відповідне стаціонарне значення

$$R_{\text{stat}} = \left(k^{\frac{\rho}{\rho-1}} + l^{\frac{\rho}{\rho-1}} + m^{\frac{\rho}{\rho-1}}\right)^{\frac{\rho-1}{\rho}} \tag{85}$$

В одномірних областях на границі трикутника (79) — $z=0$, $y=z$ та $y=1$, відповідні стаціонарні значення будуть:

$$R_{\text{stat}} = \left(k^{\frac{\rho}{\rho-1}} + l^{\frac{\rho}{\rho-1}}\right)^{\frac{\rho-1}{\rho}} \tag{86}$$

$$R_{\text{stat}} = \left[k^{\frac{\rho}{\rho-1}} + 2\left(\frac{l+m}{2}\right)^{\frac{\rho}{\rho+1}}\right]^{\frac{\rho-1}{\rho}} \tag{87}$$

$$R_{\text{stat}} = \left[2\left(\frac{k+l}{2}\right)^{\frac{\rho}{\rho-1}} + m^{\frac{\rho}{\rho-1}}\right]^{\frac{\rho-1}{\rho}} \tag{88}$$

Зіставляючи останні три значення із значенням (85), легко пересвідчитись, що вони менші від нього при $\rho > 1$ і, навпаки, більші від нього при $0 < \rho < 1$. Щодо значення (86) це видно безпосередньо. Щождо значень (87) та (88), то це випливає з порівнення величини виразу $\left(\frac{l+m}{2}\right)^{\frac{\rho}{\rho-1}}$ або $\left(\frac{k+l}{2}\right)^{\frac{\rho}{\rho-1}}$ відповідно з величиною виразу $\frac{1}{2}\left(l^{\frac{\rho}{\rho-1}} + m^{\frac{\rho}{\rho-1}}\right)$ або $\frac{1}{2}\left(k^{\frac{\rho}{\rho-1}} + l^{\frac{\rho}{\rho-1}}\right)$ і це порівнення роблять на підставі таких відомих загальних

[1]) Щодо терміну „стаціонарні точки“ (зазначення тих точок, де перетворюються на нуль частинні похідні даної функції по всіх незалежних змінних) пор. Courant-Hilbert, Meth. der Math. Physik, I, S. 143.

[2]) Легко бачити, що точка (y, z) з координатами (84) належить до області (79).

нерівностей, які стосуються зважених середніх від степенів кількох чисел:

$$\left.\begin{array}{r}\lambda x^r+\mu y^r+\nu z^r+\cdots+\pi u^r-(\lambda x+\mu y+\nu z+\cdots+\pi u)^r\geqslant 0\\ \text{при } r>1 \text{ та при } r<0\end{array}\right\} \quad (89)$$

$$\left.\begin{array}{r}\lambda x^r+\mu y^r+\nu z^r+\cdots+\pi u^r-(\lambda x+\mu y+\nu z+\cdots+\pi u)^r\leqslant 0\\ \text{при } 0<r<1,\end{array}\right\} \quad (90)$$

де числа $x, y, z, \ldots, u$, а також і вагові коефіцієнти $\lambda, \mu, \nu, \ldots, \pi$, всі, припускаємо, є додатні, при чому $\lambda+\mu+\nu+\cdots+\pi=1$ [1]).

Отже значення (85) якраз потрібне нам: воно дає $R_{\max}$ при $\rho>1$ та $R_{\min}$ при $0<\rho<1$. Оскільки точка (84), в якій це екстремальне значення має місце, знаходиться всередині, у вузькому розумінні, області (79), цим самим забезпечена і єдиність розв'язку.

Після цього ясно, на підставі (32)—(34), що шукана формула найкращого наближення виду (75) буде така:

$$(x^\rho+y^\rho+z^\rho)^{\frac{1}{\rho}}=\sigma(kx+ly+mz), \quad (91)$$

де k, l, m мають значення (81), при $\rho>0$, і

$$\sigma=\frac{2}{1+\left(k^{\frac{\rho}{\rho-1}}+l^{\frac{\rho}{\rho-1}}+m^{\frac{\rho}{\rho-1}}\right)^{\frac{\rho-1}{\rho}}} \quad (92)$$

а гранична відносна похибка є $\varepsilon_0=|1-\sigma|$.

Досліджуючи граничну відносну похибку $\varepsilon_0=|1-\sigma|$ як функцію від показника ρ, не важко встановити, що при $\rho=1$ та $\rho=\infty$ гранична похибка ε_0 стає нулем. Навпаки, при $\rho\to 0$ вона прямує до 100%. Для показників більших від одиниці точність формул взагалі не гірша, ніж для формул, що їх було подано в попередньому параграфі. Для показників менших від $\frac{1}{2}$ гранична похибка швидко наростає.

При іншому числі незалежних змінних хід міркувань та висновки цілком аналогічні [2]).

Щоб скласти уяву про порядок точності та про характер коефіцієнтів при різних значеннях ρ, наведімо тут кілька двочленних формул для області $x\geqslant y\geqslant 0$:

[1]) Теорему цю дуже легко довести шляхом індукції від n до $n+1$, встановивши її насамперед для випадку двох чисел x, y звичайними засобами диференціального числення (дорівнявши при цьому більше число до одиниці).

[2]) При N змінних $x, y, z, \ldots, v$ значення параметрів будуть:

$$k=1,\ l=2^{\frac{1}{\rho}}-1,\quad m=3^{\frac{1}{\rho}}-2^{\frac{1}{\rho}},\ldots,r=N^{\frac{1}{\rho}}-(N-1)^{\frac{1}{\rho}} \quad (81')$$

$$\left.\begin{array}{lll}
\left(x^{\frac{3}{2}}+y^{\frac{3}{2}}\right)^{\frac{2}{3}} & \approx 0{,}969\,x+0{,}569\,y & \text{з точн. до } 3{,}1\,\% \\
\sqrt{x^2+y^2} & \approx 0{,}960\,x+0{,}398\,y & \text{„ „ „ } 4\,\% \\
\sqrt[3]{x^3+y^3} & \approx 0{,}958\,x+0{,}249\,y & \text{„ „ „ } 4{,}2\,\% \\
\sqrt{x^4+y^4} & \approx 0{,}961\,x+0{,}182\,y & \text{„ „ „ } 3{,}9\,\% \\
\sqrt[10]{x^{10}+y^{10}} & \approx 0{,}976\,x+0{,}070\,y & \text{„ „ „ } 2{,}4\,\%
\end{array}\right\} \quad (93)$$

$$\left.\begin{array}{lll}
\left(x^{\frac{2}{3}}+y^{\frac{2}{3}}\right)^{\frac{3}{2}} & \approx 1{,}066\,x+1{,}949\,y & \text{„ „ „ } 7\,\% \\
\left(\sqrt{x}+\sqrt{y}\right)^2 & \approx \frac{8}{7}(x+3y) & \text{„ „ „ } 14\tfrac{1}{2}\,\% \\
\left(\sqrt[3]{x}+\sqrt[3]{y}\right)^3 & \approx 1,\ 31\,x+9,\ 17\,y & \text{„ „ „ } 31\,\%
\end{array}\right\} \quad (94)$$

Трохи інакші міркування у випадку від'ємного показника $\rho < 0$.

Особливість його в тому, що тут $\lim\limits_{z \to 0} f(y, z) \equiv 0$. Проте відповідна модифікація міркувань розділу I дозволяє і тут знайти повний і цілком обгрунтований розв'язок. Визначаючи спершу інтерполяційний вираз виду (78), відповідно до системи рівнянь (5) знайдемо значення параметрів [1]):

$$\alpha = 0 = k,\ \beta = 0 = l,\ \gamma = 3^{\frac{1}{\rho}} = m \tag{95}$$

Функція $f(y, z)$ в даному разі є кратно-супергармонічна в області (79) та на її границі; щоб у цьому переконатися, тут досить буде знову взяти на увагу вирази (80) — за винятком другого виразу, який замінюється тотожним нулем. Звідси виходить, що величина відношення

$$R(y, z;\ k, l, m) = \frac{mz}{f(y, z)} \tag{96}$$

в області (79) та на її границі не перевищує одиниці [2]). Значить, тут $R_{\max} = 1$. Шукаючи найменше значення того самого відношення, насамперед легко пересвідчитись, що всередині області (79) стаціонарних точок функції (96) немає. Крім того, на граничному відрізку $z = 0$ вираз (96) втрачає сенс, бо переводиться на $\frac{0}{0}$. Легко зрозуміти, що цікаве для нас питання щодо розшукання $R_{\min}$ в трикутнику (79) буде

[1]) Ми визначаємо функцію $f(y, z)$ при $z = 0$ за умовою неперервності: $f(y, 0) = \lim\limits_{z \to o,\ Y \to y} f(Y, z) = 0$.

[2]) В даному разі це й безпосередньо легко перевіряється.

повнотою розв'язане, коли ми визначимо найменше значення виразу (96) в трикутнику

$$1 \geqslant y \geqslant z \geqslant \delta > 0 \tag{97}$$

як певну функцію $g(\delta)$ і потім знайдемо її границю при $\delta \to 0$. Ми спершу знайдемо

$$g(\delta) = \frac{m\delta}{(2+\delta\rho)^{\frac{1}{\rho}}} = \frac{m}{(2\delta^{-\rho}+1)^{\frac{1}{\rho}}}$$

це є значення виразу (96) у вершку $y = 1$, $z = \delta$. Звідки

$$R_{\min} = m = 3^{\frac{1}{\rho}} \tag{98}$$

Зауважмо тепер, що, в зв'язку з відзначеною вище основною особливістю даного випадку, в наш розгляд входять, власне, лише наближені представлення виду

$$\varphi_1(y, z; 0, 0, \gamma) = \gamma z \tag{99}$$

виключно, бо для інших лінійних виразів відношення $\varphi_1 : f$ буде нескінченно велике при $z \to 0$. А коли так, то шуканий розв'язок можна визначити, базуючись виключно на нерівності (28) та користуючися з корективного множника (31). Це дає нам, коли ми повернемось до області (76), шукану формулу найкращого наближення у вигляді

$$(x^\rho + y^\rho + z^\rho)^{\frac{1}{\rho}} = \frac{2 \cdot 3^{\frac{1}{\rho}}}{1+3^{\frac{1}{\rho}}} z = \frac{2}{3^{-\frac{1}{\rho}}+1} z \tag{100}$$

при $x \geqslant y \geqslant z > 0$ та при $\rho < 0$, при чому гранична відносна похибка тут є

$$\varepsilon_0 = \frac{2}{1+3^{\frac{1}{\rho}}} - 1 = \frac{1-3^{\frac{1}{\rho}}}{1+3^{\frac{1}{\rho}}} \tag{101}$$

Знайдений розв'язок є, очевидно, єдиний.

Загальніше, при будьякому числі N незалежних змінних $x, y, \ldots, v$ матимемо формулу найкращого наближення

$$(x^\rho + y^\rho + z^\rho + \cdots + v^\rho)^{\frac{1}{\rho}} = \frac{2}{1+N^{-\frac{1}{\rho}}} v \tag{100'}$$

в області $x \geqslant y \geqslant z \geqslant \ldots \geqslant v > 0$ при $\rho < 0$, з граничною відносною похибкою

$$\varepsilon_0 = \frac{1-N^{\frac{1}{\rho}}}{1+N^{\frac{1}{\rho}}} \tag{101'}$$

Точність, очевидно, буде найкраща при великих абсолютних значеннях ρ ($\varepsilon_0 \to 0$ при $\rho \to -\infty$). Навпаки, при малих абсолютних значеннях ρ

похибка буде чимала, але точність істотно підвищується, коли складати формулу для області

$$x \geqslant \frac{y}{\lambda} \geqslant \frac{z}{\lambda\mu} \geqslant \ldots \geqslant \frac{v}{\lambda\mu\nu\ldots\pi} \geqslant cx \tag{102}$$

при $0 < \lambda, \mu, \nu, \ldots, \pi < 1$, $0 \leqslant c < 1$. Аналогічне можна зауважити й щодо можливого підвищення точності у випадку $0 < \rho < 1$ при малих значеннях ρ.

Приклади:

$$\begin{aligned}
&\left(\frac{1}{x^2}+\frac{1}{y^2}\right)^{-\frac{1}{2}} \approx 0{,}829y \text{ з точн. до } 17{,}2\% \text{ при } x \geqslant y > 0\\
&\left(\frac{1}{x^4}+\frac{1}{y^4}\right)^{-\frac{1}{4}} \approx 0{,}914y \quad „ \quad „ \quad „ \quad 8{,}6\% \quad „ \quad „\\
&\left(\frac{1}{x^{10}}+\frac{1}{y^{10}}\right)^{-0{,}1} \approx 0{,}965y \quad „ \quad „ \quad „ \quad 3{,}5\% \quad „ \quad „\\
&\left(\frac{1}{x}+\frac{1}{y}\right)^{-1} \approx \frac{2}{3}y \quad „ \quad „ \quad „ \quad 33{,}3\% \quad „ \quad „\\
&\left(\frac{1}{x}+\frac{1}{y}\right)^{-1} \approx \frac{4}{5}y \quad „ \quad „ \quad „ \quad 20\% \quad „ \quad x \geqslant \frac{y}{0{,}5} > 0\\
&\left(\frac{1}{x}+\frac{1}{y}\right)^{-1} \approx \frac{20}{21}y \quad „ \quad „ \quad „ \quad 4{,}8\% \quad „ \quad x \geqslant \frac{y}{0{,}1} > 0\\
&\left(\frac{1}{x}+\frac{1}{y}\right)^{-1} \approx 0{,}1013x + 0{,}456y \quad „ \quad 1{,}3\% \quad „ \quad \frac{2}{3}x \geqslant y \geqslant \frac{1}{3}x
\end{aligned} \tag{103}$$

У випадках функціональних залежностей виду

$$w^\rho = x^\rho + y^\rho + \cdots + z^\rho - t^\rho - u^\rho - \cdots - v^\rho \tag{104}$$

можуть бути корисні, в зв'язку з попередніми формулами, ще формули виду

$$\left(X^\rho - Y^\rho\right)^{\frac{1}{\rho}} = \alpha X + \beta Y \tag{105}$$

Розгляньмо це питання при $\rho > 0$, маючи на увазі область

$$X \geqslant \frac{Y}{\lambda} \geqslant 0 \quad (0 < \lambda < 1) \tag{106}$$

Зважаючи на однорідність шуканої формули, маємо еквівалентну задачу: при даному значенні $\rho > 0$ скласти формулу найкращого наближення, за критерієм відносної похибки, виду

$$\left(1 - Y^\rho\right)^{\frac{1}{\rho}} = \alpha + \beta Y$$

для інтервалу

$$0 \leqslant Y \leqslant \lambda$$

Умова (A) тут має місце, бо

$$\nabla^2\left(1-Y^\rho\right)^{\frac{1}{\rho}} = \frac{d^2}{dY^2}\left(1-Y^\rho\right)^{\frac{1}{\rho}} = -\left(\rho-1\right)Y^{\rho-2}\left(1-Y^\rho\right)^{\frac{1}{\rho}-2}\begin{cases} <0 & \text{при } \rho>1 \\ >0 & \text{„ } 0<\rho<1 \end{cases}$$

Відповідно до системи рівнянь (5), знайдемо:

$$\alpha = 1 = k, \quad \beta = -\frac{1-(1-\lambda^\rho)^{\frac{1}{\rho}}}{\lambda} = l$$

Екстремальне на розглядуваному інтервалі значення відношення

$$R(Y;\ k,\ l) = \frac{k+lY}{(1-Y^\rho)^{\frac{1}{\rho}}}$$

а саме — мінімальне при $\rho > 1$ і, навпаки, максимальне при $0 < \rho < 1$ — буде:

$$R_{\text{extr}} = \left[1-(-l)^{\frac{\rho}{\rho-1}}\right]^{\frac{\rho-1}{\rho}}$$

На підставі (32) — (34) і повертаючись до області (106), одержимо формулу найкращого наближення:

$$(X^\rho - Y^\rho)^{\frac{1}{\rho}} = \frac{2}{1+R_{\text{exrt}}}(X+lY), \tag{107}$$

де значення l та R_{exrt} дано останніми формулами, а гранична відносна похибка є

$$\varepsilon_0 = \left|1-\frac{2}{1+R_{\text{ext}}}\right| \tag{108}$$

Наведімо ще пару формул найкращого наближення, визначених також за способом розділу I, але відповідно до критерія абсолютної похибки.

$$\begin{aligned} &\sqrt{1+y^2} \approx 0{,}955+0{,}414y \text{ з точн. до } 0{,}045 \text{ при } 1 \geqslant y \geqslant 0 \\ &\sqrt{1+y^2+z^2} \approx 0{,}926+0{,}414y+0{,}318z \text{ „ } 0{,}074 \text{ „ } 1 \geqslant y \geqslant z \geqslant 0 \end{aligned} \tag{109}$$

Перетворюючи формули до однорідного вигляду, матимемо:

$$\begin{aligned} &\sqrt{x^2+y^2} \approx 0{,}955\,x+0{,}414\,y \text{ з точн. до } 4{,}5\% \text{ велич. } x \\ &\qquad \text{при } x \geqslant y \geqslant 0 \\ &\sqrt{x^2+y^2+z^2} \approx 0{,}926\,x+0{,}414\,y+0{,}318\,z \text{ з точн. до } 7{,}4\% \\ &\qquad \text{велич. } x \text{ при } x \geqslant y \geqslant z \geqslant 0 \end{aligned} \tag{110}$$

§ 6. Приклад застосування до інтеграції диференціального рівняння

Щоб дати приклад такого застосування, візьмімо диференціальне рівняння Коші-Муаньо

$$\frac{dy}{dx}=\sqrt{x}+\sqrt{y}, \tag{111}$$

що на ньому вже випробовували різні методи наближеної інтеграції [1]).

Взявши спершу початкові умови $y=0{,}03090$ при $x=0{,}1$, шукатимемо дві суворо встановлені межі для значення y при $x=0{,}5$. Для цього скористуймося з поданої вище наближеної формули (94_2):

$$(\sqrt{x}+\sqrt{y})^2 \approx \frac{8}{7}(x+3y) \text{ з точн. до } 14^1/_2\% \text{ при } x \geqslant y \geqslant 0$$

Інакше:

$$(\sqrt{x}+\sqrt{y})^2=K(x,y)\cdot\frac{8}{7}(x+3y) \text{ при } 0{,}855<K(x,y)<1{,}145, \tag{112}$$

де знак $=$ уже позначає рівність точну. Замінивши функцію $K(x,y)$ на будьяку сталу k, що її величину ми припустимо в межах інтервалу (0,855; 1,145), та підставивши в диференціальне рівняння (111), дістанемо рівняння

$$\frac{dy}{dx}=\sqrt{\frac{8k}{7}(x+3y)}, \tag{113}$$

яке інтегрується у скінченному виді. А саме, зробивши заміну змінного

$$\sqrt{\frac{8k}{7}(x+3y)}=u, \tag{114}$$

дістанемо рівняння

$$\frac{7}{12k}u\frac{du}{dx}-\frac{1}{3}=u, \tag{115}$$

що його загальний інтеграл є

$$3u-\log\left(u+\frac{1}{3}\right)=\frac{36k}{7}x+C \tag{116}$$

Виходячи із заданих початкових умов ($y=0{,}03090$ при $x=0{,}1$), визначаємо для довільної сталої C два значення: $C_1=1{,}1881$ та $C_2=1{,}0390$ відповідно до крайніх значень параметра $k_1=0{,}855$ та $k_2=1{,}145$. В такий спосіб дістанемо (в неявному виді) два окремі інтеграли u_1 та u_2 диференціального рівняння (115) відповідно до значень параметра $k=k_1$ та $k=k_2$. Дві функції

$$y_1(x)=\frac{7u_1^2}{24k_1}-\frac{1}{3}x \text{ та } y_2(x)=\frac{7u_2^2}{24k_2}-\frac{1}{3}x \tag{117}$$

[1]) Moigno. Leçons de calc. diff. et de calc. intégr., rédigées... d'après les méthodes de M. A. L. Cauchy, t. 2, pp. 432—434.

А. Н. Крылов. Прибл. числ. интегр. обыкн. диф. уравнений, Берлин, 1923.

Є. Ремез. Деякі способи чисельної інтеграції диф. рівнянь. Зап. Природн.-техн. відділу ВУАН, 1931, № 1.

дадуть нам розв'язки диференціального рівняння (113) при даних початкових умовах при $k = k_1$ та при $k = k_2$ відповідно. Якщо взяти до уваги відому теорему Чаплигіна [1]), точний розв'язок $y(x)$ диференціального рівняння Коші-Муаньо (111) при даних початкових умовах міститиметься між функціями $y_1(x)$ та $y_2(x)$, які й будуть для нього шукані межі:

$$y_1(x) < y(x) < y_2(x) \tag{118}$$

принаймні поки величина y не вийде з границь $x \geqslant y \geqslant 0$.

Переводячи обчислення, щоб визначити насамперед y_1 (0,5), маємо:

$$3u - \log\left(u + \frac{1}{3}\right) = \frac{36\,k_1}{7} \cdot 0{,}5 + C_1 = 3{,}386$$

Щоб розв'язати це трансцендентне рівняння, яке має, як легко бачити, один і тільки один корінь, покладімо

$$\Phi(u) = 3u - \log\left(u + \frac{1}{3}\right) - 3{,}386$$

Оскільки $\Phi(1{,}2) = -0{,}213$; $\Phi(1{,}3) = +0{,}023$, — шуканий корінь міститься між 1,2 та 1,3 і визначається точніше через

$$1{,}3 - \frac{\Phi(1{,}3)}{\Phi'(1{,}3)} = 1{,}3 - 0{,}010 = 1{,}290$$

Після цього (117) дає:

$$y_1(0{,}5) = \frac{7 \,.\, 1{,}290^2}{24\,k_1} - \frac{1}{3} \,.\, 0{,}5 = \underline{0{,}401}$$

Аналогічно, шукаючи y_2 (0,5), матимемо:

$$F(u) = 3u - \log\left(u + \frac{1}{3}\right) - 3{,}983 = 0$$

$$F(1{,}5) = -0{,}089;\ F(1{,}6) = 0{,}158;\ u = 1{,}6 - \frac{F(1{,}6)}{F'(1{,}6)} = 1{,}536$$

$$y_2(0{,}5) = \frac{7 \,.\, 1{,}536^2}{24\,k_2} - \frac{1}{3} \cdot 0{,}5 = \underline{0{,}434}$$

Взявши для y при $x = 0{,}5$ середню арифметичну знайдених двох меж

$$y = \frac{y_1 + y_2}{2} = 0{,}4175, \tag{119}$$

матимемо похибку напевне меншу від $\frac{y_2 - y_1}{2} = 0{,}0165$, тобто меншу від 4%. Але ця гранична похибка, звичайно, є більша від справжньої похибки: зіставляючи із значенням y при $x = 0{,}5$, що його дає деталізована

[1]) Акад. С. А. Чаплыгин. Новый метод прибл. интегр. дифф. ур-ий, ГТТИ, 1932.

інтеграція за методом Адамса-Штермера [1]), побачимо, що різниця 0,4175 — 0,4134 $<$ 1%.

Щоб оцінити цю точність, слід взяти на увагу, поперше, що застосована тут формула (94_2) належить до найгрубіших між формулами розглядуваних в § 5 типів, як це ми вже відзначали; подруге, що тут інтеграцію провадиться поблизу особливої точки даного диференціального рівняння (початку координат), де застосування інших, деталізованих методів вимагає досить великого числа інтеграційних кроків; і потрете — особливу важкість задачі щодо визначення суворо встановлених меж для інтегралів диференціальних рівнянь взагалі.

Зауважмо, далі, що ніщо не заважає застосувати даний метод і в тому разі, коли задана початкова точка інтегральної кривої збігається з самою особливою точкою $x=0$, $y=0$: довільна стала C в загальному нтегралі (116) дорівнювала б тут log 3; щодо застосованості теореми Чаплигіна в цьому випадку, то для теоретичного обгрунтування тут досить було б удатися то тієї інтерпретації, яку дав цій теоремі Жуковський [2]).

Щоб висвітлити розглядуваний шлях наближеної інтеграції ще в іншому зв'язку, можна було б спробувати скористатися розв'язком (116)—(114) для дальшого застосування ітераційного методу. Така спроба не позбавлена повчальних елементів, хоч даний приклад і не є достатньою мірою виграшний з цього погляду. Залишаючи тут осторонь ітераційний процес двобічного наближення С. А. Чаплигіна, що його в даному прикладі з різних причин було б важче застосувати, ми б удалися до відомого Пікарового ітераційного процесу [3]).

Добувши насамперед із (116) при $C=\log 3$ (відповідно до початкових умов $x=y=0$) та при $k=1$ (середнє значення інтервалу 0,855; 1,145) для досить малих значень x наближене співвідношення

$$u^2=\frac{8}{7}x+2u^3, \tag{120}$$

звідки, в першому наближенні, $u^2=\frac{8}{7}x$ і потім точніше

$$u^2=\frac{8}{7}x+2\cdot\left(\frac{8}{7}x\right)^{\frac{3}{2}}, \tag{121}$$

що дає, при підставленні до (114),

$$y=\frac{2}{3}\sqrt{\frac{8}{7}}\,x^{\frac{3}{2}}, \tag{122}$$

— ми взяли б цей явний вираз $y=\frac{2}{3}\sqrt{\frac{8}{7}}\cdot x^{\frac{3}{2}}$ за вихідне наближення y_0 в Пікаровому процесі, покладаючи

$$y_{i+1}=\int_0^x\left(\sqrt{x}+\sqrt{y_i}\right)dx,\quad i=0,\ 1,\ 2,$$

При належному заокругленні наближеного раціонального виразу $\sqrt{y_i}$ беручи в кожному

[1]) А. Н. Крылов, op. cit.

[2]) С. А. Чаплыгин, op. cit.

[3]) Було б не важко встановити a priori збіжність процесу, дарма що тут початкова точка є особлива.

наступному наближенні одним членом більше, ніж у попередньому, ми після п'ятьох[1]) ітерацій дістали б:

$$y_5 = \frac{2}{3} x^{\frac{6}{4}} + \frac{4}{7}\sqrt{\frac{2}{3}}\, x^{\frac{7}{4}} + \frac{1}{7} x^{\frac{8}{4}} + \frac{1}{49}\sqrt{\frac{2}{3}}\, x^{\frac{9}{4}} - \frac{2}{1715} x^{\frac{10}{4}} - \frac{261}{1056440}\sqrt{\frac{2}{3}}\, x^{\frac{11}{4}} \tag{123}$$

Спиняючись на цьому наближенні й підставляючи сюди, щоб спробувати його точність, навіть $x = 1$, ми дістали б $y = 1{,}29139$ — значення, що цілком збігається з тим, яке дала максимально - уточнена інтеграція за методом Адамса-Штермера в цитованій праці А. Н. Крилова [2]).

[1]) Це становить на одну ітерацію менше, ніж при застосуванні вихідного наближення $y_0 \equiv 0$. Можна було б сподіватися більш ефектного виграшу, якби не та особливість даного випадку, що тут $\frac{dy}{dx} = 0$ у початковій точці.

[2]) Що розв'язок, виведений зразу лише для малих значень x, добре справджується і для більших, очевидно, не може нас здивувати, якщо взяти на увагу аналітичну форму цього розв'язку. Цікаво відзначити, що розвинення (123) дає відрізок того самого ряду, до якого А. Н. Крилов прийшов зовсім іншим шляхом (*op. cit.*).

РОЗДІЛ III

ПРО ЗАДАЧУ ПОНСЕЛЕ

§ 7. Деякі історичні зауваження. З'ясування предмету цього розділу

Задача Понселе про наближене представлення квадратного кореня з суми квадратів двох або трьох аргументів у виді лінійної однорідної функції тих самих аргументів виникла в найщільнішому зв'язку з відповідними застосуваннями до питань технічної механіки[1]). Як приклад застосування до істотно іншого практичного питання, зазначмо метод А.С. Попова проектування копальневого провітрювання[2]), де формули Понселе двочленна та тричленна застосовуються при розв'язуванні систем алгебричних рівнянь. Якщо в застосуваннях до механіки кількість квадратів залежала від числа взаємно перпендикулярних складників вектора, в останньому випадку вона вже залежить від числа боків відповідного контура копальневої вентиляції. Виведені в § 4 даної праці нові формули з довільною кількістю квадратів обіймають формули Понселе як окремий випадок.

Історія задачі Понселе становить істотну складову частину історії тих методів апроксимації, що їх ми тепер позначаємо іменем Чебишова, який остаточно сформулював суть методу і застосував його до питань досить глибоких. П. Л. Чебишов у своєму мемуарі „Théorie des mécanismes connus sous le nom de parallélogrammes[3]), формулюючи свою задачу про визначення найкращого поліноміального наближеного представлення n-го степеня для функції kx^{n+1} на інтервалі $(-1,1)$ зазначає: „Щодо методу апроксимації, про який ми оце зараз говорили, ми маємо лише досліди Понселе, який дав лінійні формули для оцінки оцих трьох виразів:

$$\sqrt{x^2+y^2}, \qquad \sqrt{x^2-y^2}, \qquad \sqrt{x^2+y^2+z^2}$$

формули багато вживані в практичній механіці"

[1]) Крім курсів механіки самого Poncelet, див. H. Lorenz. Technishe Mechanik starrer Systeme; також J. Weisbach. Lehrbuch der Ingenieur und Maschinen-Mechanik.

[2]) А. С. Попов. Проектирование рудничной вентиляции при диагональном соединении проводов воздуха. Изд. Донугля, 1930; також Н. Протодьяконов.

Понселе мав своїх попередників. Так, Ейлер[1]) в мемуарі, присвяченому картографічній проекції Деліля, визначає найкраще наближене представлення функції $\cos x$ на інтервалі (α, β), де $0 < \alpha < \beta < \frac{\pi}{2}$, за допомогою лінійної функції $k(b-x)$, добираючи параметри k та b за умовою, щоб різниця $\cos x - k(b-x)$ набирала свойого максимального абсолютного значення з альтернуючим знаком у точках $x=\alpha$, $x=\arcsin k$, $x=\beta$.

Аналогічне питання, але по-інакшому поставлене, знаходимо у Лапласа[2]). Перевіряючи еліптичну форму земних меридіанів на підставі емпіричних даних, Лаплас добирає значення емпіричних констант у відповідній формулі, виходячи з вимоги мінімізації найбільшої з похибок, які мають місце, коли підставляти у формулу дані вимірюваня. Отже в Лапласа область апроксимації є дискретна — відповідно до скінченного числа виміряних значень незалежної змінної (географічної ширини).

Фур'є[3]) формулює аналогічну до Лапласової задачу щодо наближеного розв'язання системи m лінійних рівнянь з n невідомими $(m > n)$. Зупинившись докладніше на випадку $n=2$, Фур'є пробує, за допомогою своєрідної геометричної інтерпретації, накреслити деякий систематичний шлях, щоб знайти оптимальну систему шуканих значень.

Сам Понселе[4]), власне, не визначив найкращих лінійних наближених представлень для всіх трьох іраціональних виразів, про які згадує Чебишов: найкращі наближені вирази він знайшов лише для перших двох радикалів; а вираз, який він дає для $\sqrt{a^2+b^2+c^2}$, вже не відповідає розглядуваному принципові. Випадок, що його розглянув Понселе, стосувався по суті до функції одного змінного. Шукаючи значення параметрів α, β, при яких максимум виразу (відносної похибки)

$$\left|\frac{\sqrt{1+t^2}-(\alpha t+\beta)}{\sqrt{1+t^2}}\right|$$

при $n \leqslant t < \infty$, був би найменший, Понселе насамперед визначає в явному виді згаданий максимум, як функцію від α та β (за даного n), що в даному випадку труднощів не являє. Потім він мінімізує цей максимум, виходячи з елементарної геометричної інтуіції.

Horvath[5]) узагальнив (не досить вдало) розв'язок задачі Понселе

[1]) Euler. De projectione geographica Delisliana in mappa generali Imperii Russici usitata Acta Academiae, pro anno 1777. Цитуємо за Д. О. Граве — „Об основных задачах математической теории построения географических карт". СПБ, 1896.

[2]) Laplace. Traité de Mécanique céleste, 1798, n° 39, pp. 126—134.

[3]) Jean Baptiste Joseph Fourier. Analyse des équations déterminées. Paris, 1831. Übergesetzt und herausgegeben von Alfred Loewy. Leipzig. 1902, S. 77—79.

J. V. Poncelet. Sur la valeur approchée linéaire et rationnelle des radicaux de la forme $\sqrt{a^2-b^2}$, $\sqrt{a^2-b^2}$ etc. Crelle's Journal für reine u. angewandte Mathematik. 1835, S. 277—291.

Horvath. Sur les valeurs approximatives et rationnelles des radicaux de la forme $\sqrt{x^2+y^2}$. Доповідано було в Société Philomathique de Paris і надруковано в L'Institut, 1868, pp. 67—69.

для випадку квадратного кореня з суми трьох квадратів. Розглядаючи $\sqrt{x^2+y^2+z^2}$ як вираз довжини вектора із взаємно перпендикулярними складниками x, y, z, він обмежує напрямок вектора, за початок якого він вважає початок координат, тріедром із заданими рубами OR_1, OR_2, OR_3. Вираз відносної похибки

$$e=\left|\frac{\alpha x+\beta y+\gamma z}{\sqrt{x^2+y^2+z^2}}-1\right. \tag{124}$$

що його максимум підлягає мінімізації, Horvath інтерпретує, розглядаючи віддаль від точки перетину розглядуваного вектора (або його продовження) зі сферою одиничного радіуса — до площі $\alpha X+\beta Y+\gamma Z-1=0$.

H. Résal[1]) додав зауваження (також не досить вдале), ніби розв'язок задачі Понселе за умов Horvath-а зберігає силу й тоді, коли вектор обмежується лише конусом, описаним навколо даного тріедра. Résal користується іншою геометричною інтерпретацією, розглядаючи величину (124) як вираз різниці між радіусом-вектором сфери $X^2+Y^2+Z^2=\alpha X+\beta Y+\gamma Z$, що переходить через початок координат, та радіусом-вектором сфери одиничного радіуса з центром у початку координат, і припускаючи, що напрямки обох цих радіусів-векторів збігаються з напрямком розглядуваного вектора, який має складники x, y, z.

А. А. Марков[2]) розглянув інший випадок задачі Понселе, а саме коли напрямок вектора обмежується не тригранним, а чотиригранним кутом, який означається умовами

$$\lambda_1 y<x<\mu_1 y, \quad \lambda_2 y<z<\mu_2 y$$

Покійний академік у своєму мемуарі ніби навмисне уникає геометричної мови, подаючи всі виклади без жодної геометричної інтерпретації.

Як ми вже зауважили у вступі до даної роботи, трактування задачі про $\sqrt{x^2+y^2+z^2}$ у всіх попередніх авторів не є викінчене. Воно являє істотні прогалини як з погляду повноти розв'язання, так і з погляду точності результатів та ефективності доведення. Щодо статті Horvath-а, то вона, крім того, містить ряд числових помилок зовсім незрозумілих у роботі такого характеру. Саме від статті Horvath-а мають початок грубезні помилки, з якими досі друкується навіть двочленну формулу самого Понселе в такому поширеному виданні як технічний довідник Dubbel-я: $\sqrt{a^2+b^2}=0{,}960a+0{,}368b$ замість $0{,}398b$ (!)[3]). Таксамо формула з трьома квадратами досі подається (і в Dubbel-я, і в Hütte) з неправильними останніми цифрами в коефіцієнтах — в такому вигляді, як її подав Horvath. Щодо формули для $\sqrt{x^2+y^2}$ та $\sqrt{x^2+y^2+z^2}$ при $x\geqslant y\geqslant z\geqslant 0$,

[1]) H. Résal. Sur un théorème de Poncelet et sa généralisation par M. Horvath, Bull. de la Soc. Mathémat. de France, 1873.

[2]) А. Марков. Новый случай задачи Понселе о приближенном выражении квадратного корня из суммы квадратов. Известия Рос. Академии Наук, 1906, сс. 65—81.

[3]) В довіднику Hütte (рос. мовою) в усіх виданнях до останнього, переробленого, та сама формула подавалася з помилкою ще більш грубою: 0,906 замість 0,960.

то ми їх уже розглянули, як частинний випадок загальнішої задачі, в розділі II [див. перші дві формули (71)]. В дальших двох параграфах даного розділу буде дано закінчене трактування випадків Horvath-а та А. Маркова задачі Понселе. По суті спираючись в цьому трактуванні на теорію розділу I, ми користуємось тут стислою мовою геометрії. Незалежно від знайомства з статтями згаданих французьких авторів, ми тут також вжили деякої геометричної інтерпретації, але істотно відмінної від тих і багато зручнішої, ніж робоча інтерпретація. Наша інтерпретація виходить із спеціалізованого для випадку трьох аргументів аналогічного до 62) перетворення виразу $R(x, y, z; \alpha, \beta, \gamma) = \frac{\alpha x + \beta y + \gamma z}{\sqrt{x^2 + y^2 + z^2}}$. Інтерпретуючи змінні x, y, z та параметри α, β, γ як напрямні параметри двох променів, що виходять з початку координат (прямокутної системи) та впроваджуючи в розгляд точки перетину цих променів із сферою одиничного радіуса, яка має центр у початку координат, ми маємо змогу застосувати мову сферичної та аналітичної геометрії при трактуванні питань суто арифметичних.

Зробімо тут кілька загальних зауважень. Згадану сферу позначмо через Σ. Точку перетину сфери Σ з променем, який виходить із початку координат O і має напрямні параметри x, y, z, позначмо M або іще $M(x, y, z)$. Аналогічно, замінюючи x, y, z на α, β, γ, матимемо на сфері точку K або $K(\alpha, \beta, \gamma)$. Перетворення, аналогічне до (62), тут дає:

$$R(x, y, z; \alpha, \beta, \gamma) = \frac{\alpha x + \beta y + \gamma z}{\sqrt{x^2 + y^2 + z^2} \cdot \sqrt{\alpha^2 + \beta^2 + \gamma^2}} \cdot \sqrt{\alpha^2 + \beta^2 + \gamma^2} = \\ = \sqrt{\alpha^2 + \beta^2 + \gamma^2} \cdot \cos MK, \tag{125}$$

де MK означає сферичну віддаль, тобто найкоротшу віддаль на сфері, поміж точками M та K. Задача Понселе набирає такого узагальненого сформулювання:

Дано на сфері Σ якусь неперервну область $\mathfrak{A}$. Припускаючи, що точка M може займати будьяке положення в області $\mathfrak{A}$, треба знайти такі значення параметрів α, β, γ, щоб вираз (125) щонайменше відхилявся від одиниці при $M \subset \mathfrak{A}$.

З міркувань § 2 розділу I випливає, що параметри α, β, γ будуть відомі до спільного корективного множника σ, який і собі визначається з формули (31), якщо розв'яжемо таку задачу:

При тих самих даних знайти на сфері таку фіксовану точку $K(\alpha, \beta, \gamma)$, щоб величина

$$\Omega(\alpha, \beta, \gamma) = \frac{R_{\max}}{R_{\min}} = \frac{\max\limits_{M \subset \mathfrak{A}} [\cos MK]}{\min\limits_{M \subset \mathfrak{A}} [\cos MK]} \tag{126}$$

була щонайменша

самим сенсом задачі апроксимації, $\cos MK$ слід припускати

додатною при всіх можливих положеннях точки M в області $\mathfrak{A}$.

Отже точку K можна шукати лише поміж тих точок сфери Σ, для яких максимум сферичної віддалі до точок області $\mathfrak{A}$ не перевищує $\frac{\pi}{2}$. Зрозуміла річ, що на підставі цієї умови не кожну неперервну область на сфері можна взяти за область $\mathfrak{A}$.

Умовмося також про дальші позначення. Якщо LM та MN є дві дуги великих кіл на сфері, ми для кута їхнього перетину вживатимемо позначення $\measuredangle LMN$ або $\measuredangle M$. Сферичним центром малого кола на сфері називатимемо ближчу з двох точок на сфері, які мають сталу віддаль до точок цього кола. Якщо маємо дугу MN великого кола на сфері, вживатимемо скороченої назви — „сферичний середній перпендикуляр до MN“ для великого кола, проведеного перпендикулярно до дуги MN через її середину.

§ 8. Horvath-ів випадок задачі Понселе

Цей випадок ми матимемо тоді, коли роль області $\mathfrak{A}$ попереднього параграфа відіграє якийсь сферичний трикутник Δ, якого вершки ми позначимо через $A_1(x_1, y_1, z_1)$, $A_2(x_2, y_2, z_2)$ та $A_3(x_3, y_3, z_3)$ (див. рис. 1), при чому для більшої визначеності дальших міркувань припустімо, що $\measuredangle A_1 \leqslant \measuredangle A_2 \leqslant \measuredangle A_3$ в трикутнику Δ. Область допущенних положень шуканої точки $K(\alpha, \beta, \gamma)$ тут не може бути порожня, бо до неї напевне належать сферичний центр C описаного навколо Δ кола та деякий окіл точки C на сфері [1]).

Припустімо, що координатний тріедр орієнтований так, щоб площина YOZ була паралельна до площини описаного навколо Δ круга [2]). Тригранний кут $OA_1A_2A_3$ (при вершку O) в перерізі з площиною $x=1$ дає деяку трикутну область Δ'. Розглядувана геометрична задача („задача a“) є цілком еквівалентна арифметичній задачі („задача b“) найкращого наближеного представлення функції $\sqrt{1+y^2+z^2}$ в області трикутника Δ' за допомогою лінійного виразу виду $\alpha+\beta y+\gamma z$. Але до останньої задачі є цілком застосована теорія розділу I цієї праці: легко пересвідчитись, що функція $f(y, z)=\sqrt{1+y^2+z^2}$ є кратно-субгармонічна (за позначеннями § 3) в області будьякого трикутника площини (y, z) та на складових частинах його границі. Перекладаючи цю теорію на вживану тут геометричну мову, маємо такі висновки:

1°. Один розв'язок задачі (a) завжди дає сферичний центр C описаного навколо Δ кола [3]). Зауважимо, що C є водночас точка дотику сфери й площини $x=1$.

2°. Якщо C лежить всередині сферичного трикутника Δ, що має місце при умові $\measuredangle A_3 < \measuredangle A_1 + \measuredangle A_2$, то розв'язок є єдиний, бо водночас $R_{\max}$ в задачі (b) має тоді місце у внутрішній точці C області Δ'

[1]) Справді, відомо, що сферичний радіус описаного кола завжди менший від $\frac{\pi}{2}$

[2]) Легко зрозуміти, що припущення це аж ніяк не порушує загальності наступних висновків.

[3]) Ми цим хочемо сказати, що одна система значень параметрів α, β, γ, яка мінімізує величину (126), напевне дається умовою $K(\alpha, \beta, \gamma) \equiv C$.

3° Коли ж C лежить поза сферичним трикутником Δ або на його контурі, тобто коли $\sphericalangle A_3 \geqslant \sphericalangle A_1 + \sphericalangle A_2$ [1]), тоді R_{max} в задачі (b) має місце в тій точці області Δ', яка визначається перетином площини $x = 1$ із променем OD, позначаючи літерою D середину дуги A_1A_2 (рис. 1). Оскільки ця точка є на границі області Δ', теорія розділу I говорить за те, що тут може бути безліч розв'язків, при чому відповідні точки $K(\alpha, \beta, \gamma)$ можна шукати лише на сферичному середньому перпендикулярі BG див. до дуги A_1A_2. Ми маємо зараз довести, що розв'язків справді буде безліч, коли співвідношення

$$\sphericalangle A_3 \geqslant \sphericalangle A_1 + \sphericalangle A_2 \tag{127}$$

має місце із знаком $>$; коли ж це співвідношення має місце із знаком $=$, то розв'язок є єдиний, як і у випадку аналогічного співвідношення з знаком $<$.

Почнімо з такого зауваження. Якщо на сферичному середньому перпендикулярі BG дуги A_1A_2 візьмемо першу-ліпшу точку P (її на рисунку не позначено), то, звичайно, матимемо $PA_1 = PA_2$. Припустімо, що співвідношення (127) має місце із знаком $>$ та що точки B, G якраз обмежують ту частину розглядуваного середнього перпендикуляра, яка належить до області допущених положень шуканої точки $K(\alpha, \beta, \gamma)$ [2]).

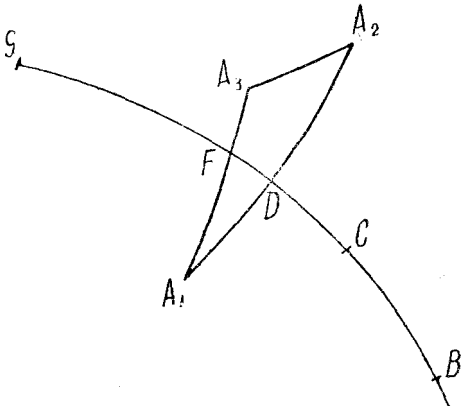

Рис. 1.

Якщо точку P взяти на частині BF дуги середнього перпендикуляра, то, як легко пересвідчитись, матимемо одно з трьох співвідношень: $PA_3 > PA_1$, $PA_3 < PA_1$ або $PA_3 = PA_1$ залежно від того, чи співвідношення

$$\sphericalangle PA_1A_3 + \sphericalangle PA_2A_3 \gtreqless \sphericalangle A_1A_3A_2 \tag{128}$$

має місце із знаком $>$, $<$ або $=$ [3]). Але коли точка P збігається з C, має місце, очевидно, знак $=$. Звідси видко (звертаючи увагу на кути), що, коли точка P припадає на дугу BC (виключаючи кінець C), матимемо $PA_3 > PA_1$, а коли точка P припадає на дугу CF (знову таки виключаючи

[1]) Співвідношення $\sphericalangle A_3 \geqslant \sphericalangle A_1 + \sphericalangle A_2$ становить умову необхідну і достатню для того, щоб центр C знаходився не всередині трикутника Δ. Додамо ще два зауваження: 1°. Кут A_3 в цьому разі є, очевидно, тупий. 2°. Легко показати, що інші два кути повинні бути гострі. Щоб це довести, скажімо щодо кута A_2, досить розглядувану нерівність зіставити з іншою нерівністю: $\sphericalangle A_3 < (\pi - \sphericalangle A_2) + \sphericalangle A_1$, яка має місце завжди, бо вона є еквівалентна до нерівності $2A_1 > A_1 + A_2 + A_3 - \pi$.

[2]) Зауважмо, що дуга DB, напевне менша від $\frac{\pi}{2}$, не може містити точок перетину з продовженнями боків A_3A_1 та A_3A_2 (взяти на увагу останню виноску).

[3]) Доводиться безпосередньо, застосовуючи окремо до сферичного трикутника PA_1A_3 та PA_2A_3 теорему: „проти більшого боку лежить більший кут".

кінець C), то матимемо $PA_3 < PA_1$. Остання нерівність повинна справджуватись і на продовженні дуги CF аж до точки G, бо порушення нерівності $PA_3 < PA_1$ могло б тут статися лише тоді, коли б на дугу FG припадала діаметрально-протилежна до C точка сфери, а цього (на підставі основних уваг попереднього параграфа щодо області допущенних значень α, β, γ) не може бути.

Ще зауважмо, що, оскільки максимум віддалі будьякої точки P дуги BG від точок сферичного трикутника Δ не перевищує $\frac{\pi}{2}$, цей максимум при зафіксованій точці P, може мати місце лише для вершків трикутника Δ, але (виключаючи випадок[1]) $P \equiv G$) не для середових точок його боків і, зрозуміла річ, також не для внутрішніх точок трикутника. Звідси, між іншим, як легко бачити, випливає безпосередньо наявність умови (A) параграфа 1 в розглядуваній задачі — незалежно від вказаного раніше аналітичного дослідження.

Тепер потрібне доведення дістаємо вже досить легко. Справді, ми вже відзначали, що найменше значення $\Omega(\alpha, \beta, \gamma)$ напевне має місце, коли точка $K(\alpha, \beta, \gamma)$ збігається з точкою C[2]) і що інші оптимальні положення точки K можуть знаходитись лише на дузі BG. При $K(\alpha, \beta, \gamma) \equiv C$, згідно з (126), маємо:

$$\Omega_{\min} = \frac{\cos CD}{\cos CA} = \operatorname{sc} A_1 D \tag{129}$$

Беручи точку $K(\alpha, \beta, \gamma)$ на дузі CD, матимемо завжди:

$$\Omega(\alpha, \beta, \gamma) = \frac{\cos KD}{\cos KA_1} = \operatorname{sc} A_1 D = \Omega_{\min},$$

тобто кожна точка дуги CD також дає нам розв'язок задачі.

При інших же положеннях точки $K(\alpha, \beta, \gamma)$ на дугах BC, DF або FG, як легко пересвідчитись, величина $\Omega(\alpha, \beta, \gamma)$ буде більша. Справді, при положенні K на дузі BC (виключаючи кінець C) матимемо:

$$\Omega = \frac{\cos KD}{\cos KA_3} > \frac{\cos KD}{\cos KA_1} = \operatorname{sc} A_1 D, \text{ тобто } \Omega > \Omega_{\min}$$

Далі при будьякому положенні точки K на дузі DF (виключаючи кінець D) матимемо:

$$\frac{\cos O}{\cos KA_1} = \operatorname{sc} KA_1 > \operatorname{sc} A_1 D, \text{ тобто } \Omega > \Omega_{\min}$$

Нарешті, коли точку K взяти на дузі FG, матимемо:

$$\Omega \geqslant \frac{\cos KF}{\cos KA_1} > \frac{\cos KD}{\cos KA_1} = \operatorname{sc} A_1 D, \text{ тобто } \Omega > \Omega_{\min}$$

Отже бачимо, в кінцевому підсумку, що розв'язок дає кожна точка дуги CD, і що цими розв'язками вичерпується вся їх сукупність.

В граничному випадку, коли співвідношення (127) має місце із знаком $=$, тобто коли точка C збігається з D, дістаємо, очевидно, знову лише один розв'язок — подібно до того випадку, коли точка C знаходиться всередині трикутника Δ.

Сам Horvath у своїй статті знаходить розв'язок, який в дійсності має силу лише при умові $\sphericalangle A_3 \leqslant \sphericalangle A_1 + \sphericalangle A_2$, в противному ж випадку він не збігається з жодним із справжніх розв'язків, відрізняючись довільно-великим множником від того окремого розв'язку, який відповідає значенням параметрів, визначеним (за нашого сформулювання) з умови $K(\alpha, \beta, \gamma) \equiv C$.

H. Résal у своїй статті, присвяченій тому самому питанню, приходить до висновку, що розв'язок задачі Понселе за умов Horvath-а має силу в поширеній області, яка визначається круговим конусом, описаним навколо тріедра Horvath-а. Ясно, що й цей висновок виявляється неспроможний при $\sphericalangle A_3 > \sphericalangle A_1 + \sphericalangle A_2$, бо тут роль області $\mathfrak{A}$ відіграє поверхня сферичного сегмента, що її обмежує описаний навколо трикутника Δ круг, і тоді, при $K \equiv C$, матимемо:

$$\mu = \frac{\cos O}{\cos CA_1} = \operatorname{sc} CA_1,$$

і ця величина напевне є більша від $\operatorname{sc} A_1 D$![1]).

§ 9. Задача Понселе при умовах А. Маркова

При умовах А. Маркова область $\mathfrak{A}$ на сфері Σ перетворюється на сферичний чотирикутник $P_1 P_2 P_3 P_4$ (див. рис. 2), що його боки є лінії перерізу першого октанта сфери Σ з чотирма площинами

$$x = \lambda_1 y, \; x = \mu_1 y, \; z = \lambda_2 y, \; z = \mu_2 y, \tag{130}$$

де $0 < \lambda_1 < \mu_1$; $0 < \lambda_2 < \mu_2$. Марков у свому мемуарі досліджує (при нашому сформулюванні) лише такі значення параметрів α, β, γ, які відповідають положенням точки $K(\alpha, \beta, \gamma)$ в самій області $\mathfrak{A}$ чотирикутника $P_1 P_2 P_3 P_4$ — в той час як, за прямим змістом поставленої ним задачі, область допущених положень точки $K(\alpha, \beta, \gamma)$ завжди є ширша від цілого першого октанта сфери Σ. В дійсності розв'язок, що його знайшов покійний академік, цілком відповідає поставленій ним задачі, і до того ж він

[1]) Легко бачити, що розв'язок, який відповідає, загальним чином, системі значень параметрів, що їх відношення визначаються з умови $K(\alpha, \beta, \gamma) \equiv P$, де P — будьяка точка дуги CD (рис. 1), в дійсності має силу лише в межах сферичної зони між двома колами, описаними на сфері з центра P сферичними радіусами PA_1 та PD. Лише в тому єдиному випадку, коли візьмемо $K(\alpha, \beta, \gamma) \equiv D$, сферична зона перетворюється на поверхню сферичного сегмента, що їй справді відповідає тоді конус. Але цей розв'язок зовсім не має уже нічого спільного з результатом Horvath-а! Зауважмо до речі, що при $K(\alpha, \beta, \gamma) \equiv D$ дістаємо на сфері область найбільшої площі.

є завжди єдиний. Але з міркувань, викладених в самому мемуарі А. Маркова, це зовсім не випливає.

На базі наших попередніх міркувань цілком обгрунтоване трактування задачі, тотожне в своїх кінцевих результатах з розв'язанням самого Маркова, добувається досить просто.

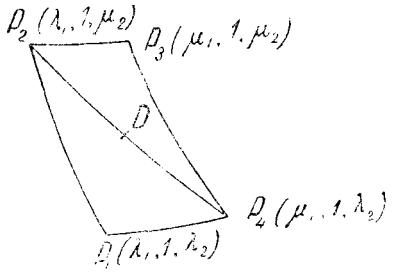

Рис. 2.

Почнімо з деякого ознайомлення з елементами сферичного чотирикутника $P_1P_2P_3P_4$. Його вершки є $P_1(\lambda_1, 1, \lambda_2)$, $P_2(\lambda_1, 1, \mu_2)$, $P_3(\mu_1, 1, \mu_2)$, $P_4(\mu_1, 1, \lambda_2)$, де в дужках указано напрямні параметри відповідних радіусів-векторів OP_i $(i = 1, 2, 3, 4)$. Ми можемо легко пересвідчитись, що кути сферичного чотирикутника при вершках P_1 та P_3 є тупі, а при вершках P_2 та P_4 — гострі. Для цього визначмо напрямні косинуси внутрішніх нормалей до граней чотиригранного кута $OP_1P_2P_3P_4$, що має вершок у початку координат. Позначаючи через M_i $(i = 1, 2, 3, 4)$ відповідні додатні нормуючі множники, матимемо:

для площі OP_1P_2: $\cos\alpha_1 = (\mu_2 - \lambda_2)M_1$, $\cos\beta_1 = -\lambda_1(\mu_2 - \lambda_2)M_1$, $\cos\gamma_1 = 0$;

" " OP_2P_3: $\cos\alpha_2 = 0$, $\cos\beta_2 = \mu_2(\mu_1 - \lambda_1)M_2$, $\cos\gamma_2 = -(\mu_1 - \lambda_1)M_2$;

" " OP_3P_4: $\cos\alpha_3 = -(\mu_2 - \lambda_2)M_3$, $\cos\beta_3 = \mu_1(\mu_2 - \lambda_2)M_3$, $\cos\gamma_3 = 0$;

" " OP_4P_1: $\cos\alpha_4 = 0$, $\cos\beta_4 = -\lambda_2(\mu_1 - \lambda_1)M_4$, $\cos\gamma_4 = (\mu_1 - \lambda_1)M_4$.

Звідси безпосередньо встановлюються знаки косинусів кутів між внутрішніми нормалями до суміжних граней чотиригранного кута. Знаки ж косинусів кутів між самими гранями (тобто також і косинусів кутів сферичного чотирикутника $P_1P_2P_3P_4$) будуть, очевидно, протилежні до них.

Після того, як це встановлено, геометричне уявлення, в зв'язку з результатами попереднього параграфа, безпосередньо підказує дальший шлях дослідження розглядуваного питання. Коли позначити літерою D середину діагоналі P_2P_4 (яка сполучає вершки гострих кутів), можливі лише такі два випадки:

1°. Жодна з двох сферичних віддалей DP_1 та DP_3 не перевищує величину $DP_2 = DP_4 = \frac{1}{2}P_2P_4$

2°. Бодай одна із згаданих двох сферичних віддалей, наприклад, DP_1, є більша від DP_2.

В першому випадку, згідно з міркуваннями параграфа 8, матимемо водночас

$$\measuredangle P_1 \geqslant \measuredangle P_1P_2P_4 + \measuredangle P_1P_4P_2 \text{ та } \measuredangle P_3 \geqslant \measuredangle P_3P_2P_4 + \measuredangle P_3P_4P_2$$

А це значить, що задача [1]) мінімізації величини (126), якщо поставити її

[1]) *Ми пам'ятаємо, що, знайшовши систему значень* α, β, γ [„точку $K(\alpha, \beta, \gamma)$"], *яка мінімізує величину* $\Omega(\alpha, \beta, \gamma)$ *для даної області* $\mathfrak{A}$, *ми завжди добуваємо негайно шукане найкраще наближене представлення за допомогою корективного множника* (31).

для сферичних трикутників $P_1 P_2 P_4$ та $P_3 P_2 P_4$, для кожного окремо, призводить (згідно з висновками § 8) до однієї і тільки однієї спільної системи значень параметрів, яка визначається з умови

$$K(\alpha, \beta, \gamma) \equiv D, \tag{131}$$

при чому величина (125) для одного і для другого сферичного трикутника хитається в тих самих межах:

$$R = \sqrt{\alpha^2 + \beta^2 + \gamma^2} \cdot \cos 0 \quad \text{та} \quad R = \sqrt{\alpha^2 + \beta^2 + \gamma^2} \cdot \cos DP_2 \tag{132}$$

Ясно, що й для області $P_1 P_2 P_3 P_4$ в цілому матимемо той самий єдиний розв'язок, який визначається з умови (131), при чому показник коливання відносної точності $\Omega(\alpha, \beta, \gamma) = \mathrm{sc}\, DP_2$.

В другому ж випадку ми маємо

$$\measuredangle P_1 < \measuredangle P_1 P_2 P_4 + \measuredangle P_1 P_4 P_2,$$

отже центр C_1 описаного круга сферичного трикутника $P_1 P_2 P_4$ знаходиться всередині цього трикутника (на рис. точку C_1 не позначено). Ми маємо, звичайно, $C_1 P_1 = C_1 P_2 = C_1 P_4$. Умова

$$K'(\alpha, \beta, \gamma) \equiv C_1 \tag{133}$$

визначає єдиний розв'язок розглядуваної задачі щодо мінімізації $\Omega(\alpha, \beta, \gamma)$, коли за область $\mathfrak{A}$ взяти лише трикутник $P_1 P_2 P_4$, при чому величина (125) тут хитається в границях

$$R = \sqrt{\alpha^2 + \beta^2 + \gamma^2} \cos 0 \quad \text{та} \quad R = \sqrt{\alpha^2 + \beta^2 + \gamma^2} \cos C_1 P_2 \tag{134}$$

Якби виявилося, що водночас $C_1 P_3 \leqslant C_1 P_2$, то той самий розв'язок і при тому самому значенні $\Omega = \mathrm{sc}\, C_1 P_2$ матиме силу й для всього сферичного чотирикутника $P_1 P_2 P_3 P_4$, і він тут напевне буде єдиний. Але це справді й є так, оскільки із співвідношень

$$\measuredangle P_1 + \measuredangle P_3 > \measuredangle P_2 + \measuredangle P_4 \quad \text{та} \quad \measuredangle P_1 = \measuredangle C_1 P_2 P_1 + \measuredangle C_1 P_4 P_1$$

з необхідністю виходить нерівність

$$\measuredangle P_3 > \measuredangle C_1 P_2 P_3 + \measuredangle C_1 P_4 P_3,$$

яка й тягне за собою нерівність $C_1 P_3 < C_1 P_2$.

Умови (131) та (133) або аналогічна до останньої (якби, замісць $DP_1 > DP_2$, мали $DP_3 > DP_2$), з дальшим застосуванням корективного множника (31) розділу I, і дають повне розв'язання задачі А. Маркова в обох можливих випадках.

Арифметична реалізація розв'язань задачі Понселе за умов Horvath-а та Маркова, поданих нами в геометричній формі, виконується, очевидно, безпосередньо за допомогою елементарного апарату аналітичної геометрії.

РОЗДІЛ IV

НАБЛИЖЕНІ ЛІНІЙНІ ВИРАЗИ ДЛЯ ДЕЯКИХ ЯВНИХ АЛГЕБРИЧНИХ ФУНКЦІЙ, ДРОБОВИХ ТА ІРАЦІОНАЛЬНИХ — ВІД КІЛЬКОХ ЗМІННИХ. ЗАСТОСУВАННЯ ДО РОЗВ'ЯЗУВАННЯ АЛГЕБРИЧНИХ РІВНЯНЬ

§ 10. Нехай спершу треба визначити параметри α, β, γ в наближеній формулі

$$\frac{Ax^{m+1}+By^{m+1}+Cz^{m+1}}{Ax^m+By^m+Cz^m}=\alpha x+\beta y+\gamma z \tag{135}$$

так, щоб вона давала для дробу, написаного на лівій стороні, найкраще, за критерієм відносної похибки, лінійне наближене представлення в області

$$x \geqslant y \geqslant z \geqslant 0, \quad x \neq 0, \tag{136}$$

при чому показник m та коефіцієнти A, B, C позначають якісь дані додатні числа.

Подібно до того, як в розділі II, ми приходимо до задачі найкращого наближеного представлення функції двох змінних

$$f_m(y, z)=\frac{A+By^{m+1}+Cz^{m+1}}{A+By^m+Cz^m} \tag{137}$$

в області типу (T)

$$1 \geqslant y \geqslant z \geqslant 0 \tag{136'}$$

за допомогою лінійного виразу

$$\varphi_1(y, z;\ \alpha, \beta, \gamma)=\alpha+\beta y+\gamma z \tag{138}$$

Система (5) тут дає:

$$\alpha=1,\ \beta=\gamma=0; \quad \varphi_1(y, z;\ 1, 0, 0) \equiv 1 \tag{139}$$

В даному разі безпосередньо видно, що $f_m(y, z)<1$ в області (136′), отже умова (A) розділу I має місце.

Визначаючи maximum виразу

$$R(y, z; 1, 0, 0)=\frac{1}{f_m(y, z)}=\frac{A+By^m+Cz^m}{A+By^{m+1}+Cz^{m+1}} \tag{140}$$

в області (136′), приходимо насамперед до системи рівнянь

$$\left.\begin{aligned} \frac{my^{m-1}}{A+By^m+Cz^m}&=\frac{(m+1)y^m}{A+By^{m+1}+Cz^{m+1}} \\ \frac{mz^{m-1}}{A+By^m+Cz^m}&=\frac{(m+1)z^m}{A+By^{m+1}+Cz^{m+1}} \end{aligned}\right\} \tag{141}$$

яка показує, що стаціонарних точок розглядуваного виразу (140) всередині області (136′) немає: справді, якщо $y, z \neq 0$, із системи (141) виходить $y = z$. Отож, вираз (140) набирає свойого найбільшого значення на якійсь із сторін [1]) трикутника (136′): $y = z$ або $y = 1$, або, нарешті, $z = 0$

При $y = z$, шукаючи maximum функції

$$R(y, y; 1, 0, 0) = \frac{A + (B + C) y^m}{A + (B + C) y^{m+1}} \tag{140′}$$

при $0 \leqslant y \leqslant 1$, приходимо до рівняння виду

$$\lambda y^{m+1} = m - (m + 1) y, \tag{142}$$

де $\lambda = \frac{B + C}{A} = \lambda_1$ (індекс запроваджений для зручності дальших позначень)

Для дальшого мають значення, крім випадку $\lambda > 0$, ще випадки, де $0 \geqslant \lambda > -1$. Рівняння (142) в зазначених випадках має завжди точно один корінь y_0 на інтервалі (0,1) (див. рис. 3). Корінь цей, як легко бачити, то більший, що більше m (при даному λ) або що менше λ (при даному m). Зближаючи межі $0 < y_0 < 1$ однократним застосуванням regula falsi та поправки Ньютона, знайдемо нові межі $\frac{m}{m + 1 + \lambda}$ та $\frac{m}{m + 1}$. Отже матимемо:

$$\left.\begin{aligned} &\frac{m}{m + 1 + \lambda} < y_0 < \frac{m}{m + 1} \quad \text{при } \lambda > 0 \\ &\frac{m}{m + 1} < y_0 < \frac{m}{m + 1 + \lambda} \quad \text{при } 0 > \lambda > -1 \end{aligned}\right\} \tag{143}$$

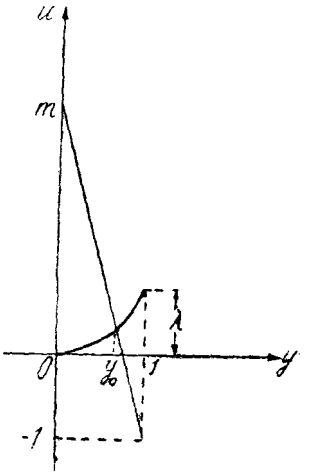

Рис. 3.

Підставляючи у вираз (140′), замість y, корінь y_0 рівняння (142) при $\lambda = \lambda_1 = \frac{B + C}{A}$, знайдемо значення

$$R_{\max} = \frac{1 + \lambda_1 y_0^m}{1 + \lambda_1 y_0^{m+1}} = \frac{y_0 + \lambda_1 y_0^{m+1}}{y_0 (1 + \lambda_1 y_0^{m+1})} =$$

$$= \frac{y_0 + m - (m + 1) y_0}{y_0 [1 + m - (m + 1) y_0]} = \frac{m}{(m + 1) y_0} \tag{144}$$

Шукаючи далі maximum виразів $R(1, z; 1, 0, 0)$ та $R(y, 0; 1, 0, 0)$ при $0 \leqslant z \leqslant 1$ або $0 \leqslant y \leqslant 1$, ми знову прийдемо до рівнянь типу (142), але при менших значеннях параметра λ, а саме $\lambda_2 = \frac{C}{A + B}$ та $\lambda_3 = \frac{B}{A}$ отже відповідні значення коренів z_0 або y_0 тут будуть більші, а для R дістанемо значення менші, ніж при дослідженні виразу (140′). Отож

[1]) Тут і в дальшому інтерпретуємо y, z як декартові координати точки на площині.

вираз (144), де y_0 є корінь рівняння (142) при $\lambda = \lambda_1$, виражає maximum maximorum функції $R(y, z; 1, 0, 0)$ в усій області (136').

Отже, зважаючи на (143), маємо:

$$1 < R_{\max} < 1 + \frac{\lambda_1}{m+1} \tag{144'}$$

Остаточна формула найкращого наближення буде (при $x \geqslant y \geqslant z \geqslant 0$, $x \neq 0$):

$$\frac{Ax^{m+1} + By^{m+1} + Cz^{m+1}}{Ax^m + By^m + Cz^m} = \sigma x \tag{145}$$

де, згідно з (33),

$$\sigma = \frac{2}{1 + R_{\max}}$$

Максимальна відносна похибка при цьому буде $1 - \sigma$ [1])

При даних λ та m можна було б, очевидно, в кожному окремому випадку обчислити y_0 і далі $R_{\max}$ та σ з якою завгодно точністю. Коли ж візьмемо навіть явно перебільшене значення $R_{\max} = 1 + \frac{\lambda_1}{m+1}$, максимальна відносна похибка формули (145) при відповідному значенні $\sigma = \frac{1}{1 + \frac{\lambda_1}{2(m+1)}}$ буде у всякому разі менша за $\frac{\lambda_1}{2(m+1)}$, де $\lambda_1 = \frac{B+C}{A}$

Зауважмо також, що при $\sigma = 1$ максимальна відносна похибка наближеної формули (145) буде напевне менша за $\frac{\lambda_1}{m+1}$

Розгляньмо тепер певну модифікацію попередньої задачі.

Якщо, при $m > 0$, коефіцієнти B та C не мають однакового знака, умова (A) розділу I вже не справджується. Для тих застосувань, які ми тут маємо на увазі [2]), являють інтерес наближені формули типу (145) також при неоднакових знаках коефіцієнтів A, B, C, при чому ми припустимо в такому разі $|B| + |C| < |A|$ [3]). Найкращу формулу цього

[1]) Не важко пересвідчитись, що знайдений розв'язок є єдиний, дарма що точка $M(\eta, \zeta)$, в якій $R(y, z; 1, 0, 0)$ набирає найбільшого значення, знаходиться на границі трикутника (136'), а саме на стороні $y = z$. Для цього досить взяти на увагу, що, згідно з загальними висновками основної теореми § 1, будьякий інший розв'язок, в неоднорідній формі, міг би бути лише виду $f_m(y, z) \approx \sigma + \mu(y - z)$. Для такої наближеної формули, позначаючи відносну похибку через $e(y, z)$, нерівність $e(1,0) \leqslant 1 - \sigma$ має місце лише при $\mu \geqslant 0$. З другого ж боку, звертаючи увагу на те, що градієнт функції $f_m(y, z)$, як легко перевірити, перетворюється на нуль у точці $M(\eta, \zeta)$, візьмімо на внутрішній нормалі до сторони $y = z$ точку (η', ζ') безмежно близьку до (η, ζ). Тоді ми матимемо $e(\eta', \zeta') \geqslant -(1 - \sigma)$ лише при $\mu \leqslant 0$. Зіставляючи, бачимо, що найкраще наближення маємо виключно при $\mu = 0$.

[2]) Див. нижче.

[3]) За цього припущення знаменник та чисельник дробу (137) напевне відмінні від нуля скрізь в області (136'). Крім того, вирази $\lambda_1 = \frac{B+C}{A}$, $\lambda_2 = \frac{C}{A+B}$, $\lambda_3 = \frac{B}{A}$ будуть своєю абсолютною величиною менші за одиницю.

типу одержимо щоразу, згідно з (31), взявши $\sigma = \frac{2}{R_{\max} + R_{\min}}$ Значення виразу (140) в стаціонарних точках на боках трикутника (136′), як і в попередній задачі, матимуть вид $\frac{m}{(m+1)y_0}$ або $\frac{m}{(m+1)z_0}$ де y_0 або z_0 корінь рівняння типу (142), якщо замінимо в останньому λ на одно із значень λ_1, λ_2 або λ_3. Тут буде, як легко бачити, $R_{\max} \geqslant 1$, $R_{\min} < 1$. Максимальна відносна похибка одержаної так найкращої наближеної формули типу (145) [1]) буде, згідно з (28), $\varepsilon = \frac{R_{\max} - R_{\min}}{R_{\max} + R_{\min}}$. Коли ж ми взяли б навіть $\sigma = 1$, то максимальна відносна похибка формули (145) при цьому (не оптимальному) значенні σ буде, як легко бачити [2]), у всякому разі менша за величину

$$\max \left| \frac{\lambda_i}{m-1} \right| \text{ при } i = 1, 2, 3 \tag{146}$$

До аналогічних результатів приводить розгляд загальнішого дробу

$$\frac{A_1 x_1^{m+1} + A_2 x_2^{m+1} + \cdots + A_n x_n^{m+1}}{A_1 x_1^m + A_2 x_2^m + \cdots + A_n x_n^m} \quad \text{при } x_1 \geqslant x_2 \geqslant \ldots \geqslant x_n \geqslant 0,\ x_1 \neq 0$$

де, при $m > 0$, коефіцієнти $A_1, A_2, \ldots, A_n$ або всі одного знаку, або різних знаків, але в останньому разі ми припустимо $|A_2| + |A_3| + \cdots + |A_n| < |A_1|$. Роль сторін трикутника (136′) тут належить „рубам“ понадтетраедра, тобто відповідним одномірним частинам його границі. Замість трьох значень λ_1, λ_2, λ_3, матимемо тут $\frac{n(n-1)}{2}$ значень параметра λ, що їхні вирази будуть виду

$$\frac{A_r + A_{r+1} + \cdots + A_s}{A_1 + A_2 + \cdots + A_{r-1}}, \quad 1 < r \leqslant s \leqslant n$$

Розглянуті задачі пов'язуються із своєрідним методом обчислення найбільшого кореня алгебричного рівняння, що його ідея належить Даніелеві Бернуллі [3]).

За методом Бернуллі, найбільший за модулем корінь рівняння обчисляється як границя, при $m \to \infty$, відношення $\frac{S_{m+1}}{S_m}$, де S_m, як функція індекса m,

[1]) Проте, це, при $B : C < 0$, вже не буде найкраща наближена формула ширшого типу (135), хоч відміна в певних випадках може бути дуже мала.

[2]) Враховуючи нерівності, подібні до (144′).

[3]) Commentarii Acad. Sc. Petropol. 1732. Цитуємо за книгою Whittaker and Robinson — The calculus of observations, 1928 (є рос. переклад).

задовольняє рекурентне співвідношення (різницеве рівняння)

$$S_{m+n} + a_1 S_{m+n-1} + a_2 S_{m+n-2} + \cdots + a_n S_m = 0, \tag{147}$$

при чому $a_1, a_2, \ldots, a_n$ є коефіцієнти даного алгебричного рівняння

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n = 0 \tag{148}$$

Відомо, що S_m при таких умовах виражається завжди як певна лінійна однорідна комбінація степенів коренів рівняння:

$$S_m = A_1 x_1^m + A_2 x_2^m + \cdots + A_n x_n^m, \tag{149}$$

де коефіцієнти $A_1, A_2, \ldots, A_n$ при цілочислових значеннях індекса m, можемо розглядати як певні сталі, які визначаються добором значень S_0, $S_1, \ldots, S_{n-1}$.

Наведені в цьому параграфі міркування дозволяють насамперед, для рівняння з додатними коренями, одержати в деяких випадках оцінку точності обчисленого за методом Бернуллі найбільшого кореня x_1, незалежно від значень відношень між коренями $x_1 : x_2 : \ldots : x_n$; а в першому з розглядуваних двох випадків можна також із попередніх міркувань дістати деякі вказівки, щоб трохи скорегувати значення кореня, обчисленого за методом Бернуллі:

1°. Коли, нічого ближче не знаючи про значення коренів $x_1 \geqslant x_2 \geqslant x_3 \geqslant \ldots \geqslant x_n \geqslant 0$, ми надаємо $S_0, S_1, \ldots, S_{n-1}$ значення відповідних сум Ньютона. В цьому разі $A_1 = A_2 = \ldots = A_n = 1$, отже маємо розглянутий вище випадок додатних коефіцієнтів.

2°. Коли, при $x_1 > x_2 > x_3 > \ldots > x_n > 0$, ми вже маємо одержане за будьяким іншим або за цим самим методом (варіант 1) наближене значення $x_1 \approx \xi_1$. Тоді, взявши $S_0 = 1$, $S_1 = \xi_1$, $S_2 = \xi_1^2, \ldots, S_{n-1} = \xi_1^{n-1}$, ми знаходитимемось, взагалі, в умовах другого розглянутого вище випадку, а саме: A_1 близьке до одиниці, а інші коефіцієнти $A_2, \ldots, A_n$ є малі своєю абсолютною величиною в порівненні з A_1. Проте, не знаючи ближче значень $A_1, A_2, \ldots, A_n$, в цьому разі оцінка точності поліпшеного за методом Бернуллі значення x_1 матиме переважно якісний характер.

В слідуючому параграфі ми даємо детальне трактування деяких аналогічного характеру питань, що стосуються практично важливішого методу Греффе.

§ 11. Взявши, для конкретності, рівняння 3-го степеня

$$x^3 + k_1 x^2 + k_2 x + k_3 = 0, \tag{150}$$

припустімо, що всі три його корені x_1, x_2 та x_3 є дійсні і, крім того, для першого міркування, нехай вони будуть також усі додатні.

При застосуванні методу Греффе, за наближені абсолютні значення коренів, як відомо, беруть такі симетричні функції коренів:

$$F_m(x_1, x_2, x_3) = \sqrt[m]{x_1^m + x_2^m + x_3^m} \approx |x_1|$$

$$\Phi_m(x_1, x_2, x_3) = \sqrt[m]{\frac{x_1^m x_2^m + x_1^m x_3^m + x_2^m x_3^m}{x_1^m + x_2^m + x_3^m}} \approx |x_2| \qquad (151)$$

$$\Psi_m(x_1, x_2, x_3) = \sqrt[m]{\frac{x_1^m x_2^m x_3^m}{x_1^m x_2^m + x_1^m x_3^m + x_2^m x_3^m}} \approx |x_3|$$

де $m = 2^k$, k — число ціле додатне. Застосування методу Греффе натрапляє, як відомо, на істотні утруднення, коли модулі деяких коренів рівняння є дуже близькі між собою. Маючи на увазі, з одного боку, трохи поліпшити точність методу для таких випадків, а з другого боку — визначити граничні похибки, незалежні од відношень $x_1 : x_2 : x_3$, які a priori не можна вважати відомими, ми шукатимемо насамперед для виразів F_m, Φ_m та Ψ_m найкращі лінійні наближені представлення виду $\alpha x_1 + \beta x_2 + \gamma x_3$ з оцінкою їхньої точності.

а) Щодо функції $F_m(x_1, x_2, x_3) = \sqrt[m]{x_1^m + x_2^m + x_3^m}$ при $x_1 \geqslant x_2 \geqslant x_3 \geqslant 0$ то розв'язання даного питання вже маємо в § 5 розділу II цієї роботи, а саме: найкраще наближене представлення буде:

$$\sqrt[m]{x_1^m + x_2^m + x_3^m} = \frac{2}{1 + R_{\max}}\left[x_1 + (\sqrt[m]{2} - 1)\, x_2 + (\sqrt[m]{3} - \sqrt[m]{2})\, x_3\right] \qquad (152)$$

де

$$R_{\max} = \left[1 + (\sqrt[m]{2} - 1)^{\frac{m}{m-1}} + (\sqrt[m]{3} - \sqrt[m]{2})^{\frac{m}{m-1}}\right]^{\frac{m-1}{m}}$$

при чому максимальна відносна похибка тут $\varepsilon_0 = 1 - \frac{2}{1 + R_{\max}} = \frac{R_{\max} - 1}{R_{\max} + 1}$

При великих значеннях m ми матимемо з точністю до величин вищого порядку малості щодо $\frac{1}{m}$ [1]):

$$R_{\max} \approx \sqrt[m]{3} \approx 1 + \frac{1}{m}\log 3; \quad \varepsilon_0 \approx \frac{1}{2m}\log 3 \qquad (153)$$

б) Розглядаючи, в другу чергу, функцію Ψ_m, шукаємо формулу найкращого наближення виду

$$\Psi_m(x_1, x_2, x_3) = \sqrt[m]{\frac{x_1^m x_2^m x_3^m}{x_1^m x_2^m + x_1^m x_3^m + x_2^m x_3^m}} = \alpha x_1 + \beta x_2 + \gamma x_3$$

[1]) Рівності $R_{\max} - 1 \approx \sqrt[m]{3} - 1 \approx \frac{1}{m}\log 3$; $\varepsilon_0 \approx \frac{1}{2m}\log 3$ можна розглядати як асимптотичні при $m \to \infty$. Зауважмо ще, що значення $R_{\max} \approx \sqrt[m]{3}$, $\varepsilon_0 \approx \frac{\sqrt[m]{3} - 1}{\sqrt[m]{3} + 1}$ є наближені значення з перебільшенням.

при $x_1 \geqslant x_2 \geqslant x_3 > 0$, або найкраще наближене представлення функції двох змінних

$$\psi_m(x_2, x_3) = \sqrt[m]{\frac{x_2^m x_3^m}{x_2^m + x_3^m + x_2^m x_3^m}} \tag{154}$$

в області

$$1 \geqslant x_2 \geqslant x_3 > 0 \tag{155}$$

за допомогою лінійного виразу $\alpha + \beta x_2 + \gamma x_3$. Легко бачити, що вимога скінченності відносної похибки, з одного боку, при $x_2 = x_3 \to 0$, а з другого — при $x_2 = \text{const}$, $x_3 \to 0$, визначає значення $\alpha = 0$, $\beta = 0$. Отже шуканий лінійний вираз може відрізнятись лише сталим множником σ від x_3. Розглядаючи відношення

$$R(x_2, x_3; 0, 0, 1) = \sqrt[m]{\frac{x_2^m + x_3^m - x_2^m x_3^m}{x_2^m}} = \sqrt[m]{1 + \left(\frac{x_3}{x_2}\right)^m + x_3^m}, \tag{156}$$

безпосередньо бачимо, що воно в області (155) має своїми точними межами числа 1 та $\sqrt[m]{3}$. Отже, $R_{\max} = \sqrt[m]{3}$, $R_{\min} = 1$. Згідно з (31), приходимо до формули найкращого наближення

$$\sqrt[m]{\frac{x_1^m x_2^m x_3^m}{x_1^m x_2^m + x_1^m x_3^m + x_2^m x_3^m}} = \frac{2}{\sqrt[m]{3} + 1} \cdot x_3 \tag{157}$$

при $x_1 \geqslant x_2 \geqslant x_3 > 0$, при чому максимальна відносна похибка, згідно з (28), буде:

$$\varepsilon_0 = \frac{\sqrt[m]{3} - 1}{\sqrt[m]{3} + 1} \approx \frac{1}{2m} \log 3 \tag{158}$$

в) Трохи делікатнішого дослідження вимагає випадок функції Φ_m. Шукаючи, при $x_1 \geqslant x_2 \geqslant x_3 \geqslant 0$, формулу найкращого наближення виду

$$\Phi_m(x_1, x_2, x_3) = \sqrt[m]{\frac{x_1^m x_2^m + x_1^m x_3^m + x_2^m x_3^m}{x_1^m + x_2^m + x_3^m}} = \alpha x_1 + \beta x_2 + \gamma x_3,$$

або найкраще наближене представлення функції двох змінних

$$\varphi_m(x_2, x_3) = \sqrt[m]{\frac{x_2^m + x_3^m + x_2^m x_3^m}{1 + x_2^m + x_3^m}} \tag{159}$$

в області

$$1 \geqslant x_2 \geqslant x_3 \geqslant 0 \tag{160}$$

за допомогою лінійного виразу $\alpha + \beta x_2 + \gamma x_3$, насамперед знаходимо, відповідно до системи (5), такі значення параметрів:

$$\alpha = 0, \ \beta = \sqrt[m]{\frac{1}{2}} = L, \ \gamma = 1 - \sqrt[m]{\frac{1}{2}} = M \tag{161}$$

Проте в дальшому схема дослідження розділу I вимагатиме деяких істотних модифікацій через те, що умова (A) параграфа 1 тут не має місця.

Шукаючи точні межі відношення

$$R(x_2,\ x_3;\ 0,\ L,\ M) = (Lx_2 + Mx_3)\sqrt[m]{\frac{1 + x_2^m + x_3^m}{x_2^m + x_3^m + x_2^m x_3^m}} \tag{162}$$

в області (160), ми повинні визначити екстремальні значення цієї функції всередині трикутника[1]) (160), якщо такі є; потім встановити межі її в одномірних областях сторін того самого трикутника: $x_3 = 0$, $x_2 = x_3$ та $x_2 = 1$ і нарешті окремо дослідити поводження функції поблизу особливої точки $x_2 = x_3 = 0$.

Для скорочення запису умовмося замість $R(x_2,\ x_3;\ 0,\ L,\ M)$ писати [при незмінних L, M, що означають числа (161)] коротше $R(x_2,\ x_3)$.

Система рівнянь, що визначає стаціонарні точки виразу $R(x_2, x_3)$ в області (160), приводить до умови $x_2^{2m} + x_3^{2m} + x_2^{2m} x_3^{2m} = 0$, яка показує, що стаціонарних точок функції $R(x_2,\ x_3)$ всередині області (160) немає.

На відрізку $x_3 = 0$, $0 \leqslant x_2 \leqslant 1$ маємо:

$$R(x_2,\ 0) = Lx_2\sqrt[m]{\frac{1 + x_2^m}{x_2^m}} = L\sqrt[m]{1 + x_2^m}$$

Отже

$$L \leqslant R(x_2,\ 0) \leqslant L\sqrt[m]{2}$$

або

$$\sqrt[m]{\frac{1}{2}} \leqslant R(x_2,\ 0) \leqslant 1$$

Далі, на відрізку $x_2 = x_3$, $0 \leqslant x_2 \leqslant 1$:

$$R(x_2,\ x_2) = (L + M)\,x_2\sqrt[m]{\frac{1 + 2x_2^m}{2x_2^m + x_2^{2m}}} = (L + M)\sqrt[m]{\frac{1 + 2x_2^m}{2 + x_2^m}} =$$

$$= (L + M)\sqrt[m]{2 - \frac{3}{2 + x_2^m}}$$

Звідси:

$$(L + M)\sqrt[m]{\frac{1}{2}} \leqslant R(x_2,\ x_2) \leqslant L + M$$

або:

$$\sqrt[m]{\frac{1}{2}} \leqslant R(x_2,\ x_2) \leqslant 1$$

На відрізку $x_2 = 1$, $0 \leqslant x_3 \leqslant 1$ матимемо:

$$R(1,\ x_3) = (L + Mx_3)\sqrt[m]{\frac{2 + x_3^m}{1 + 2x_3^m}}$$

Шукаючи extrema цієї функції, приходимо до рівняння

$$3L\,x_3^{m-1} = 2M(1 + x_3^m + x_3^{2m}) \tag{163}$$

[1]) Змінні x_2 та x_3 інтерпретуємо і тут як декартові координати точки на площині.

Легко пересвідчитись (хоч би користуючись теоремою Декарта), що це рівняння має, при $m \geqslant 2$, на інтервалі $0 \leqslant x_3 \leqslant 1$ точно один корінь, який ми позначимо через z.

Це вже дає змогу зробити певні висновки. А саме, оскільки на кінцях інтервалу $R(1,0) = R(1,1) = 1$, а з другого боку $\frac{d}{dx_3} R(1, x_3) > 0$ при $x_3 = 0$, то при наявності одного лише екстремального значення ми повинні мати на всьому інтервалі $R(1, x_3) \geqslant 1$, при чому значення $R(1, z)$ є максимальне.

Фактичне визначення z є для нас актуальне лише для великих значень m, які є звичайні при застосуванні методу Греффе. Як легко бачити, асимптотичне значення z^{m-1} при $m \to \infty$ буде:

$$z^{m-1} \sim \frac{2}{3} \frac{M}{L} = \frac{2}{3} \left(\sqrt[m]{2} - 1 \right) \sim \frac{2}{3m} \log 2 \tag{164}$$

Отже, вживаючи відомих позначень Bachmann-a — Landau, матимемо:

$$z = 1 - O\left(\frac{\log m}{m}\right); \; z^m \sim z^{m-1} \sim \frac{2}{3m} \log 2$$

Беручи ще до уваги, що $M = 1 - \sqrt[m]{\frac{1}{2}} \sim \frac{1}{m} \log 2$, дійдемо висновку:

$$R(1, z) = (L + Mz) \sqrt[m]{2 - \frac{3z^m}{1 + 2z^m}} = (L + M) \sqrt[m]{2} - \delta_m = \sqrt[m]{2} - \delta_m,$$

де $\delta_m = O\left(\frac{\log m}{m^2}\right) = o\left(\frac{1}{m}\right)$ означає (при $m \geqslant 2$) додатну величину меншу за $\sqrt[m]{2} - 1$, і яка при $m \to \infty$ являє собою нескінченно малу вищого порядку щодо $\frac{1}{m}$.

Звідси маємо на інтервалі $0 \leqslant x_3 \leqslant 1$ такі межі для функції $R(1, x_3)$:

$$1 \leqslant R(1, x_3) \leqslant \sqrt[m]{2} - \delta_m = \sqrt[m]{2} - O\left(\frac{\log m}{m^2}\right) \tag{165}$$

Отже ми пересвідчились, що, поперше, всередині трикутника (160) точок maximum-у або minimum-у виразу (162) зовсім немає і, подруге, що на сторонах трикутника функція $R(x_2, x_3)$ має своїми межами

$$\sqrt[m]{\frac{1}{2}} \leqslant R(x_2, x_3) \leqslant \sqrt[m]{2} - \delta_m = \sqrt[m]{2} - O\left(\frac{\log m}{m^2}\right)^{1)} \tag{166}$$

Нам треба ще пересвідчитись, що і в околі особливої точки $x_2 = x_3 = 0$ величина $R(x_2, x_3)$ ніде не виходить за межі (166).

[1]) Верхню межу $\sqrt[m]{2} - \delta_m$ для кожного даного m можна обчислити з якою завгодно точністю, визначивши корінь $x_3 = z$ рівняння (163) з відповідним ступенем точності В практиці ж застосування методу Греффе величиною δ_m можемо взагалі нехтувати.

Застосовуючи для цього перетворення змінних

$$x_2 = s, \qquad x_3 = st, \tag{167}$$

ми, замість задачі дослідження меж виразу (162) в області (160), маємо задачу дослідити межі перетвореного виразу

$$r(s, t) = R(s, st) = (L + Mt)\sqrt[m]{\frac{1 + s^m + s^m t^m}{1 + t^m + s^m t^m}} \tag{162'}$$

в області прямокутника

$$0 \leqslant s \leqslant 1; \qquad 0 \leqslant t \leqslant 1 \tag{160'}$$

на площі (s, t). Цей вираз являє собою функцію неперервну у всій розглядуваній області. Попереднє дослідження показує, що функція $r(s, t)$ не має екстремальних значень всередині прямокутника (160'), і що на трьох сторонах $t = 0$, $t = 1$, $s = 1$ її величина міститься в межах (166). Нам тут залишається пересвідчитись, що і на останній стороні $s = 0$ величина $r(s, t)$ не виходить за ці межі. Але

$$r(0, t) = (L + Mt)\sqrt[m]{\frac{1}{1 + t^m}}, \tag{168}$$

де, нагадаємо про це, $L = \sqrt[m]{\frac{1}{2}}$, $M = 1 - L$. Шукаючи extrema цього виразу при $0 \leqslant t \leqslant 1$, приходимо до рівняння $Lt^{m-1} = M$, звідки, позначаючи через τ корінь цього рівняння, маємо:

$$\tau = \sqrt[m-1]{\frac{M}{L}}$$

і потім

$$r(0, \tau) = (L + M\tau)^{1 - \frac{1}{m}}\sqrt[m]{L} \tag{169}$$

Оскільки вираз (168) набирає на кінцях інтервалу однакового значення $r(0, 0) = r(0, 1) = L = \sqrt[m]{\frac{1}{2}}$, а з другого боку — безпосередньо видко, що похідна $\frac{d}{dt} r(0, t)$ є додатна при $t = 0$ (та при $m > 1$), то бачимо, як вище, що значення (169) є найбільше, якого вираз (168) досягає при $0 \leqslant t \leqslant 1$. Але значення (169) є, очевидно, менше за одиницю. Найменше ж значення виразу (168), яке має місце на кінцях інтервалу, дорівнює $\sqrt[m]{\frac{1}{2}}$. Тим самим цілком доведено, що вираз $r(s, t)$ не виходить за межі (166) у всій області (160'), і те саме є справедливе щодо виразу (162) у всій області (160).

Отже маємо у всій області (160):

$$R_{\min} = \sqrt[m]{\frac{1}{2}}, \quad R_{\max} = \sqrt[m]{2} - \delta_m = \sqrt[m]{2} - O\left(\frac{\log m}{m^2}\right) \tag{170}$$

при чому R_{max} має місце в точці $(1, z)$ області (160), де z означає корінь рівняння (163).

Тепер ми твердимо: не зважаючи на те, що умова (A) розділу I в даній задачі не справджується щодо функції (159) та лінійного виразу Lx_2+Mx_3, останній вираз має основну екстремальну властивість (9), до того ж він становить єдиний (до довільного сталого множника) розв'язок відповідної extremum-задачі („попередньої задачі" § 1).

Щоб це довести, треба насамперед взяти на увагу, що лише лінійні вирази виду $\beta x_2+\gamma x_3$ (тобто без вільного члена) ми тут маємо розглянути, бо при $\alpha\neq 0$ відносна похибка наближеного представлення $\alpha+$ $+\beta x_2+\gamma x_3$ не залишається навіть скінченна. Досліджуючи відношення $R(x_2, x_3;\ 0,\ \beta,\ \gamma)=\dfrac{\beta x_2+\gamma x_3}{\varphi_m(x_2,\ x_3)}$ на границі області (160), ми насамперед за допомогою викладок цілком аналогічних тим, що були застосовані при розгляді $R(x_2, 0)=R(x_2, 0; 0, L, M)$ та $R(x_2, x_2)=R(x_2, x_2;\ 0, L, M)$, встановлюємо, що, поперше:

$$\frac{R(1,\ 0;\ \ 0,\ \beta,\ \gamma)}{R(0,\ 0;\ \ 0,\ \beta,\ \gamma)}=\frac{\beta\sqrt[m]{2}}{\beta}=\sqrt[m]{2} \tag{171}$$

розуміючи $x_2=x_3=0$ в знаменнику як граничну точку на відрізку $x_3=0$, а подруге:

$$\frac{R(1,\ 1;\ \ 0,\ \beta,\ \gamma)}{R(0,\ 0;\ \ 0,\ \beta,\ \gamma)}=\frac{\beta+\gamma}{(\beta+\gamma)\sqrt[m]{\frac{1}{2}}}=\sqrt[m]{2} \tag{172}$$

де $x_2=x_3=0$ в знаменнику ми вже розуміємо, як граничну точку на відрізку $x_2=x_3$.

Далі, розглядаючи на відрізку $x_2=1$ функцію одного змінного $\varphi_m(1, x_3)=\sqrt[m]{\dfrac{1+2x_3{}^m}{2+x_3{}^m}}$ та її наближене представлення за допомогою лінійного виразу $L+Mx_3$, де значення $L=\sqrt[m]{\dfrac{1}{2}}$, $M=1-L$ відповідають системі (5) для цього випадку, бачимо з одержаних раніше висновків щодо $R(1, x_3)=R(1, x_3;\ 0,\ L,\ M)$, що в даному разі [тобто в задачі, що стосується наближеного лінійного представлення функції одного змінного $\varphi_m(1, x_3)$] умова (A) розділу I задовольняється. На підставі міркувань основної теореми параграфа 1 можемо твердити, що за кожного іншого добору коефіцієнтів β, γ, не пропорціональних до L, M, одно з двох відношень

$$\frac{R(1,\ z;\ 0,\ \beta,\ \gamma)}{R(1,\ 0;\ 0,\ \beta,\ \gamma)}\quad\text{або}\quad\frac{R(1,\ z;0,\ \beta,\ \gamma)}{R(1,\ 1;0,\ \beta,\ \gamma)}\ ^{1)}$$

буде напевне більше за визначене при розгляді $R(1, x_3)=R(1, x_3;\ 0,\ L,\ M)$ значення $R_{max}=\sqrt[m]{2}-\delta_m$.

[1]) z у чисельниках означає корінь рівняння (163).

Зіставляючи це з (171) та (172), приходимо до висновку, що за кожного добору β, γ, не пропорціональних до L, M, граничне значення одного з двох відношень

$$\frac{R(1,\ z;\ 0,\ \beta,\ \gamma)}{R(\delta,\ 0;\ 0,\ \beta,\ \gamma)} \quad \text{або} \quad \frac{R(1,\ z;\ 0,\ \beta,\ \gamma)}{R(\delta,\ \delta;\ 0,\ \beta,\ \gamma)}$$

при $\delta \to 0$ буде більше за $\frac{R_{\max}}{R_{\min}} = \frac{\sqrt[m]{2} - \delta_m}{\sqrt[m]{\frac{1}{2}}}$, отже показник Ω коливання відносної точності наближеного представлення функції $\varphi_m(x_2,\ x_3)$ за допомогою лінійного виразу $\alpha + \beta x_2 + \gamma x_3$ в області (160) має мінімальне значення лише при $\alpha = 0, \beta = L = \sqrt[m]{\frac{1}{2}}$, $\gamma = M = 1 - L$ або при значеннях параметрів пропорціональних до цих, а це ми й хотіли довести.

Згідно з (32), (31), (28), ми тепер одержуємо шукане найкраще наближене представлення у виді:

$$\sqrt[m]{\frac{x_1^m x_2^m + x_1^m x_3^m + x_2^m x_3^m}{x_1^m + x_2^m + x_3^m}} = \frac{2}{\sqrt[m]{\frac{1}{2}} + \sqrt[m]{2} - \delta_m}\left[\sqrt[m]{\frac{1}{2}}\, x_2 + \left(1 - \sqrt[m]{\frac{1}{2}}\right) x_3\right] \quad (173)$$

з максимальною відносною похибкою

$$\varepsilon_0 = \frac{\sqrt[m]{2} - \delta_m - \sqrt[m]{\frac{1}{2}}}{\sqrt[m]{2} - \delta_m + \sqrt[m]{\frac{1}{2}}} \sim \frac{\sqrt[m]{2} - \sqrt[m]{\frac{1}{2}}}{\sqrt[m]{2} + \sqrt[m]{\frac{1}{2}}} \sim \frac{1}{m} \log 2, \quad (174)$$

де, як вище було зазначено,

$$\delta_m = O\left(\frac{\log m}{m^2}\right) = o\left(\frac{1}{m}\right)^{1)} \quad (175)$$

З одержаних лінійних формул найкращого наближення (152), (157) та (173) безпосередньо випливають деякі практичні формули, придатні для обчислення коренів рівняння (150), якщо між ними є корені дуже близькі [2]. Коли, наприклад, в наслідок 9-кратного перетворення Греффе з рівняння (150) одержано рівняння

$$y^3 + c_1 y^2 + c_2 y + c_3 = 0, \quad (176)$$

то, підставляючи на правих сторонах (152), (157) та (173) $m = 2^9 = 512$, а на лівих сторонах відповідні функції від коефіцієнтів рівняння (176), для яких приймемо позначення

$$\sqrt[512]{c_1} = \xi_1, \quad \sqrt[512]{\frac{c_2}{c_1}} = \xi_2, \quad \sqrt[512]{\frac{c_3}{c_2}} = \xi_3, \quad (177)$$

[1]) Пор. попередню виноску щодо $\sqrt[m]{2} - \delta_m$.

[2]) Останній вислів буде уточнено нижче.

матимемо:

$$\left.\begin{aligned}&\xi_3 = 0{,}9989x_3 \text{ (з точн. до } 1{,}1^0/_{00}),\ \xi_2 = 9{,}9986x_2 + 0{,}0014x_3 \text{ (до } 1{,}4^0/_{00}),\\ &\xi_1 = 0{,}9989x_1 + 0{,}0014x_2 + 0{,}0007x_3 \text{ (до } 1{,}1^0/_{00}),\end{aligned}\right\} \quad (178)$$

звідки й дістанемо наближені вирази для коренів:

$$\left.\begin{aligned}&x_3 = 1{,}0011\xi_3 \text{ (до } 1{,}1^0/_{00}),\ x_2 = 1{,}0014\xi_2 - 0{,}0014x_3 \text{ (до } 1{,}4^0/_{00}),\\ &x_1 = 1{,}0011\xi_1 - 0{,}0014x_2 - 0{,}0007x_3 \text{ (до } 1{,}1^0/_{00})\end{aligned}\right\} \quad (179)$$

Бачимо, що ці значення коренів одержуються з ξ_1, ξ_2, ξ_3 (тобто із значень коренів за способом Греффе), за допомогою малих поправок, що їх обчисляти слід у мислі, виражаючи їх однією-двома цифрами. Вживати формули типу (179) при обчисленні кореня x_i (i — одно із значень 1, 2, 3) є рація лише при наявності іншого дуже близького кореня — в даному разі, конкретно кажучи, коли суміжний корінь (один із суміжних коренів) відрізняється від розглядуваного не більше як на 2—3$^0/_{00}$.

Дослідження, що його ми тут провели для рівняння (150), узагальнюється і для рівнянь інших степенів. Загальний результат маємо такий.

Якщо $\xi_1, \xi_2, \ldots, \xi_n$ є наближені значення коренів за методом Греффе для рівняння

$$x^n + k_1x^{n-1} + k_2x^{n-2} + \cdots + k_n = 0 \quad (180)$$

з дійсними (але не обов'язково додатними) коренями одержані після k-кратного перетворення Греффе, то, позначаючи $2 = m$ та розуміючи під $|x_i|$ абсолютне значення кореня x_i, $i, = 1, 2, \ldots, n$, матимемо:

$$\left.\begin{aligned}&\xi_i = \sigma \sqrt[m]{\frac{1}{i}}\Big[\,|x_i| + \left(\sqrt[m]{2-1}\right)|x_{i+1}| + \left(\sqrt[m]{3} - \sqrt[m]{2}\right)|x_{i+2}| + \\ &\qquad + \cdots + \left(\sqrt[m]{n+1-i} - \sqrt[m]{n-i}\right)|x_n|\,\Big]\\ &\qquad (i = 1, 2, \ldots, n)\end{aligned}\right\} \quad (181)$$

де

$$\sigma_i = \frac{2}{\sqrt[m]{\frac{1}{i}} + \sqrt[m]{n+1-i} - (\delta_i)_m} \approx \frac{2}{\sqrt[m]{\frac{1}{i}} + \sqrt[m]{n+1-i}}, \quad (182)$$

$$(\delta_i)_m = O\left(\frac{\log m}{m^2}\right) = o\left(\frac{1}{m}\right), \quad (183)$$

при чому максимальна відносна похибка для ξ_i є:

$$(\varepsilon_i)_0 = 1 - \sigma_i \sqrt[m]{\frac{1}{i}} \sim \frac{1}{2m} \log\left[i(n+1-i)\right] \quad (184)$$

Розв'язуючи (181) щодо $|x_i|$, $i=n, n-1, \ldots, 2, 1$, дістанемо такі наближені вирази для абсолютних значень коренів:

$$\left.\begin{aligned} |x_i| = \frac{\sqrt{i}}{\sigma_i}\xi_i - \Big[\Big(\sqrt[m]{2}-1\Big)|x_{i+1}| + \Big(\sqrt[m]{3}-\sqrt[m]{2}\Big)|x_{i+2}| + \cdots \\ \cdots + \Big(\sqrt[m]{n+1-i}-\sqrt[m]{n-i}\Big)|x_n|\Big] \\ (i=n,\ n-1, \ldots, 2, 1)\ ^{1)} \end{aligned}\right\} \quad (185)$$

при максимальній відносній похибці (184)[2]) і при тих самих значеннях σ_i та $(\delta_i)_m$.

Нижче подаємо маленьку допомічну таблицю, щоб швидко обчисляти радикали, які входять і в ці формули, і в ті, які подаємо далі.

Тепер ми хочемо показати, що максимальна похибка формул (185) практично збігається з максимальною похибкою наступних простіших формул типу $x_i = \mu_i \xi_i$, які ми зараз виведемо.

Взявши знову конкретний випадок рівняння 3-го степеня (150) з дійсними і (для першого міркування) додатними коренями, ставимо собі завдання вивести можливо точніші формули виду

$$x_1 = \mu_1\xi_1,\ x_2 = \mu_2\xi_2,\ x_3 = \mu_3\xi_3, \quad (186)$$

де ξ_1, ξ_2, ξ_3 — ідентичні з виразами $F_m(x_1, x_2, x_3)$, $\Phi_m(x_1, x_2, x_3)$, $\Psi_m(x_1, x_2, x_3)$ [див. (151)].

Щодо кореня x_3, таку формулу фактично вже одержано: вона безпосередньо випливає з (157).

Переходячи далі до кореня x_1, пишемо насамперед наближену формулу

$$\sqrt[m]{x_1^m + x_2^m + x_3^m} \approx x_1 \quad (187)$$

Щоб дослідити її точність, складаємо відношення

$$R_1 = \frac{x_1}{\sqrt[m]{x_1^m + x_2^m + x_3^m}} = \sqrt[m]{\frac{x_1^m}{x_1^m + x_2^m + x_3^m}}$$

Безпосередньо ясно, що це відношення при $x_1 \geqslant x_2 \geqslant x_3 \geqslant 0$, $x_1 \neq 0$ міститься в межах $\sqrt[m]{\frac{1}{3}} \leqslant R_1 \leqslant 1$.

Проте, при визначенні найкращої формули типу (186) безпосереднє значення має для нас обернене відношення $\frac{1}{R_1}$, яке в даному разі міститься

[1]) При $i=n$ вираз в квадратних дужках зовсім відпадає.

[2]) Тут відносну похибку розраховується щодо ξ_i.

між точними межами 1 та $\sqrt[m]{3}$ Запроваджуючи корективного множника, згідно з (31), одержуємо шукану формулу найкращого наближення типу (186) у виді

$$x_1 = \frac{2}{\sqrt[m]{3}+1}\xi_1 \tag{188}$$

при максимальній відносній похибці $\frac{\sqrt[m]{3}-1}{\sqrt[m]{3}+1}$

Трохи детальніше треба і тут спинитися на випадку x_2. Написавши спершу наближену формулу

$$\sqrt[m]{\frac{x_1^m x_2^m + x_1^m x_3^m + x_2^m x_3^m}{x_1^m + x_2^m + x_3^m}} \approx x_2, \tag{189}$$

досліджуємо знову таки її точність, розглядаючи відношення

$$R_2 = x_2 \sqrt[m]{\frac{x_1^m + x_2^m + x_3^m}{x_1^m x_2^m + x_1^m x_3^m + x_2^m x_3^m}}$$

Позначаючи

$$x_2 = x_1 s,\ x_3 = x_2 t$$

де, очевидно,

$$0 \leqslant s \leqslant 1,\ 0 \leqslant t \leqslant 1 \tag{190}$$

матимемо:

$$\left.\begin{aligned} R_2 &= \sqrt[m]{\frac{1+s^m+s^m t^m}{1+t^m+s^m t^m}} = \sqrt[m]{1+\frac{s^m-t^m}{1+t^m+s^m t^m}} \\ \frac{1}{R_2} &= \cdots\cdots\cdots = \sqrt[m]{1+\frac{t^m-s^m}{1+s^m+s^m t^m}} \end{aligned}\right\} \tag{191}$$

Безпосередньо видко, що $\frac{\partial}{\partial t}R_2 < 0, \frac{\partial}{\partial s}\left(\frac{1}{R_2}\right) < 0$ у всій області (190). Звідси:

$$(R_2)_{\max} = (\sqrt[m]{1+s^m})_{\max} = \sqrt[m]{2} \text{ при } t=0,\ s=1$$

$$\left(\frac{1}{R_2}\right)_{\max} = (\sqrt[m]{1+t^m})_{\max} = \sqrt[m]{2} \text{ при } s=0,\ t=1$$

Отже, маємо для $\frac{1}{R_2}$ такі точні межі:

$$\sqrt[m]{\frac{1}{2}} \leqslant \frac{1}{R_2} \leqslant \sqrt[m]{2}$$

в області (190).

А це, згідно з (31), дає нам негайно для формули (189) корективного множника $\frac{2}{\sqrt[m]{2}+\sqrt[m]{\frac{1}{2}}}$ і ми одержуємо потрібну нам формулу найкра-

щого наближення типу (186):

$$x_2 = \frac{2}{\sqrt[m]{2} + \sqrt[m]{\frac{1}{2}}}\,\xi_2 \tag{192}$$

при максимальній відносній похибці $\dfrac{\sqrt[m]{2} - \sqrt[m]{\frac{1}{2}}}{\sqrt[m]{2} + \sqrt[m]{\frac{1}{2}}}$

Наведене дослідження узагальнюється і для рівнянь інших степенів і приводить до такої загальної формули (позначення попередні, див. с. 68)

$$\left.\begin{aligned} &|x_i| = \frac{2}{\sqrt[m]{\frac{1}{i}} + \sqrt[m]{n+1-i}}\,\xi_i, \qquad i = 1, 2, \ldots, n \\ &\text{при максимальній відносній похибці:} \\ &(\varepsilon_i)_0 = \frac{\sqrt[m]{n+1-i} - \sqrt[m]{\frac{1}{i}}}{\sqrt[m]{n+1-i} + \sqrt[m]{\frac{1}{i}}} \sim \frac{1}{2m}\log\left[i(n+1-i)\right] \end{aligned}\right\} \tag{193}$$

Бачимо, що справді точність цієї надзвичайно простої формули практично збігається з точністю формул (185), принаймні оскільки мовиться про максимум похибки.

Наведена нижче невеличка таблиця дозволяє швидко обчисляти радикали, які входять у формули (193), і подані вище формули (185).

Допомічна таблиця

n	2	3	4	5	6
$\sqrt[1024]{n}$	1,000 677	1,001 073	1,001 355	1,001 573	1,001 751
n	7	8	9	10	11
$\sqrt[1024]{n}$	1,001 902	1,002 033	1,002 148	1,002 251	1,002 344

Користуючись цією таблицею, дуже легко обчисляти (в мислі) й різні інші аналогічні радикали з потрібною, для практичного застосування поданих вище формул, точністю:

$$\sqrt[1024]{\frac{1}{6}} = \frac{1}{1{,}0018} \approx 1 - 0{,}0018; \quad \sqrt[1024]{18} = \sqrt[1024]{6} \cdot \sqrt[1024]{3} \approx 1{,}0028$$

$$\sqrt[512]{9} = 1{,}00215^2 \approx 1{,}0043 \text{ і т. д.}$$

Застосовувати формули (193) [або (185)] є рація лише в тих випадках і до тих коренів x_i, які дуже близькі бодай до одного з суміжних (за абсолютним значенням) коренів рівняння. Критерій „близькості“ визначається можливим впливом суміжного кореня на точність наближеного значення, що його дає для обчислюваного кореня метод Греффе. Одержані тут формули мають значення лише для тих випадків, коли вплив суміжного кореня (або сукупний вплив двох чи кількох суміжних коренів) може спричинитись у методі Греффе до похибки порядку $(\varepsilon_i)_0$. Враховуючи можливий вплив суміжних коренів на підставі безпосереднього аналізу виразів типу (151), ми, власне, замість коренів x_i, x_{i+1}, $x_{i-1}, \ldots$, маємо в свойому розпорядженні наближені абсолютні значення ξ_i, ξ_{i+1}, $\xi_{i-1}, \ldots$, і слід мати на увазі, що, скажімо, різниця $|x_i| - |x_{i+1}|$ може бути менша за $\xi_i - \xi_{i+1}$ на величину, що про її порядок можна собі скласти уявлення на підставі одержаних тільки но результатів. Для ближчої оцінки можливого відхилення між $|x_i| - |x_{i+1}|$ та $\xi_i - \xi_{i+1}$ становить істотний інтерес один своєрідний феномен, який виявляється при дослідженні цього питання. Феномен цей можна було б характеризувати як певний „закон розсіяння“ безмежно близьких (за абсолютним значенням) коренів, коли їх визначають методом Греффе. Виявляється, що відношення $\frac{\xi_1}{\xi_{i+1}}$ (при зафіксованому m) не може ніколи упасти нижче певної межі, більшої за одиницю, хоч би які близькі були точні абсолютні значення коренів.

Розгляньмо детальніше це питання у випадку обчислення кореня x_1 рівняння n-го степеня (180).

Позначаючи для скорочення запису $x_i^m = a_i$ $(i = 1, 2, \ldots, n)$, маємо:

$$\left.\begin{aligned} \frac{\xi_1}{\xi_2} &= \sqrt[m]{\frac{(a_1 + a_2 + \cdots + a_n)^2}{a_1 a_2 + a_1 a_3 + \cdots + a_{n-1} a_n}} = \sqrt[m]{\frac{\sum\limits_{r,s}\left[2a_r a_s + \frac{1}{n-1}(a_r^2 + a_s^2)\right]}{\sum\limits_{r,s} a_r a_s}} \\ &\geqslant \sqrt[m]{\frac{\sum\limits_{r,s} 2\left(1 + \frac{1}{n-1}\right) a_r a_s}{\sum\limits_{r,s} a_r a_s}} = \sqrt[m]{\frac{2n}{n-1}} \end{aligned}\right\} \quad (194)$$

де

$s = r + 1, r + 2, \ldots, n; r = 1, 2, \ldots, n - 1$. Величина

$$\sqrt[m]{\frac{2n}{n-1}}$$

являє собою нижню межу відношення $\frac{\xi_1}{\xi_2}$, хоч би які були значення коренів $x_1, x_2, \ldots, x_n$ і відношення між ними[1])

[1]) Виключаючи, звичайно, випадок $x_1 = 0$.

Аналогічно нижні межі (більші від одиниці) існують і для всіх інших відношень $\frac{\xi_2}{\xi_3}, \frac{\xi_3}{\xi_4}, \ldots, \frac{\xi_{n-1}}{\xi_n}$, при чому мають місце такі основні закономірності:

1°. $$\left(\frac{\xi_{n-i}}{\xi_{n-i+1}}\right)_{\min} = \left(\frac{\xi_i}{\xi_{i+1}}\right)_{\min}$$

2°. Крайні відношення $\frac{\xi_1}{\xi_2}$ та $\frac{\xi_{n-1}}{\xi_n}$ мають найбільшу нижню межу $\left(\text{рівну } \sqrt[m]{\frac{2n}{n-1}}\right)$, для інших відношень (пересуваючись від країв до середини в ланцюзі відношень $\xi_1 : \xi_2 : \xi_3 : \ldots : \xi_{n-2} : \xi_{n-1} : \xi_n$) нижня межа трохи зменшується.

РОЗДІЛ V

ДЕЯКІ ІНШІ ЗАДАЧІ

12. Про найкраще наближене представлення конвексної функції за допомогою ординати ламаної

Функція конвексна (опукла вниз) на даному інтервалі — це є субгармонічна[1]) функція у випадку одного незалежного змінного. „Кожна дуга геометричного образу конвексної функції $y=f(x)$ (трактуючи x та y як декартові координати точки на площині) не має жодної точки понад хордою, що сполучає її кінці“ — таке є трохи умовне формою, але у всякому разі досить стисле, визначення конвексної функції[2]). Функція $f(x)$ зветься конкавною (опуклою вгору), якщо функція $-f(x)$ є конвексна. В тому разі, коли функція $f(x)$ має другу похідну, вона є конвексна, якщо $f''(x)\geqslant 0$, і конкавна, якщо $f''(x)\leqslant 0$ в розглядуваному інтервалі.

Безпосереднім предметом дальшого нашого розгляду є функції конвексні, але цілком очевидно, що тут випливають відповідні наслідки і для функцій конкавних.

Як відомо, конвексна на інтервалі (a, b) функція є неперервна на цьому інтервалі або принаймні півнеперервна зверху. Проте, в останньому разі точками розриву можуть бути лише кінці інтервалу. Залишаючи осторонь останній випадок, ми в дальшому маємо на увазі виключно функцію $f(x)$ неперервну на даному інтервалі (a, b). Крім того, щоб уникнути зайвих ускладнень міркування, ми припустимо, що геометричне зображення даної функції не включає жодних прямолінійних частин, але остаточні результати дальшого дослідження мають силу і для таких випадків, коли геометричне зображення включає будьяку кількість — скінченну або нескінченну — прямолінійних відрізків; при цьому ми можемо, звичайно, залишити осторонь тривіальний випадок, коли в задачі апроксимації за допомогою ординати m — ланкової ламаної $(m\leqslant n)$ геометричне зображення самої даної функції являє собою ламану зазначеного типу.

З інших загальних властивостей конвексних функцій зазначмо ще такі: конвексна функція має в кожній точці цілком визначену (узагальнену) похідну праворуч, а також визначену похідну ліворуч, при чому перша не менша за другу. Не важко до того ж встановити, що нерівність правої та лівої похідних може мати місце щонайбільше в счисленній множині

[1]) Див. означення в § 3.

[2]) Див. P. Montel. Sur les fonct. convexes et les fonct. sousharmoniques. Journal de Mathématiques pures et appliquées, 1928, fac. № 1.

точок[1]). Якщо $x_1 < x_2$, то похідна ліворуч у точці x_2 не менша за похідну праворуч у точці x_1. Конвексна функція може мати всередині інтервалу не більше як один мінімум і не може мати жодного максимуму. Сума або різниця конвексної функції і функції лінійної є знову функція конвексна. Кожна ламана, вписана в дугу геометричного зображення конвексної функції, сама являє опуклу вниз ламану.

Виходячи з цих загальних властивостей конвексних функцій, розвиваємо дальші міркування в такій послідовності.

1°. Якщо (α, β) є якийсь інтервал, що міститься в (a, b), сполучімо хордою MN (див. рис. 4) точки з абсцисами α, β на геометричному зображенні конвексної функції $y = f(x)$. Максимум абсолютної величини вертикальної віддалі між дугою геометричного зображення та хордою на інтервалі (α, β) хай буде $h = h(\alpha, \beta)$. Легко зрозуміти на підставі самого визначення конвексної функції, що h є неперервна ростуча функція від β при зафіксованому α, неперервна спадна функція від α при зафіксованому β. Аналогічно залежить від β та α максимальний відхил найкращого лінійного представлення функції $f(x)$ за допомогою лінійної функції на інтервалі (α, β): справді, згідно з висновками розділу I, ясно, що цей максимальний відхил дорівнює $\frac{h}{2}$.

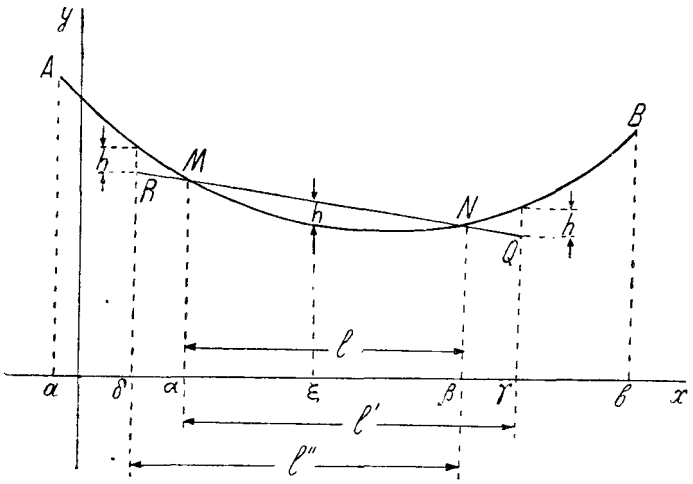

Рис. 4.

Позначмо через l горизонтальну проекцію хорди: $l = \beta - \alpha$. З попереднього легко бачити, що при даному α величина l є однозначна [визначена в межах між $h = 0$ та $h = h(\alpha, b)$], до того ж ростуча й неперервна[2]) функція від h. Залежність величини l від α та від h виражатимемо записом $l = l(\alpha, h)$. Безпосередньо ясне є існування аналогічної залежності величини l від β та h.

2 . Продовжимо хорду MN за точку N (той самий рис. 4). Оскільки вертикальна віддаль між дугою AB та продовженою хордою сама являє конвексну функцію абсциси[3]), що її мінімум [єдиний на інтервалі (a, b)] має місце при $x = \xi$, праворуч від точки N ця віддаль монотонно росте. Якщо $b - \beta \geqslant \beta - \xi$, то вона напевне досягне знову величини h — уже із

[1]) Див., напр. P. Montel, op. cit.

[2]) Слід пам'ятати, що $h(\alpha, \beta)$ є, за даного α, неперервна й монотонна функція від β, а це і має своїм наслідком аналогічну залежність β від h.

[3]) Як різниця між конвексною та лінійною функціями.

знаком плюс — при якомусь значенні $x=\gamma<b$ [1]). Не важко зрозуміти, що проекція відрізка продовженої ламаної $\gamma-\alpha=l'$ при даному α є однозначна функція від h, визначена в межах від $h=0$ до якогось значення, напевне більшого за $h\left(\alpha, \frac{\alpha+b}{2}\right)$, ростуча та неперервна [2]). Ми писатимемо $l'=l'(\alpha, h)$. Із сказаного ясно, що

$$l(\alpha, h)<l'(\alpha, h)<2l(\alpha, h) \tag{195}$$

Ордината відрізка MQ, як легко бачити, дає найкраще наближене представлення функції $y=f(x)$ на інтервалі (α, γ) за допомогою лінійної функції, підпорядкованої додатковій умові — збігатися точно з функцією $f(x)$ при $x=\alpha$ [3]). Звідси випливає, що

$$h(\alpha, \beta)>\frac{1}{2}h(\alpha, \gamma) \tag{196}$$

Аналогічні міркування мають силу і при продовженні хорди MN за лівий кінець M. Позначаючи в цьому разі горизонтальну проекцію $\beta-\delta$ відрізка NR через l'', ми можемо розглядати l'' або як функцію від β та h, що допускало б точно таке саме трактування, як раніше $l'(\alpha, h)$, абож як функцію від δ та h. В останньому разі писатимемо $l''=l''(\delta, h)$. Це знову буде, при даному δ, однозначна, ростуча та неперервна функція від h, визначена в межах від $h=0$ до якогось значення напевне більшого за $\frac{1}{2}h(\delta, b)$.

3°. Задавши собі певні значення α та h, впишімо в дугу геометричного зображення конвексної функції $y=f(x)$ ламану з вершками, що відповідають таким абсцисам:

$$\alpha;\ \alpha+l(\alpha, h)=\alpha_1;\ \alpha_1+l(\alpha_1, h)=\alpha_2; \ldots; \alpha_{m-1}+l(\alpha_{m-1}, h)=\alpha_m \tag{197}$$

Цю m-ланкову ламану позначмо $\mathfrak{P}_m[\alpha, h]$, а її горизонтальна проекція нехай буде $L_m(\alpha, h)$, тобто:

$$l(\alpha, h)+l(\alpha_1, h)+\cdots+l(\alpha_{m-1}, h)=L_m(\alpha, h) \tag{198}$$

Для задачі, що її розв'язання буде дано в наступному n° 4, важливо встановити існування такого значення $h=h_0$, щоб задовольнялася рівність

$$L_m(a, h_0)=b-a, \tag{199}$$

тобто щоб вписана ламана $\mathfrak{P}_m[a, h_0]$ точно відповідала повній дузі AB рис. 4.

[1]) Щоб у цьому пересвідчитись, досить уявити собі окремо звичайне графічне зображення розглядуваної вертикальної віддалі як функції від абсциси та застосувати до неї останню з вищеперелічених загальних властивостей конвексних функцій.

[2]) Це випливає з аналогічних властивостей функції $h=\psi(\alpha, l')$ при зафіксованому α. Щодо вивчення останньої функції, то воно досить легке в тому разі, коли при збільшенні l' ми здійснюємо визначення $\psi(\alpha, l')$, продовжуючи попередній відрізок MQ і потім повертаючи його навколо точки M.

[3]) Справді, будьяке повертання прямої MQ навколо точки M неминуче збільшує вертикальний відхил або в точці $x=\xi$, або в точці $x=\gamma$.

Почнімо з розгляду деяких властивостей функції $L_m(\alpha, h)$. ми твердимо, що нерівність $h_1 < h_2$ завжди має своїм ом $L_m(\alpha, h_1) < L_m(\alpha, h_2)$ [1]). Справді, якщо для якихось значень α, h_1 при $h_1 < h_2$, ми мали б, супротивно до нашого твердження, $L_m(\alpha, h_1) \geqq L_m(\alpha, h_2)$, то проекція бодай однієї з ланок ламаної $\mathfrak{P}_m[\alpha, h_1]$ обіймала б проекцію якоїсь ланки ламаної $\mathfrak{P}_m[\alpha, h_2]$, але це явно суперечило б відомому нам характерові залежності функції $h(\alpha, \beta)$ від її аргументів.

Розвиваючи далі основну думку останнього міркування, припустімо, що маємо якусь вписану m-ланкову ламану з абсцисами вершків $\alpha, \alpha_1, \alpha_2, \ldots, \alpha_m$, яка не справджує умов типу (197). Нехай будуть $h_{\min}$ та $h_{\max}$ найменша та найбільша з m величин $h(\alpha, \alpha_1)$, $h(\alpha_1, \alpha_2), \ldots, h(\alpha_{m-1}, \alpha_m)$.

Ми тепер твердимо, що

$$\left.\begin{aligned} L_m(\alpha, h_{\min}) &< \alpha_m - \alpha \\ L_m(\alpha, h_{\max}) &> \alpha_m - \alpha \text{ [2])} \end{aligned}\right\} \tag{200}$$

Доводження здійснюється цілком аналогічно до попереднього твердження.

Нарешті покажімо, що при нескінченно малому збільшенні величини h, при даному α, збільшення $L_m(\alpha, h)$ теж є нескінченно мале. Справді, нехай будуть $\alpha, \alpha_1, \alpha_2, \ldots, \alpha_m$ абсциси вершків ламаної $\mathfrak{P}_m[\alpha, \underline{h}]$, де $\underline{h}$ — якесь дане значення величини h, при чому, звичайно, припускається $\alpha_m < b$. Задавши довільно малу додатну величину ε, при єдиній умові $\alpha_m + \varepsilon < b$, ми можемо послідовно пересунути трохи праворуч точки $\alpha_{m-1}, \alpha_{m-2}, \ldots, \alpha_1$ так, щоб усі m величин

$$h(\alpha'_{m-1}, \alpha_m + \varepsilon),\ h(\alpha'_{m-2}, \alpha'_{m-1}),\ h(\alpha'_{m-3}, \alpha'_{m-2}), \ldots, h(\alpha, \alpha'_1)$$

були трохи більші від попередньої величини $\underline{h}$, при чому $\alpha_m + \varepsilon$, α'_{m-1}, $\alpha'_{m-2}, \ldots, \alpha'_1$, α є нові абсциси вершків. Якщо $h_{\min}$ є найменша із зазначених m величин, то досить буде взяти додатне число $\eta < h_{\min} - \underline{h}$, щоб було $L_m(\alpha, \underline{h} + \eta) < L_m(\alpha, \underline{h}) + \varepsilon$, а це й закінчує потрібне доведення.

На підставі одержаних висновків ми маємо змогу легко встановити основний факт існування значення h_0, яке точно задовольняє рівність (199), і водночас з'ясувати шлях, щоб визначити значення h_0 з довільним ступенем точності.

Виходячи спершу з припущення, що значення h_0 існує, ми можемо встановити для нього дві вихідні межі, вписавши в дугу AB (рис. 4) будьяку m-ланкову ламану з абсцисами вершків $a, \alpha_1, \alpha_2, \ldots, \alpha_{m-1}, b$. Найменша і найбільша із m величин

$$h(a, \alpha_1),\ h(\alpha_1, \alpha_2), \ldots, h(\alpha_{m-1}, b)$$

[1]) Припускаючи, звичайно, що збудування відповідних ламаних є можливе, коли не виходити за межі інтервалу (a, b).

[2]) Припускаючи, що $\mathfrak{P}_m[\alpha, h_{\max}]$ ще „уміщається" на інтервалі (a, b).

дадуть нам згадані дві межі: нехай буде $h_{\min} = \underline{h}$, $h_{\max} = \overline{h}$. Отже, маємо припускаючи існування h_0):

$$\underline{h} < h_0 < \overline{h} \tag{201}$$

Далі поступово зближаємо межі шляхом послідовних випробувань. Взявши значення $\frac{\underline{h}+\overline{h}}{2}$, будуємо $\mathfrak{P}_m \left| a, \frac{\underline{h}+\overline{h}}{2} \right|$. Якщо збудування можливе й відповідне значення $L_m < b - a$, то $\underline{h}' = \frac{\underline{h}+\overline{h}}{2}$ дає нову нижню межу для h_0. Навпаки, якщо збудування не можливе [1]), то $\overline{h}' = \frac{\underline{h}+\overline{h}}{2}$ дає нову верхню межу. В обох випадках дістаємо нову пару меж для h_0, що їх різниця вдвічі зменшена. Продовжуючи аналогічно, приходимо до нескінченного процесу, якому відповідає, незалежно від правдивості припущення щодо існування h_0, якесь граничне значення H такої властивості: при $h = H - \varepsilon$ збудування $\mathfrak{P}_m [a, h]$ є можливе і $L_m (a, h) < b - a$; при $h = H + \varepsilon$ зазначене збудування не можливе; при цьому ε означає довільно мале додатне число. Абсциси

$$a + L_1 (a, H - \varepsilon) = a + l (a, H - \varepsilon);\ a + L_2 (a, H - \varepsilon);$$
$$a + L_3 (a, H - \varepsilon); \ldots; a + L_m (a, H - \varepsilon)$$

при $\varepsilon \to 0$, монотонно змінюючись (зростаючи), наближаються до певних границь $a_1, a_2, a_3, \ldots, a_m$, які не виходять за межі інтервалу (a, b). Абсциса a_m повинна спадатись з b, бо при $a_m < b$ ламана $\mathfrak{P}_m [a, H + \varepsilon]$ була б ще здійсненна при досить малому додатному ε. А коли так, то $h_0 = H$: ми водночас довели існування значення h_0 і з'ясували шлях послідовного наближення, щоб його визначити.

4°. **Задача 1.** Дано конвексну на інтервалі (a, b) функцію $y = f(x)$. Треба збудувати на тому самому інтервалі таку m-ланкову ламану, де $m \leqslant n$ (n означає дане число), щоб максимум абсолютного значення різниці, на інтервалі (a, b), між $f(x)$ та ординатою ламаної був щонайменший.

Розв'язання. Будуємо на інтервалі (a, b) вписану ламану $\mathfrak{P}_m [a, h_0]$ попереднього п° [2]). Пересунувши цю ламану вертикально вниз на величину $\frac{1}{2} h_0$, ми й одержимо шукану ламану. Розв'язок є єдиний.

Справді, кожна окрема ланка одержаної ламаної, згідно з висновками розділу I, дає найкраще лінійне наближення на відповідній частині інтервалу (a, b). Кожна інша ламана матиме хоча б одну таку ланку, горизонтальна проекція якої обійматиме [3]) проекцію якоїсь ланки одержа-

[1]) Тобто ламана $\mathfrak{P}_m$ „не уміщається" на інтервалі (a, b).

[2]) Тобто таку вписану ламану, кожна сторона якої має однаковий максимум вертикального відхилення h_0 від дуги геометричного образу функції $y = f(x)$.

[3]) Обіймання частинного інтервалу (α, β) інтервалом (α', β') ми тут розуміємо широко: $\alpha' \leqslant \alpha < \beta \leqslant \beta'$.

ної нами ламаної, в той час як сами розглядувані дві ланки не збігатимуться. Отже кожна інша ламана бодай на цій ділянці інтервалу (a, b) матиме максимум вертикального відхилення від дуги AB більший за $\frac{h_0}{2}$, в той час як для здобутої нами ламаної максимум відхилення на всьому інтервалі (a, b) точно дорівнює $\frac{h_0}{2}$

5°. Розгляньмо тепер три інших гатунків ламані, що залежать від значень α та h, а саме:

поперше, ламану (типу $MM_1M_2M_3$ на рис. 5), яку ми позначимо через $\mathfrak{P}_m^*[\alpha, h]$; її вершки мають своїми абсцисами значення

$$\left.\begin{array}{l} \alpha;\ \alpha + l'(\alpha, h) = \alpha_1;\ \alpha_1 + l(\alpha_1, 2h) = \alpha_2;\ \alpha_2 + l(\alpha_2, 2h) = \alpha_3; \\ \ldots;\ \alpha_{m-2} + l(\alpha_{m-2}, 2h) = \alpha_{m-1};\ \alpha_{m-1} + l''(\alpha_{m-1}, h) = \alpha_m \end{array}\right\} \quad (202)$$

а ординатами крайні два вершки мають ординати самої дуги AB, решта ж вершків — ординати дуги AB зменшені на величину h;

подруге, ламану $\mathfrak{P}_m'[\alpha, h]$ (типу MM_1M_2 на рис. 5), яка відрізняється від попередньої тим, що останній вершок тепер матиме абсцису $\alpha_{m-1} + l(\alpha_{m-1}, 2h)$ і ординатою — ординату дуги AB, зменшену на величину h;

нарешті, ламану $\mathfrak{P}_m''[\alpha, h]$ (типу $M_1M_2M_3$ на рис. 5), яка відрізняється від $\mathfrak{P}_m^*[\alpha, h]$ тим, що ордината першого вершка дорівнює ординаті дуги AB, зменшеній на h, і, крім того, абсциса другого вершка дорівнює $a + l(a, 2h)$ з відповідним переміщенням інших вершків.

Горизонтальні проекції цих ламаних позначмо відповідно чечерез $L^*_m(\alpha, h)$, $L'_m(\alpha, h)$ $L''_m(\alpha, h)$. Основне значення і тут має встановлення того факта, що для кожної з цих трьох ламаних можна дібрати таке значення $h = h_0$, щоб задовольнялася одна із трьох рівностей:

$$L^*_m(a, h_0) = b - a \quad (199^*)$$

або

$$L'_m(a, h_0) = b - a \quad (199')$$

або, нарешті,

$$L''_m(a, h_0) = b - a \quad (199'')$$

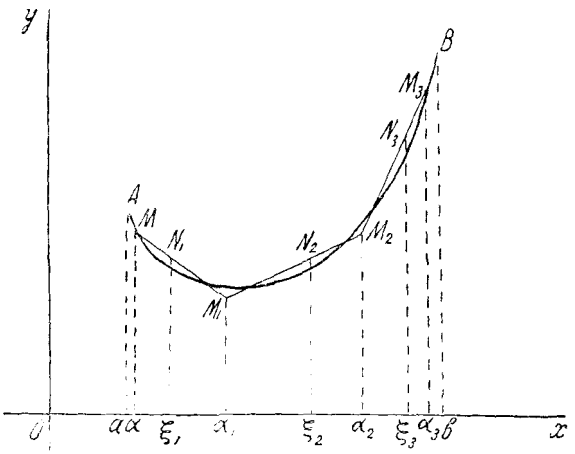

Рис. 5.

Потрібне для цього дослідження здійснюється за таким самим планом як в п° 3, але з деякими модифікаціями міркування. Ми тут зупинимось для прикладу лише на виводі нерівностей, аналогічних до (200), у випадку ламаної $\mathfrak{P}^*_m$. В цьому разі ми повинні мати на увазі якусь дану m-ланкову ламану, що крайні її вершки лежать на дузі AB геометричного образу конвексної функції, а ланки ламаної перетинають дугу максимально можливе число разів, тобто крайні ланки — кожна в одній точці, а про-

міжні — кожна в двох точках. Інакше кажучи, це буде ламана виду аналогічного[1]) до $MM_1M_2M_3$ рис. 5, але з тією істотною відміною, що для неї абсолютні вертикальні відхилення $\mathfrak{h}$ від дуги AB при $x = \xi_1, \alpha_1, \xi_2, \alpha_2, \ldots, \alpha_{m-1}, \xi_m$ будуть між собою взагалі нерівні. Позначмо через $\mathfrak{h}_{\min}$ та $\mathfrak{h}_{\max}$ точні межі, між якими містяться всі ці $2m-1$ відхилів за абсолютною величиною. Тоді, позначаючи через α та α_m абсциси крайніх вершків розглядуваної ламаної, матимемо такі нерівності:

$$\left.\begin{aligned} L^*_m(\alpha, \mathfrak{h}_{\min}) &< \alpha_m - \alpha \\ L^*_m(\alpha, \mathfrak{h}_{\max}) &> \alpha_m - \alpha \end{aligned}\right\} \qquad (200^*)$$ [2])

Доводячи, наприклад, першу з цих двох нерівностей, досить зауважити, що в противному разі ламана $\mathfrak{P}^*_m[\alpha, \mathfrak{h}_{\min}]$ мала б бодай одну ланку, що її проекція обіймала б[3]) проекцію якоїсь ланки, скажімо, $M_iN_{i+1}M_{i+1}$ розглядуваної ламаної, в той час як сами ланки не збігалися б, але це означало б, що максимальний відхил найкращого наближеного лінійного представлення функції $f(x)$ на інтервалі (α_i, α_{i+1}), величиною не більший за $\mathfrak{h}_{\min}$, не містився б у вузькому розумінні між найменшим та найбільшим (за абсолютним значенням) із трьох[4]) вертикальних відхилів відрізка $M_iN_{i+1}M_{i+1}$ від дуги AB в точках M_i, N_{i+1} та M_{i+1}, що буде суперечити відомій теоремі de la Vallée-Poussin-a[5]).

6°. Аналогічно до n° 4 можна сформулювати три інші задачі (назвімо їх задача II, задача III та задача IV) щодо найкращої[6]) апроксимації на інтервалі (a, b) конвексної функції $f(x)$ за допомогою ординати m-ланкової ламаної, де $m \leqslant n$, підпорядкованої додатковій умові, згідно з якою або обидва крайні вершки ламаної повинні лежати на дузі AB геометричного образу конвексної функції або лише один із двох крайніх вершків. Розв'язок дають у всіх трьох задачах відповідно ламані типу $\mathfrak{P}^*_m[a, h_0]$, $\mathfrak{P}_m[a, h_0]$ або $\mathfrak{P}''_m[a, h_0]$, де h_0 задовольняє відповідну рівність (199*), (199') або (199''). Доводження, цілком аналогічне до поданого в n° 4, ми опускаємо.

[1]) Це дозволяє нам далі вжити аналогічних позначень для вершків та точок екстремального відхилу розглядуваної ламаної ($MN_1M_1N_2M_2 \ldots M_{m-1}N_mM_m$) та їх абсцис.

[2]) Припускаючи, що друга з цих двох ламаних ще „уміщається" на інтервалі (a, b).

[3]) Принаймні в широкому розумінні, як це пояснює друга виноска n° 4.

[4]) Кажучи про три відхили, припускаємо, що відрізок $M_iN_{i+1}M_{i+1}$ не є один із крайніх відрізків ламаної. В супротивному разі ми тут говорили б про два відмінні від нуля відхили.

[5]) Про цю теорему див., у більш загальному сформулюванні, розд. VI, § 16, друге твердження теореми I. Теорема є застосувальна, mutatis mutandis, і до того випадку лінійної апроксимації (при додатковій умові), про який говорилося в n° 2 даного параграфа.

[6]) За критерієм абсолютної похибки. Проте можливість узагальнити на випадок критерія відносної похибки, якщо дуга AB не перетинає осі абсцис, безпосередньо очевидна.

Розв'язання задачі I n° 4 поширюється, при очевидній модифікації сформулювання, і на випадок функції конкавної на даному інтервалі. Задачі ж II, III та IV даного n° можуть мати значення для того загальнішого випадку, коли даний інтервал поділяється на скінченне число ділянок, на яких функція $y=f(x)$ є альтернативно конвексна та конкавна.

Взагалі ж практичний сенс розглянутих чотирьох задач щодо апроксимації даної неперервної функції за допомогою ординати ламаної можна собі з'ясувати, поставивши їх у зв'язок, з одного боку, з ідеєю, яка належить H. Lebesgue [1]), щодо способу наближеного представлення неперервних функцій у виді поліномів через попередню наближену заміну даної функції на ординату ламаної, а з другого боку — з задачею найкращого поліноміального наближеного представлення функції $|x|$, яка стала широко відома після визначних теоретичних дослідів C. de la Vallée-Poussin-a, С. Н. Бернштейна [2]) та інших видатних сучасних математиків і на якій в аспекті фактичного обчислення ми ще зупинимось в § 21 цієї роботи.

§ 13. Деякі питання апроксимації, що стосуються елементарних трансцендентних виразів з одним змінним і мають значення при розв'язуванні відповідних трансцендентних рівнянь

1°. Виведімо насамперед дві формули найкращого наближення, за критерієм відносної похибки та за критерієм абсолютної похибки, для функції $f(x)=e^x$ з допомогою лінійного виразу

$$\varphi(x;\ \alpha,\ \beta)=\alpha+\beta x \tag{203}$$

на інтервалі (0,1). Оскільки $f''(x)>0$, умова (A) розділу I тут напевне має місце. Застосовуючи загальний шлях, з'ясований в §§ 1, 2, знайдемо такі формули найкращого наближення:

а)
$$e^x=\frac{2}{1+R_{\max}}\left[1+(e-1)x\right] \tag{204}$$

де $R_{\max}=\dfrac{e-1}{e^{\frac{e-2}{e-1}}}\approx 1,131$, — при максим. відносній похибці $\varepsilon_0=\dfrac{R_{\max}-1}{R_{\max}+1}$,

б)
$$e^x=1+(e-1)x-\frac{1}{2}D_{\max} \tag{205}$$

де $D_{\max}=1+(e-1)\log(e-1)-(e-1)\approx 0{,}212$, — при максимальній абсолютній похибці $\delta_0=\dfrac{1}{2}D_{\max}$

[1]) Bulletin des Sciences Mathématiques, 1898. Також: E. Borel. Leç. sur les fonct. de var. réelles et les dév. en séries de polyn.; E. Goursat. Cours d'Analyse Mathém., 1910, t. I, n° 206 (є рос. перекл.).

[2]) Див. особливо фундаментального значення мемуар — S. Bernstein. Sur la meill. approx. de $|x|$ par des polyn. de degrés donnés. Acta Mathematica, 1913.

Замінюючи константи їх наближеними значеннями, матимемо:

а') $$\left.\begin{array}{l} e^x = 0{,}939 + 1{,}613x \approx 0{,}94 + 1{,}61x \\ \text{з точн. до } 6\frac{1}{4}\ \% \text{ при } 0 \leqslant x \leqslant 1 \end{array}\right\} \qquad (204')$$

б') $$\left.\begin{array}{l} e^x = 0{,}894 + 1{,}718x \approx 0{,}89 + 1{,}72x \\ \text{з граничною абсолютною похибкою} \\ \delta_0 = 0{,}11 \text{ при } 0 \leqslant x \leqslant 1 \end{array}\right\} \qquad (205')$$

Множачи обидві сторони одержаних наближених формул на e^k та замінюючи x на $x-k$, ми дістаємо формули найкращого наближеного представлення для тієї самої функції e^x на інтервалі $k \leqslant x \leqslant k+1$. При цьому максимальна абсолютна похибка помножиться на e^k, а максимальна відносна похибка не зміниться.

Зауважмо, що при розв'язуванні трансцендентних рівнянь, крім лінійних формул, можуть знайти застосування й поліноміальні наближені представлення вищих степенів. Така є формула

$$\boxed{\left.\begin{array}{l} e^x = 0{,}8459x^2 + 0{,}8549x + 1{,}0087 \\ \text{з точн. до } 0{,}0088 \text{ при } 0 \leqslant x \leqslant 1 \end{array}\right\} \quad (206)}$$

складена за методом розділу VII цієї роботи, яка дає найкраще наближене представлення e^x у виді полінома другого степеня при $0 \leqslant x \leqslant 1$.

2°. Найкраще наближене представлення x через лінійну однорідну комбінацію $\sin x$ та sinus versus x.

Розв'язуючи трансцендентні рівняння типу

$$Ax + B\sin x + C\cos x + D = 0\ ^{1)}, \qquad (207)$$

які часто зустрічаються в астрономії та механіці [2]), може бути доцільно, для першого наближення, перетворити таке рівняння на інше, що безпосередньо розв'язується в скінченному виді, користуючись наближеною формулою виду

$$x = \alpha \sin x + \beta(1 - \cos x) \qquad (208)$$

Ми хочемо визначити параметри α та β при такій умові, щоб максимальна відносна похибка формули при $0 < x \leqslant \frac{\pi}{2}$ була щонайменша, розраховуючи відносну похибку або щодо величини самого x, або щодо величини $\sin x$.

[1]) A, B, C, D — сталі; $A \neq 0$, $|B| + |C| \neq 0$.

[2]) До цього типу належить, наприклад, відоме рівняння Кеплера.

Запровадивши нову незалежну змінну $\operatorname{tg}\frac{x}{2}=t$, зводимо формулу (208) до виду

$$f(t)=\frac{(1+t^2)\operatorname{arctg} t}{t}=\alpha+\beta t \tag{209}$$

при $0\leqslant t\leqslant 1$. Не важко пересвідчитись, що $f''(t)>0$ при $0<t<1$, отже умова (A) розділу I має місце. Загальним шляхом розділу I знайдемо дві формули найкращого лінійного наближеного представлення — одну за критерієм відносної похибки і другу — за критерієм абсолютної похибки:

$$f(t)=\frac{2}{1+R_{\max}}\left[1+\left(\frac{\pi}{2}-1\right)t\right] \tag{210}$$

та

$$f(t)=1+\left(\frac{\pi}{2}-1\right)t-\frac{1}{2}D_{\max} \tag{211}$$

Відношення $R=\dfrac{\left[1+\left(\frac{\pi}{2}-1\right)t\right]t}{(1+t^2)\operatorname{arctg} t}$ набирає найбільшого значення при $t=\omega$, де ω — корінь трансцендентного рівняння

$$\left(\pi-2+\frac{1}{\omega}-\omega\right)\operatorname{arctg}\omega=1+\left(\frac{\pi}{2}-1\right)\omega \tag{212}$$

Розв'язуючи його таблично-графічним способом із дальшим застосуванням regula falsi, знаходимо $\omega=0{,}409$, звідки $R\approx 1{,}115$. Отже маємо першу формулу найкращого наближення:

$$\frac{(1+t^2)\operatorname{arctg} t}{t}=\frac{2}{2{,}115}(1+0{,}5708\,t)=0{,}946+0{,}540\,t \tag{210'}$$

з точн. до $5^1/_2\%$ при $0\leqslant t\leqslant 1$.

Щодо виразу $D=1+\left(\frac{\pi}{2}-1\right)t-\frac{(1+t^2)\operatorname{arctg} t}{t}$, то він набирає найбільшого значення при $t=\upsilon$, де υ задовольняє рівняння

$$\left(1-\frac{1}{\upsilon^2}\right)\operatorname{arctg}\upsilon+\frac{1}{\upsilon}=\frac{\pi}{2}-1 \tag{213}$$

Розв'язуючи його таким самим способом, як (212), знайдемо $\upsilon=0{,}463$, звідки $D_{\max}\approx 0{,}127$, що призводить до другої формули найкращого наближення:

$$\frac{(1+t^2)\operatorname{arctg} t}{t}=0{,}936+0{,}571\,t \tag{211'}$$

граничною абсолютною похибкою 0,064 при $0\leqslant t\leqslant 1$.

Повертаючись до первісної змінної x, дістанемо із (210') та (211') дві остаточні формули:

$$x=0{,}946\sin x+0{,}540(1-\cos x) \tag{214}$$

з точн. до $5^1/_2\%$ величини x при $0 < x \leqslant \frac{\pi}{2}$, та

$$x = 0{,}936 \sin x + 0{,}571 \ (1 - \cos x) \tag{215}$$

з точністю до 6,4% величини $\sin x$ при $0 < x \leqslant \frac{\pi}{2}$.

Перша з цих двох формул дає найбільшу точність приблизно в межах $0 \leqslant x \leqslant \frac{\pi}{6}$, а для інтервалу $0 \leqslant x \leqslant \frac{\pi}{2}$ в цілому друга формула взагалі краща.

Наближені формули (214)—(215) дають змогу, коли шукаємо перше наближення для кореня рівняння (207), перетворити його до виду

$$a \sin x + b \cos x + c = 0, \tag{216}$$

що розв'язується безпосередньо. Але в рівнянні (216) можна ще позбутися й вільного члена c за допомогою однієї з таких двох формул найкращого наближення, які виражають 1 як лінійну однорідну комбінацію синуса φ та косинуса φ:

$$1 = 0{,}960 \cos \varphi + 0{,}398 \sin \varphi, \tag{217}$$

з точн. до 4% при $0 \leqslant \varphi \leqslant \frac{\pi}{4}$, та

$$1 = 0{,}955 \cos \varphi + 0{,}414 \sin \varphi \tag{218}$$

з точн. до 4,5% величини $\cos \varphi$ при $0 \leqslant \varphi \leqslant \frac{\pi}{4}$

Ці дві формули можна одержати безпосередньо з формул (71_1) та (110_1) розділу II, поклавши $\frac{x}{\sqrt{x^2+y^2}} = \cos \varphi$, $\frac{y}{\sqrt{x^2+y^2}} = \sin \varphi$.

§ 14. Лінійні формули найкращого [1]) наближення для x^n, xy, xyz і т. д. та деякі їх застосування

1°. Почнімо з формул наближеного представлення для x^n.

Нехай буде спершу $-1 \leqslant x \leqslant 1$.

Коли $n = 2k$ — ціле додатне парне число, шукана формула найкращого наближення є, очевидно,

$$x^{2k} = \frac{1}{2} \text{ з точністю до } \pm \frac{1}{2} \tag{219}$$

Коли ж $n > 1$ — число непарне, умова (A) розділу I не має місця, але розв'язок легко знайти з дальших міркувань. Поперше, шуканий вираз

[1]) За критерієм абсолютної похибки.

виду $\alpha+\beta x$ повинен тут являти собою непарну функцію, подібно до даної функції x^n, тобто $\alpha=0$ [1]). До того ж, визначаючи другий параметр β, можна за симетрією обмежитись інтервалом $0\leqslant x\leqslant 1$. Згідно з попереднім зауваженням (§ 12, n° 2), для найкращого наближеного представлення максимум вертикального відхилу між кривою та прямою при $x=\xi$ (див. рис. 6) повинен дорівнювати, абсолютним своїм значенням, вертикальному відхилові при $x=1$. Ця умова приводить до рівняння

$$\beta\left(\frac{\beta}{n}\right)^{\frac{1}{n-1}}-\left(\frac{\beta}{n}\right)^{\frac{n}{n-1}}=1-\beta \qquad (220)$$

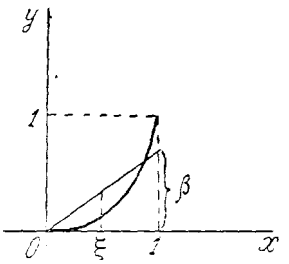

Рис. 6.

Позначаючи ж $\frac{\beta}{n}=z$, маємо:

$$(n-1)z^{\frac{n}{n-1}}=1-nz \qquad (220')$$

Графічне представлення безпосередньо показує, що рівняння (220′) має точно один корінь на інтервалі $0\leqslant z\leqslant\frac{1}{n}$. Корінь цей з достатньою для нашої мети точністю знайдемо шляхом двократного застосування поправки Ньютона, виходячи із першого наближеного значення $z\approx\frac{1}{n}$. В такий спосіб, позначаючи індексом залежність від показника n, одержимо:

$$\left.\begin{aligned}\beta_5&=5z_5=0{,}67355\approx\frac{2}{3}\\ \beta_7&=7z_7=0{,}63511\approx 0{,}64\\ \beta_9&=9z_9=0{,}61155\approx 0{,}61\end{aligned}\right\} \qquad (221)$$

Ми пропустили тут значення $\beta_3=\frac{3}{4}$, добре відоме a priori як коефіцієнт відповідного члена в кубічному поліномі Чебишова $\frac{1}{4}\cos(3\arccos x)$. Зауважмо також, що $\lim\limits_{n\to\infty}\beta_n=0{,}5$

Максимальна абсолютна похибка наближеної формули $x^n=\beta_n x$ на інтервалі $-1\leqslant x\leqslant 1$ (при непарному додатному n) є $\delta_0=1-\beta_n$.

Проте для застосувань, які ми маємо на увазі, буде далеко зручніше мати готові формули найкращого лінійного наближеного представлення функції x^n на інтервалі $-h\leqslant x\leqslant h$ при довільному h. Зробивши заміну змінного

$$x=ht \qquad (222)$$

[1]) З міркувань симетрії легко бачити, що при заміні будьякого наближеного представлення x^n у виді βx на $\alpha+\beta x$ при $\alpha\neq 0$ максимум абсолютної похибки неминуче повинен збільшитись або на інтервалі $(-1,0)$, або на інтервалі $(0,1)$.

при $-1 \leqslant t \leqslant 1$, з формули $t^n = \beta_n t$ (з точн. до $1-\beta_n$, при непарному n), одержимо перетворену формулу

$$\left(\frac{x}{h}\right)^n = \beta_n \frac{x}{h} \text{ з точн. до } 1-\beta_n, \tag{223}$$

звідки маємо таку загальну формулу для випадку непарного показника:

$$\boxed{x_n = \beta_n h^{n-1} x \text{ з точн. до } (1-\beta_n) h^n \text{ при } -h \leqslant x \leqslant h \quad (224)}$$

Аналогічно для випадку парного показника:

$$\boxed{x^n = \frac{1}{2} h^n \text{ з точн. до } \frac{1}{2} h^n \text{ при } -h \leqslant x \leqslant h \quad (225)}$$

В розгорненому виді матимемо ряд формул:

$$\left.\begin{array}{lllll}
x^2 = \frac{1}{2} h^2 & \text{з точн. до} & \frac{1}{2} h^2 & \text{при} & -h \leqslant x \leqslant h \\
x^3 = \frac{3}{4} h^2 x & \text{„ „} & \text{„ } \frac{1}{4} h^3 & \text{„ „ „ „} & \\
x^4 = \frac{1}{2} h^4 & \text{„ „} & \text{„ } \frac{1}{2} h^4 & \text{„ „ „ „} & \\
x^5 = \frac{2}{3} h^4 x & \text{„ „} & \text{„ } \frac{1}{3} h^5 & \text{„ „ „ „} & \\
\cdots & & & & \\
x^7 = 0{,}64\, h^6 x & \text{„ „} & \text{„ } 0{,}37\, h^7 & \text{„ „ „ „} & \\
\cdots & & & & \\
x^9 = 0{,}61\, h^8 x & \text{„ „} & \text{„ } 0{,}39\, h^9 & \text{„ „ „ „} &
\end{array}\right\} \tag{226}$$

2°. Лінійні формули найкращого наближення для xy, xyz, $xyzt$ і т. д.

Нехай спершу треба знайти найкраще наближене представлення для функції xy в області $1 \geqslant x \geqslant y \geqslant 0$ у виді лінійного виразу $\alpha + \beta x + \gamma y$. Функція xy, як легко бачити, є кратно-субгармонічна в розглядуваній області і на трьох складових частинах її границі, отже умова (A) розділу I напевне має місце. Система (5) тут дає: $\alpha = \beta = 0$, $\gamma = 1$. Вираз $D = y - xy = y(1-x)$ доходить свойого максимального значення $D_{\max} = \frac{1}{4}$ в єдиній точці $x = y = \frac{1}{2}$, яка лежить на границі області апроксимації. Як найпростіший розв'язок, одержуємо таку формулу найкращого наближення:

$$xy = y - \frac{1}{8} \tag{227}$$

з точн. до $\frac{1}{8}$ при $1 \geqslant x \geqslant y \geqslant 0$.

Згідно з загальними висновками розділу I, крім зазначеного розв'язку, матимемо тут і інші розв'язки, які охоплюються загальною формулою

$$xy = -\frac{1}{8} + y + \lambda(x - y), \tag{227'}$$

де, як показує відповідне дослідження, максимальна похибка зберігає оптимальну величину $\delta_0 = \frac{1}{8}$ лише при $0 \leqslant \lambda \leqslant \frac{1}{4}$.

Аналогічно у випадку трьох незалежних змінних знайдемо формулу найкращого наближення у виді

$$xyz = -\frac{1}{3}\sqrt{\frac{1}{3}} + z + \lambda(x - z) + \mu(y - z) \tag{228}$$

з точн. до $\frac{1}{3}\sqrt{\frac{1}{3}}$ при $1 \geqslant x \geqslant y \geqslant z \geqslant 0$ при умові, що параметри λ та μ задовольняють систему нерівностей:

$$\left.\begin{aligned} 0 \leqslant \lambda \leqslant 2\sqrt{\frac{2}{3}\sqrt{\frac{1}{3}} - 1} \\ -\lambda \leqslant \mu \leqslant \frac{2}{3}\sqrt{\frac{1}{3}} - \lambda \end{aligned}\right\} \tag{229}$$

В загальнішому випадку n незалежних змінних, обмежуючись розв'язком найпростішим, матимемо лінійну формулу найкращого наближення

$$x_1 \cdot x_2 \cdot x_3 \ldots x_n = x_n - \frac{n-1}{2n}\sqrt[n-1]{\frac{1}{n}} \tag{230}$$

з точн. до $\frac{n-1}{2n}\sqrt[n-1]{\frac{1}{n}}$ при $1 \geqslant x_1 \geqslant x_2 \geqslant \ldots \geqslant x_n \geqslant 0$.

Як окремі випадки, маємо звідси:

$$\left.\begin{array}{llll} xy = y - \frac{1}{8} & \text{з точн. до } \frac{1}{8} & \text{при} & 1 \geqslant x \geqslant y \geqslant 0 \\ xyz = z - 0{,}192 & \text{„ „ „ } 0{,}192 & \text{„} & 1 \geqslant x \geqslant y \geqslant z \geqslant 0 \\ xyzt = t - 0{,}236 & \text{„ „ „ } 0{,}236 & \text{„} & 1 \geqslant x \geqslant y \geqslant z \geqslant t \geqslant 0 \\ zyztu = u - 0{,}267 & \text{„ „ „ } 0{,}267 & \text{„} & 1 \geqslant x \geqslant y \geqslant z \geqslant t \geqslant u \geqslant 0 \end{array}\right\} \tag{231}$$

Покажімо ще на прикладі добутку трьох співмножників, як перетворюються формули (231), коли область апроксимації є загальнішого типу (T) розділу I.

Отже, нехай треба визначити формулу найкращого наближення виду

$$xyz = \alpha + \beta x + \gamma y + \delta z$$

для області

$$c \geqslant x \geqslant \frac{y}{l} \geqslant \frac{z}{lm} \geqslant 0, \tag{232}$$

де c, l, та m означають додатні сталі. Зробивши заміну змінних

$$x = c\xi, \quad y = cl\eta, \quad z = clm\zeta \tag{233}$$

матимемо для нових змінних область

$$1 \geqslant \xi \geqslant \eta \geqslant \zeta \geqslant 0$$

Для цих змінних має силу формула

$$\xi\eta\zeta = \zeta - 0{,}192 \text{ з точн. до } 0{,}192$$

Повертаючись до первісних змінних, одержимо

$$xyz = c^2lz - 0{,}192\ c^3l^2m \tag{234}$$

з точн. до 0,192 c^3l^2m в області (232).

3°. Важкість задачі Чебишова взагалі, як відомо, зв'язана насамперед з тим, що залежність наближеного представлення функції від самої функції не має лінійного характеру [1]); найкраще наближене представлення суми зовсім не дорівнює взагалі сумі найкращих наближених представлень окремих доданків. Проте в деяких випадках, зокрема коли треба знайти тільки перше наближення для дальшого уточнення за тими методами, що складають предмет розділів VI—VIII цієї роботи, можливо дістати непогане наближення (хоч і не найкраще в розумінні Чебишова), застосовуючи принцип апроксимації Чебишова не до самої даної функції, а до окремих складових частин [2]) її аналітичного виразу — скінченного або нескінченного.

Здобуті в n°n° 1—2 цього параграфа формули дозволяють здійснити цей шлях — у випадку розглядуваних тут наближених представлень типів (1) та (2) розділу I, якщо для даної функції попереду знайдено або розвинення Тайлора, доведене до досить малої остачі, або інше досить точне поліноміальне наближене представлення вищого степеня, добуте будьяким способом. Вони дають водночас і оцінку похибки.

Пояснімо це на кількох простих прикладах.

Приклад I. Нехай треба для функції $e^{\frac{1}{2}(x+y)}$ знайти наближене представлення типу

$$\varphi_2(x, y; \alpha, \beta, \gamma, \delta) = \alpha + \beta x + \gamma y + \delta xy$$

в області

$$0 \leqslant x \leqslant 1, \quad 0 \leqslant y \leqslant 1$$

Виходячи з розвинення Тайлора по степенях $x - \frac{1}{2}$, $y - \frac{1}{2}$, якщо обмежимося в цьому розвиненні лише членами до 3-го степеня включно, ми

[1]) Пор. відповідні висловлювання в E. Borel-я — op. cit., а також у Dunham-Jackson-a — The general theory of approximation by polyn. and trigon. sums. Bulletin of the American Mathematical Society, 1921, p. 417.

[2]) Маємо тут на увазі не тільки окремі доданки суми, але й, наприклад, окремі співмножники добутку тощо.

дістанемо після лінеаризації окремих степенів $\left(x-\frac{1}{2}\right)^2$, $\left(x-\frac{1}{2}\right)^3$ $\left(y-\frac{1}{2}\right)^2$, $\left(y-\frac{1}{2}\right)^3$ за формулами (226) при $h=\frac{1}{2}$:

$$\left.\begin{aligned} e^{\frac{1}{2}(x+y)} &= e^{\frac{1}{2}}\left\{\frac{149}{256}+\frac{99}{256}(x+y)+\frac{1}{4}xy\right\} \\ &= 0{,}960+0{,}638\,(x+y)+0{,}412\,xy \end{aligned}\right\} \tag{235}$$

з граничною абсолютною похибкою

$\delta < 0{,}066$ [гран. пох. від лінеаризації членів 2 та 3 степеня]

$+0{,}004$ [„ „ при відкиданні вищих степенів розвинення Тайлора]

$+0{,}002$ [„ „ від заокруглення коефіцієнтів]

$=$ **0,072**

Розв'язуючи те саме питання за принципом Чебишова, згідно з розділом I цієї роботи [1]), ми знайшли б таку формулу найкращого наближення:

$$e^{\frac{1}{2}(x+y)} = 0{,}946+0{,}649\,(x+y)+0{,}421\,xy \tag{236}$$

з точн. до 0,054 при $0 \leqslant x, y \leqslant 1$, що її коефіцієнти відрізняються від коефіцієнтів формули (235) у межах $\sim 2\%$.

Приклад II. Для тієї самої функції $e^{\frac{1}{2}(x+y)}$ знайти наближене представлення типу

$$\varphi_1(x,\ y;\ \alpha,\ \beta,\ \gamma) = \alpha+\beta x+\gamma y$$

в області

$$1 \geqslant x \geqslant y \geqslant 0$$

Користуючись результатом (235) попереднього прикладу, нам залишається лише зробити заміну: $xy = y-\frac{1}{8}$ з точн. до $\frac{1}{8}$, згідно з формулою (231). Одержуємо:

$$e^{\frac{1}{2}(x+y)} = 0{,}908+0{,}638\,x+1{,}050\,y \tag{237}$$

з граничною абсолютною похибкою

$\delta < 0{,}072$ [гран. пох. попереднього прикладу]

$+0{,}052$ [„ „ від лінеаризації члена $0{,}412\,xy$]

$=$ **0,124**

Наведімо і тут, для порівняння, формулу найкращого наближення, знайдену згідно з розділом I цієї роботи:

$$e^{\frac{1}{2}(x+y)} = 894+0{,}649\,x+1{,}070\,y+\lambda\,(x-y) \tag{238}$$

[1]) Умова (A) розділу I тут має місце, бо функція $e^{\frac{1}{2}(x+y)}$ є, очевидно, кратно-субгармонічна в області апроксимації та на всіх складових частинах її границі. Аналогічне стосується і наступного прикладу II.

ЧАСТИНА ДРУГА

МЕТОДИ БІЛЬШ ЗАГАЛЬНОГО ХАРАКТЕРУ

РОЗДІЛ VI

ОСНОВНІ ОЗНАЧЕННЯ Й ТЕОРЕМИ, ЩО СТОСУЮТЬСЯ ЗАДАЧІ ЧЕБИШОВА У ВИПАДКУ НАБЛИЖЕНОГО ПРЕДСТАВЛЕННЯ ФУНКЦІЇ ОДНОГО НЕЗАЛЕЖНОГО ЗМІННОГО ЗА ДОПОМОГОЮ ПОЛІНОМА

§ 15. Поставлення задачі. Існування розв'язку. Критерій точності наближеного розв'язку

Припустімо, що на обмеженій точковій множині [1]) E числової прямої, себто такій множині, що розміщена на скінченному сегменті (a, b) [2]), задано якусь функцію $f(x)$, для якої шукається найкраще, в розумінні Чебишова, наближене представлення у виді полінома

$$K_0 x^n + K_1 x^{n-1} + K_2 x^{n-2} + \cdots + K_n = P_n(x; K_0, K_1, \ldots, K_n) \tag{240}$$

степеня не вищого від n. Саму дану функцію $f(x)$ ми припускаємо лише обмеженою, себто

$$|f(x)| \leqslant M \tag{241}$$

на множині E. Ніяких інших обмежних припущень щодо цієї функції та щодо множини E ми тут не введемо. Ми не залишаємо осторонь і такі випадки (що трапляються і в застосуваннях, і в самому аналізі) [3]), коли функцію $f(x)$ не задано як цілком означену (однозначну) функцію, але коли кожній точці $x \subset E$ поставлено у відповідність деяку множину числових значень Γ_x, що кожне з них можна приписати функції $f(x)$ при даному значенні x. В останньому випадку точні нижня та верхня границі числової множини Γ_x не перевищують числа M абсолютною величиною.

Позначмо через E° замкнену множину [розміщену, певна річ, на тому

[1]) Вживаючи і тут, і в дальшому звичайної геометричної мови, ми в наших позначеннях не відрізнятимемо взагалі числових множин од відповідних точкових множин. Не виключаючи з розгляду й таких випадків, коли множина E складається із скінченного числа точок (ці випадки мають навіть досить важливе значення — і практичне і теоретичне), ми, проте, скрізь залишаємо осторонь випадок тривіальний, коли число точок $\leqslant n+1$.

[2]) Припускається, що (a, b) є найменший сегмент, що містить у собі E.

Зауважмо, що для замкненого інтервалу в дальшому викладі вживається скрізь терміну „сегмент" і позначення ( ) і лише в деяких спеціальних випадках — позначення $\langle\ \rangle$.

[3]) Пор., напр., з одного боку, F. Klein — Präzisions-und Approximationsmathematik, а з другого боку, H. Lebesgue — Leçons sur l'intégration et la recherche des fonct. primitives (є рос. перекл.)

самому сегменті (a,b)], утворену через сполучення даної множини E з її похідною множиною E'. Означмо для кожної точки $x \subset E^\circ$ два числа — границі неозначеності[1]) функції $f(x)$ у точці x:

$$\varphi(x) = \lim_{\varepsilon \to 0} \left[\operatorname{borne\,inf.}_{x-\varepsilon < \xi < x+\varepsilon} f(\xi)\right] \tag{242}$$

$$\Phi(x) = \lim_{\varepsilon \to 0} \left[\operatorname{borne\,sup.}_{x-\varepsilon < \xi < x+\varepsilon} f(\xi)\right] \tag{243}$$

де, зрозуміла річ, припускаємо $\xi \subset E$ [2]).

Цілком означені в такий спосіб дві функції $\varphi(x)$ та $\Phi(x)$ є півнеперервні[3]), відповідно знизу та зверху, на замкненій множині E°. Таксамо „коливання“ $\Phi(x) - \varphi(x)$ є функція півнеперервна зверху.

Якщо маємо тепер якесь наближене представлення функції $f(x)$ на точковій множині E у виді полінома P_n, за міру наближення, що його дає цей поліном, приймаємо величину

$$\operatorname{borne\,sup.}_{x \subset E} |f - P_n| \tag{244}$$

Легко бачити, що ця величина завжди дорівнює більшій з двох величин

$$\max_{x \subset E^\circ} (P_n - \varphi) \tag{245}$$

$$\max_{x \subset E^\circ} (\Phi - P_n) \tag{246}$$

Це важливе зауваження дозволяє сформулювати задачу Чебишова щодо найкращого наближеного представлення функції $f(x)$ на точковій множині E з допомогою полінома P_n степеня $\leqslant n$ так:

Поміж всіма поліномами $P_n = P_n(x; K_0, K_1, \ldots, K_n)$ визначити такий, для якого більша з двох величин (245) та (246) була б якнайменша.

[1]) Дальше точне означення, дане за допомогою рівностей (242) та (243), виключає можливість будьякої неясності, що могла б мати місце у зв'язку з недосить усталеним змістом терміну „границі неозначеності“ в літературі (P. du Bois-Reymond, H. Lebesgue).

[2]) Це є ті числа, що R. Baire у випадку однозначно-означеної функції f, позначає як minimum та maximum функції f у даній точці x щодо множини E. Див. Sur la représentation des fonctions discontinues, chap. III (Acta Mathematica, t. 30).

В тому випадку, коли функція $f(x)$ є однозначно-означена в E і має коливання рівне нулеві в кожній точці E° щодо E (за термінологією R. Baire op. cit.) і тільки в цьому випадку — функції $\varphi(x)$ та $\Phi(x)$ збігатимуться між собою на множині E° і збігатимуться з $f(x)$ на множині E. Зокрема ці умови мають місце, якщо сама множина E є замкнена, $E = E^\circ$ і функція $f(x)$ є однозначно-означена й неперервна на цій множині, або ще — коли множина E складається лише із скінченного числа точок, і функція $f(x)$ є однозначно-означена в них.

[3]) R. Baire, loc. cit. Та обставина, що у нас $f(x)$ може означати функцію не однозначно-означену, як не важко бачити, справи не міняє.

Всякий поліном, що відповідає цій вимозі, ми називатимемо поліномом найкращого наближення степеня $\leqslant n$ для функції $f(x)$ на точковій множині E.

Довести існування розв'язку для розглядуваної задачі [1]) можна виходячи з дальших міркувань.

Число L, яке дорівнює більшому з двох чисел (245) та (246), являє собою, очевидно, деяку неперервну функцію параметрів (коефіцієнтів) полінома $P_n(x; K_0, K_1, \ldots, K_n)$:

$$L = L(K_0, K_1, \ldots, K_n), \tag{247}$$

що її ми можемо трактувати як функцію точки у відповідному $n+1$-мірному просторі. Обмежуючись тими поліномами P_n, які на точковій множині E не перевищують [2]) абсолютною величиною $2M$ [див. нерівн. (241)], нам досить буде мати на увазі лише деяку обмежену та замкнену точкову множину $\mathfrak{C}$ у цьому просторі: межі параметрів $K_0, K_1, \ldots, K_n$ можливо встановити в явному виді [3]) за допомогою Лягранжової інтерполяційної формули.

Згідно з добре відомою теоремою Вайєрштраса, неперервна на замкненій обмеженій точковій множині $\mathfrak{C}$ функція $L(K_0, K_1, \ldots, K_n)$ досягає свого minimum-у бодай в одній точці цієї множини. Цей minimum являє собою величину найкращого наближення ρ для даної задачі. Позначмо через $\mathfrak{C}_0$ ту замкнену частину множини $\mathfrak{C}$, що складається з точок, в яких цей minimun досягається. Поліноми P_n, які відповідають точкам цієї множини $\mathfrak{C}_0$, і будуть шукані поліноми найкращого наближення. Для цих поліномів ми вживатимемо позначення Π_n в тих випадках, коли треба буде їх виділити. Щождо числа розв'язків, зауважимо поки, що лише два випадки є можливі: або задача має єдиний розв'язок, або безліч розв'язків [4]).

В тому випадку, коли поліном найкращого наближення визначається за допомогою нескінченного процесу, в справі судження про збіжність процесу має істотне значення така теорема:

Теорема. *Кожному довільно малому додатному числу* ε *можна протиставити таке додатне число* η, *щоб нерівність*

$$L(K_0, K_1, \ldots, K_n) < \rho + \eta \tag{248}$$

[1]) Зразки суворого доведення для теорем існування, що стосуються задачі Чебишова, уперше дали, як відомо, P. Kirchberger у своїй дисертації — Über Tchebychefsche Annäherungsmethoden (див. Mathem. Annalen, Bd. 57), а надто E. Borel у своїх — Leçons sur les fonct. de variables réelles et les développements en séries de polynomes, pp. 82—92. Paris, 1905, 1928.

[2]) Всякий поліном, який цю вимогу не задовольняє, дає наближення напевне за M, в той час як поліном тотожно рівний нулеві дає наближення, що не
Пор. E. Borel, op. cit., pp. 82—92.

[3]) E. Borel, ibidem.

[4]) При наявності бодай двох розв'язків Π_{n1} та Π_{n2} поліном $\Pi_{n1}\cos^2\alpha + \Pi_{n2}\sin^2\alpha$ становить розв'язок за кожного дійсного значення α. Пор. L. Tone
prossimazione di Tchebychef (Annali di Matematica pura ed applicata
неперервні функції кількох змінних.

мала своїм наслідком, одностайно на всій точковій множині E, нерівність

$$|P_n(x;\ K_0,\ K_1, \ldots,\ K_n) - P_n(x;\ K_{00},\ K_{10}, \ldots,\ K_{n0})| < \varepsilon, \qquad (249)$$

де $P_n(x; K_{00}, K_{10}, \ldots, K_{n0})$ *є якийсь із поліномів найкращого наближення* Π_n [1].

Щоб довести теорему, нам досить буде [2], очевидно, встановити, що коли має місце (248) із досить малим η, то віддаль точки $(K_0, K_1, \ldots, K_n)$ від вищезазначеної точкової множини $\mathfrak{C}_0$ буде менша за яке завгодно наперед задане додатне число δ. І справді, оскільки точки множини $\mathfrak{C}$, для яких віддаль до множини $\mathfrak{C}_0$ є більша або дорівнює δ, складають, очевидно, деяку замкнену множину [3] $\mathfrak{C}_\delta \subset \mathfrak{C} - \mathfrak{C}_0$, в якій неперервна функція $L(K_0, K_1, \ldots, K_n)$ досягає певного minimum-у $\rho_\delta > \rho$, то досить буде взяти $\eta < \rho_\delta - \rho$, щоб дійти бажаного результату.

Увага I. В дальшому, оцінюючи точність наближеного розв'язку задачі Чебишова, одержаного у виді якогось полінома $P_n(x;\ K_0,\ K_1, \ldots,\ K_n)$, ми критерій точності встановлюватимемо взагалі з допомогою величини різниці $L(K_0,\ K_1, \ldots,\ K_n) - \rho$.

Увага II. Найкраще наближення ρ можна, очевидно, визначити як мінімальне значення величини

$$\max_{(x,\ y) \subset \mathfrak{A}} |y - P_n(x;\ K_0,\ K_1, \ldots,\ K_n)|$$

за оптимального добору параметрів $K_0,\ K_1, \ldots,\ K_n$, якщо замкнена область $\mathfrak{A}$ визначається умовами

$$\left.\begin{array}{c} x \subset E^0 \\ \varphi(x) \leqslant y \leqslant \Phi(x) \end{array}\right\} \qquad \text{(область } \mathfrak{A}\text{)}$$

Важливо звернути увагу на ту очевидну обставину, що величина зазначеного maximum-у, отже і величина ρ, може лише збільшитись, або в крайньому разі залишитись без зміни при кожному поширенні точкової множини $\mathfrak{A}$ і навпаки.

§ 16. Дві теореми, що встановлюють нижню межу для найкращого наближення. Необхідні і достатні умови для того, щоб поліном $p_n(x)$ давав найкраще наближення

Теорема I, яка відповідає відомому твердженню de la Valée-Poussin-a [4], але з точнішою оцінкою нижньої межі найкращого наближення.

[1]) Для випадку неперервної функції $f(x)$, заданої на неперервному сегменті, аналогічну теорему було встановлено іншим міркуванням, яке істотно виходило з неперервності функції $f(x)$ і яке, саме через це, не є застосувальне до нашої загальнішої задачі. Див. C. de la Vallée-Poussin. Leçons sur l'approximation des fonctions d'une variable réelle. Paris, 1919, n° n° 61, 65.

[2]) Слід пам'ятати, що точкова множина E лежить на скінченному сегменті (a, b)

[3]) Припускаємо δ досить малим для того, щоб множина $\mathfrak{C}_\delta$ не була порожня.

[4]) C. de la Valée-Poussin. Sur les pol. d'approx. et la représent. approchée d'un angle, n° n° 15, 17 (Bulletins de l'Acad. Roy. de Belgique, 1910), а також — Leçons sur l'approximation etc., p. 78. Там мова про функцію $f(x)$ однозначно-означену та неперервну на даному неперервному сегменті.

Нехай буде $Q_n(x)$ якийсь поліном степеня $\leqslant n$. Якщо для $n+2$ точок $x_0 < x_1 < x_2 < \ldots < x_{n+1}$ точкової множини E^0 має місце один із таких двох рядів нерівностей:

$$\left.\begin{array}{ll} \Phi(x_0) - Q_n(x_0) > 0, & \varphi(x_1) - Q_n(x_1) < 0, \\ \Phi(x_2) - Q_n(x_2) > 0, & \varphi(x_3) - Q_n(x_3) < 0, \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \end{array}\right\} \tag{250}$$

або:

$$\left.\begin{array}{ll} \varphi(x_0) - Q_n(x_0) < 0, & \Phi(x_1) - Q_n(x_1) > 0 \\ \varphi(x_2) - Q_n(x_2) < 0, & \Phi(x_3) - Q_n(x_3) > 0 \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \end{array}\right\} \tag{251}$$

то, коли позначити в одному або в другому випадку найменшу з розглядуваних додатних різниць через μ', найменшу з абсолютних величин від'ємних різниць через μ'' — найкраще наближення ρ для даної функції $f(x)$ на точковій множині E за допомогою полінома степеня $\leqslant n$ не може бути менше за $\frac{\mu' + \mu''}{2}$; разом з тим воно напевне більше за менше з обох чисел μ' та μ'', якщо тільки даний поліном $Q_n(x)$ сам не являє собою полінома найкращого наближення.

Перше твердження теореми дає, очевидно, взагалі більш високу нижню межу для ρ, дозволяючи при оцінці точності того чи того наближеного розв'язку задачі Чебишова зменшувати верхню межу величини $L - \rho$ часом аж до 50%. Але, як буде видко з дальших розглядів цієї роботи, і друге твердження теореми зберігає своє самостійне значення в багатьох питаннях, де саме величина $\min\{\mu', \mu''\}$ відіграє дуже істотну роль.

Доведення першого твердження одержуємо з того міркування, що в противному разі різниця $\Pi_n(x) - \left[Q_n(x) + \frac{\mu' - \mu''}{2}\right]$, де $\Pi_n(x)$ означає, як ми вище умовилися, поліном найкращого наближення, мала б альтернуючий знак у послідовних точках $x_0, x_1, x_2, \ldots, x_{n+1}$, отже вона мала б щонайменше $n+1$ нульових точок, не бувши тотожним нулем, що для полінома степеня $\leqslant n$ не можливе.

Щождо другого твердження теореми, то, доводячи його також від супротивного, ми приходимо лише до необхідності припустити, що для різниці $\Pi_n(x) - Q_n(x)$ повинні задовольнятися дальші нерівності в широкому розумінні:

$$\left.\begin{array}{ll} \Pi_n(x_0) - Q_n(x_0) \geqslant 0, & \Pi_n(x_1) - Q_n(x_1) \leqslant 0, \\ \Pi_n(x_2) - Q_n(x_2) \geqslant 0, & \Pi_n(x_3) - Q_n(x_3) \leqslant 0 \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \end{array}\right\} \tag{252}$$

або протилежні нерівності, що ми їх одержуємо з написаних, міняючи скрізь знак $>$ на $<$ і навпаки. Припустімо для означеності, що нерівності (252) мають місце саме з тими знаками, з якими вони написані. Оскільки $Q_n(x)$, не бувши тотожне з поліномом $\Pi_n(x)$, не може, певна

річ, збігатися з ним при всіх $n+2$ значеннях $x=x_0, x_1, x_2, \ldots, x_{n+1}$, припустімо, що x_i є одно з тих значень, для яких

$$(-1)^i [\Pi_n(x_i) - Q_n(x_i)] > 0, \tag{253}$$

виключаючи знак $=$. Беручи на увагу останні $n+1$ нерівностей (252), ми на підставі (253) приходимо послідовно до необхідності таких висновків:

а) що на інтервалі $(x_i, x_{i+1}>$ [1]) міститься принаймні один корінь полінома $\Pi_n(x) - Q_n(x)$;

б) що на інтервалі $(x_i, x_{i+2}>$ міститься бодай два простих корені, або один корінь парної кратності того самого полінома;

в) що, взагалі, на інтервалі $(x_i, x_{i+r}>$, де $r=1, 2, \ldots, n+1-i$, міститься не менш як r коренів, рахуючи при цьому кратні корені (якщо є такі) відповідно до їх кратності;

г) що, аналогічно, в кожному з інтервалів $<x_{i-s}, x_i)$, де $s=1, 2, \ldots, i$, міститься не менш як s коренів того самого полінома $\Pi_n - Q_n$.

Але коли так, то поліном $\Pi_n - Q_n$ має на інтервалі (a, b) не менш як $(n+1-i)+i=n+1$ коренів, тобто він повинен бути тотожним нулем, а це суперечить прийнятому припущенню щодо Q_n.

Теорема II. *Якщо $2k$ означає максимум різниці $\Phi(x) - \varphi(x)$ на точковій множині E°, тобто, інакше кажучи, максимум коливання функції $f(x)$*[2]), *то найкраще наближення ρ не може бути менше, ніж k.*

Щоб цю теорему довести, досить взяти на увагу, що коливання різниці $f - P_n$ у будьякій точці множини E° дорівнює коливанню самої функції f, та що наближення (244), очевидно, не може бути менше за половину максимального коливання цієї різниці.

Теорема III. *Для того, щоб поліном $p_n(x)$ давав найкраще наближення для функції $f(x)$ на точковій множині E, досить і необхідно, щоб задовольнялася принаймні одна з таких двох умов:*

1° або для якогось ансамбля $n+2$ точок $x_0 < x_1 < x_2, < \ldots < x_{n+1}$, що належать до точкової множини E°, має місце один із двох рядів рівностей:

$$\left.\begin{array}{ll} \Phi(x_0) - p_n(x_0) = \quad L, & \varphi(x_1) - p_n(x_1) = -L, \\ \Phi(x_2) - p_n(x_2) = \quad L, & \varphi(x_3) - p_n(x_3) = -L \\ \cdots\cdots\cdots\cdots & \end{array}\right\} \tag{254}$$

або

$$\left.\begin{array}{ll} \varphi(x_0) - p_n(x_0) = -L, & \Phi(x_1) - p_n(x_1) = \quad L, \\ \varphi(x_2) - p_n(x_2) = -L, & \Phi(x_3) - p_n(x_3) = \quad L \\ \cdots\cdots\cdots\cdots & \end{array}\right\} \tag{255}$$

[1]) $(\alpha, \beta>$ означає інтервал замкнений на правому кінці, тобто множину числових значень x, що справджують нерівності $\alpha < x \leqslant \beta$; в аналогічному розумінні вживаємо й позначення $<\alpha, \beta)$.

[2]) Цього максимуму досягають принаймні в одній точці множини E° (пор. R. Baire, op. cit.).

де L означає, як і вище, наближення, що його дає поліном $p_n(x)$, тобто в даному разі спільну величину обох чисел (245) та (246)[1]);

2° або наближення L, що його дає поліном $p_n(x)$ [тобто величина (244)], дорівнює точно половині максимального коливання $2k$ функції $f(x)$ на точковій множині E°.

Справді. достатність зазначених умов безпосередньо виходить з доведених у попередньому параграфі двох основних теорем. Щождо їх необхідності, то в розділі VII буде виявлено, що в тому випадку, коли жодна із зазначених двох умов не задовольняється, до полінома $p_n(x)$ можна додати таку поправку у виді полінома $\omega_n(x)$ степеня $\leqslant n$, що скорегований поліном $p_n(x)+\omega_n(x)$ даватиме менше наближення, а це і доводить необхідність розглядуваних умов[2]).

Із цієї теореми можна зробити деякі загальні висновки також щодо єдиності[3]) розв'язку, а саме:

а) Якщо величина найкращого наближення ρ перевищує половину максимального коливання $2k$ функції $f(x)$ (на точковій множині E° щодо E), то задача напевне має лише один розв'язок.

Справді, в цьому випадку всякий поліном найкращого наближення неодмінно задовольняє умову 1° розглядуваної теореми. А коли так, то єдиність розв'язку встановлюється точно так само, як у випадку однозначно-означеної та неперервної функції $f(x)$, що розглядається на неперервному сегменті (a, b)[4]).

б) Якщо ж $\rho=k$, задача може мати або безліч розв'язків, або лише один розв'язок[5]).

Розв'язок напевне буде єдиний, якщо для нього (разом з умовою 2°) виконується й умова 1° даної теореми: справді, спосіб міркування, аналогічний до застосованого при доводженні другого твердження теореми I цього розділу, дозволяє в цьому випадку встановити єдиність розв'язку. Розв'язок буде, очевидно, також єдиний, якщо, при $\rho=k$, функція $f(x)$ має максимальне коливання, рівне $2k$, хоча б у $n+1$ точках. Але поданий нижче приклад показує, що єдиність розв'язку, при $\rho=k$, цілком може бути забезпечена і в таких випадках, де maximum різниць $\Phi-p_n, p_n-\varphi$ має місце всього в одній точці.

Справа в тому, що у випадку $\rho=k$ значення полінома Π_n в кожній точці, де коливання $\Phi(x)-\varphi(x)$ досягає максимальної величини $2k$,

[1]) При умові, що P_n замінюється на p_n в цих виразах.

[2]) Подібний тип міркування, щоб встановити необхідні умови для найкращого наближеного представлення, є систематично застосований уперше Чебишовим. Див. P. L. Tchebychef. Sur les questions de minima qui se rattachent à la representation approximative des fonctions, §§ 6–16 (Oeuvres, t. 1, p. 273–378).

[3]) Існування розв'язку встановлено в § 15 даної роботи.

[4]) Див., напр., S. Bernstein. Leçons sur les propriétés extrémales des fonct. Paris, 1926, p. 4 або E. Borel, op. cit.

[5]) Пор. вище відповідну виноску в § 15, а також приклади в § 22 п. 3 цього параграфа.

повинно дорівнювати $\frac{\Phi(x)+\varphi(x)}{2}$ а крім того в кожній такій точці, якщо вона є гранична точка множини E^0 маємо ще взагалі певні обмеження для першої похідної, а в деяких випадках і для двох або кількох перших похідних від Π_n. Урахування цих умов і визначає основні особливості трактування питання у випадку $\rho = k$.

Варто звернути особливу увагу на таке досить зручне сформулювання задачі щодо розглядуваного випадку:

Визначити параболічні криві $y = P_n(x; K_0, K_1, \ldots, K_n)$ порядку $\leqslant n$, які при $x \subset E^0$ містяться в області

$$\left.\begin{array}{c}\Phi(x) - k \leqslant y \leqslant \varphi(x) + k \\ x \subset E^0\end{array}\right\} \qquad \text{(область } \mathfrak{B}\text{)}$$

або пересвідчитись (напр., відповідним застосуванням теореми I) в тому, що такої параболічної кривої немає.

Точкові множини $y = \Phi(x) - k$ та $y = \varphi(x) + k$, очевидно, перекриваються в точках максимального коливання функції $f(x)$.

Приклад. Припустімо, що треба знайти найкраще наближене представлення за допомогою полінома $p_n(x)$ степеня $\leqslant n$ на якомусь даному сегменті (a, b), який містить точку $x = 0$ $(a, b \neq 0)$ для функції

$$f(x) = \Psi(x) + k\left(1 - e^{-\frac{1}{x^2}}\right)[1 - 2\chi(x)], \tag{256}$$

де $\Psi(x)$ означає якусь аналітичну функцію, регулярну на сегменті (a, b), а $\chi(x) = \lim\limits_{m \to \infty} [\lim\limits_{n \to \infty} (\cos m!\pi x)^{2n}]$ є розривна у всіх точках функція Dirichlet.

Тут ми маємо:

$$\left.\begin{array}{l}\Phi(x) = \Psi(x) + k\left(1 - e^{-\frac{1}{x^2}}\right) \\ \varphi(x) = \Psi(x) - k\left(1 - e^{-\frac{1}{x^2}}\right)\end{array}\right\} \tag{257}$$

Максимальне коливання функції $f(x)$ на сегменті (a, b) є $2k$. Область $\mathfrak{B}$ тут є

$$\left.\begin{array}{c}\Psi(x) - ke^{-\frac{1}{x^2}} \leqslant y \leqslant \Psi(x) + ke^{-\frac{1}{x^2}} \\ a \leqslant x \leqslant b\end{array}\right\} \tag{258}$$

Оскільки дві криві $y = \Psi(x) \mp k \cdot e^{-\frac{1}{x^2}}$ мають при $x = 0$ контакт нескінченно-високого порядку, то ясно, що, коли існує параболічна крива $y = p_n(x)$, яка при $a \leqslant x \leqslant b$ міститься в області (258), функції $\Psi(x)$ та $p_n(x)$ повинні спадатися при $x = 0$, так само як їхні відповідні похідні всіх порядків. Це ж припускає $p_n(x) \equiv \Psi(x)$. Звідси — такі висновки щодо розглядуваної задачі найкращого наближення:

1) Якщо $\Psi(x)$ сама не є поліном степеня $\leqslant n$, найкраще наближення $f(x)$ на сегменті (a, b) є більше, ніж k. Задача має, певна річ, лише один розв'язок.

2) У випадках, коли $\Psi(x)$ є поліном степеня $\leqslant n$, цей поліном і дає найкраще наближення, рівне k. Розв'язок і тут є єдиний, хоч максимальний відхил між $p_n(x) \equiv \Psi(x)$ та границями неозначеності функції $f(x)$ має місце тільки в одній точці.

§ 17. Деякі дальші теореми

Лема I. *Припустімо, що дано сегмент (a, b) на дійсній осі та зазначено верхню межу n степеня полінома $P(x)$. Тоді кожному як завгодно малому $\sigma > 0$ та кожному як завгодно великому $N > 0$ можна протиставити таке число $R > N$, що при* $\max\limits_{a \leqslant x \leqslant b} |P(x)| > R$ *нерівність* $|P(x)| < N$ *може мати місце, при $a \leqslant x \leqslant b$, щонайбільше — на обмеженому числі $(< n + 1)$ інтервалів, що їх загальна довжина не перевищує σ.*

Доведення. Збудуймо на площині комплексного змінного x коло $\mathfrak{K}$ на сегменті $(2a - b, 2b - a)$ дійсної осі як на діаметрі. Поліном $P(x)$ можна представити у виді добутку $P(x) = q_\alpha(x) . r_\beta(x)$, де поліном степеня $\alpha (0 \leqslant \alpha \leqslant n)$ $q_\alpha(x)$ має в найвищому члені коефіцієнт одиницю і не має коренів поза кругом $\mathfrak{K}$, ані на самому колі, а поліном степеня $\beta (0 \leqslant \beta \leqslant n - \alpha)$ $r_\beta(x)$, навпаки, не має коренів на полі круга $\mathfrak{K}$. Оскільки $|q_\alpha(x)| < [2(b-a)]^\alpha$ при $a \leqslant x \leqslant b$, то при $\max\limits_{a \leqslant x \leqslant b} |P(x)| > R$ ми повинні мати $\max |r_\beta(x)| > \dfrac{R}{[2(b-a)]^\alpha}$, а це має своїм наслідком, як не важко бачити,

$$\min_{a \leqslant x \leqslant b} |r_\beta(x)| > \frac{R}{[2(b-a)]^\alpha \cdot 2^\beta} \geqslant \frac{R}{2^n (b-a)^\alpha}$$

Навколо кожної точки, що відповідає на комплексній площині якому-небудь з коренів полінома $q_\alpha(x)$, опишімо круг радіуса $\dfrac{\sigma}{2n}$. В тих точках сегмента (a, b), які не потрапляють на поле якогонебудь із цих кругів, ми маємо $|q_\alpha(x)| > \left(\dfrac{\sigma}{2n}\right)^\alpha$. Загальна довжина тих частин сегмента (a, b), які можуть потрапити на поле збудованих нами кругів радіуса $\dfrac{\sigma}{2n}$, не перевищує $\dfrac{\sigma}{n} \cdot \alpha \leqslant \sigma$. На останніх точках сегмента (a, b) матимемо $|P(x)| > N$, якщо взяти

$$R > 2^n (b-a)^\alpha \cdot \left(\frac{2n}{\sigma}\right)^\alpha . N \tag{259}$$

Останню умову ми лише зміцнимо, якщо візьмемо

$$R > \left(4n \cdot \frac{b-a}{\sigma}\right)^n \cdot N \tag{259'}$$

Увага. Наведене дослідження мало встановити лише скінченність відношення $R:N$. Інший метод дослідження дозволяє встановити мінімальне значення того самого відношення, але на останньому питанні ми тут не зупинимось.

Лема II. *Нехай маємо будьяку обмежену функцію $f(x)$, задану на обмеженій точковій множині E.* Нехай x' та x'' позначають дві будьякі точки множини E. *Тоді кожному довільно малому $\varepsilon > 0$ можна протипоставити $\eta > 0$ так, щоб нерівність $|x' - x''| < \eta$ мала своїм наслідком* $|f(x') - f(x'')| < 2k + \varepsilon$, позначаючи, як і раніше, через $2k$ максимум коливання функції $f(x)$ на множині E° [1]) щодо E.

Доведення. Доводячи від супротивного, ми повинні припустити, що при певному $\varepsilon > 0$ існує така нескінченна послідовність пар точок $x_1', x_1''; x_2', x_2''; x'_3, x_3''; \ldots$ на множині E, що для них $|f(x') - f(x'')| \geqslant 2k + \varepsilon$ в той час як $\lim x_i' = \lim x_i'' = x_0$ (якась точка множини E°). Але тоді в цій граничній точці x_0 коливання функції $f(x)$ не могло б бути менше, ніж $2k + \varepsilon$, а це суперечить припущенню леми.

Теорема I. *Нехай буде $p_n(x)$ будьякий поліном степеня $\leqslant n$, що розглядається як наближене представлення (не найкраще) даної обмеженої функції $f(x)$ на обмеженій точковій множині E. Позначимо через A найбільше значення нижньої границі, що її можливо встановити для найкращого наближення ρ на підставі дослідження різниць $\Phi - p_n$, $\varphi - p_n$ за допомогою другої частини теореми I параграфа 16, а також теореми II того самого параграфа* [2]). *Ми твердимо, що кожному довільно малому $\varepsilon > 0$ можна буде протипоставити таке число $S > 0$, щоб нерівність*

$$\max_{a \leqslant x \leqslant b} |p_n(x)| > S$$

мала своїм наслідком $A \leqslant k + \varepsilon$.

Доведення. Насамперед зауважмо, що необмежене зростання величини $\max\limits_{a \leqslant x \leqslant b} |p_n(x)|$ в той час, як знак різниці $f - p_n$ не залишається сталим на точковій множині E, можливе, очевидно, не інакше, як при одночасному

[1]) Ця лема, являючи собою узагальнення добре відомої теореми про одностайну неперервність, не є істотно нова: для випадку функції однозначно-означеної на неперервному сегменті доведення аналогічного твердження можна знайти, наприклад, у H. Lebesgue. Leçons sur l'intégration ... 1904, p. 21 (є рос. перекл.). Ми тут розглядаємо обмежену функцію $f(x)$ в такій самій широкій концепції як у §§ 15 і 16.

[2]) Це значить, що A дорівнює, при позначеннях параграфа 16, найбільшому значенню величини $\min\{\mu', \mu''\}$, що його вона набирає при „оптимальному" доборі ансамблю $n+2$ точок $x_0 < x_1 < x_2 < \ldots < x_{n+1}$ на точковій множині E°, якщо це найбільше значення перевищує k. Коли ж воно менше ніж k, або коли для різниць $\Phi - p_n, \varphi - p_n$ зовсім не існує такого ансамблю $n+2$ точок, який потрібний для застосування теореми I § 16, ми покладаємо $A = k$.

необмеженому зростанні також і максимуму абсолютної величини похідної $\max\limits_{a \leqslant x \leqslant b} |p'_n(x)|$. Отже, допускаючи, що теорема несправедлива, нам доведеться припустити, що при певному значенні ε можна вказати нескінченну послідовність поліномів

$$p_{n1}(x),\ p_{n2}(x),\ p_{n3}(x), \tag{260}$$

таку, що

$$\max_{a \leqslant x \leqslant b} |p'_{nj}(x)| = T_j,\ \lim_{j \to \infty} T_j = \infty \tag{261}$$

і водночас, для всіх значень індекса j,

$$A_j > k + \varepsilon \tag{262}$$

позначаючи через A_j число A, яке відповідає поліномові p_{nj}.

Якщо число A_j встановлюється, на підставі другого твердження теореми I § 16, за допомогою ансамблю $n+2$ точок $x_{0j}, x_{1j}, \ldots, x_{n+1,j}$ позначмо через $\Delta_j(x_{\nu j}), \nu = 0, 1, 2, \ldots, n+1$, величини різниць $\Phi(x_{\nu j}) - p_{nj}(x_{\nu j})$ та $\varphi(x_{\nu j}) - p_{nj}(x_{\nu j})$, що фігурують на лівих сторонах системи нерівностей типу (250) або (251), а саме — тієї з цих двох систем, яку беруть до уваги, досліджуючи різниці $\Phi - p_{nj}$, $\varphi - p_{nj}$ при даному значенні індекса j.

Розгляньмо сукупність $2(n+2)$ чисел

$$x_{0j}, \frac{1}{\Delta_j(x_{0j})};\ x_{1j}, \frac{1}{\Delta_j(x_{1j})};\ \ldots;\ x_{n+1j}, \frac{1}{\Delta_j(x_{n+1,j})} \tag{263}$$

як координати точки в $2(n+2)$-мірнім просторі. Оскільки всі $2(n+2)$ координати є обмежені, повинна існувати принаймні одна гранична точка з координатами, скажімо,

$$x_{00}, \frac{1}{\Delta^{(0)}};\ x_{10}, \frac{1}{\Delta^{(1)}};\ \ldots;\ x_{n+1,0}, \frac{1}{\Delta^{(n+1)}} \tag{264}$$

при чому координати $x_{00}, x_{10}, \ldots, x_{n+1,0}$ дорівнюють якимсь числам множини E^0, а останні координати абсолютною величиною не перевищують $\frac{1}{k+\varepsilon}$

З послідовності (260) можна буде вилучити частинну нескінченну послідовність поліномів, що їх ми для спрощення позначень заново занумеруємо

$$p_{ni}, \quad i = 1, 2, 3, \tag{260'}$$

і які матимуть такі властивості:

1°. $\max\limits_{a \leqslant x \leqslant b} |p'_{ni}(x)| = T_i,\ \lim\limits_{i \to \infty} T_i = \infty$ (261′)

2. $A_i > k + \varepsilon$ (262′)

3°. $\lim\limits_{i \to \infty} x_{\nu i} = x_{\nu 0},\ \nu = 0, 1, 2, \ldots, n+1$ (265)

$$4^\circ.\ \operatorname{sg}\Delta_i(x_{0i}) = -\operatorname{sg}\Delta_i(x_{1i}) = \ldots = (-1)^{n+1}\operatorname{sg}\Delta_i(x_{n+1,i}) \tag{266}$$
$$= \text{const при } i = 1, 2, 3, \ldots\ ^{1)}$$

$$5^\circ.\ \lim_{i\to\infty}\Delta_i(x_{\nu i}) = \Delta^{(\nu)},\ \nu = 0, 1, \ldots, n+1\ ^{2)}. \tag{267}$$

Візьмімо тепер у послідовності (260′) індекс i досить великий, щоб задовольнялися такі вимоги:

а) щоб усі $n+2$ різниці $x_{\nu i} - x_{\nu 0}$ були абсолютною величиною менші, ніж $\frac{\eta}{2}$, де величині η надамо такого самого значення, яке визначається вимогою леми II, але відповідно до того зафіксованого значення ε, яке фігурує тепер у нашому міркуванні; до того ж ми підпорядкуємо ще величину η такій додатковій умові, щоб у всякому разі було $\eta < \delta$, позначаючи через δ найменшу відмінну від нуля (припускаючи, що такі є) з різниць $x_{\nu+1,0} - x_{\nu 0}$, $\nu = 0, 1, 2, \ldots, n$;

б) щоб T_i було більше, ніж число R, яке ми означимо, згідно з лемою I даного параграфа, але відповідно до полінома $p'_n(x)$, замінивши в (259′) n на $n-1$, ε — на $\frac{\delta-\eta}{2}$, N — на $\frac{4M}{\delta-\eta}$, де M означає, згідно з (241), верхню границю $|f(x)|$ на точковій множині E.

Щоб не залишити жодної можливості не розгляненою, додамо, що умову $\eta < \delta$, а також вимогу б) в цілому ми б зовсім опустили, як би трапилося, що всі $n+2$ значення $x_{\nu 0}$, $\nu = 0, 1, \ldots, n+1$ збігаються між собою.

Ми виявимо суперечність зроблених припущень, показавши, що при цих умовах поліном $p'_{ni}(x)$, ступінь якого не перевищує $n-1$, мусів би мати альтернуючий знак в $n+1$ послідовних точках, що належать відповідно до інтервалів:

$$(x_{0i},\ x_{1i}),\ (x_{1i},\ x_{2i}), \ldots, (x_{ni},\ x_{n+1,i})$$

Справді, якщо, для якогонебудь значення ν, $x_{\nu 0} = x_{\nu+1,0}$ то (додатна) різниця $x_{\nu+1,i} - x_{\nu i}$ є менша, ніж η [зважаючи на умову (a)]; отже $|\Phi(x_{\nu+1,i}) - \varphi(x_{\nu i})| \leqslant 2k+\varepsilon$, $|\varphi(x_{\nu+1,i}) - \Phi(x_{\nu i})| \leqslant 2k+\varepsilon$ в той час, як $|\Delta_i(x_{\nu+1,i}) - \Delta_i(x_{\nu i})| > 2k+\varepsilon$. Це, очевидно, є можливе лише тоді, коли

$$\operatorname{sg}[p_{ni}(x_{\nu+1,i}) - p_{ni}(x_{\nu i})] = \operatorname{sg}[\Delta_i(x_{\nu i}) - \Delta_i(x_{\nu+1,i})] = \operatorname{sg}[\Delta^{(\nu)} - \Delta^{(\nu+1)}] \tag{268}$$

Алеж

$$\operatorname{sg}[p_{ni}(x_{\nu+1,i}) - p_{ni}(x_{\nu i})] = \operatorname{sg}\int_{x_{\nu i}}^{x_{\nu+1,i}} p'_{ni}(x)\,dx \tag{269}$$

Отже, є необхідне в цьому разі, щоб хоч на деякій частині інтервалу $(x_{\nu i},\ x_{\nu+1,i})$ знак $p'_n(x)$ збігався із знаком $[\Delta^{(\nu)} - \Delta^{(\nu+1)}]$.

[1]) $\operatorname{sg}\alpha$ означає $\frac{\alpha}{|\alpha|}$

[2]) Деякі з чисел $\Delta^{(\nu)}$ можуть дорівнювати нескінченності, але у всякому разі з певним знаком.

А коли для якогонебудь значення ν буде $x_{\nu 0} \neq x_{\nu+1,0}$ і, отже, $x_{\nu+1,i} - x_{\nu i} > \delta - \eta$, то знак $p'_{ni}(x)$ знову таки не може на всьому інтервалі $(x_{\nu i}, x_{\nu+1,i})$ бути протилежний до знаку $[\Delta^{(\nu)} - \Delta^{(\nu+1)}]$. Справді бо, при сталому знакові $p'_{ni}(x)$ в зазначеному інтервалі ми мали б:

$$|p_{ni}(x_{\nu+1,i}) - p_{ni}(x_{\nu i})| = \int_{x_{\nu i}}^{x_{\nu+1,i}} |p'_{ni}(x)|\,dx > \frac{4M}{\delta - \eta} \cdot \frac{\delta - \eta}{2} = 2M \tag{270}$$

і, виходить, оскільки $|f(x)| < M$, мали б

$$\operatorname{sg}[\Delta_i(x_{\nu i}) - \Delta_i(x_{\nu+1,i})] = \operatorname{sg} p'_{ni}(x), \tag{271}$$

Отже: або на всьому інтервалі $(x_{\nu i}, x_{\nu+1,i})$

$$\operatorname{sg} p'_{ni}(x) = \operatorname{sg}[\Delta_i(x_{\nu i}) - \Delta_i(x_{\nu+1,i})] = \operatorname{sg}[\Delta^{(\nu)} - \Delta^{(\nu+1)}]. \tag{272}$$

або знак $p'_n(x)$ не є сталий в розглядуваному інтервалі, і тоді співвідношення (272) між знаками має місце принаймні на деякій частині цього інтервалу.

Підсумовуючи, бачимо, що всередині (у вузькому розумінні) кожного з $n+1$ послідовних інтервалів $(x_{\nu i}, x_{\nu+1,i})$, $\nu = 0, 1, \ldots, n$, напевне можна знайти бодай одну точку x, в якій $\operatorname{sg} p'_{ni}(x) = \operatorname{sg}[\Delta^{(\nu)} - \Delta^{(\nu+1)}]$. В цих $n+1$ послідовних точках поліном $p'_{ni}(x)$ степеня $\leqslant n-1$ повинен мати альтернуючий знак на підставі (266), (267), (262'). Це завершує доведення від супротивного.

Означена в попередній теоремі величина A — найбільше значення нижньої границі найкращого наближення ρ, що її можна встановити для функції $f(x)$ на точковій множині E на підставі дослідження різниць

$$\Phi(x) - P_n(x;\ K_0, K_1, \ldots, K_n), \quad \varphi(x) - P_n(x;\ K_0, K_1, \ldots, K_n)$$

за допомогою другої частини теореми I § 16 та теореми II того самого параграфа, є, очевидно, при $\rho > k$, певна неперервна функція від параметрів полінома:

$$A = A(K_0, K_1, \ldots, K_n) \tag{273}$$

Наступна теорема, поруч тільки но доведеної теореми I, має істотне значення для тих методів, які є предметом дальшого розділу.

Теорема II. *Коли $\rho > k$, то кожному довільно малому додатному числу ε можна протипоставити додатне число $\eta < \rho - k$ таке, щоб нерівність*

$$A(K_0, K_1, \ldots, K_n) > \rho - \eta \tag{274}$$

мала наслідком, одностайно на всій точковій множині E, нерівність:

$$|P_n(x;\ K_0, K_1, \ldots, K_n) - P_n(x;\ K_{00}, K_{10}, \ldots, K_{n0})| < \varepsilon, \tag{275}$$

де

$$P_n(x;\ K_{00}, K_{10}, \ldots, (K_{n0}) = \Pi_n(x)$$

є (*єдиний в даному разі* — згідно з теоремою III § 16) *поліном найкращого наближення.*

Доведення цієї теореми становить майже дослівне повторення міркування, застосованого при доводженні аналогічної теореми § 15, де замість $A(K_0, K_1, \ldots, K_n)$ фігурувало $L(K_0, K_1, \ldots, K_n)$. В даному разі ми насамперед зауважимо, що, згідно з доведеною тільки но теоремою I, ми тут можемо мати на увазі лише поліноми P_n обмежені в своїй сукупності абсолютною величиною. Потім тут матиме істотне значення те, що, за прямим сенсом теореми I § 16, для кожного полінома $P_n \neq \Pi_n$ величина A є напевне менша, ніж ρ. Точкова множина $\mathfrak{S}$ в даному разі, очевидно, складається з однієї лише точки.

§ 18. Зауваження до можливої алгебризації границь неозначеності функції $f(x)$ в розглядуваній задачі

Означені вище на точковій множині E° дві функції $\Phi(x)$ та $\varphi(x)$ — границі неозначеності функції $f(x)$, є півнеперервні, відповідно зверху та знизу, на цій множині. Звідси виникає можливість розглядати кожну з них як границю монотонної (відповідно спадної або ростучої) послідовності поліномів. Щоб такі послідовності здійснити, насамперед доповнимо означення функцій $\Phi(x)$ та $\varphi(x)$, поширивши його на весь сегмент (a, b)[1]) так, щоб і при поширеному означенні їх властивість півнеперервності збереглася. Це, звичайно, можливе при дуже багатьох способах. Наприклад, можна застосувати, однаково для $\Phi(x)$ та $\varphi(x)$, лінійну інтерполяцію на кожному з тих (незамкнених) „суміжних" інтервалів (що їх множина може бути або скінченна, або счисленна), з яких складається доповнююча до E° частина сегмента (a, b), відповідно до значень $\frac{\Phi(x)+\varphi(x)}{2}$ на кінцях інтервалу. А в точках самої множини E° попередні означення (242) — (243) залишаються без зміни. Після цього, взявши для конкретності, наприклад, функцію $\Phi(x)$, означення якої поширено на ввесь сегмент (a, b), зробімо ось що.

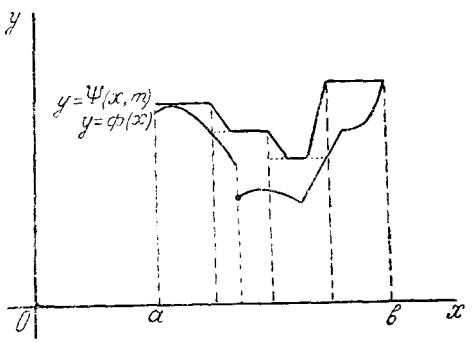

Рис. 7.

Поділивши сегмент (a, b) на 2^m рівних частин, побудуймо насамперед „східчасту" лінію, що її ордината на кожному з розглядуваних 2^m частинних сегментів зберігає стале значення, рівне максимуму функції $\Phi(x)$ на тому самому частинному сегменті. Потім замінюємо цю східчасту лінію, не зменшуючи ніде її ординати, на неперервну ламану

[1]) Нагадаємо, що, згідно з попереднім припущенням, кінці сегмента (a, b) спадаються з крайніми точками множини E°.

яку показано на рис. 7. Остання відрізняється від попередньої східчастої лінії наявністю злучних похилих сторін, уведених для заміни розміщених безпосередньо під ними (на рис. 7 показаних пунктиром) частин східчастої лінії, при чому, для означеності, горизонтальна проекція кожної з цих похилих сторін нехай дорівнює, наприклад, $\frac{1}{3} \cdot \frac{b-a}{2^m}$. Нарешті, позначивши ординату одержаної неперевної ламаної, як функцію абсциси за даного m, через $\Psi(x;\ m)$ та призначивши якусь залежну від m додатну величину δ_m, яка монотонно йде до нуля при $m \to \infty$, визначимо такий поліном $G^{(m)}(x)$, щоб задовольнялися на всьому сегменті $(a,\ b)$ нерівності

$$\Psi(x;\ m) + \delta_{m+1} < G^{(m)}(x) < \Psi(x;\ m) + \delta_m \tag{276}$$

а це завжди можливе, згідно з теоремою Вайерштраса щодо апроксимації неперервних функцій за допомогою поліномів [1]).

Не важко зрозуміти, базуючись на самому означенні півнеперервних функцій, що при $m \to \infty$ матимемо на всьому сегменті $(a,\ b)$, хоч взагалі і не одностайно,

$$\lim_{m \to \infty} G^{(m)}(x) = \Phi(x) \tag{277}$$

монотонно для кожного даного x на сегменті $(a,\ b)$.

Такий самий алгорифм, застосований з безпосередньо очевидними модифікаціями до функції $\varphi(x)$, приводить до визначення поліному $g^{(m)}(x)$, який задовольняє аналогічну умову

$$\lim_{m \to \infty} g^{(m)}(x) = \varphi(x) \tag{278}$$

Поліноми $G^{(m)}(x)$ та $g^{(m)}(x)$ при $m = 1,\ 2,\ 3, \ldots$ і здійснюють вищезгадані дві послідовності поліномів.

При кожному значенні m матимемо:

$$g^{(m)}(x) < \varphi(x) \leqslant \Phi(x) < G^{(m)}(x) \tag{279}$$

Впровадьмо позначення $2k_m = \max\limits_{x \subset E^\circ} (G^{(m)} - g^{(m)})$, $\Pi_n^{(m)}$, ρ_m, замість попередніх $2k$, Π_n, ρ, для того випадку, коли Φ та φ у виразах (245) та (246) замінюються відповідно на $G^{(m)}$ та $g^{(m)}$, при чому точкова множина E° зберігає своє значення як область зміни аргумента x.

Якщо не з безпосередньо практичного, то у всякому разі з погляду теоретичного не можна обминути питання про можливу заміну довільних півнеперервних функцій $\varphi(x)$ та $\Phi(x)$ на поліноми $g^{(m)}(x)$ та $G^{(m)}(x)$ в задачі найкращої поліноміальної апроксимації за принципом Чебишова.

Основне значення для нас тут має характер зв'язку між поліномами $\Pi_n^{(m)}$ та шуканим поліномом Π_n при $m \to \infty$.

[1]) Щоб уникнути будьякого свавілля в доборі $G^{(m)}(x)$, можна було б вимагати, щоб степінь цього полінома був щонайнижчий і додати ще якінебудь додаткові умови характеру більш-менш штучного. В цьому насправді тут немає потреби: в дальших наших міркуваннях ми лише припустимо, що для кожного значення m визначено в певний спосіб один поліном $G^{(m)}(x)$, який задовольняє нерівності (276).

Позначивши через L_m наближення, що його дає поліном $\Pi_n^{(m)}$ для функції f на точковій множині E, ми, очевидно, матимемо завжди[1]):

$$\rho_m > L_m \geqslant \rho \tag{280}$$

В тому окремому випадку, коли функції Φ та φ є неперервні, поліноми $G^{(m)}$ та $g^{(m)}$ прямуватимуть відповідно до Φ та φ одностайно на множині E°, отже тоді, як легко зрозуміти, матимемо

$$\lim_{m\to\infty} \rho_m = \rho, \tag{281}$$

а це має наслідком також

$$\lim_{m\to\infty} L_m = \rho \tag{282}$$

Отже, при $m \to \infty$ в цьому разі, згідно з теоремою параграфа 15, поліном $\Pi_n^{(m)}$ буде нескінченно близький до шуканого полінома (або якогось із поліномів) Π_n.

Але характеристичні властивості півнеперервних функцій дозволяють довести, що граничні рівності (281) та (282) зберігають силу і в загальному випадку функцій Φ та φ півнеперервних у власному розумінні, не зважаючи на те, що поліноми $G^{(m)}$ та $g^{(m)}$ ідуть у даному разі до відповідних граничних функцій неодностайно на точковій множині E°.

До цього доведення ми й переходимо.

Почнімо з зауваження, що при досить великих значеннях m матимемо

$$k_m < k + \varepsilon, \tag{283}$$

позначаючи через ε довільно мале додатне число. Справді, доводячи це від супротивного, приходимо на підставі звичайно вживаного міркування до необхідності, для певного значення $\varepsilon > 0$, припустити існування на точковій множині E° якоїсь точки $x = \xi$, що має таку властивість: при якому завгодно великому m і при якому завгодно малому δ найбільше значення різниці $G^{(m)} - g^{(m)}$ на інтервалі $(\xi - \delta, \xi + \delta)$ перевищує $2(k + \varepsilon)$. Але це неможливе з такого міркування: з одного боку, при досить великому $m = m_\varepsilon$, зважаючи на (277) та (278), матимемо:

$$\left.\begin{aligned} G^{(m_\varepsilon)}(\xi) &< \Phi(\xi) + \frac{\varepsilon}{2} \\ g^{(m_\varepsilon)}(\xi) &> \varphi(\xi) - \frac{\varepsilon}{2} \end{aligned}\right| \tag{284}$$

З другого ж боку, на підставі неперервності $G^{(m_\varepsilon)}$ та $g^{(m_\varepsilon)}$ а також монотонного прямування поліномів $G^{(m)}$ та $g^{(m)}$ відповідно до Φ та φ, ми при досить малому $\delta = \delta_\varepsilon$ та при $m \geqslant m_\varepsilon$ матимемо:

[1]) Пор. увагу II наприкінці § 15.

$$\left.\begin{aligned} G^{(m)}(x) &< G^{(m_\varepsilon)}(\xi) + \frac{\varepsilon}{2} \\ g^{(m)}(x) &> g^{(m_\varepsilon)}(\xi) - \frac{\varepsilon}{2} \end{aligned}\right| \text{ при } \xi - \delta_\varepsilon < x < \xi + \delta_\varepsilon \tag{285}$$

А це разом з (284) показує неможливість нерівності

$$G^{(m)}(x) - g^{(m)}(x) \geqslant 2k + 2\varepsilon \tag{286}$$

при $m \geqslant m_\varepsilon$ та $|x - \xi| \leqslant \delta_\varepsilon$, що й завершує доводження від супротивного.

Далі, відповідно до зауваження II наприкінці § 15, матимемо

$$\rho_1 > \rho_2 > \rho_3 > \tag{287}$$

Звідси, маючи ще на увазі (280), бачимо, що повинна існувати границя

$$\lim_{m \to \infty} \rho_m = \rho' \geqslant \rho \geqslant k \tag{288}$$

Можливі два випадки.

1°. $\rho' = k$. В такому разі граничні рівності (281) та (282), певна річ, справджуються.

2°. $\rho' > k$. Ми покажемо, що і в цьому випадку задовольняється (281) отже і (282).

Справді, для досить великих значень m ми, зважаючи на (283), матимемо в цьому разі також $\rho_m > k_m$, отже поліном $\Pi_n^{(m)}$ неодмінно повинен тоді задовольняти умову 1° теореми III параграфа 16, при чому ми замінимо у ряді рівностей (254) або (255) відповідно Φ, φ, p_n, L на $G^{(m)}$, $g^{(m)}$, $\Pi_n^{(m)}, \rho_m$, а також $x_0, x_1, \ldots, x_{n+1}$ на $x_{0m}, x_{1m}, \ldots, x_{n+1,m}$. Розглядаючи $x_{0m}, x_{1m}, \ldots$ $\ldots, x_{n+1,m}$ як координати точки в $n+2$-мірному просторі та беручи до уваги обмеженість цих чисел, приходимо до необхідності існування принаймні однієї граничної точки $(x'_0, x'_1, \ldots, x'_{n+1})$, що її координати, зрозуміла річ, збігаються з деякими числами тієї самої замкненої множини E°. З послідовності $m = 1, 2, 3, \ldots$ можна буде виділити частинну нескінченну ростучу послідовність чисел

$$m_1 < m_2 < m_3 < \tag{289}$$

так, щоб мали місце граничні рівності:

$$\lim_{m \to \infty} x_{im} = x'_i, \quad i = 0, 1, 2, \ldots, n+1 \tag{290}$$

надаючи надалі індексові m лише тих числових значень, які входять у послідовність (289). До того ж ми маємо, очевидно, право припустити, що коли m перебігає послідовність значень (289), система рівностей (254) або (255), яка відповідає поліномові $\Pi_n^{(m)}$, починається кожного разу з різниці зафіксованого знака[1]), напр. із $G^{(m)}(x_{0m}) - \Pi_n^{(m)}(x_{0m}) = + \rho_m$.

[1]) На останню вимогу можна, звичайно, дивитися як на додаткову умову, яка повинна бути врахована при самому доборі послідовності значень (289).

Можна тепер насамперед пересвідчитись, що координати граничної точки $x'_0, x'_1, x'_2, \ldots, x'_{n+1}$ повинні бути всі різні між собою. Щоб це довести, ми візьмемо на увагу, що поліноми $\Pi_n^{(m)}$ при $m \to \infty$ залишаються обмеженими в своїй сукупності на точковій множині E^0, отже кожний із $n+1$ коефіцієнтів полінома $\Pi_n^{(m)}$ залишається абсолютною величиною нижче певної сталої межі, яку можна було б явно зазначити за допомогою Лягранжової інтерполяційної формули. Звідси виходить, що поліноми $\Pi_n^{(m)}$ при $m \to \infty$ (вважаючи n зафіксованим), є в своїй сукупності рівно-неперервні (egualmente continui) на сегменті (a, b) [1]. Коли ж так, то, припустивши, що принаймні дві координати x'_r та x'_s граничної точки мають однакове значення, яке позначимо чере ξ, ми на підставі умови 1° теореми III параграфа 16 негайно прийдемо до суперечності між нерівністю $\rho' > k$ та нерівностями типу (284) — (285).

Зауваживши це, ми можемо далі аналогічним шляхом прийти до того вирішального висновку, що при досить великому m у послідовності (289) та при досить малих δ_i в системі нерівностей $x'_i - \delta_i < x_{im} < x'_i + \delta_i$ $(i = 0, 1, \ldots, n+1)$ застосування теореми III (умова 1°) параграфа 16 до $\Pi_n^{(m)}$, $G^{(m)}$, $g^{(m)}$ призводить до системи $n+2$ нерівностей

$$\left.\begin{array}{l} \Phi(x'_0) - \Pi_n^{(m)}(x'_0) > \rho' - \varepsilon, \quad \Pi^{(m)}(x'_1) - \varphi(x'_1) > \rho' - \varepsilon, \\ \Phi(x'_2) - \Pi_n^{(m)}(x'_2) > \rho' - \varepsilon, \quad \Pi_n^{(m)}(x'_3) - \varphi(x'_3) > \rho' - \varepsilon \\ \cdots\cdots\cdots\cdots\cdots\cdots \end{array}\right\} \quad (291)$$

при довільно малому додатному ε. А на підставі теореми I параграфа 16 це приводить до твердження, що $\rho > \rho' - \varepsilon$ при довільно малому додатному ε. Отож, з огляду на (288), $\rho' = \rho$, а це й треба було довести.

Отже у всіх випадках поліном $\Pi_n^{(m)}(x)$ при досить великому значенні m буде скільки завгодно близький до шуканого полінома (або одного з поліномів) Π_n.

Щодо коефіцієнтів полінома $\Pi_n^{(m)}(x)$ та величини ρ_m, то в тому випадку, коли множина E^0 являє собою неперервний сегмент, або два чи кілька відокремлених сегментів, вони задовольняють певну систему алгебричних рівнянь, які можна одержати шляхом, аналогічним до того, який уперше вказав П. Л. Чебишов для окремого простого випадку, що стосувався функції $f(x) = x^{n+1}$. Докладніше на останньому питанні ми тут не зупинимось, бо прямого зв'язку з викладеними в дальших розділах методами воно не має.

[1]) Тобто, за даного n, кожному $\varepsilon > 0$ можна протипоставити таке $\eta > 0$, що, при $a \leqslant x \leqslant b$ та $|\theta| < 1$, нерівність $|\Pi_n^{(m)}(x + \theta\eta) - \Pi_n^{(m)}(x)| < \varepsilon$ буде забезпечена незалежно від m. Пор. відповідні міркування в §§ 19, 22 п. 1.

РОЗДІЛ VII

ПРО ДЕЯКИЙ МЕТОД ПОСЛІДОВНИХ НАБЛИЖЕНЬ, ЩОБ ВИЗНАЧАТИ ПОЛІНОМИ НАЙКРАЩОГО НАБЛИЖЕННЯ

Ми тут насамперед (§§ 19, 20) означимо два алгорифми, що кожен з них може стати за основу для збіжного процесу послідовного наближення. Далі (§ 21) ми зупинимось на застосуванні методу до задачі обчислення поліномів найкращого наближення для $|x|$. У перших параграфах, для більшої конкретності викладу, трактується випадок однозначної та неперервної функції, заданої на неперервному сегменті. В параграфі 22 результати поширюються на загальну задачу, сформульовану на початку розділу VI, і розглядаються деякі дальші застосування.

§ 19. Попередні зауваження. Перший алгорифм

Задача, яку ми тут, у першому викладі, матимемо на увазі, полягає в тому, щоб визначити для даної неперервної функції $f(x)$ поліном найкращого наближення $\pi_n(x)$ степеня не вище за n, на множині C усіх точок неперервного сегмента (a, b). Нехай буде $p_n(x)$ будь-який поліном (не найкращого наближення) степеня $\leqslant n$, що розглядається як наближене представлення функції $f(x)$ на точковій множині C.

Позначимо

$$f(x) - p_n(x) = \Delta(x) \tag{292}$$

і, крім того, відповідно до наших попередніх позначень,

$$\max_{x \subset C} |\Delta(x)| = L \tag{293}$$

Нехай будуть $G \subset C$ та $H \subset C$ замкнені точкові множини[1]), на яких задовольняються відповідно нерівності

$$\Delta(x) \geqslant R \quad \text{та} \quad \Delta(x) \leqslant -R, \tag{294}$$

позначаючи через R будьяке додатне число, менше за L. Нехай D буде найменший сегмент, що обіймає $G + H$.

Точкова множина $D - (G + H)$[2]) складається із скінченної або счисленної кількості інтервалів (незамкнених)[3]). Позначимо

$$I_1, \ I_2, \ I_3, \tag{295}$$

[1]) Може трапитись, що одна з цих двох множин виявиться порожня.

[2]) Якщо вона не є порожня, звичайно.

[3]) Див., напр., П. Александров и А. Колмогоров. Введ. в теор. функц. действ. перем., ГТТИ, 1933.

ті з них, які водночас межують із точковими множинами G та H. На властивостях цих інтервалів ми повинні спинитися трохи докладніше.

Поперше, припускаючи лише $|p_n(x)| \leqslant 2M$ на сегменті (a, b), де M означає, відповідно до (241), $\max|f(x)|$, можна буде для довжин інтервалів (295) a priori визначити певну додатну нижню межу $2l$, яка, при даних функції $f(x)$, сегменті (a, b) та числі n, залежить іще лише від R — і це в такий спосіб, що при зміні R величина $2l$ не може йти до нуля інакше, як одночасно із самим R. Щоб у цьому переконатись, досить взяти до уваги, що оскільки $f(x)$ є неперервна на сегменті (a, b), включаючи його кінці, а з другого боку максимальне абсолютне значення похідної $p'_n(x)$ на сегменті (a, b), згідно з відомою теоремою А. Маркова[1]), за даних умов не може перевищувати величину

$$\frac{2}{b-a} n^2 \cdot 2M = \frac{4n^2 M}{b-a} \tag{296}$$

то функції $\Delta(x) = f(x) - p_n(x)$ в своїй сукупності[2]) є рівно неперервні (egualmente continue)[3]) на сегменті (a, b).

Подруге, щодо числа інтервалів (295), то, якщо воно є більше, ніж n, найкраще наближення ρ напевне є більше, ніж R — на підставі теореми I § 16. Ми й тут позначимо через A (пор. § 17, теор. I) найбільше значення нижньої межі, яку можна встановити для найкращого наближення ρ на підставі дослідження різниці $\Delta = f - p_n$ — припускаючи вже тут, зрозуміла річ, не лише $f(x)$, але й поліном $p_n(x)$ фіксованим[4]). Коли ж число інтервалів (295) було б менше, ніж $n+1$, при якому завгодно малому R, ми поклали б $A = 0$. У всіх випадках ми матимемо, очевидно,

$$A < \rho < L \text{ [5])} \tag{297}$$

Після цих попередніх зауважень переходимо до з'ясування суті самого алгорифма.

Визначивши (точно або з певним наближенням) числа L та A, які відповідають поліномові $p_n(x)$, візьмімо два числа, що містяться між A та L:

$$\left.\begin{aligned} R_1 = A + \vartheta_1 (L - A), \quad R_2 = A + \vartheta_2 (L - A), \\ \text{де} \quad 0 < \vartheta_1 < \vartheta_2 < 1 \end{aligned}\right\} \tag{298}$$

[1]) А. Марков. Об одном вопросе Д. И. Менделеева (Записки. Рос. Акад. Наук, 1889, т. 62). В. Марков. О функциях наименее уклоняющихся от нуля в данном промежутке, СПБ, 1892.

[2]) Вважаючи за дані $f(x)$, сегмент (a, b) й число n та фіксуючи лише верхню межу $2M$ абсолютної величини полінома $p_n(x)$ на (a, b).

[3]) Пор. останню виноску § 18. Нам немає потреби тут писати в явній формі вираз величини $2l$: це буде зроблено, за умов більш загальних, в § 22.

[4]) В звичайних випадках відшукання екстремальних точок функції $\Delta(x) = f(x) - p_n(x)$ на сегменті (a, b) являє природний шлях, щоб визначити число A.

[5]) Тривіальний випадок, коли $\rho = 0$, залишаємо осторонь; $L > \rho$, бо даний поліном $p_n(x)$ не є поліном найкращого наближення.

Позначмо через G_1, H_1; G_2, H_2 чотири точкові множини, які одержуємо замість раніше розглянутих точкових множин G та H, замінюючи R один раз на R_1 та другий раз на R_2. Нехай будуть

$$I_1,\ I_2, \ldots, I_m\,(m \leqslant n) \tag{299}$$

інтервали (295), які відповідають випадкові $R = R_1$.

Ми можемо понизити величину наближення, що його дає розглядуваний поліном $P_n(x)$, додаючи до нього поправку $\omega_n(x)$ в такій формі [1]):

$$\omega_n(x) = \lambda q_{n-m}(x)\ (x - \xi_1)\ (x - \xi_2) \ldots (x - \xi_m) \tag{300}$$

Тут $\xi_1, \xi_2, \ldots, \xi_m$ є деякі числа, що їх беремо відповідно на інтервалах $I_1, I_2, \ldots, I_m$ за тією єдиною умовою, щоб мінімальна віддаль δ між точками $\xi_1, \xi_2, \ldots, \xi_m$ та точковою множиною $G_2 + H_2$ навіть при повторному застосуванні поправки виду (300) залишалася кожного разу вища якоїсь додатної фіксованої межі [2]). $q_{n-m}(x)$ є якийсь поліном степеня $\leqslant n - m$, який зберігає сталий знак на точковій множині $G_1 + H_1$ і, крім того, є підпорядкований вимозі

$$\min_{x \subset G_2 + H_2} |q_{n-m}(x)| : \max_{x \subset c} |q_{n-m}(x)| > r > 0, \tag{301}$$

позначаючи через r якесь фіксоване число [3]). Нарешті, щодо сталого множника λ, то він підпорядковується таким умовам.

1°. Знак λ повинен бути такий, щоб задовольнялася вимога $\omega_n(x) > 0$ на всій точковій множині G_1, отже $\omega_n(x) < 0$ на H_1.

2°. Абсолютна величина λ повинна відповідати вимозі

$$\max_{x \subset C - (G_1 + H_1)} |\Delta(x) - \omega_n(x)| < R_2, \tag{302}$$

[1]) Якби виявилося, що інтервалів (299) зовсім немає, ми взяли б, звичайно, $\omega_n(x) = \lambda \cdot q_n(x)$.

[2]) Цю вимогу можна буде напевне задовольнити, залишаючи осторонь тривіальний випадок, коли $\rho = 0$, тобто коли $f(x)$ сама є поліном степеня $\leqslant n$. Беручи, наприклад, точку $\xi_i\ (i = 1, 2, \ldots, m)$ в середині відповідного інтервалу I_i, ми мали б $\delta > l$, отже розглядуваної умови було б напевне додержано — надаючи числу R значення не надто малі при повторному застосуванні поправки виду (300). Але те саме можна забезпечити, очевидно, і при іншому доборі точок ξ_i на інтервалах I_i, і центральне положення точок ξ_i на цих інтервалах не є взагалі оптимальне.

[3]) Для більшої означеності коефіцієнт найвищого члена в поліномі $q_{n-m}(x)$ можемо вважати рівним одиниці. Зауважмо далі, що можна було б завжди брати $q_{n-m}(x) \equiv 1$: при цьому ми мали б, певна річ, також $r = 1$. Але за іншого належного добору множника $q_{n-m}(x)$, як величини змінної, можна взагалі значно поліпшити збіжність розглядуваного процесу послідовного наближення.

що її можна у всякому разі сполучити з такою умовою:

$$|\lambda| \geqslant \frac{R_2 - R_1}{(b-a)^m \max\limits_{x \subset \sigma} |q_{n-m}(x)|} \text{ [1]} \tag{303}$$

Не важко пересвідчитись, що, додаючи поправку $\omega_n(x)$ до полінома $p_n(x)$, ми знижуємо наближення L щонайменше на величину, яка дорівнює меншій із таких двох:

$$L - R_2 = (1 - \vartheta_2)(L - A) > (1 - \vartheta_2)(L - \rho) \tag{304}$$

та

$$\left.\begin{aligned} &\frac{R_2 - R_1}{(b-a)^m} \cdot r \cdot \delta^m = (\vartheta_2 - \vartheta_1)(L - A) \cdot r \left(\frac{\delta}{b-a}\right)^m \\ &\geqslant (\vartheta_2 - \vartheta_1) r \left(\frac{\delta}{b-a}\right)^n (L - A) > (\vartheta_2 - \vartheta_1) r \left(\frac{\delta}{b-a}\right)^n (L - \rho) \end{aligned}\right\} \tag{305}$$

Отже, позначаючи через L_1 наближення, що його дає скорегований поліном $p_n(x) + \omega_n(x)$, матимемо:

$$\left.\begin{aligned} L_1 - A &< \theta(L - A) \\ L_1 - \rho &< \theta(L - \rho) \end{aligned}\right\} \tag{306}$$

де правильний дріб θ дорівнює більшому з двох таких:

$$\vartheta_2$$

та

$$1 - (\vartheta_2 - \vartheta_1) r \cdot \left(\frac{\delta}{b-a}\right)^n \tag{307}$$

Тепер ясно, що при повторному застосуванні вказаної тут поправки, тобто коли аналогічний алгорифм застосовується до скорегованого полінома $p_n + \omega_n$ і т. д., різниця $L - \rho$ прямуватиме до нуля принаймні як член геометричної спадної прогресії, якщо числа δ, r, $1 - \vartheta_2$, $\vartheta_2 - \vartheta_1$, оскільки вони є в нашому розпорядженні, залишаються більші, ніж якісь додатні фіксовані границі. Але при $L \to \rho$ відповідний поліном, як ми знаємо (§ 15), сам одностайно йде до шуканого $\pi_n(x)$ на сегменті (a, b).

Збіжність означеного тут процесу послідовних наближень є таким чином забезпечена при досить широких умовах щодо добору ϑ_1, ϑ_2, ξ_1, ξ_2, ..., ξ_m, $q_{n-m}(x)$, λ. Це, при фактичному здійсненні алгорифма, дозволяє з одного боку, допускати досить значні наближення в деяких обчисленнях, а з другого боку, ми маємо змогу запровадити деякі додаткові умови, щоб забезпечити збіжність досить швидку. До цього питання ми повернемось в слідуючому параграфі.

[1]) Взагалі ж, маючи на оці можливо швидку збіжність розглядуваного процесу послідовних наближень, беремо абсолютну величину λ значно більшу цієї величини, оскільки це сумісне з вимогою (302). Крім інтуітивного врахування загального перебігу функції $y = \Delta(x)$ на сегменті (a, b), тут багато важать ті вказівки, що їх можна собі здобути з другого алгорифма, який становить об'єкт наступного параграфа.

§ 20. Другий алгорифм. Про деяке поєднання обох алгорифмів

Ми тут також починаємо з того, що визначаємо числа L та A для полінома $p_n(x)$, що розглядається як вихідне наближене представлення (не найкраще) функції $f(x)$ на сегменті (a, b), при чому припустімо, що $A > 0$. В такому разі напевне можна на сегменті (a, b) зазначити $n+2$ послідовні точки $x_0 < x_1 < x_2 < \ldots < x_{n+1}$ за такими умовами:

1°. Різниця $\Delta = f - p_n$ має альтернуючий знак у цих точках:

$$\operatorname{sg} \Delta(x_0) = -\operatorname{sg} \Delta(x_1) = \operatorname{sg} \Delta(x_2) = \ldots = (-1)^{n+1} \operatorname{sg} \Delta(x_{n+1}) \tag{308}$$

2°. Найменше із $n+2$ значень $|\Delta(x_i)|$ дорівнює A.

3°. Найбільше із тих самих $n+2$ значень дорівнює L, або принаймні, загальнішим чином, найбільше із $n+2$ відношень $\frac{\Delta(x_i) - A}{L - A}$ залишається вище якоїсь додатної фіксованої границі.

Як поправку $\Omega_n(x)$ до полінома $p_n(x)$ ми тепер візьмемо поліном найкращого наближення, степеня $\leqslant n$, щодо функції $\Delta(x) = f(x) - p_n(x)$ на точковій множині, яка складається із зазначених $n+2$ точок.

Згідно з теоремою III § 16 (умова 1°), коефіцієнти $K_0, K_1, K_2, \ldots, K_n$ полінома $\Omega_n(x)$ та величину ρ' наближення, яке він дає для $\Delta(x)$ на множині $n+2$ точок $x_0, x_1, x_2, \ldots, x_{n+1}$, ми визначимо, розв'язавши таку систему $n+2$ лінійних рівнянь (де невідомі $n+1$ коефіцієнтів полінома $\Omega_n(x)$ покищо явно не написані) [1]):

$$\left.\begin{aligned} \Delta(x_0) - \Omega_n(x_0) &= -[\Delta(x_1) - \Omega_n(x_1)] = \Delta(x_2) - \Omega_n(x_2) = \\ &= -[\Delta(x_3) - \Omega_n(x_3)] = \ldots = (-1)^{n+1}[\Delta(x_{n+1}) - \Omega_n(x_{n+1})] = \pm\rho' \end{aligned}\right\} \tag{309}$$

Щодо знака перед ρ', він визначається з умови [2])

$$\operatorname{sg}(\pm\rho') = \operatorname{sg} \Delta(x_0) \tag{310}$$

Розгортаючи систему (309), ми дістали б $n+2$ рівняння такого виду:

$$\left.\begin{array}{l} \pm\rho' + K_0 + K_1 x_0 + K_2 x_0^2 + \cdots + K_n x_0^n = \Delta(x_0) \\ \mp\rho' + K_0 + K_1 x_1 + K_2 x_1^2 + \cdots + K_n x_1^n = \Delta(x_1) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ (-1)^{n+1} \cdot (\pm\rho') + K_0 + K_1 x_{n+1} + K_2 x_{n+1}^2 + \cdots + K_n x_{n+1}^n = \Delta(x_{n+1}) \end{array}\right\} \tag{309'}$$

[1]) P. Kirchberger в своїй цитованій вище дисертації, слід думати, перший зайнявся вивченням властивостей найкращого наближеного представлення даної функції $f(x)$ за допомогою полінома степеня $\leqslant n$ на множині $n+2$ точок (нам відома лише та частина його дисертації, яка була надрукована в Mathem. Annalen). Але найбільш детально дослідив ці властивості та поклав їх в основу систематичного трактування поліномів найкращого наближення C. de la Vallée-Poussin — див. цитовані вище дві праці.

[2]) В противному разі, як легко бачити, поліном $\Omega_n(x)$ мусів би мати в усіх $n+2$ точках $x_0, x_1, \ldots, x_{n+1}$ такі самі знаки, як $\Delta(x)$, що, очевидно, не можливе.

Визначаючи звідси невідому $\pm\rho'$, ми після простого перетворення що сходить до поділу детермінантів Вандермонда на один і той самий вираз, знаходимо:

$$\pm\rho' = \frac{a_0\Delta(x_0) - a_1\Delta(x_1) + a_2\Delta(x_2) + \cdots + (-1)^{n+1} a_{n+1}\Delta(x_{n+1})}{a_0 + a_1 + a_2 + \cdots + a_{n+1}} \tag{311}$$

де, при $i = 0, 1, 2, \ldots, n+1$,

$$a_i = \frac{1}{(x_i - x_0)(x_i - x_1)\ldots(x_i - x_{i-1})(x_{i+1} - x_i)\ldots(x_{n+1} - x_i)} \tag{312}$$

Отже, зважаючи на (310) та (308),

$$\rho' = \frac{a_0|\Delta(x_0)| + a_1|\Delta(x_1)| + \cdots + a_{n+1}|\Delta(x_{n+1})|}{a_0 + a_1 + \cdots + a_{n+1}} \tag{313}$$

Щодо самого полінома $\Omega_n(x)$, який виражає шукану поправку до $p_n(x)$, ми тепер можемо визначити його безпосередньо на підставі (309) за допомогою Лягранжової інтерполяційної формули, за $n+1$ точками [1]).

Щоб вияснити роль поправки $\Omega_n(x)$ при повторному застосуванні цього алгорифма, звернімо увагу насамперед на те, що величина ρ' виражається, згідно з (313), як узагальнена середня арифметична, з додатними вагами $a_0, a_1, \ldots, a_{n+1}$, додатних величин $|\Delta(x_0)|, \ldots, |\Delta(x_{n+1})|$, отже величина ρ' є проміжна між найменшою та найбільшою з цих $n+2$ величин і, значить, напевне $\rho' > A$. Коли ж позначимо через A_1 значення числа A для скорегованого полінома $p_n + \Omega_n$, то матимемо, очевидно [2]),

$$A_1 \geqslant \rho' > A \tag{314}$$

Щоб прийти до більш точних висновків, візьмімо до уваги, що на підставі теореми I § 17 поліноми p_n, $p_n + \Omega_n$ і т. д., які ми послідовно одержуватимемо при повторному застосуванні розглядуваного алгорифма, будуть в своїй сукупності обмежені на сегменти (a, b), отже і рівно неперервні (egualmente continui) на цьому сегменті. А коли так, то легко зрозуміти, що всі різниці $x_{i+1} - x_i$ $(i = 0, 1, \ldots, n)$ для будьякого з цих послідовно одержаних поліномів будуть більші за якесь фіксоване додатне число $2l'$, отже всі різниці $x_i - x_j$ $(i > j)$ для будьякого з цих поліномів міститимуться між двома фіксованими межами:

$$2l' < x_i - x_j \leqslant b - a \tag{315}$$

Значить, і числа $a_0, a_1, \ldots, a_{n+1}$ у формулі (313) також усі міститимуться між певними двома фіксованими додатними межами. Після ж цього, беручи на увагу умови 2° та 3°, яким підпорядкований ансамбль $n+2$ точок $x_0, x_1, \ldots, x_{n+1}$, ясно, що ми матимемо:

$$\rho' - A > \theta(L - A), \tag{316}$$

[1]) Зайва $n+2$-га точка буде нам для перевірки.

[2]) Бо різниця $f - (p_n + \Omega_n)$ набуває за чергою значення $\pm\rho'$ та $\mp\rho'$ в $n+2$ послідовних точках $x_0, x_1, \ldots, x_{n+1}$, але цей ансамбль $n+2$ точок для полінома $p_n + \Omega_n$ взагалі не є оптимальний щодо визначення відповідного значення A_1.

позначаючи через θ певне фіксоване додатне число, менше одиниці. Нерівність же (314) заміниться на більш уточнену:

$$A_1 - A \geqslant \rho' - A > \theta(L - A) > \theta(\rho - A) \tag{317}$$

звідси:

$$\rho - A_1 < (1 - \theta)(\rho - A) \tag{318}$$

Це показує, що при повторному застосуванні поданого тут алгорифма величина різниці $\rho - A$ наближатиметься до нуля принаймні як член геометричної спадної прогресії.

Отже матимемо:

$$\lim A = \rho \tag{319}$$

і, значить, на підставі теореми II § 17 матимемо водночас

$$\lim L = \rho \tag{320}$$

і відповідний поліном знову наближатиметьея до $\pi_n(x)$, як до границі одностайно на сегменті (a, b).

Порівнюючи між собою два алгорифми, означені один в попередньому параграфі, а другий в даному, слід зауважити:

а) що при повторному застосуванні означеного тут другого алгорифма, на відміну від першого, величина L, хоч і наближається до ρ як до границі, але останній процес тут уже не буде неодмінно монотонний.

б) будуючи поправку $\Omega_n(x)$, ми виходимо власне не з картини перебігу функції $\Delta(x)$ на всьому сегменті (a, b), а із врахування тих значень, яких ця функція набирає лише в $n+2$ точках. Але остання хиба другого алгорифма водночас становить його серйозну перевагу, яка зумовлює порівняльну простоту розрахунку.

Метод, який ми беремо на себе сміливість пропонувати для практичного застосування при обчисленні поліномів найкращого наближення, полягає в певному поєднанні обох тут означених алгорифмів, з метою дійти можливо швидкої збіжності процесу послідовних наближень. Справді, і один і другий алгорифм залишає певний простір щодо добору деяких елементів при визначенні поправки $\omega_n(x)$ чи $\Omega_n(x)$. Це дозволяє, беручи за основу один із двох алгорифмів, разом з тим задовольнити істотні вимоги другого алгорифма. Інакше кажучи, будуючи поправку типу $\omega_n(x)$ чи $\Omega_n(x)$, ми можемо забезпечити, щоб величина L кожного разу знижалася при одночасному збільшенні величини A.

Другий алгорифм може бути рекомендований, взагалі, для тих випадків, коли при його застосуванні величина L зменшується монотонно[1]). Вихідне наближене представлення $p_n(x)$ можна при цьому одержати, наприклад, будуючи поліном найкращого наближення для самої функції $f(x)$

[1]) Слід підкреслити, що з практичного погляду доводиться мати на увазі, власне, процес послідовних наближень, який складається із скінченного числа операцій — іноді однієї-двох.

на якомунебудь ансамблі $n+2$ точок сегмента (a, b) [1]): для здобутого в такий спосіб полінома $p_n(x)$ величина A буде взагалі більша за нуль [2]), як це й вимагається для застосування другого алгорифма. Загальний характер цього алгорифма, який допускає великою мірою стандартизацію обчислень, дозволяє доручати його реалізацію обчислювачам звичайної кваліфікації.

Коли ж бажано уникнути деяких „трудомістких" обчислень, зв'язаних, при точному здійсненні другого алгорифма, з застосуванням Лягранжової формули, або коли з інших причин доцільність точного застосування другого алгорифма викликає ті чи ті сумніви, — можна практикувати в певних випадках таку модифікацію цього алгорифма, коли попереду оцінюється більш або менш приблизно величина ρ', і потім добираються, згідно із схемою першого алгорифма, абсциси $\xi_1, \xi_2, \ldots, \xi_m$, а також і множник $q_{n-m}(x)$ в такий спосіб, щоб з певним наближенням реалізувати пропорції:

$$\begin{gathered}\omega_n(x_0)\,\omega_n(x_1):\omega_n(x_2):\omega_n(x_3):\ldots= \\ =[\Delta(x_0)\mp\rho']:[\Delta(x_1)\pm\rho']:[\Delta(x_2)\mp\rho']:[\Delta(x_3)\pm\rho']:\ldots\end{gathered} \tag{321}$$

нарешті добирається множник λ так, щоб співвідношення (309) наближено задовольнялися.

§ 21. Застосування до задачі наближеного поліноміального представлення $|x|$ на сегменті $(-1,1)$

Поліноми найкращого наближення $\pi_n(x)$ парної функції $|x|$ на сегменті $(-1, 1)$ можуть містити (це добре відомо) лише парні степені x: в противному разі поліном $\pi_n(x)$ не був би єдиний, бо поліном $\pi_n(-x)$ давав би однакове наближення, не бувши тотожним із $\pi_n(x)$, а це при неперервності функції $|x|$ не є можливе — зважаючи на висновки теореми III § 16. Замінюючи x^2 на x, бачимо, що задача найкращого наближеного представлення $|x|$ на сегменті $(-1, 1)$ за допомогою полінома степеня $\leqslant 2\nu+1$ еквівалентна до задачі найкращого наближеного представлення $\sqrt{x}$ на сегменті $(0,1)$ за допомогою полінома степеня $\leqslant \nu$.

Зупинімось насамперед більш докладно на прикладі обчислення полінома $\pi_5(x)$ для $|x|$ на сегменті $(-1, 1)$, при чому ми здійснимо це обчислення, для порівнення, двома різними способами, відповідно до двох схем, зазначених наприкінці попереднього параграфа.

В розглядуваному випадку $\nu=2$, тобто питання сходить до визначення найкращого наближеного представлення функції $\sqrt{x}$ на сегменті $(0,1)$ за допомогою полінома виду $\alpha+\beta x+\gamma x^2$.

[1]) Можна тут також взяти для вихідного наближення поліном наближення за принципом найменших квадратів: різниця $\Delta(x)$ при цьому напевне міняє знак не менше як $n+1$ разів згідно із загальним зауваженням, що належить D. Jackson-ові, On functions of closest approximation (Transact. of the Amer. Math. Soc., 1921, p. 126).

[2]) Бо вона не може бути менша за величину найкращого наближення для функції $f(x)$ на розглядуваному ансамблі $n+2$ точок. Ми залишаємо, як і раніше, осторонь тривіальний випадок, коли $f(x)$ сама є поліном степеня $\leqslant n$.

Перше розв'язання. Як вихідне наближене представлення, збудуємо для $\sqrt{x}$ інтерполяційний поліном Лянгранжа з вузловими точками Чебишова [1]):

$$0{,}5+0{,}5\cos\frac{5\pi}{6}=0{,}0670;\ 0{,}5+0{,}5\cos\frac{\pi}{2}=0{,}5;\ 0{,}5+0{,}5\cos\frac{\pi}{6}=0{,}9330 \quad (322)$$

Поліном цей буде

$$p(x)=0{,}173+1{,}322x-0{,}505x^2 \quad (323)$$

Досліджуючи різницю $\Delta(x)=\sqrt{x}-p(x)$ на extrema на сегменті (0, 1), визначаємо ансамбль $\nu+2=4$ точок x_0, x_1, x_2, x_3:

$$x_0=0;\ x_1=0{,}199;\ x_2=0{,}729;\ x_3=1$$

$$\Delta(x_0)=-0{,}173;\ \Delta(x_1)=+0{,}030;\ \Delta(x_2)=-0{,}014;\ \Delta(x_3)=+0{,}010$$

Тут $L=0{,}173$; $A=0{,}010$.

Пробуючи наближено здійснити (без застосування формули Лянгранжа) алгорифм параграфа 20 на підставі врахування ансамбля точок (x_i, y_i), де $y=\Delta(x)$, $i=0, 1, 2, 3$ на площині (x, y) [2]), після пари простих спроб, які майже не потребують письмового обчислення, підібрано такий вираз поправки:

$$\omega_I(x)=-0{,}5(x-0{,}25)(x-0{,}9)$$

Значення, яких різниця $\Delta(x)-\omega_I(x)$ набирає при $x=x_0$, x_1, x_2, x_3, містяться в доволі близьких границях між 0,048 та 0,06, тобто вимога алгорифма параграфа 20 наближено задовольняється. Не важко з'ясувати собі, що поправка ця відповідає також і всім основним вимогам алгорифма параграфа 19 при деяких додатних ϑ_1 та ϑ_2 $\left(\frac{0{,}030-0{,}010}{0{,}173-0{,}010}<\vartheta_1<\vartheta_2<1\right)$ що їхні точні значення нас тут ближче не цікавлять, при чому в даному разі

$$\lambda=-0{,}5;\ q_{n-m}(x)=q_{2-0}(x)=(x-0{,}25)(x-0{,}9)$$

Інтервалів типу $I_1, I_2, \ldots, I_m$ в даному разі зовсім немає.

Додавши поправку $\omega_I(x)$ до полінома $p(x)$, дістанемо наступне наближене представлення:

$$p_I(x)=0{,}060+1{,}897x-1{,}005x^2 \quad (324)$$

Поступаючи з цим поліномом аналогічно до того, як раніше із $p(x)$, матимемо:

$$x_0=0;\ x_1=0{,}084;\ x_2=0{,}62;\ x_3=1$$

$$\Delta_I(x_0)=-0{,}060;\ \Delta_I(x_1)=+0{,}078;\ \Delta_I(x_2)=-0{,}063;\ \Delta_I(x_3)=+0{,}048$$

$$\omega_{II}(x)=-0{,}068(x+0{,}12)(x-0{,}72)$$

[1]) Див. P. L. Tchebychef. Théorie des mécanismes connus sous le nom de parallélogrammes, § 10. Пор. дальшу виноску щодо „точок відхилу Чебишова"

[2]) Графічне представлення зазначеного ансамбля точок дуже корисне.

(кількість m інтервалів $I_1, \ldots, I_m$ і тут дорівнює нулеві)

$$p_{II}(x) = 0{,}0659 + 1{,}938x - 1{,}073x^2 \tag{325}$$

Досліджуючи різницю $\Delta_{II}(x) = \sqrt{x} - p_{II}(x)$, маємо:

$$x_0 = 0;\ x_1 = 0{,}0801;\ x_2 = 0{,}6037;\ x_3 = 1$$

$$\Delta_{II}(x_0) = -0{,}0659;\ \Delta_{II}(x_1) = 0{,}0688;\ \Delta_{II}(x_2) = -0{,}0676;\ \Delta_{II}(x_3) = 0{,}0691$$

Як би ми хотіли на цьому спинитися, наближення 0,0691, що дає його поліном $p_{II}(x)$, можна зменшити на величину $\frac{0{,}0691 - 0{,}0676}{2} \approx 0{,}0008$, якщо замінити вільний член 0,0659 на $0{,}0659 + 0{,}0008$. Таким чином, замість (325), добудемо:

$$p_{II}(x) = 0{,}0667 + 1{,}938x - 1{,}073x^2, \tag{325'}$$

при чому тепер

$$\Delta_{II}(x_0) = -0{,}0667;\ \Delta_{II}(x_1) = 0{,}0680;\ \Delta_{II}(x_2) = -0{,}0684;\ \Delta_{II}(x_3) = 0{,}0683$$

Тут $L_{II} = 0{,}0684$; $A_{II} = 0{,}0667$; точніша нижня границя найкращого наближення ρ згідно з першим твердженням теореми I § 16, є $\frac{0{,}0667 + 0{,}0680}{2} \approx$ $\approx 0{,}0674$. Отже наближення $L_{II} = 0{,}0684$, що його дає поліном (325′), відрізняється від найкращого наближення не більше як на 0,001.

Проте ми складемо ще одну поправку:

$$\omega_{III}(x) = 0{,}0065(x - 0{,}15)(x - 0{,}88)$$

[тут $0{,}15 = \xi_1$; $0{,}88 = \xi_2$; $m = 2$; $q_{n-m}(x) = q_0(x) = 1$; $\lambda = 0{,}0065$]

$$p_{III}(x) = 0{,}0676 + 1{,}9313x - 1{,}0665x^2 \tag{326}$$

Досліджуючи різницю $\Delta_{III}(x) = \sqrt{x} - p_{III}(x)$, знаходимо:

$$x_0 = 0;\ x_1 = 0{,}0883;\ x_2 = 0{,}6037;\ x_3 = 1$$

$$\Delta_{III}(x_0) = -0{,}0676;\ \Delta_{III}(x_1) = 0{,}0676;\ \Delta_{III}(x_2) = -0{,}0678;\ \Delta_{III}(x_3) = 0{,}0676$$

Скорегувавши ще $p_{III}(x)$ з допомогою аддитивної сталої $\frac{-0{,}0678 + 0{,}0676}{2} =$ $= -0{,}0001$, візьмімо за остаточне наближене представлення поліном

$$p_{III}(x) = 0{,}0675 + 1{,}9313x - 1{,}0665x^2, \tag{326'}$$

при чому

$$\Delta_{III}(x_0) = -0{,}0675;\ \Delta_{III}(x_1) = 0{,}0677;\ \Delta_{III}(x_2) = -0{,}0677;\ \Delta_{III}(x_3) = 0{,}0677,$$

тобто $L_{III} = 0{,}0677$; $A_{III} = 0{,}0675$; точніша нижня границя для ρ є $\frac{0{,}0675 + 0{,}0677}{2} = 0{,}0676$ (теорема I § 16). Інакше кажучи,

$$0{,}0676 < \rho < 0{,}0677$$

Отже формула

$$|x| = 0{,}0675 + 1{,}9313x^2 - 1{,}0665x^4 \tag{327}$$

дає на сегменті $(-1, 1)$ наближення рівне 0,0677, яке перевищує найкраще наближення менше, ніж на 10^{-4}.

Друге розв'язання. Щоб одержати вихідне наближене представлення, будуємо поліном найкращого наближення степеня $\leqslant 2$ для функції $\sqrt{x}$[1]) на множині $\nu+2=4$ наступних точок (точок відхилу Чебишова[2]):

$$\left.\begin{aligned} x_0 &= \frac{1}{2}(1+\cos\pi)=0, \quad x_1=\frac{1}{2}\left(1+\cos\frac{2\pi}{3}\right)=\frac{1}{4}, \\ x_2 &= \frac{1}{2}\left(1+\cos\frac{\pi}{3}\right)=\frac{3}{4}, \; x_3=\frac{1}{2}(1+\cos 0)=1 \end{aligned}\right\} \tag{331}$$

При позначеннях § 20 матимемо тут:

$$\rho'=\left|\frac{a_0\sqrt{x_0}-a_1\sqrt{x_1}+a_2\sqrt{x_2}-a_3\sqrt{x_3}}{a_0+a_1+a_2+a_3}\right|=0{,}0447,$$

після чого визначаємо потрібний нам для першого наближення поліном $p(x)$ як інтерполяційний поліном Лягранжа з трьох його значень: $p(0)=$ $=0{,}0447$; $p\left(\frac{1}{4}\right)=0{,}4553$; $p(1)=0{,}9553$. Ми знайдемо:

$$p(x)=0{,}0447+1{,}887x-0{,}976x^2 \tag{332}$$

Далі застосовуємо алгорифм параграфа 20 в його точній формі.

Досліджуючи на extrema різницю $\Delta(x)=\sqrt{x}-p(x)$, знаходимо як раніше:

$$x_0=0;\; x_1=0{,}0819;\; x_2=0{,}6623;\; x_3=1$$

$$\Delta(x_0)=-0{,}0447;\; \Delta(x_1)=+0{,}0935;\; \Delta(x_2)=-0{,}0526;\; \Delta(x_3)=0{,}0443$$

Після цього з формули (313) маємо $\rho'_1=0{,}0672$ і потім за інтерполяційною формулою Лягранжа, беручи на увагу точки x_0, x_2, x_3:

$$\Omega_1(x)=0{,}0225+0{,}0537x-0{,}0991x^2$$

[1]) Для цього ми скористуємося рівняннями типу (309′), але замінивши $\Delta(x)$ на $f(x)=\sqrt{x}$.

[2]) Якщо йдеться про поліноміальне наближене представлення степеня n, точками відхилу Чебишова на інтервалі (a, b), називаємо точки

$$x_i=\frac{b+a}{2}+\frac{b-a}{2}\cos\frac{(n+1-i)\pi}{n+1},\; i=0,1,\ldots,n+1, \tag{328}$$

що в них набирає максимального абсолютного значення з альтернуючим знаком тригонометричний поліном

$$T_{n+1}\left(\frac{2x-a-b}{b-a}\right)=\cos\overline{n+1}\arccos\frac{2x-a-b}{b-a} \tag{329}$$

З другого боку, вузлові точки Чебишова є точки

$$x_i=\frac{b+a}{2}+\frac{b-a}{2}\cos\frac{(2n+1-2i)\pi}{2(n+1)},\; i=0,1,\ldots,n, \tag{330}$$

що в них той самий тригонометричний поліном дорівнює нулеві (пор. попередню виноску на с. 117).

Додамо, що за неплавного перебігу кривої $y=f(x)$ на інтервалі (a, b) вживання точок (328) та (330) взагалі не буде обгрунтоване.

Додаючи цю поправку до $p(x)$, одержуємо:

$$p_I(x) = 0{,}0672 + 1{,}9407x - 1{,}0751x^2 \tag{333}$$

Ще раз застосовуючи алгорифм параграфа 20, матимемо:

$$x_0 = 0;\ x_1 = 0{,}07988;\ x_2 = 0{,}6036;\ x_3 = 1$$

$$\Delta(x_0) = -0{,}0672;\ \Delta(x_1) = 0{,}067267;\ \Delta(x_2) = -0{,}069996;\ \Delta(x_3) = 0{,}0672$$

$$\rho'_{II} = 0{,}067\,618$$

$$\Omega_{II}(x) = 0{,}000\,418 - 0{,}010\,414x + 0{,}009\,578x^2$$

$$p_{II}(x) = 0{,}067\,618 + 1{,}930\,286x - 1{,}065\,522x^2 \tag{334}$$

Досліджуючи тепер різницю $\Delta_{II}(x) = \sqrt{x} - p_{II}(x)$, матимемо:

$$x_0 = 0;\ x_1 = 0{,}080\,902;\ x_2 = 0{,}603\,863;\ x_3 = 1$$

$$\Delta_{II}(x_0) = -0{,}067\,618;\ \Delta_{II}(x_1) = 0{,}067\,625;\ \Delta_I(x_2) = -0{,}067\,617;$$

$$_{II}(x_3) = 0{,}067\,618$$

Скорегувши ще поліном (334) за допомогою аддитивної сталої $\frac{0{,}067\,625 - 0{,}067\,618}{2} \approx 3 \cdot 10^{-}$, одержимо:

$$p_{II}(x) = 0{,}067\,621 + 1{,}930\,286x - 1{,}065\,522x^2, \tag{334'}$$

при чому тепер

$$\Delta_{II}(x_0) = -0{,}067\,621;\ \Delta_{II}(x_1) = 0{,}067\,622;\ \Delta_{II}(x_2) = -0{,}067\,620;\ \Delta_{II}(x_3) = 0{,}067\,615$$

$$\text{Тут } L_{II} = 0{,}067\,622 > \rho > \frac{0{,}067\,615 + 0{,}067\,620}{2} = 0{,}067\,6175$$

Отже, остаточно

$$|x| = 0{,}067\,621 + 1{,}930\,286x^2 - 1{,}065\,522x^4 \tag{335}$$

на сегменті (—1,1) з наближенням 0,067 622, яке перевищує найкраще наближення ρ менше, ніж на $5 \cdot 10^{-6}$.

Нижче [див. зведення (336)] подається для функції $|x|$ на сегменті (—1,1) поліноми найкращого наближення до 10-го степеня включно. Поліноми $\pi_7(x)$, $\pi_9(x)$, $\pi_{11}(x)$ обчисляли за алгорифмом § 20 в його точній формі — аналогічно до того, як це тільки но було з'ясовано щодо $\pi_5(x)$. За вихідне наближене представлення взято в кожному з цих випадків поліном найкращого наближення функції $\sqrt{x}$ на відповідному ансамблі (328) $\nu + 2$ точок відхилу Чебишова. Після цього застосовували послідовно дві поправки, складаючи кожного разу ансамбль $\nu + 2$ точок $x_0, x_1, \ldots, x_{\nu+1}$ з точок maxima та minima $\Delta(x)$ із включенням кінців сегмента (0,1). Щождо $\pi_3(x)$, то ясно, що цей поліном безпосередньо визначається в скінченному виді, застосовуючи до функції $\sqrt{x}$ на сег-

менті (0,1) спосіб розділу I. Таким чином і одержано такі вирази поліномів найкращого наближення[1]), які можуть мати застосування при розв'язуванні задач щодо поліноміальної апроксимації неперервних функцій[2]) зокрема ж коли бажано одержати для неперервної функції вихідне наближене представлення, яке підлягає дальшому уточненню за методом послідовних наближень параграфів 19 та 20:

$$\left.\begin{aligned}
&\pi_3(x)=\frac{1}{8}+x^2;\ L=\frac{1}{8}=\rho\\
&\pi_5(x)=0{,}067\,621+1{,}930\,286\,x^2-1{,}065\,522x^4;\ L=0{,}067\,622\\
&\pi_7(x)=0{,}045\,928+2{,}867\,830x^2-4{,}177\,453x^4+\\
&\qquad\qquad+2{,}309\,623x^6;\ L=0{,}045\,936\\
&\pi_9(x)=0{,}034\,680+3{,}809\,830x^2-10{,}363\,280x^4+\\
&\qquad+13{,}719\,487x^6-6{,}235\,392x^8;\ L=0{,}034\,699\\
&\pi_{11}(x)=0{,}027\,837+4{,}753\,770x^2-20{,}646\,839x^4+\\
&\qquad+47{,}776\,685\,x^6-49{,}593\,272\,x^8+18{,}709\,656\,x^{10};\\
&\qquad\qquad L=0{,}027\,853
\end{aligned}\right\}\tag{336}$$

Різниця $L-\rho$ для $\pi_7(x)$, $\pi_9(x)$, $\pi_{11}(x)$ менша, ніж 10^{-5}

§ 22. Поширення двох алгорифмів, зазначених у параграфах 19—20 на загальний випадок, трактований в розділі VI. Приклад

1°. Узагальнення першого алгорифма. В загальному випадку, коли на обмеженій точковій множині E числової прямої дано будьяку обмежену функцію $f(x)$ із границями неозначеності $\varphi(x)$ та $\Phi(x)$ і максимальним коливанням $2k$ на відповідній замкненій множині E° щодо E, — величина k, згідно з теоремою II § 16, а priori являє собою нижню межу для найкращого наближення ρ. Якщо одержано наближене представлення (не найкраще) функції $f(x)$ за допомогою якогось полінома $p_n(x)$ степеня $\leqslant n$, ми й тут користуватимемося числом A, означеним у § 17 (теорема I). Ми матимемо, очевидно,

$$0\leqslant k\leqslant A\leqslant\rho<L\leqslant M\tag{337}$$

Позначаючи

$$\Phi(x)-p_n(x)=\overline{\Delta}(x),\ \varphi(x)-p_n(x)=\underline{\Delta}(x),\tag{338}$$

ми, замість нерівностей (294) § 19, матимемо такі:

$$\overline{\Delta}(x)\geqslant R,\ \underline{\Delta}(x)\leqslant -R,\tag{339}$$

де тепер $k<R<L$. Для ширини відповідних інтервалів (295)[3]) можна

[1]) Обчислення переводили студентки Київського державного університету З. А. Линник, К. С. Кунявська та Т. Ю. Мазановська.

[2]) Пор. відповідні зауваження наприкінці § 12 цієї роботи.

[3]) Замість $G\subset C$, $H\subset C$, ми тут маємо $G\subset E^\circ$, $H\subset E^\circ$

буде і тут встановити існування певної додатної нижньої границі $2l$, яка має властивості, аналогічні до зазначених в § 19, і може наближатися безмежно до нуля не інакше, як водночас із різницею $R-k$. Справді, різниці $\Delta(x)=f(x)-p_n(x)$, при умові $\max\limits_{x\subset E^\circ}|p_n(x)|\leqslant 2M$, матимуть у своїй сукупності деяку властивість, аналогічну до характеристичної властивості функцій рівно неперервних, а саме: кожному довільно малому числу $\varepsilon<0$ можна протипоставити таке число $\bar{\eta}<0$, щоб нерівність $|x'-x''|<\bar{\eta}$ при $x', x''\subset E$ мала своїм наслідком $|\Delta(x')-\Delta(x'')|<2k+\varepsilon$ незалежно від окремого добору полінома $p_n(x)$, підпорядкованого лише згаданій умові щодо максимуму модуля. Ми матимемо при цьому

$$\bar{\eta}=\min\left\{\frac{\varepsilon}{2M_n'},\ \psi\left(\frac{\varepsilon}{2}\right)\right\}, \tag{340}$$

якщо позначимо, з одного боку, через M_n' верхню границю[1]) модуля похідної $p_n'(x)$ на сегменті (a, b) при умові $\max\limits_{x\subset E^\circ}|p_n(x)|\leqslant 2M$, — що встановлюється в даному разі за допомогою інтерполяційної формули Лягранжа[2]), а з другого боку, — позначимо через $\eta=\psi(\varepsilon)$ функціональну залежність величини η від ε в лемі II § 17, умовившись надавати η найбільше можливе значення, що задовольняє вимогу леми II — проте у всякому разі не більше за $b-a$. Відповідно до цього ми можемо написати такий явний вираз для величини $2l$:

$$2l=\min\left\{\frac{R-k}{M_n'},\ \psi(R-k)\right\} \tag{341}$$

Задавши два числа R_1 та R_2 за умовами (298), ми утворюємо точно так само, як в § 19, поправку $\omega_n(x)$ до полінома $p_n(x)$, яка приводить до зниження величини L[3]). При повторному застосуванні алгорифма величина L, невпинно меншаючи, наближається до якоїсь границі $\Lambda\geqslant k$. У випадку $\Lambda>k$ означена в § 19 величина δ[4]) напевне залишається вище за якусь додатну фіксовану границю, якщо належним чином добирати кожного разу абсциси $\xi_1, \xi_2, \ldots, \xi_m$. А коли так, то при таких самих умовах щодо δ, r, ϑ_1, ϑ_2, як у § 19, виявляється, що різниця $L-\rho$ наближається до нуля принаймні як член геометричної спадної прогресії. Щождо другого можливого випадку, коли $\Lambda=k$, то водночас матимемо, очевидно, й $\rho=k$, $\lim L=\rho$; отже і в цьому разі збіжність процесу

[1]) Щоб далі не залишити місця для будьякої неясності, ми припустимо, що число M_n' є більше за точну верхню границю модуля $p_n'(x)$ при зазначених умовах.

[2]) Беручи на увагу обмеженість коефіцієнтів полінома $p_n(x)$ при розглядуваній умові (пор. § 15).

[3]) Подробиці міркування для розглядуваного випадку можна знайти в нашій роботі, яка незабаром вийде друком, — „Sur la détermination des polynomes d'approximation de degré donné" (Communications de la Soc. Math. de Kharkow, t. X).

[4]) Мінімальна віддаль поміж точками $\xi_1, \xi_2, \ldots, \xi_m$ та точковою множиною G_2+H_2.

забезпечена[1]). Зауважмо зокрема, що у випадку однозначної функції $f(x)$, заданої на дискретній точковій множині E, розміщеній на сегменті (a, b), так само як у випадку, трактованому в § 19, різниця $L-\rho$ напевне наближається до нуля принаймні як член геометричної спадної прогресії, якщо залишити осторонь тривіальний випадок $\rho=0$. Інше наше зауваження стосуватиметься того випадку, коли, при $\rho=k$, кількість розв'язків нескінченна. Поліноми $p_{n1}(x)$, $p_{n2}(x)$, $p_{n3}(x), \ldots$, які послідовно одержують при повторному застосуванні даного алгорифма, є рівно неперервні (egualmente continui) в своїй сукупності на сегменті (a, b), оскільки вони в сукупності є обмежені на точковій множині E°. Як добре відомо[2]), кожна частинна нескінченна множина поліномів, вилучена з розглядуваної послідовності, матиме принаймні одну граничну функцію. Кожна така гранична функція являтиме собою також поліном степеня $\leqslant n$[3]), і це буде, як легко зрозуміти, один із поліномів найкращого наближення $\pi_n(x)$. Якщо гранична функція єдина, сама послідовність поліномів $p_{n1}(x)$, $p_{n2}(x)$, $p_{n3}(x), \ldots$ збігається. Проте, з практичного погляду для нас найбільш істотне є те, що $\lim L=\rho$ — незалежно від того, чи збігається сама розглядувана послідовність поліномів в цілому.

2°. **Застосування другого алгорифма.** Припустімо, що попереду вже знайдено тим чи тим способом вихідне наближене представлення функції $f(x)$ на точковій множині E за допомогою якогось полінома $p_n(x)$, таке, що $A>k$[4]). Тоді всі міркування параграфа 20 залишаються в силі без істотних змін. Лише замість значень $\Delta(x_0)$, $\Delta(x_1), \ldots, \Delta(x_{n+1})$ ми тут матимемо один із двох рядів значень

$$\overline{\Delta}(x_0),\ \underline{\Delta}(x_1),\ \overline{\Delta}(x_2),\ \underline{\Delta}(x_3), \ldots \tag{342}$$

абож

$$\underline{\Delta}(x_0),\ \overline{\Delta}(x_1),\ \underline{\Delta}(x_2),\ \overline{\Delta}(x_3), \ldots, \tag{343}$$

а саме той із цих двох рядів значень, що за його допомогою встановлено величину A. Залишаться в силі й загальні уваги параграфа 20 щодо можливого поєднання істотних моментів обох алгорифмів.

Але розглядуваний алгорифм, із деякими можливими модифікаціями, застосований також у випадках, де $A=k$ $(\rho \geqslant k)$, якщо одержано попереду будьякий поліном $p_n(x)$, що допускає застосування теореми I § 16 при $0<\min\{\mu', \mu''\} \leqslant k$[5]).

[1]) В обох випадках, очевидно, $A=\rho$.

[2]) C. Arzela. „Funzioni di linee" (Rendiconti della R. Accademia dei Lincei, 1889).

[3]) Пор. L. Tonelli, op. cit., pp. 63—65.

[4]) Це можливе, очевидно, при $\rho>k$.

[5]) Користуючися лемою 1 § 17 та деякими іншими міркуваннями параграфів 15—21, можна було б, справді, показати можливість наблизити величину L скільки завгодно до ρ за допомогою скінченного числа операцій, аналогічних до тієї, що відповідає алгорифмові параграфа 20. Проте докладніше ми тут на цьому питанні не спинимось.

3°. **Приклад.** Наближене представлення, за допомогою полінома даного степеня, функції $f(x)=|x|+k\,\mathrm{sg}\,x$ $(k>0)$[1]) на сегменті $<-1,+1>$ без точки $x=0$.

Тут точкова множина E складається з точок двох півзамкнених інтервалів $<-1,0)$ та $(0,+1>$. Замкнена точкова множина E° є повний сегмент $<-1,+1>$. Функції $\Phi(x)$ та $\varphi(x)$ означаються умовами

$$\left.\begin{aligned}\Phi(x)&=\begin{cases}-x-k & \text{при } -1\leqslant x<0\\ x+k & \text{„}\quad 0\leqslant x\leqslant 1\end{cases}\\ \varphi(x)&=\begin{cases}-x-k & \text{„}\quad -1\leqslant x\leqslant 0\\ x+k & \text{„}\quad 0<x\leqslant 1\end{cases}\end{aligned}\right\}\tag{344}$$

Почнімо з докладного розгляду одного з простіших випадків, коли шуканий поліном $\alpha x^2+\beta x+\gamma$ є степеня $\leqslant 2$, маючи при цьому на увазі виявити деякі принципові особливості задачі.

За вихідне наближене представлення візьмімо

$$p(x)=x^2+\frac{1}{8},\tag{345}$$

тобто поліном найкращого наближення функції $|x|$ на сегменті $<-1,+1>$. Досліджуючи на extrema обидві різниці $\overline{\Delta}(x)=\Phi(x)-p(x)$ і $\underline{\Delta}(x)=\varphi(x)-p(x)$, знаходимо (включаючи кінці інтервалу та точки розриву):

$$\Delta(-1)=-k-\frac{1}{8},\quad \Delta\left(-\frac{1}{2}\right)=-k+\frac{1}{8},\ \underline{\Delta}(0)=-k-\frac{1}{8}$$

$$\overline{\Delta}(0)=k-\frac{1}{8},\quad \Delta\left(\frac{1}{2}\right)=k+\frac{1}{8},\quad \Delta(1)=k-\frac{1}{8},$$

де ми пишемо просто $\Delta(x)$ у тих випадках, коли $\overline{\Delta}(x)=\underline{\Delta}(x)$. Бачимо, що найбільший абсолютний відхил $L=k+\frac{1}{8}$ має місце в точках $x=-1$, $x=0$ із знаком $-$ та в точці $x=\frac{1}{2}$ із знаком $+$. Беручи за основу перший із з'ясованих вище двох основних алгоритмів, число m інтервалів $I_1, I_2,\ldots$ тут дорівнює 1. Виходячи з міркувань симетрії, шукаємо поправку $\omega_I(x)$ у виді $\lambda\left(x-\frac{1}{4}\right)$. Маючи на увазі зрівняти за абсолютною величиною $\overline{\Delta}\left(-\frac{1}{2}\right)$, $\underline{\Delta}(0)$, $\overline{\Delta}\left(\frac{1}{2}\right)$ та $\underline{\Delta}(1)$, щоб добути ансамбль $n+2=4$ точок, який допускав би застосування теореми I § 16, беремо $\lambda=2k$, отже

$$\omega_I(x)=2k\left(x-\frac{1}{4}\right)$$

$$p_I(x)=x^2+2kx+\frac{1}{8}-\frac{1}{2}k\tag{346}$$

[1]) $\mathrm{sg}\,x=\frac{x}{|x|}=\begin{cases}-1 & \text{при } x<0\\ 1 & \text{„}\quad x>0\end{cases}$

Дослідження різниць $\overline{\Delta}_I = \Phi - p_I$, $\underline{\Delta}_I = \varphi - p_I$ приводить до різних результатів залежно від того, чи $0 < k < \frac{1}{2}$ або $k \geqslant \frac{1}{2}$

При $0 < k < \frac{1}{2}$ матимемо:

$$\Delta_I(-1) = \frac{3}{2}k - \frac{1}{8}; \quad \Delta_I\left(-k - \frac{1}{2}\right) = k^2 + \frac{1}{2}k + \frac{1}{8}; \quad \underline{\Delta}_I(0) = -\frac{1}{2}k - \frac{1}{8}$$

$$\overline{\Delta}_I(0) = \frac{3}{2}k - \frac{1}{8}; \quad \Delta_I\left(-k + \frac{1}{2}\right) = k^2 + \frac{1}{2}k + \frac{1}{8}; \quad \Delta_I(1) = -\frac{1}{2}k - \frac{1}{8}$$

Звідси легко бачити, що ми одержимо безпосередньо точний розв'язок задачі, додаючи до $p_I(x)$ сталу, рівну $\frac{1}{2}k^2$

Ми матимемо тоді:

$$\pi(x) = x^2 + 2kx + \frac{1}{2}\left(k - \frac{1}{2}\right)^2 \tag{347}$$

при $\rho = \frac{1}{2}\left(k + \frac{1}{2}\right)^2 > k$.

В другому ж випадку, коли $k \geqslant \frac{1}{2}$, маємо:

$$\Delta_I(-1) = \frac{3}{2}k - \frac{1}{8}, \quad \underline{\Delta}_I(0) = -\frac{1}{2}k - \frac{1}{8}$$

$$\overline{\Delta}_I(0) = \frac{3}{2}k - \frac{1}{8}, \quad \Delta_I(1) = -\frac{1}{2}k - \frac{1}{8}$$

В цьому разі ми додамо до $p_I(x)$ сталу $\frac{1}{2}k - \frac{1}{8}$, і це знову дасть точний розв'язок

$$\pi(x) = x^2 + 2kx \tag{348}$$

при $\rho = k$. Але цей розв'язок буде єдиний лише при $k = \frac{1}{2}$. Коли ж $k > \frac{1}{2}$, матимемо безліч точних розв'язків задачі, що обіймаються (як показує нескладне дослідження) формулою

$$\left.\begin{aligned} &\pi(x) = \alpha x^2 + \beta x \\ \text{при} \quad &1 \leqslant \beta \leqslant 2k \\ &\max\{1 - \beta, 1 + \beta - 2k\} \leqslant \alpha \leqslant \min\{1 + \beta, 2k + 1 - \beta\} \end{aligned}\right\} \tag{349}$$

Тут буде доречно зазначити, що шляхом граничного переходу звідси можна добути й відповідну формулу найкращого наближення для функції

$$\operatorname{sg} x = \lim_{k \to \infty}\left[\frac{1}{k}(|x| + k \operatorname{sg} x)\right] \tag{350}$$

Ця формула буде:

$$\left.\begin{array}{c} \operatorname{sg} x \approx ax^2 + bx;\ \rho = 1 \\ 0 \leqslant b \leqslant 2 \\ \max\{-b, b-2\} \leqslant a \leqslant \min\{b, 2-b\} \end{array}\right\} \tag{351}$$

В загальному випадку, коли треба знайти найкраще наближене представлення функції $|x| + k \operatorname{sg} x$ в тій самій області за допомогою полінома степеня $\leqslant n$, існує ще, для кожного значення n, граничне додатне число $k_0 = k_0(n)$, яке посідає таку властивість: задача має єдиний розв'язок при $k \leqslant k_0$ і безліч розв'язків при $k > k_0$.

Задача визначити $k_0(n)$ для даного значення n, що на ній ми хочемо тут трохи спинитись, сходить, як це легко бачити, до такого питання: між усіма поліномами $p_n(x)$ степеня $\leqslant n$, що їхній відхил[1]) від функції $|x|$ є додатний при $-1 \leqslant x < 0$ та від'ємний при $0 < x \leqslant 1$, визначити той, для якого максимальний абсолютний відхил на сегменті $\langle -1, +1 \rangle$

$$\max_{-1 \leqslant x \leqslant 1} \big| |x| - p_n(x) \big| \tag{352}$$

є щонайменший[2]). Оце найменше значення величини (352) і буде якраз подвоєне значення $k_0(n)$.

Розглядуване питання є особливого типу, відмінного від задачі найкращого наближення в її звичайному розумінні. Замість осцилювати між значеннями $\pm\rho$ тут різниця

$$\delta(x) = |x| - p_n(x) \tag{353}$$

осцилюватиме між 0 та $2k_0$ на сегменті $\langle -1, 0 \rangle$ та між 0 і $-2k_0$ на сегменті $\langle 0, 1 \rangle$. Проте, метод, з'ясований в §§ 19—22, може бути застосований mutatis mutandis і до цієї задачі. Так, у випадку $n = 3$, беручи за вихідне наближене представлення

$$p(x) = x^2 + x, \tag{354}$$

тобто роз'язок аналогічної задачі при $n = 2$, ми матимемо далі, вживаючи позначень аналогічних до тих, з яких ми вище користувалися:

[1]) Тобто різниця $|x| - p_n(x)$.

[2]) Слід мати на увазі, що при будьякому значенні k відповідна область $\mathfrak{B}$ (див. кінець § 16) складається тут з двох незамкнених паралелограмів із спільним вершком у початку координат

$$\begin{cases} -x - 2k \leqslant y \leqslant -x \\ -1 \leqslant x < 0 \end{cases} \quad \text{та} \quad \begin{cases} x \leqslant y \leqslant x + 2k \\ 0 < x \leqslant 1 \end{cases}$$

до яких додається сама точка (0,0). Задача полягає у визначенні найменшого значення k, при якому в цій області $\mathfrak{B}$ переходить при $-1 \leqslant x \leqslant 1$ принаймні одна параболічна крива степеня $\leqslant n$.

$$\delta(x) = \begin{cases} -2x - x^2 & \text{при} -1 \leqslant x \leqslant 0 \\ -x^2 & \text{„} \quad 0 \leqslant x \leqslant 1 \end{cases}$$

$$\delta(-1) = 1,\ \delta(0) = \delta'(+0) = 0\ ^{1)},\ \delta(1) = -1$$

$$\omega_I(x) = -\frac{4}{5}x^2\left(x - \frac{1}{4}\right)$$

$$p_I(x) = -\frac{4}{5}x^3 + \frac{6}{5}x^2 + x \qquad 355$$

$$\delta_I(x) = \begin{cases} -2x - \frac{6}{5}x^2 + \frac{4}{5}x^3 & \text{при} -1 \leqslant x \leqslant 0 \\ -\frac{6}{5}x^2 + \frac{4}{5}x^3 & \text{„} \quad 0 \leqslant x \leqslant 1 \end{cases}$$

$$\delta_I(-1) = 0;\ \delta_I(-0{,}538) = 0{,}604;\ \delta_I(0) = \delta_I'(+0) = 0;\ \delta_I(1) = -0{,}4$$

$$\omega_{II}(x) = 0{,}096\,x^2(x+1)$$

$$p_{II}(x) = -0{,}704\,x^3 + 1{,}296\,x^2 + x \qquad (356)$$

$$\delta_{II}(-1) = 0;\ \delta_{II}(-0{,}537) = 0{,}591;\ \delta_{II}(0) = \delta_{II}'(+0) = 0;\ \delta_{II}(1) = -0{,}592$$

Для того, щоб дальша поправка $\omega_{III}(x)$ могла понизити максимальний відхил до величини меншої за 0,591, вона повинна була б задовольняти одночасно такі умови:

$$\omega_{III}(-1) \leqslant 0,\ \omega_{III}(-0{,}537) > 0,\ \omega_{III}(0) = 0,\ \omega_{III}'(0) \geqslant 0,\ \omega_{III}(1) < 0$$

Але легко бачити, що ці вимоги несумісні, коли степінь полінома $\omega_{III}(x)$ не повинен перевищувати 3. Отже, величина $2k_0$ міститься між 0,591 та 0,592, тому

$$k_0(3) = 0{,}296 \qquad (357)$$

з похибкою меншою за 0,001.

У випадку $n = 4$, взявши (356) за вихідне наближене представлення, знаходимо аналогічним шляхом

$$p(x) = -1{,}12\,x^4 - x^3 + 2{,}12\,x^2 + x \qquad (358)$$

при

$$\delta(-1) = \delta(0) = \delta'(+0) = \delta(1) = 0;\ \delta(-0{,}425) = 0{,}427;\ \delta(0{,}694) = -0{,}427$$

Отже

$$k_0(4) = 0{,}2135 \qquad (359)$$

з точністю до кількох десятитисячних, і аналогічно можна було б продовжувати послідовно при $n = 5$, 6,

[1]) $\delta'(+0) = \left[\frac{d}{dx}\delta(x)\right]_{x=+0}$

$k_0(n)$ є, очевидно, монотонно-спадна функція від n. Не важко також пересвідчитись a priori, що

$$\lim_{n\to\infty} k_0(n) = 0 \tag{360}$$

Справді, зважаючи на те, що $p_n(0) = 0$ та позначаючи поліном $\frac{p_n(x)}{x}$ через $q_{n-1}(x)$, задача визначення $k_0(n)$ сходить, очевидно, до питання: визначити найменше додатне значення k, при якому існує хоч одна параболічна крива $y = q_{n-1}(x)$, що при $-1 \leqslant x < 0$ та при $0 < x \leqslant 1$ переходить відповідно в областях

$$\begin{cases} -1 \leqslant y \leqslant -1 - \frac{2k}{x} \\ -1 \leqslant x < 0 \end{cases} \quad \text{та} \quad \begin{cases} 1 \leqslant y \leqslant 1 + \frac{2k}{x} \\ 0 < x \leqslant 1 \end{cases}$$

Але при такому поставленні питання добре відома теорема Вайєрштраса дозволяє легко довести граничну рівність (360)

РОЗДІЛ VIII

СИСТЕМИ ЛІНІЙНИХ РІВНЯНЬ ТА ЗАДАЧА АПРОКСИМАЦІЇ ФУНКЦІЙ ВІД БУДЬЯКОЇ КІЛЬКОСТІ НЕЗАЛЕЖНИХ ЗМІННИХ

§ 23. Задача „мінімального наближення“ для системи лінійних суперечних рівнянь[1]). Деякі основні теореми, що стосуються її

1°. Нехай маємо систему m лінійних суперечних рівнянь із N невідомими $x, y, z, \ldots, u$:

$$\left.\begin{array}{c} f_i(x, y, \ldots, u) \equiv a_i x + b_i y + \cdots + h_i u = l_i \\ (i = 1, 2, \ldots, m) \end{array}\right\} \tag{361}$$

Позначаючи через Δ_i відповідні відхили

$$l_i - (a_i x + b_i y + \cdots + h_i u) = \Delta_i; \; i = 1, 2, \ldots, m, \tag{362}$$

задача мінімального наближення[2]) полягає в доборі такої системи значень $x, y, \ldots, u$, щоб величина

$$\max \{ |\Delta_1|, |\Delta_2|, \ldots, |\Delta_m| \} = L, \tag{363}$$

яка береться за міру наближення, була щонайменша.

2°. В дальшому викладі нам доведеться вдаватися неодноразово до поняття лінійної залежності.

Означення I[3]). *Лінійні форми $f_1, f_2, \ldots, f_r$ ми називатимемо лінійно залежними (в широкому розумінні), якщо існує система констант $C_1, C_2, \ldots, C_r$, з яких принаймні одна відмінна від нуля, і таких що*

$$C_1 f_1 + C_2 f_2 + \cdots + C_r f_r \equiv 0 \tag{364}$$

Необхідна і достатня умова для існування лінійної залежності поміж $f_1, f_2, \ldots, f_r$ полягає в тому, щоб ранг матриці коефіцієнтів системи цих r лінійних форм був менший, ніж r.

[1]) Це питання, що обіймає як окремий випадок (при $a_i = 1$, $i = 1, 2, \ldots, m$) задачу, трактовану P. Kirchberger-ом у його дисертації (op. cit.), привертало до себе увагу ще Лапласа та Фур'є (див. історичні зауваження в розд. III), а поточного століття було за об'єкт дослідів Goedseels-a та de la Vallée-Poussin-a (Annales de la Société Scientifique de Bruxelles, 1910—1911), яким належить сама назва задачі. Проте важливі досліди названих учених залишалися досі невикінчені в ряді істотних пунктів.

[2]) Термін запропонований Goedseels-ом у другому виданні його Théorie des erreurs d'observation (Louvain, 1909).

[3]) Пор. M. Bôcher. Einführung in die höhere Algebra, Leipzig—Berlin, 1932.

Означення II. *Ті самі r лінійних форм ми називатимемо лінійно залежними у вузькому розумінні, якщо в тотожності* (374) *усі r констант $C_1, C_2, \ldots, C_r$ є відмінні від нуля, і якщо не має місця ніяка аналогічна тотожність із меншим числом членів.*

Необхідна і достатня умова для існування лінійної залежності у вузькому розумінні між $f_1, f_2, \ldots, f_r$ полягає в тому, щоб ранг матриці коефіцієнтів дорівнював $r-1$ і щоб з неї можна було виділити таку („вкорочену") матрицю з $r-1$ стовпцями, що в ній усі r детермінантів $r-1$-го порядку є відмінні від нуля [1]).

Легко зрозуміти, що в даному разі константи $C_1, C_2, \ldots, C_r$ є однозначно-означені до спільного множника.

3°. Припустімо, що ранг матриці коефіцієнтів форм $f_1, f_2, \ldots, f_m$ в системі (361) дорівнює n, при чому, певна річ,

$$n \leqslant N, \quad n < m \tag{365}$$

В такому разі серед форм $f_1, f_2, \ldots, f_m$ є n лінійно незалежних між собою. Для простоти позначень — нехай це будуть $f_1, f_2, \ldots, f_n$. Решта ж $f_{n+1}, f_{n+2}, \ldots, f_m$ виражатимуться лінійно через n перших.

Введімо нові n змінних:

$$f_i(x, y, \ldots, u) = x_i; \quad i = 1, 2, \ldots, n \qquad 366$$

Тоді система (361) перетвориться на таку:

$$\begin{array}{lll}
F_1 & \equiv & x_1 = l_1 \\
F_2 & \equiv & x_2 = l_2 \\
\cdots & \cdots & \cdots \\
F_n & \equiv & x_n = l_n \\
F_{n+1} & \equiv & a_{n+1,1} x_1 + a_{n+1,2} x_2 + \cdots + a_{n+1,n} x_n = l_{n+1} \\
\cdots & \cdots & \cdots \\
F_m & \equiv & a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n = l_m
\end{array} \tag{367}$$

де $F_i = F_i(x_1, x_2, \ldots, x_n)$, при $i = 1, 2, \ldots, m$, є результат підставлення $x_1, x_2, \ldots, x_n$ замість $f_1, f_2, \ldots, f_n$ в лінійні форми, що виражають f_i через $f_1, f_2, \ldots, f_n$.

Кожній системі значень $x_1, x_2, \ldots, x_n$ відповідає така сумісна система точних рівностей, що в ній $\Delta_1, \Delta_2, \ldots, \Delta_m$ позначають не що інше, як величини відхилів $l_i - F_i$ для відповідних рівнянь несумісної системи (367):

$$\left.\begin{array}{lll}
F_1 + \Delta_1 & \equiv & x_1 + \Delta_1 = l_1 \\
\cdots & \cdots & \cdots \\
F_n + \Delta_n & \equiv & x_n + \Delta_n = l_n \\
F_{n+1} + \Delta_{n+1} & \equiv & a_{n+1,1} x_1 + a_{n+1,2} x_2 + \cdots + a_{n+1,n} x_n + \Delta_{n+1} = l_{n+1} \\
\cdots & \cdots & \cdots \\
F_m + \Delta_m & \equiv & a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n + \Delta_m = l_m
\end{array}\right\} \tag{368}$$

[1]) Не важко бачити, що при цих умовах два будьякі детермінанти $r-1$-го порядку, утворені з елементів одних і тих самих стовпців, можуть лише одночасно дорівнювати нулеві або бути відмінні від нуля.

Задачу мінімального наближення для системи суперечних рівнянь (367) ми тепер можемо сформулювати так: дібрати змінні $x_1, x_2, \ldots, x_n$ в системі (368) так, щоб величина

$$\max \{ |\Delta_1|, |\Delta_2|, \ldots, |\Delta_m| \} = L \tag{363}$$

була щонайменша.

Між попередньою задачею, що стосувалася системи (361) та новою задачею, що стосується системи (367), існує, очевидно, така залежність:

а) Кожному розв'язкові задачі мінімального наближення для системи (361) відповідає на підставі (366) один і тільки один розв'язок задачі щодо системи (367).

б) Кожному розв'язкові задачі мінімального наближення для системи (367) відповідає один розв'язок задачі щодо системи (361) лише при $n = N$ і безліч розв'язків (∞^{N-n} розв'язків) при $n < N$.

Це зауваження дозволяє нам надалі обмежитись виключно розглядом задачі мінімального наближення для системи суперечних рівнянь (367).

4°. Встановлюючи існування розв'язків для задачі щодо системи (367), можна, як показують перші n рівнянь, мати на увазі лише ту обмежену й замкнену область n-мірного простору $(x_1, x_2, \ldots, x_n)$, що означається умовами

$$\left. \begin{array}{c} |x_i| \leqslant |l_i| + \max \{ |l_1|, |l_2|, \ldots, |l_m| \} \\ (i = 1, 2, \ldots, n) \end{array} \right\} \tag{369}$$

бо кожна система значень $x_1, x_2, \ldots, x_n$, яка не задовольняє цих нерівностей, дає більше наближення, ніж система значень $x_1 = x_2 = \ldots = x_n = 0$.

Це дає змогу довести існування розв'язків $(x_{10}, x_{20}, \ldots, x_{n0})$, число яких може бути або один або безліч, а також довести теорему аналогічну до тієї, яка була встановлена наприкінці § 15, таким самим шляхом, який був застосований в § 15. Найменшу величину максимального відхилу L позначатимемо і тут через ρ і називатимемо „мінімальне наближення".

5°. **Задача.** Із системи (367) виділено якусь підсистему r рівнянь $(r \leqslant n+1)$ [1])

$$F_\alpha = l_\alpha, \; F_\beta = l_\beta, \ldots, \; F_\varkappa = l_\varkappa \tag{370}$$

таку, що форми $F_\alpha, F_\beta, \ldots, F_\varkappa$ лінійно залежні між собою у вузькому розумінні (n° 2, означення II). Розв'язати задачу мінімального наближення щодо цієї системи (370) [2]).

[1]) При $m = n+1$ ця підсистема може обійняти і всю систему (367).

[2]) В аналогічному порядку міркувань легко можна одержати розв'язання задачі мінімального наближення за критерієм відносного відхилу (відносної похибки) — при наявності тих умов, за яких застосування цього критерія має рацію, а також розв'язання тієї особливої відміни задачі мінімального наближення, яка відповідає зустрінутому в § 22 n° 3 питанню щодо визначення $k_0(n)$ і характеризується в даному разі вимогою, щоб відхили $\Delta_\alpha, \ldots, \Delta_\varkappa$ мали наперед приписані знаки. Ми тут обмежимось цим зауваженням, щоб у дальшому викладі не відтягати уваги від основної послідовності міркувань.

Розв'язання. Напишімо відповідну до (370) систему точних рівностей:

$$F_\alpha + \Delta_\alpha = l_\alpha,\ F_\beta + \Delta_\beta = l_\beta, \ldots,\ F_\varkappa + \Delta_\varkappa = l_\varkappa \qquad 371$$

Нехай будуть $C_\alpha, C_\beta, \ldots, C_\varkappa$ такі константи, що

$$C_\alpha F_\alpha + C_\beta F_\beta + \cdots + C_\varkappa F_\varkappa \equiv 0 \qquad 372$$

Тоді, очевидно, матимемо також

$$C_\alpha \Delta_\alpha + C_\beta \Delta_\beta + \cdots + C_\varkappa \Delta_\varkappa = C_\alpha l_\alpha + C_\beta l_\beta + \cdots + C_\varkappa l_\varkappa \qquad 373$$

для кожної системи значень $x_1, x_2, \ldots, x_n$. Запишімо це інакше:

$$|C_\alpha| \Delta_\alpha \cdot \operatorname{sg} C_\alpha + \cdots + |C_\varkappa| \Delta_\varkappa \operatorname{sg} C_\varkappa = C_\alpha l_\alpha + \cdots + C_\varkappa l_\varkappa \qquad 374$$

Розгляньмо систему значень (або одну із систем значень)

$$x_1 = \xi_1, x_2 = \xi_2, \ldots,\ x_n = \xi_n,$$

при якій

$$\left.\begin{aligned} &\Delta_\alpha \operatorname{sg} C_\alpha = \Delta_\beta \operatorname{sg} C_\beta = \cdots = \Delta_\varkappa \operatorname{sg} C_\varkappa = L' \\ &L' = \frac{C_\alpha l_\alpha + C_\beta l_\beta + \cdots + C_\varkappa l_\varkappa}{|C_\alpha| + |C_\beta| + \cdots + |C_\varkappa|} \end{aligned}\right\} \qquad (375)$$

Така система значень напевне існує і визначається з системи сумісних лінійних рівнянь:

$$l_\alpha - F_\alpha = L' \operatorname{sg} C_\alpha,\ l_\beta - F_\beta = L' \operatorname{sg} C_\beta, \ldots,\ l_\varkappa - F_\varkappa = L' \operatorname{sg} C_\varkappa, \qquad (376)$$

що в ній перші $r-1$ рівнянь незалежні, а останнє — наслідок з них. Система значень $\xi_1, \xi_2, \ldots, \xi_n$ буде єдина лише при $r = n+1$, а в інших випадках якісь $n+1-r$ невідомих із числа $x_1, x_2, \ldots, x_n$ залишаються довільні.

Ми твердимо, що для системи значень $\xi_1, \xi_2, \ldots, \xi_n$ наближення, рівне $|L'|$, є менше, ніж для всякої іншої системи значень $x_1, x_2, \ldots, x_n$, що не задовольняє систему рівнянь (376).

Справді, міркуючи точно так, як у розділі I § 1 з приводу лінійної форми (16), бачимо, що при кожній системі значень $x_1, x_2, \ldots, x_n$, яка не задовольняє (376), в рівності (374) принаймні одна з величин $\Delta_\alpha \operatorname{sg} C_\alpha$, $\Delta_\beta \operatorname{sg} C_\beta, \ldots, \Delta_\varkappa \operatorname{sg} C_\varkappa$ буде більша за L' і принаймні одна менша за L' — отже, незалежно від знака L', принаймні одна з цих величин буде абсолютним значенням більша за $|L'|$, а це й доводить наше твердження.

Таким чином, розглянута задача має один або ∞^{n+1-r} розв'язків, які визначаються із системи лінійних рівнянь (376). Для кожного із цих розв'язків усі r відхилів $\Delta_\alpha, \Delta_\beta, \ldots, \Delta_\varkappa$ абсолютним значенням дорівнюють самій величині мінімального наближення ρ', де:

$$\rho' = |L'| = \left| \frac{C_\alpha l_\alpha + C_\beta l_\beta + \cdots + C_\varkappa l_\varkappa}{|C_\alpha| + |C_\beta| + \cdots + |C_\varkappa|} \right| \qquad (377)$$

а знаки відхилів збігаються із знаками $C_\alpha L', C_\beta L', \ldots, C_\varkappa L'$

Увага. З формального погляду важливо зауважити, що аналогічне трактування допускає і задача мінімального наближення для одного рівняння, яке має на лівій стороні тотожний нуль.

6 Повертаючись до загального випадку системи (367), взятої в цілому, нехай буде $(x_{10}, x_{20}, \ldots, x_{n0})$ розв'язок або один із розв'язків задачі мінімального наближення для розглядуваної системи рівнянь, при чому найбільший абсолютний відхил (363), рівний ρ, має напевне місце принаймні для одного (взагалі ж для кількох) із рівнянь (367).

Теорема. *Мінімальне наближення ρ для системи рівнянь* (367) *в цілому дорівнює мінімальному наближенню ρ_1 для підсистеми тих рівнянь, для яких відхил при $x_1 = x_{10}, \ldots, x_n = x_{n0}$ дорівнює $\pm\rho$.*

Доведення. В супротивному припущенні матимемо $\rho_1 < \rho$. Нехай буде $x_1', x_2', \ldots, x_n'$ розв'язок задачі мінімального наближення для розглядуваної підсистеми рівнянь. Розгляньмо систему значень

$$x_j'' = x_{j0}\cos^2\varphi + x_j'\sin^2\varphi;\ j = 1, 2, \ldots, n;\ 0 < \varphi < \frac{\pi}{2} \tag{378}$$

Якщо Δ_{i0}, Δ_i' та Δ_i'' позначають, при $i = 1, 2, \ldots, m$, відхили рівняння $F_i = l_i$ відповідно для трьох систем значень невідомих $x_j = x_{j0}$, $x_j = x_j'$ та $x_j = x_j''$, $j = 1, 2, \ldots, n$, матимемо на підставі (378):

$$\Delta_i'' = \Delta_{i0}\cos^2\varphi + \Delta_i'\sin^2\varphi,\ i = 1, 2, \ldots, m \tag{379}$$

Бачимо, що відхили Δ_i'' для рівнянь розглядуваної підсистеми будуть абсолютним значенням менші за ρ при кожному φ $\left(\text{при } 0 < \varphi < \frac{\pi}{2}\right)$, а для решти рівнянь системи (367) те саме буде забезпечене при досить малому φ. А коли так, то ρ не могло б виражати мінімальне наближення для системи (367) всупереч умовам теореми.

7. **Лема I.** *Нехай маємо $\mu + 1$ нерівностей* $(\mu \geqslant n \geqslant 1)$:

$$\left.\begin{array}{lll}
F_1 & \equiv & x_1 > 0 \\
F_2 & \equiv & x_2 > 0 \\
\cdots & \cdots & \cdots \\
F_n & \equiv & x_n > 0 \\
F_{n+1} & \equiv & a_{n+1,1}x_1 + a_{n+1,2}x_2 + \cdots + a_{n+1,n}x_n > 0 \\
\cdots & \cdots & \cdots \\
F_\mu & \equiv & a_{\mu 1}x_1 + a_{\mu 2}x_2 + \cdots + a_{\mu n}x_n > 0 \\
F_{\mu+1} & \equiv & a_{\mu+1,1}x_1 + a_{\mu+1,2}x_2 + \cdots + a_{\mu+1,n}x_n \geqslant 0
\end{array}\right\} \tag{380}$$

Припустімо, що перші μ нерівностей сумісні (тобто існують системи значень $x_1, x_2, \ldots, x_n$, які задовольняють водночас усі μ нерівностей), *і що остання $\mu + 1$-ша нерівність являє собою наслідок перших μ нерівностей* (тобто нерівність $F_{\mu+1} < 0$ несумісна з першими μ нерівностями). *Ми твердимо, що в такому разі існує система невід'ємних констант $C_1, C_2, \ldots, C_\mu$ — таких, що*

$$\left.\begin{array}{c}
F_{\mu+1} \equiv C_1F_1 + C_2F_2 + \cdots + C_\mu F_\mu \\
C_1 \geqslant 0,\ C_2 \geqslant 0, \ldots, C_\mu \geqslant 0
\end{array}\right\} \tag{381}$$

Доведення. У випадку $n=1$ справедливість твердження очевидна. Отже досить, припустивши справедливість леми для $n=\nu$, довести, що вона зберігає силу також при $n=\nu+1$.

Співвідношення (381) напевне має місце, як би трапилось, що $F_{\mu+1}\equiv 0$. Виключаючи цей випадок, $F_{\mu+1}$ міститиме принаймні одну із змінних $x_1, x_2, \ldots, x_{\nu+1}$. Припустімо для простоти запису, що вона містить у всякому разі $x_{\nu+1}$, себто $a_{\mu+1,\,\nu+1}\neq 0$. Взявши за нові незалежні змінні

$$x_1'=x_1,\ x_2'=x_2,\ \ldots,\ x_\nu'=x_\nu,\ x_{\nu+1}'=F_{\mu+1} \tag{382}$$

та міняючи належним чином нумерацію нерівностей (380) (при $n=\nu+1$), ми можемо записати систему перетворених нерівностей в такому виді:

$$\begin{array}{llll}
F_1' & =F_1 & = x_1' & >0\\
F_2' & =F_2 & = x_2' & >0\\
\cdots & \cdots & \cdots & \cdots\\
F_\nu' & =F_\nu & = x_\nu' & >0\\
F_{\nu+1}' & =F_\alpha & = a_{\alpha 1}\,x_1'+a_{\alpha 2}\,x_2'+\cdots+a_{\alpha\nu}\,x_\nu' & >0\\
\cdots & \cdots & \cdots & \cdots\\
F_{\nu+r}' & =F_\beta & = a_{\beta 1}\,x_1'+a_{\beta 2}\,x_2'+\cdots+a_{\beta\nu}\,x_\nu' & >0\\
F_{\nu+r+1}' & =F_\gamma & = \left|\dfrac{a_{\gamma,\,\nu+1}}{a_{\mu+1,\,\nu+1}}\right|\left[f_{\nu+r+1}(x_1',x_2',\ldots,x_\nu')-x_{\nu+1}'\right] & >0\\
\cdots & \cdots & \cdots & \cdots\\
F_{\nu+r+s}' & =F_\delta & = \left|\dfrac{a_{\delta,\,\nu+1}}{a_{\mu+1,\,\nu+1}}\right|\left[f_{\nu+r+s}(x_1',x_2',\ldots,x_\nu')-x_{\nu+1}'\right] & >0\\
F_{\nu+r+s+1}' & =F_\varepsilon & = \dfrac{a_{\varepsilon,\,\nu+1}}{a_{\mu+1,\,\nu+1}}\left[g_{\nu+r+s+1}(x_1',x_2',\ldots,x_\nu')+x_{\nu+1}'\right] & >0\\
\cdots & \cdots & \cdots & \cdots\\
F_\mu' & =F_\eta & = \dfrac{a_{\eta,\,\nu+1}}{a_{\mu+1,\,\nu+1}}\left[g_\mu(x_1',x_2',\ldots,x_\nu')+x_{\nu+1}'\right] & >0\\
F_{\mu+1}' & =F_{\mu+1} & = x_{\nu+1}' & \geqslant 0
\end{array} \tag{383}$$

Тут:

$\alpha, \ldots, \beta, \gamma, \ldots, \delta, \varepsilon, \ldots, \eta$ — числа $\nu+1, \nu+2, \ldots, \mu$ в якомусь іншому порядку;

$F_\alpha, \ldots, F_\beta$ — такі лінійні форми, що не містили в собі $x_{\nu+1}$;

$F_\gamma, \ldots, F_\delta$ — такі лінійні форми, в яких знак коефіцієнта при $x_{\nu+1}$ був протилежний до знака $a_{\mu+1,\,\nu+1}$;

$F_\varepsilon, \ldots, F_\eta$ — такі лінійні форми, в яких знак коефіцієнта при $x_{\nu+1}$ був однаковий із знаком $a_{\mu+1,\,\nu+1}$;

нарешті $f_{\nu+r+1}, \ldots, f_{\nu+r+s}, g_{\nu+r+s+1}, \ldots, g_\mu$ — певні лінійні форми від $x_1', x_2', \ldots, x_\nu'$, що їхні коефіцієнти було б не важко виразити через

відповідні коефіцієнти форм $F_\gamma, \ldots, F_\delta, F_\varepsilon, \ldots, F_\eta$ та $F_{\mu+1}$. Може трапитись, зрозуміла річ, що $r=0$ або $s=0$, або $\mu-(\nu+r+s)=0$, — тоді із системи (383) відповідна категорія рівнянь випадає.

Оскільки нерівність $x'_{\nu+1}<0$ повинна бути несумісна з μ першими нерівностями системи (383), то, як легко зрозуміти, якась із нерівностей

$$-g_{\nu+r+s+t}\,(x'_1,\ x'_2,\ \ldots,\ x'_\nu)\geqslant 0, \tag{384}$$

де $1\leqslant t\leqslant \mu-(\nu+r+s)$, повинна бути наслідком такої системи сумісних нерівностей, числом $\nu+r+s(\mu-\nu-r-s)+t-1$:

$$\left.\begin{array}{c} F'_1>0, F'_2>0,\ldots,\ F'_{\nu+r}>0,\ f_i+g_j>0,\ g_{\nu+r+s+k}>0 \\ (i=\nu+r+1,\ldots,\nu+r+s;\ j=\nu+r+s+1,\ldots,\mu;\ k=1,\ldots,t-1) \end{array}\right\} \tag{385}$$

Алеж при $n=\nu$ справедливість леми встановлена. Отже існує система таких невід'ємних констант $C'_1, C'_2,\ldots, C'_{\nu+r}, C'_{i,j}, C''_k$, що

$$\left.\begin{array}{c} -g\equiv C'_1F'_1+\cdots+C'_{\nu+r}F'_{\nu+r}+\sum\limits_{i,j}C'_{i,j}(f_i+g_j)+\sum\limits_k C''_k g_{\nu+r+s+k} \\ (i=\nu+r+1,\ldots,\ \nu+r+s;\ j=\nu+r+s+1,\ldots,\ \mu;\ k=1,\ldots,t-1) \end{array}\right\} \tag{386}$$

Взявши ж на увагу, що

$$f_i\equiv F_{\mu+1}+k_i^2F'_i,\ g_j\equiv -F_{\mu+1}+h_j^2F'_j, \tag{387}$$

де ми позначаємо через k_i^2 та h_j^2 додатні величини

$$\left|\frac{a_{\mu+1,\,\nu+1}}{a_{\gamma,\,\nu+1}}\right|,\ldots,\ \left|\frac{a_{\mu+1,\,\nu+1}}{a_{\delta,\,\nu+1}}\right|,\ \frac{a_{\mu+1,\,\nu+1}}{a_{\varepsilon,\,\nu+1}},\ldots,\ \frac{a_{\mu+1,\,\nu+1}}{a_{\eta,\,\nu+1}}$$

бачимо, що лема доведена цілком.

Лема II. *При умовах попередньої леми, якщо $F_{\mu+1}$ не являє собою тотожний нуль, має місце тотожність виду*

$$\left.\begin{array}{c} F_{\mu+1}\equiv C_1F_{i_1}+C_2F_{i_2}+\cdots+C_\sigma F_{i_\sigma} \\ C_1>0,\ C_2>0,\ldots,C_\sigma>0 \end{array}\right\} \tag{388}$$

де $F_{i_1}, F_{i_2},\ldots, F_{i_\sigma}$ є деякі з форм $F_1, F_2,\ldots, F_\mu$, лінійно незалежні між собою.

Доведення. Припустімо, що вже встановлено існування тотожності виду

$$\left.\begin{array}{c} F_{\mu+1}\equiv c_1\Phi_1+c_2\Phi_2+\cdots+c_\tau\Phi_\tau \\ c_1>0,\ c_2>0,\ldots,\ c_\tau>0 \end{array}\right\} \tag{389}$$

де $\Phi_i=\Phi_i(x_1,\ x_2,\ldots,\ x_n)$, $i=1,2,\ldots,\tau$, є деякі з форм $F_1,\ F_2,\ldots,\ F_\mu$, отже це є напевне такі лінійні форми, що τ нерівностей

$$\Phi_1>0,\ \Phi_2>0,\ldots,\ \Phi_\tau>0 \tag{390}$$

є сумісні. Така є зокрема тотожність (381), що її існування доведене в попередній лемі, якщо опустити на правій стороні ті члени, які являють тотожний нуль. Для нашої мети досить, очевидно, показати, що тотожність

(389) можна замінити на аналогічну тотожність, але з числом членів меншим принаймні на одиницю, кожного разу, коли форми $\Phi_1, \Phi_2, \ldots, \Phi_\tau$ між собою лінійно залежні.

Нехай буде тут ν ранг матриці коефіцієнтів форм $\Phi_1, \Phi_2, \ldots, \Phi_\tau$, при чому, звичайно, $\nu \leqslant n$. Припустімо, для спрощення запису, що якраз перші ν форм $\Phi_1, \Phi_2, \ldots, \Phi_\nu$ є між собою лінійно незалежні. Якщо ввести нові незалежні змінні

$$x_1' = \Phi_1,\ x_2' = \Phi_2,\ \ldots,\ x_\nu' = \Phi_\nu, \tag{391}$$

система сумісних нерівностей (390) перетвориться на таку:

$$\begin{array}{lllll}
\Phi_1 & = \Phi_1' & = & x_1' & > 0 \\
\Phi_2 & = \Phi_2' & = & x_2' & > 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\Phi_\nu & = \Phi_\nu' & = & x_\nu' & > 0 \\
\Phi_{\nu+1} & = \Phi_{\nu+1}' & = & a_{\nu+1,1}' x_1' + a_{\nu+1,2}' x_2 + \cdots + a_{\nu+1,\nu}' x_\nu' & > 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\Phi_\tau & = \Phi_\tau' & = & a_{\tau 1} x_1' + a_{\tau 2}' x_2' + \cdots + a_{\tau\nu}' x_\nu' & > 0
\end{array} \tag{392}$$

розуміючи під $\Phi_1', \ldots, \Phi_\tau'$ (і аналогічно під $F_{\mu+1}'$) ті самі форми $\Phi_1, \ldots, \Phi_\tau$ (та $F_{\mu+1}$), виражені через нові незалежні змінні. Водночас тотожність (389) заміниться на таку:

$$\left.\begin{array}{c}
F_{\mu+1}' \equiv c_1 \Phi_1' + c_2 \Phi_2' + \cdots + c_\tau \Phi_\tau' \\
c_1 > 0,\ c_2 > 0,\ \ldots,\ c_\tau > 0
\end{array}\right\} \tag{389'}$$

Звернімо нашу увагу на одну з останніх $\tau - \nu$ нерівностей системи (392), наприклад, на нерівність $\Phi_\tau' > 0$. Сумісність системи (392) припускає, що принаймні один із коефіцієнтів форми Φ_τ' є додатний, але вона може мати й кілька додатних коефіцієнтів. Нехай буде для конкретності

$$\left.\begin{array}{c}
a_{\tau 1}' > 0,\ a_{\tau 2}' > 0, \ldots, a_{\tau\lambda}' > 0 \\
a_{\tau,\lambda+1}' < 0, \ldots, a_{\tau\nu}' < 0
\end{array}\right\} \tag{393}$$

Позначмо через $k\ (k > 0)$ найменше з λ чисел

$$\frac{c_1}{a_{\tau 1}'},\ \frac{c_2}{a_{\tau 2}'}, \ldots,\ \frac{c_\lambda}{a_{\tau\lambda}'} \tag{394}$$

Нехай буде, наприклад, $k = \dfrac{c_1}{a_{\tau 1}'}$.

Тоді, як легко бачити, тотожність

$$\left.\begin{array}{l}
F_{\mu+1} \equiv (c_2 - k a_{\tau 2}') \Phi_2 + \cdots + (c_\lambda - k a_{\tau\lambda}') \Phi_\lambda + \\
\quad + (c_{\lambda+1} - k a_{\tau,\lambda+1}') \Phi_{\lambda+1} + \cdots + (c_\nu - k a_{\tau\nu}') \Phi_\nu + \\
\quad + c_{\nu+1} \Phi_{\nu+1} + \cdots + c_{\tau-1} \Phi_{\tau-1} + \\
\quad + (c_\tau + k) \Phi_\tau
\end{array}\right\} \tag{395}$$

має потрібні властивості. Тим самим лему доведено.

Увага. Доведені тільки но дві леми залишаються в силі і тоді, коли система нерівностей задана не в формі (380), а в загальнішій формі:

$$\begin{gathered} f_i = a_i x + b_i y + \cdots + h_i u > 0 \\ (i = 1, 2, \ldots, \mu) \\ f_{\mu+1} = a_{\mu+1} x + b_{\mu+1} y + \cdots + h_{\mu+1} u \geqslant 0 \end{gathered} \tag{380'}$$

від якої можна перейти до форми (380), перетворюючи змінні з допомогою формул типу (366), при чому, з одного боку, ранг матриці коефіцієнтів перших μ форм у системі (380') [що з їх числа добирають тут нові незалежні змінні (366)] не може, очевидно, бути менший за одиницю, а з другого боку, як легко зрозуміти, ранг матриці не може, за умов леми, підвищитись при доданні $\mu+1$-шої форми.

8°. **Теорема.** *Мінімальне наближення для системи рівнянь* (367) *є водночас мінімальне наближення для деякої підсистеми r рівнянь* ($r \leqslant n+1$), *що їхні ліві сторони є лінійно залежні у вузькому розумінні* (n° 2, означення II), *або для одного рівняння, що його ліва сторона є тотожний нуль* [1]).

Доведення. На підставі теореми n° 6, ρ дорівнює мінімальному наближенню для підсистеми тих рівнянь, що для них відхил при $x_i = x_{i0}$ [2]) ($i = 1, 2, \ldots, n$) дорівнює $\pm\rho$. Нехай це будуть рівняння

$$F_I = l_I,\ F_{II} = l_{II}, \ldots,\ F_{(s)} = l_{(s)} \tag{396}$$

[запроваджуємо для цих рівнянь окрему (римську) нумерацію]. Множачи в разі потреби обидві сторони деяких рівнянь на -1, можемо також завжди припустити, що відхили $\Delta_{(i)}$, [$(i) = I, II, \ldots (s)$] усі дорівнюють $+\rho$. Із того, що ρ являє собою величину мінімального наближення для підсистеми рівнянь (396) і, з другого боку,

$$F_I = l_I - \rho,\ F_{II} = l_{II} - \rho, \ldots,\ F_{(s)} = l_{(s)} - \rho \tag{397}$$

при $x_i = x_{i0}$, $i = 1, 2, \ldots, n$, виходить, що система нерівностей

$$\Delta F_I > 0,\ \Delta F_{II} > 0, \ldots,\ \Delta F_{(s)} > 0 \tag{398}$$

несумісна. Отже, заміняючи в лінійних формах $\Delta F_I, \Delta F_{II}, \ldots, \Delta F_{(s)}$ аргументи $\Delta x_1, \Delta x_2, \ldots, \Delta x_n$ відповідно на $x_1, x_2, \ldots, x_n$, приходимо до висновку, що й система нерівностей

$$F_I > 0,\ F_{II} > 0, \ldots,\ F_{(s)} > 0 \tag{399}$$

також повинна бути несумісна.

[1]) Оскільки додання нових рівнянь до будьякої системи рівнянь не може зменшити величину мінімального наближення, можна, очевидно, твердити, якщо дана теорема справедлива, що мінімальне наближення для всієї системи (367) є водночас мінімальне наближення для якоїсь підсистеми з $n+1$ рівнянь. Проте це твердження висловлює, очевидно, менше, ніж сама теорема.

[2]) $x_i = x_{i0}$, при позначеннях n° 6, є та система значень невідомих, яка дає мінімальне наближення для системи рівнянь (367) в цілому.

Можливі два випадки:

а) Яканебудь із форм $F_I, F_{II}, \ldots, F_{(s)}$ є тотожний нуль. Тоді, очевидно, теорема безпосередньо стверджується, якщо взяти на увагу відповідну рівність у системі (397).

б) Жодна із форм $F_I, F_{II}, \ldots, F_{(s)}$ не є тотожний нуль. Беручи тоді із системи (399) спочатку одну першу нерівність, потім дві перші, три перші і т. д., нехай буде t таке число $(1 < t \leqslant s)$, що система перших $t-1$ нерівностей (399) є сумісна, а система перших t нерівностей уже несумісна. Тоді в цій системі нерівностей

$$F_I > 0,\ F_{II} > 0, \ldots,\ F_{(t-1)} > 0;\ -F_{(t)} \geqslant 0 \tag{400}$$

перші $t-1$ нерівностей сумісні, а остання нерівність є наслідок із них. В такому разі, на підставі лем I та II n° 7 та уваги наприкінці того самого n°, ми повинні мати (оскільки, згідно з розглядуваним припущенням, тотожність $-F_{(t)} \equiv 0$ виключена) таку тотожність:

$$\left.\begin{array}{c} -F_{(t)} \equiv C_1 F_{(\alpha)} + C_2 F_{(\beta)} + \cdots + C_{r-1} F_{(\varkappa)} \\ C_1 > 0,\ C_2 > 0, \ldots, C_{r-1} > 0 \end{array}\right\} \tag{401}$$

де $F_{(\alpha)}, F_{(\beta)}, \ldots, F_{(\varkappa)}$ є якісь $r-1$ $(2 \leqslant r \leqslant n+1)$ лінійно незалежні форми із числа $F_I, F_{II}, \ldots, F_{(t-1)}$. Але ця тотожність має наслідком своїм несумісність системи нерівностей

$$F_{(\alpha)} > 0,\ F_{(\beta)} > 0, \ldots,\ F_{(\varkappa)} > 0,\ F_{(t)} > 0 \tag{402}$$

отже й несумісність системи нерівностей

$$\Delta F_{(\alpha)} > 0,\ \Delta F_{(\beta)} > 0, \ldots,\ \Delta F_{(\varkappa)} > 0,\ \Delta F_{(t)} > 0 \tag{403}$$

А це значить, що ρ є мінімальне наближення для системи r рівнянь

$$F_{(\alpha)} = l_{(\alpha)},\ F_{(\beta)} = l_{(\beta)}, \ldots,\ F_{(\varkappa)} = l_{(\varkappa)},\ F_{(t)} = l_{(t)}, \tag{404}$$

яка, очевидно, цілком відповідає твердженню теореми.

З приводу доведеної тільки но важливої теореми зробімо деякі зауваження.

Припустімо, що з рівнянь системи (367) ми утворили всі можливі комбінації[1]) по r рівнянь $(r \leqslant n+1)$ з лінійно залежними у вузькому розумінні лівими сторонами, а також одібрали й окремі рівняння, що мають на лівій стороні тотожний нуль. Ці окремі рівняння нам зручно буде трактувати як комбінації з $r=1$ рівнянь. Кожній із цих комбінацій r рівнянь відповідає певна величина ρ' мінімального наближення, яка визначається з формули (377) n° 5 цього параграфа. Звернімо нашу увагу на ті комбінації рівнянь, що їм відповідає найбільше значення ρ'. Згідно з доведеною теоремою, це найбільше значення дорівнює ρ, тобто мінімальному наближенню всієї системи (367). Тут можливі три випадки:

[1]) Це, звичайно, шлях теоретичний. Практичне розв'язання розглядуваних задач ми уявляємо в загальному випадку на шляхах застосування того методу послідовних наближень, який є предмет розділу VII цієї праці, а також § 25 даного розділу.

α) Якщо між останніми комбінаціями рівнянь є хоча б одна, що містить $n+1$ рівнянь, то існує (n° 5) лише одна система значень $x_{10}, x_{20}, \ldots, x_{n0}$, яка дає мінімальне наближення для цієї комбінації рівнянь, і це водночас буде єдиний [1]) розв'язок задачі мінімального наближення для системи (367) в цілому.

β) Лише для окремих рівнянь, що мають на лівій стороні тотожний нуль (а на правій $\pm\rho$), мінімальне наближення досягає величини ρ. Тоді задача мінімального наближення для системи (367) в цілому допускає ∞^n розв'язків: точки $(x_{10}, x_{20}, \ldots, x_{n0})$ заповнюють деяку n-мірну замкнену й обмежену (n° 4) конвексну [2]) область. Границі області визначаються відповідною системою лінійних нерівностей

$$l_i - \rho \leqslant F_i \leqslant l_i + \rho; \quad i = 1, 2, \ldots, m \tag{405}$$

γ) Найбільше значення r, при якому мінімальне наближення досягає величини ρ, є проміжне якесь число між 1 та $n+1$. Тоді, згідно з n° 5, ми маємо шукати розв'язки задачі мінімального наближення для системи (367) в цілому між ∞^{n+1-r} розв'язків аналогічної задачі щодо розглядуваної комбінації r рівнянь. Виразивши $r-1$ із числа невідомих $x_1, x_2, \ldots, x_n$ з допомогою системи (376), через інші $n+1-r$ невідомих та підставивши в систему (367), дістаємо аналогічну задачу мінімального наближення при меншій кількості $n+1-r$ невідомих; при тому деякі рівняння перетвореної системи мають на лівій стороні тотожний нуль, а на правій $\pm\rho$. Над одержаною перетвореною системою m рівнянь із $n+1-r$ невідомими повторюємо аналогічне дослідження [3]) і т. д. Кінець-кінцем, приходимо до одного з двох попередніх випадків: або до випадку α) — тоді задача мінімального наближення для системи (367) має лише один розв'язок; або до випадку β) — і тоді задача має ∞^ν розв'язків, якщо ν означає кількість невідомих в останній одержаній нами при послідовних перетвореннях системі рівнянь; при цьому точки $(x_{10}, x_{20}, \ldots, x_{n0})$ заповнюють обмежену, замкнену, конвексну ν-мірну область [4]).

§ 24. Задача мінімального наближення для нескінченної множини лінійних суперечних рівнянь. Застосування до задачі Чебишова щодо апроксимації функцій від довільного числа незалежних змінних

1°. Розгляньмо несумісну систему, що складається з нескінченної множини лінійних рівнянь:

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = l, \tag{406}$$

[1]) Щождо задачі мінімального наближення для системи (361), то вона при $n < N$ матиме в цьому разі не один, а ∞^{N-n} розв'язків.

[2]) Термін „конвексна область" тут узагальнює ту властивість тримірних або двомірних конвексних областей, за якою кожний відрізок, що сполучає дві точки області, належить всіма точками до цієї області.

[3]) При цьому тепер наперед відома величина ρ.

[4]) Щождо відповідної системи (361), то для неї задача мінімального наближення при $n < N$ матиме $\infty^{\nu+N-n}$ розв'язків, при чому відповідні точки $(x_0, y_0, \ldots, u_0)$ заповнюють необмежену, замкнену, конвексну $\nu+N-n$-мірну область.

де $x_1, x_2, \ldots, x_n$ є n невідомих (n — дане скінченне число), що підлягають визначенню, а коефіцієнти $a_1, a_2, a_3, \ldots, a_n, l$ є змінні, при чому допущенні системи значень цих змінних означаються деякою обмеженою точковою множиною G в $n+1$-мірному просторі $(a_1, a_2, \ldots, a_n, l)$. Позначаючи через Q точку $(a_1, a_2, \ldots, a_n, l)$ і через Δ відхил

$$l-(a_1 x_1+a_2 x_2+\cdots+a_n x_n)=\Delta, \tag{407}$$

задача мінімального наближення тут полягає у відшуканні такої системи значень невідомих $x_1=x_{10}, x_2=x_{20}, \ldots, x_n=x_{n0}$, щоб величина

$$\operatorname*{borne\,sup}_{Q\subset G}|\Delta|=L \tag{408}$$

була якнайменша.

Ми припустимо точкову множину G такою, що коефіцієнти лівої сторони $a_1, a_2, \ldots, a_n$ не є лінійно залежні між собою, тобто не є зв'язані між собою лінійним однорідним співвідношенням із сталими коефіцієнтами виду

$$C_1 a_1+C_2 a_2+\cdots+C_n a_n=0$$

або, інакше кажучи, що з рівнянь нескінченної системи (406) можна виділити одну принаймні скінченну підсистему з n рівнянь з відмінним від нуля детермінантом [1]).

Поставлена тут задача мінімального наближення обіймає, зокрема, своєю загальною схемою задачу найкращого наближеного представлення якоїсь обмеженої функції $f(x, y, \ldots, u)$ в будьякій області E, N-мірного простору $(x, y, \ldots, u)$, з допомогою лінійної комбінації n даних обмежених лінійно незалежних функцій

$$a\psi_1(x, y, \ldots, u)+b\psi_2(x, y, \ldots, u)+\cdots+h\psi_n(x, y, \ldots, u) \tag{409}$$

Тут як невідомі, що їх треба визначити, матимемо, очевидно, параметри $a, b, \ldots, h$. Змінні ж коефіцієнти є $\psi_1, \psi_2, \ldots, \psi_n, f$. Функція f, як в розділах VI, VII, може бути розривна і навіть многозначна. Більш того, від функцій $\psi_1, \psi_2, \ldots, \psi_n$ ми таксамо вимагаємо лише, щоб вони були обмежені в області E, не виключаючи, наприклад, випадку функцій ψ_i з розривами неперервності. Нарешті, область E може бути сама необмежена, припускаючи лише обмеженість функцій f та ψ_i.

2°. Зважаючи на неперервність лівої сторони (407) щодо $a_1, a_2, \ldots, a_n, l$, величина (408), очевидно, є тотожна з

$$\max_{Q\subset G^\circ}|\Delta|=L, \tag{410}$$

де $G^\circ=G+G'$ є замкнена точкова множина, утворена через сполучення [2]) G з похідною множиною G'. В дальшому ми й виходитимемо з такого поставлення задачі, що полягає в мінімізації виразу (410).

[1]) В противному разі нескінченну систему (406) можна було б, очевидно, звести до аналогічної з меншою кількістю невідомих, так щоб умова лінійної незалежності коефіцієнтів уже задовольнялася.

[2]) Звичайним знаком + позначаємо те, що Caratheodory називає Vereinigungsmenge.

Ми позначатимемо як поширену систему лінійних рівнянь сукупність рівнянь (406) у тому випадку, коли коефіцієнти підпорядковуються умові $Q \subset G^\circ$, замість первісної умови $Q \subset G$. Поширена система не відрізняється від первісної у випадку $G = G^\circ$. Тотожність величин (408) та (410) показує, що задача мінімального наближення має тотожний розв'язок для системи лінійних рівнянь (406) — первісної та поширеної.

Мінімальне значення ρ величини L, очевидно, не може перевищувати $\max\limits_{Q \subset G^\circ} l$, що являє собою значення L при $x_1 = x_1 = \cdots = x_n = 0$. Це, у зв'язку з лінійною незалежністю величин $a_1, a_2, \ldots, a_n$, встановлює, очевидно, a priori певні скінченні межі для шуканих значень $x_1, x_2, \ldots, x_n$. Коли ж так, то існування розв'язків $(x_{10}, x_{20}, \ldots, x_{n0})$, числом один або безліч, встановлюється подібно до того як у § 15, і так само доводиться теорема щодо критерія точності наближеного розв'язку, аналогічна до встановленої наприкінці § 15.

3°. **Теорема.** *Якщо точкова множина G є замкнена ($G = G$), то мінімальне наближення ρ для нескінченної системи рівнянь* (406) *є водночас мінімальне наближення для деякої скінченної підсистеми з r рівнянь* ($r \leqslant n+1$), *ліві сторони яких є лінійно залежні у вузькому розумінні* (§ 23 n° 2, означення II) *абож для одного рівняння, яке має тотожний нуль на лівій стороні. Аналогічне твердження має силу і в тому разі, коли точкова множина G не є замкнена ($G \neq G^\circ$) — з тією лише відміною, що в цьому разі підсистема r рівнянь* ($1 \leqslant r \leqslant n+1$) *складається з рівнянь поширеної системи* (406), *в яку включено рівняння, що відповідають точкам множини G', похідної від G.*

Доведення. Зважаючи на теорему n° 8 § 23, нам досить, очевидно, пересвідчитись, що ρ дорівнює найкращому наближенню якоїсь скінченної підсистеми рівнянь, наприклад, підсистеми $n+1$ рівнянь. Потрібне для цього міркування складатиметься з двох частин.

Насамперед, покажімо, що кожному числу $\rho_1 < \rho$ можна протипоставити таку скінченну підсистему рівнянь, що для неї мінімальне наближення буде більше за ρ_1. До цієї підсистеми ми, поперше, включимо n рівнянь $\overline{Q_1}, \overline{Q_2}, \ldots, \overline{Q_n}$ [1]), дібраних із системи (406) за такою умовою, щоб відповідний детермінант n-го порядку був відмінний від нуля (n° 1). Вимагаючи, щоб відхили для цих n рівнянь не перевищували абсолютним значенням величину $\max\limits_{Q \subset G^\circ} l \geqslant \rho$, ми, очевидно, можемо встановити певні межі для значень невідомих:

$$|x_i| \leqslant m_i; \quad i = 1, 2, \ldots, n, \tag{411}$$

поза якими цієї вимоги не можна задовольнити. Щодо точок $(x_1, x_2, \ldots, x_n)$ області (411), то кожній із них відповідає певне значення (410) $L \geqslant \rho$

[1]) Ми позначаємо кожне з рівнянь системи (406) тією самою літерою Q, як і відповідну точку $(a_1, a_2, \ldots, a_n, l)$. Нам доведеться в дальшому міркуванні, користуючись геометричною мовою, говорити, з одного боку, про точки $Q \subset G^\circ$ в $n+1$-мірному просторі $(a_1, a_2, \ldots, a_n, l)$, а з другого боку — про точки n-мірного простору $(x_1, x_2, \ldots, x_n)$.

і певне рівняння $\overline{\overline{Q}}$ [1]), для якого $|\Delta| = L$. Ясно, що кожну таку точку $(x_1, x_2, \ldots, x_n)$ можна буде оточити досить малою понадсферою $(X_1 - x_1)^2 + (X_2 - x_2)^2 + \cdots + (X_n - x_n)^2 \leqslant r^2$ так, щоб для значень $X_1, X_2, \ldots, X_n$, що належать до області цієї понадсфери, відхил того самого рівняння $\overline{\overline{Q}}$ був більший від ρ_1 (пам'ятаючи, що $\rho_1 < \rho$). Кожній точці області (411) відповідає, отже, певна точка $\overline{\overline{Q}}$ і певний понадсферичний окіл. Згідно з відомою лемою Бореля, відповідно узагальненою для області кількох вимірів, область (411) можна буде обійняти скінченною кількістю таких понадсфер. Нехай будуть $\overline{\overline{Q}}_1, \overline{\overline{Q}}_2, \ldots, \overline{\overline{Q}}_m$ рівняння, які відповідають (у вищез'ясованому розумінні) центрам цих понадсфер. Не важко зрозуміти, що система $m + n$ рівнянь

$$\overline{Q}_1, \overline{Q}_2, \ldots, \overline{Q}_n; \qquad \overline{\overline{Q}}_1, \overline{\overline{Q}}_2, \ldots, \overline{\overline{Q}}_m \tag{412}$$

відповідає поставленій вимозі: при яких завгодно значеннях $x_1, x_2, \ldots, x_n$ найбільший з відхилів для системи (412) буде напевне більший за ρ_1. Отже мінімальне наближення системи (412) є більше за ρ_1. А згідно з теоремою п° 8 § 23 (та відповідною виноскою до неї), можна із системи (412) виділити підсистему $n + 1$ рівнянь, яка матиме таку саму властивість.

Переходячи до другої (і останньої) частини міркування, задамо собі послідовність чисел

$$\rho_1 < \rho_2 < \rho_3 < \ldots; \quad \lim_{i \to \infty} \rho_i = \rho \tag{413}$$

З кожним із цих чисел ρ_i можна буде зіставити підсистему $n + 1$ рівнянь $Q_{1i}, Q_{2i}, \ldots, Q_{n+1,i}$ таку, для якої мінімальне наближення ρ' більше за ρ_i. Розглядаючи $(n + 1)^2$ координат відповідної комбінації $n + 1$ точок простору $(a_1, a_2, \ldots, a_n, l)$ як координати однієї точки в $(n + 1)^2$-мірнім просторі, ми матимемо для неї при $i = 1, 2, 3, \ldots$ принаймні одну граничну точку, що відповідає якійсь частинній послідовності значень індекса i. Нехай буде $Q_{10}, Q_{20}, \ldots, Q_{n+1,0}$ відповідна гранична комбінація $n + 1$ точок простору $(a_1, a_2, \ldots, a_n, l)$ [2]). Беручи на увагу неперервну [3]) залежність величини ρ' від згаданих $(n + 1)^2$ координат, легко зрозуміти, що для системи рівнянь $Q_{10}, Q_{20}, \ldots, Q_{n+1,0}$ мінімальне наближення ρ' точно дорівнює ρ, а це й завершує доведення.

Наслідок. Доведена теорема має безпосереднє застосування до сформульованої в п° 1 цього параграфа задачі найкращого наближеного представлення функції $f(x, y, \ldots, u)$ в якійсь області E простору $(x, y, \ldots, u)$ за допомогою лінійної комбінації, із сталими коефіцієнтами $a, b, \ldots, h$, даних функцій $\psi_i(x, y, \ldots, u)$, $i = 1, 2, \ldots, n$, — даючи певні необхідні

[1]) Точка Q належить до замкненої точкової множини G°

[2]) Точки $Q_{j0} \subset G^\circ$, $j = 1, 2, \ldots, n + 1$.

[3]) Неперервність цієї залежності безпосередньо ясна на підставі неперервної залежності величин відхилів (407) од відповідних коефіцієнтів $a_1, a_2, \ldots, a_n, l$ розглядуваної скінченної підсистеми рівнянь, — хоч, з другого боку, варто зауважити, що кількість розв'язків $(x_{10}, x_{20}, \ldots, x_{n0})$ для тієї самої скінченної системи рівнянь може мінятися скоком при нескінченно малій зміні коефіцієнтів.

умови для найкращого наближеного представлення. Зокрема, у випадку однозначних та неперервних f, ψ_i та замкненої області E ми маємо, як наслідок доведеної теореми, таке твердження: *найкраще наближене представлення в області E є водночас найкраще наближене представлення на множині якихось $r \leqslant n+1$ точок максимального відхилу в області E, для яких згадані лінійні комбінації, розглядувані як форми від коефіцієнтів $a, b, \ldots, h$, є лінійно залежні у вузькому розумінні* (§ 23, n° 2, означення II), *абож для однієї точки максимального відхилу, що в ній $\psi_i = 0$, $i = 1, 2, \ldots, n$.*

У більш частинному випадку поліноміального наближеного представлення аналогічне твердження (в менш викінченому сформулюванні) у виді припущення висловив у своїй дисертації P. Kirchberger (op. cit.). Нам невідомо, щоб досі хтонебудь довів твердження Kirchberger-а. Доведена ж у даному n° теорема обіймає згадане твердження як досить частинний випадок.

4°. **Лема.** *Нехай маємо систему $r \leqslant n+1$ лінійних рівнянь*

$$\left.\begin{array}{c} F_i(x_1, x_2, \ldots, x_n) \equiv a_{1i} x_1 + a_{2i} x_2 + \cdots + a_{ni} x_n = l_i \\ (i = 1, 2, \ldots, r) \end{array}\right\} \tag{414}$$

з лінійно залежними у вузькому розумінні (§ 23, n° 2, означення II) *лівими сторонами. Нехай будуть $C_1, C_2, \ldots, C_r$* такі константи, що

$$C_1 F_1 + C_2 F_2 + \cdots + C_r F_r \equiv 0 \tag{415}$$

Позначмо $l_i - F_i = \Delta_i$; $i = 1, 2, \ldots, r$. Якщо для певної системи значень невідомих $x_1, x_2, \ldots, x_n$

$$\operatorname{sg} \Delta_1 \cdot \operatorname{sg} C_1 = \operatorname{sg} \Delta_2 \cdot \operatorname{sg} C_2 = \cdots = \operatorname{sg} \Delta_r \cdot \operatorname{sg} C_r, \tag{416}$$

то мінімальне наближення ρ' системи (414) *не може бути менше за величину*

$$\min\{|\Delta_1|, |\Delta_2|, \ldots, |\Delta_r|\} \tag{417}$$

Воно не може й дорівнювати цій величині, крім випадку $|\Delta_1| = |\Delta_2| = \cdots = |\Delta_r|$.

Доведення. Так само як в § 23 n° 5, насамперед встановлюємо:

$$|C_1| \Delta_1 \operatorname{sg} C_1 + \cdots + |C_r| \Delta_r \operatorname{sg} C_r = C_1 l_1 + \cdots + C_r l_r \tag{418}$$

Беручи на увагу ще аналогічну рівність для тієї системи значень x_{10}, $x_{20}, \ldots, x_{n0}$, яка дає мінімальне наближення для нашої системи рівнянь (414), та позначаючи через Δ_{i0}, $i = 1, 2, \ldots, r$ відповідні відхили, через віднімання дістанемо:

$$\sum_{i=1}^{r} |C_i| (\Delta_i - \Delta_{i0}) \operatorname{sg} C_i = 0 \tag{419}$$

Якщо припустити супротивно до твердження леми, що ρ' є менше, ніж величини (417), то матимемо $\operatorname{sg}(\Delta_i - \Delta_{i0}) = \operatorname{sg} \Delta_i$, $i = 1, 2, 3, \ldots, r$ —

отже, на підставі умови (416), всі члени на лівій стороні (419) мали б однаковий знак, що, очевидно, не можливе. Таксамо бачимо, що ρ' не може й дорівнювати величині (417), якщо не маємо $|\Delta_1| = |\Delta_2| = \cdots = |\Delta_r|$.

Теорема. *Якщо із системи лінійних рівнянь* (406) (*при* $G = G^0$), *або з поширеної системи* (*при* $G \neq G^0$) *можна виділити підсистему* $r \leqslant n+1$ *рівнянь виду* (414) *з лінійно залежними у вузькому розумінні лівими сторонами, для якої, за певної системи значень невідомих* $x_1, x_2, \ldots, x_n$, *мають місце співвідношення* (416) *між знаками відхилів, то мінімальне наближення* ρ *розглядуваної системи* (406) *не може ніколи бути менше за величину* (417) — *і не може дорівнювати їй, крім випадку, коли* $|\Delta_1| = |\Delta_2| = \cdots = |\Delta_r|$. Теорема залишається в силі й тоді, коли точкова множина G є дискретна, тобто коли маємо систему скінченного числа суперечних лінійних рівнянь.

Доведення. Теорема ця є безпосередній наслідок попередньої леми: треба взяти лише до уваги, що мінімальне наближення ρ для всієї системи рівнянь (406), рівне мінімальному наближенню поширеної системи (при $G \neq G^0$), не може бути менше за мінімальне наближення ρ' виділеної підсистеми r рівнянь.

Увага. При $r = 1$ маємо, при узагальненому розумінні теореми, випадок, коли одно з рівнянь системи (406) або поширеної системи має на лівій стороні тотожний нуль.

5°. **Теорема.** *Якщо, за умов попередньої теореми, всі* r *відхилів абсолютною величиною дорівнюють максимальному відхилові* L (*див.* 410), *тоді* $\rho = L$, *і розглядувана система значень невідомих* $x_1, x_2, \ldots, x_n$, *дає шуканий розв'язок* (*або один із розв'язків*) *задачі мінімального наближення для системи рівнянь* (406).

Доведення. Справді, з одного боку, $\rho \geqslant L$ на підставі попередньої теореми, а з другого боку $\rho \leqslant L$, оскільки ρ є найменше можливе значення величини L. Отже $\rho = L$.

Увага попереднього n° є й тут застосувальна.

6°. Доведені в n°n° 5 та 3 [1]) цього параграфа теореми дають в своїй сукупності достатні і водночас необхідні умови для того, щоб дана система значень $x_1, x_2, \ldots, x_n$ давала мінімальне наближення для системи рівнянь (406). Теорема ж n° 4 дозволяє встановити нижню границю для мінімального наближення ρ кожного разу, коли знайдено будьяку систему значень $x_1, x_2, \ldots, x_n$, при якій мають місце умови, потрібні для застосування цієї теореми: з цього погляду вона відповідає Валле-Пусеновій теоремі щодо найкращого наближеного представлення неперервної функції $f(x)$ з допомогою полінома даного степеня.

Прикладання цих результатів до сформульованої в n° 1 цього параграфа задачі апроксимації функцій від довільного числа незалежних змінних є безпосереднє [2]).

[1]) Застосовуючи теорему n° 3 цього §, зважаємо на § 23, n° 5.

[2]) Пор. наслідок з теореми n° 3 цього §, а також три теореми § 16 даної праці.

§ 25. Поняття про застосування методу послідовних наближень. Наближені та емпіричні формули з двома — трьома — чотирма параметрами, лінійні щодо параметрів

Практичне здійснення методу послідовних наближень розділу VII в загальному випадку задач мінімального наближення або задачі найкращого наближеного представлення, які були сформульовані в §§ 23 та 24, залежить значною мірою від того, оскільки застосування основних теорем останніх двох параграфів може бути полегшене з допомогою засобів геометричних. Останнє у всякому разі можливе у випадку наближених або емпіричних формул з трьома, почасти з чотирма параметрами, лінійних щодо параметрів.

Нехай маємо спершу систему чотирьох рівнянь

$$\left.\begin{array}{c}\alpha\xi_i+\beta\eta_i+\gamma=l_i\\(i=1,\ 2,\ 3,\ 4)\end{array}\right\}\qquad(420)$$

з невідомими α, β γ. Якщо інтерпретувати $(\xi_i,\ \eta_i)$, $i=1,\ 2,\ 3,\ 4$, як чотири точки на площині, ліві сторони рівнянь будуть попарно лінійно незалежні, коли між зазначеними чотирма точками немає тотожних. Три рівняння матимуть ліві сторони лінійно залежні у вузькому розумінні (§ 23 n° 2, означ. II), якщо відповідні три точки, відмінні між собою, лежать на одній прямій. При цьому множник C_i, який відповідає середній точці, матиме, очевидно, протилежний знак до інших двох множників. У загальному ж випадку, коли ніякі три точки не лежать на одній прямій, всі чотири рівняння разом мають ліві сторони лінійно залежні у вузькому розумінні. Легко бачити, що всі чотири множники C_i не можуть мати однаковий знак. Виражаючи ці множники з допомогою детермінантів та беручи на увагу геометричне значення цих детермінантів, легко прийти до таких висновків. Взаємне розміщення розглядуваних чотирьох точок може являти собою лише два такі випадки: або одна із чотирьох точок міститься всередині трикутника трьох останніх точок, — тоді знак відповідного множника C_i є протилежний до спільного знака трьох інших множників; абож чотири точки допускають такий розподіл на дві пари точок, що відповідні два відрізки, які сполучають точки однієї і другої пари, перетинаються між собою, утворюючи дві діагоналі опуклого чотирикутника, — і в такому разі множники C_i, що відповідають одній парі точок, мають додатний знак, а інші два множники — від'ємний.

Синтезуючи всі розглянуті випадки, легко встановити таке загальне положення, яке має найближчий зв'язок із змістом леми § 24, n° 4: хоч би яке було число рівнянь (4, 3 або 2), що їхні ліві сторони є лінійно залежні у вузькому розумінні, дві групи точок, які відповідають множникам C_i з різними знаками, завжди так розміщені, що нема на площині такої прямої, яка б розділяла обидві групи точок [1]).

[1]) Пор. деякі уваги у P. Kirchberger-a, op. cit.

Аналогічно, коли маємо систему п'ятьох рівнянь

$$\left.\begin{array}{c}\alpha\xi_i+\beta\eta_i+\gamma\zeta_i+\delta=l_i\\ (i=1,\ 2,\ \ldots,\ 5)\end{array}\right\} \tag{421}$$

з невідомими α, β, γ, δ, інтерпретуючи $(\xi_i,\ \eta_i,\ \zeta_i)$, $i=1,\ 2,\ \ldots,\ 5$ як п'ять точок у просторі та залишаючи осторонь випадки, що сходять до вже розглянутих, коли якісь чотири точки знаходяться в одній площині [1], — ми маємо тут розглянути лише той загальний випадок, коли всі п'ять рівнянь разом мають ліві сторони лінійно залежні у вузькому розумінні. Тут, подібно до попереднього, виявляється, що можливі лише два такі випадки: або одна із п'ятьох точок міститься всередині тетраедра чотирьох інших точок, і тоді знак відповідного множника C_i є протилежний до спільного знака останніх множників; абож відрізок, що з'єднує дві якісь точки, перетинає площу трикутника трьох інших точок, і в такому разі множники, які відповідають двом групам точок, мають протилежний знак. Загальне положення тут буде таке: **дві групи точок, які відповідають множникам C_i з різними знаками, завжди так розміщені, що їх розділити площиною не можна.**

Ми тут зупинимось на задачі апроксимації обмеженої функції $f(x, y, \ldots, u)$ за допомогою лінійної комбінації трьох даних обмежених функцій

$$a\psi_1(x,\ y,\ \ldots,\ u)+b\psi_2(x,\ y,\ \ldots,\ u)+c\psi_3(x,\ y,\ \ldots,\ u)$$

на обмеженій точковій множині S простору $(x,\ y,\ \ldots,\ u)$. Ми тут обмежимося лише тим випадком, коли одна із функцій, наприклад $\psi_3(x, y, \ldots, u)$ залишається абсолютною величиною вище якоїсь додатної нижньої границі на точковій множині S. Щоб не ускладнювати міркувань, ми, крім того, зробимо такі припущення: 1°. Точкова множина S є замкнена, а функції f, ψ_1, ψ_2, ψ_3 однозначні та неперервні [2]); 2°. Коли позначити $\frac{\psi_1}{\psi_3}=\Psi_1$, $\frac{\psi_2}{\psi_3}=\Psi_2$, якщо $(x',\ y',\ \ldots,\ u')$ та $(x''\ y'',\ \ldots,\ u'')$ є дві **різні** точки множини S ніколи рівності

$$\Psi_1(x',\ y',\ \ldots,\ u')=\Psi_1(x'',\ y'',\ \ldots,\ u'') \text{ та } \Psi_2(x',\ y',\ \ldots,\ u')=\Psi_2(x'',\ y'',\ \ldots,\ u''),$$

не мають місця одночасно. При цих умовах, запровадивши нові змінні:

$$\frac{\psi_1}{\psi_3}=\xi,\ \frac{\psi_2}{\psi_3}=\eta,\ \frac{f}{\psi_3}=F(\xi,\ \eta), \tag{422}$$

ми поставимо собі задачу найкращого (за критерієм абс. пох.) наближе-

[1]) Ліві сторони відповідних чотирьох рівнянь будуть, очевидно, лінійно залежні між собою. Знаки множників C_i визначаються згідно з попередніми геометричними правилами.

[2]) Термін „неперервні" вживаємо тут (відповідно до позначень Baire — op. cit.) в тому розумінні, що коливання функції дорівнює нулеві в кожній точці множини S, яка може містити й точки ізольовані.

ного представлення однозначної та неперервної функції $F(\xi, \eta)$ за допомогою лінійної функції

$$\varphi(\xi, \eta; a, b, c) = a\xi + b\eta + c \tag{423}$$

на обмеженій та замкненій точковій множині E площі (ξ, η), що відповідає точковій множині S простору $(x, y, \ldots, u)$. Щоб уникнути в дальших міркуваннях будьякої неясності, зауважмо, що випадок, коли всі точки множини E лежать на одній прямій, ми тут можемо виключити з розгляду: інакше ми приходимо до простішої аналогічної задачі, яка стосується функції одного змінного.

Нехай буде φ будьяке наближене представлення (не найкраще) типу (423), яке задовольняє умову

$$\max_{(\xi, \eta) \subset E} |F - \varphi| = L(a, b, c) \leqslant M, \tag{424}$$

де

$$M = \max_{(\xi, \eta) \subset E} |F(\xi, \eta)| \tag{425}$$

Розуміючи під R будьяке додатне число, менше за $L = L(a, b, c,)$, та позначаючи $F - \varphi = \Delta$, розгляньмо дві замкнені точкові множини[1]) $G \subset E$ та $H \subset E$, на яких задовольняються відповідно нерівності

$$\Delta(\xi, \eta) \geqslant R, \quad \Delta(\xi, \eta) \leqslant -R \tag{426}$$

Можливі два випадки: або точкові множини G та H розділяються одна від однієї якоюсь прямою[2]); абож такої прямої нема, і тоді, як це легко встановити[3]), базуючись на основній теоремі § 24 n° 3, на точкових множинах G та H можна знайти такі дві групи точок, загальним числом 4 або 3, яких уже не можна розділити прямою. В останньому разі, беручи на увагу теорему § 24 n° 4, можна буде твердити, що

$$\rho \geqslant R, \tag{427}$$

позначаючи через ρ, як і досі, величину найкращого наближення. Ми й тут позначимо через A найбільше значення нижньої границі, яку можна встановити таким чином для найкращого наближення ρ на підставі дослідження різниці $F - \varphi = \Delta$, коли вираз φ є фіксований.

Кожного разу, коли візьмемо $R > A$ (залишаючи в силі умову $R < L$), точкові множини G та H можна розділити прямою. Поміж прямими, які розділяють точкові множини G та H, ми умовно позначатимемо як оптимальну (при даних E, F, φ та R) ту пряму, для якої мінімум d віддалі від замкненої точкової множини $G + H$ є щонайбільший. Не важко

[1]) Одна із цих двох множин може виявитись порожня. Легкі модифікації дальших міркувань, що відповідають таким випадкам, ми залишаємо осторонь.

[2]) Ці слова слід розуміти у вузькому сенсі: розглядувана пряма не повинна мати спільних точок із точковими множинами G та H.

[3]) Уводячи до розгляду функцію від змінних ξ та η, рівну, наприклад, $+1$ на точковій множині G та -1 на точковій множині H, і беручи на увагу, що за найкраще наближене представлення виду $a\xi + b\eta + c$ для такої функції на точковій множині $G + H$ правитиме, очевидно, тотожний нуль.

довести існування принаймні однієї „оптимальної“ прямої. Справді, нехай $d_{\max}$ позначає точну верхню границю [1]) мінімальної віддалі d при даних E, F, φ та R. А що в нормальному рівнянні ($x\cos\alpha + y\sin\alpha - p = 0$) прямої, яка розділяє точкові множини G та H, параметр p є, очевидно, число обмежене, то з кожної нескінченної послідовності цих прямих, для якої $\lim d = d_{\max}$, можна буде [2]) виділити таку нескінченну частинну послідовність, щоб водночас параметри p та α прямували до граничних якихсь значень $\bar{p}$ та $\bar{\alpha}$. Легко бачити, що визначена в такий спосіб гранична пряма є „оптимальна“.

Відзначімо деякі властивості „оптимальної“ (при даних E, F, φ та R) прямої, що її ми позначимо літерою U.

Поперше, вона править за вісь симетрії такої смуги між двома паралельними прямими U' та U'', що всередині її немає жодної точки множини $G+H$, а на її граничних прямих лежать принаймні дві точки [3]) цієї множини — по одній від G та від H відповідно: як би цього не було, мінімальну віддаль d можна було б, очевидно, збільшити паралельним переміщенням розглядуваної прямої U.

Подруге, вживаючи літери P для позначення точок множини G, які лежать на прямій U', та літери Q для позначення точок множини H, які лежать на прямій U'', і позначаючи літерами p та q ортогональні проекції на вісь U точок P та Q відповідно, — точкові множини $\{p\}$ та $\{q\}$ повинні мати таке розміщення на прямій U, щоб їх не можна було відокремити (у вузькому розумінні — пор. виноску вище) одну від однієї якоюсь точкою: якби така точка існувала, то нескінченно малим повертанням прямої U навколо цієї точки можна було б, як це легко зрозуміти, збільшити мінімум віддалі d. В тому разі, коли точкові множини $\{p\}$ та $\{q\}$ не мають жодної спільної точки, повинні існувати на цих точкових множинах дві групи точок, загальним числом три, що уже їх не можна відокремити точкою одну від однієї [4]): хоч одна із точок множини $\{p\}$ повинна знаходитись поміж двома точками множини $\{q\}$ або навпаки.

Тепер легко встановлюється і єдиність „оптимальної“ прямої (при даних E, F, φ та R). Справді, якщо точкові множини $\{p\}$ та $\{q\}$ мають спільну точку, яка є водночас проекцією для двох точок P_0 та Q_0, то при кожному іншому положенні прямої U, яка розділяє точки P_0 та Q_0, повинна зменшитись віддаль принаймні однієї із цих двох точок до прямої U, отже і мінімальна віддаль d від $G+H$ до U не може не зменшитись. Коли ж точкові множини $\{p\}$ та $\{q\}$ не мають жодної спільної точки, то, припустивши, що згадані три точки є P_1, Q_1, Q_2, ми прийдемо до аналогічного висновку, замінивши в останньому міркуванні точки P_0 та

[1] Ця верхня границя, очевидно, менш за половину максимальної віддалі двох точок множини E.

[2]) Пор. кінець § 22, n° 1.

[3]) Не слід забувати, що точкові множини G та H є замкнені.

[4]) Це твердження знову таки слід поставити в зв'язок з теоремою § 24 n° 3.

Q_0 на P_1 і ту точку відрізка $Q_1 Q_2$, яка є основою перпендикуляра, спущеного з точки P_1, при чому ми візьмемо до уваги, що віддаль від точки відрізка $Q_1 Q_2$ до прямої, якої він не перетинає, змінюється монотонно, коли точка перебігає відрізок $Q_1 Q_2$ в одному напрямку.

Нехай тепер, при даних — точковій множині E та функції F, лінійний вираз φ не є більше фіксований, а підпорядкований лише умові (424). Нехай ρ_1 означає максимальну величину найкращого наближення при заміні всієї точкової множини E будьякими трьома точками цієї множини[1]). Ми твердимо, що величина $d_{\max}$ залишатиметься вище якоїсь фіксованої додатної нижньої межі δ, якщо число R задовольняє умову

$$R > \rho_1 + \varepsilon, \tag{428}$$

позначаючи через ε будьяке фіксоване додатне число.

Справді в супротивному разі можна буде зазначити таку послідовність виразів

$$\varphi_1, \varphi_2, \varphi_3, \ldots, \varphi_i, \ldots \tag{429}$$

та відповідних чисел

$$R_1, R_2, R_3, \ldots, R_i, \ldots \tag{430}$$

підпорядкованих відповідно умовам (424) та (428), щоб водночас

$$\lim_{i \to \infty} (d_{\max})_i = 0 \tag{431}$$

Зважаючи на те, що функції $F - \varphi_i$ при даних умовах є рівно неперервні (egualmente continue) в своїй сукупності[2]) на точковій множині E, ясно насамперед, що віддаль між означеними вище двома точковими множинами $\{p\}$ та $\{q\}$ на „оптимальній" прямій U_i залишатиметься вище певної додатної границі при досить малих значеннях $\{d_{\max}\}$, тобто при досить великих значеннях індекса i в послідовностях (429) та (430). Отже, при досить великих значеннях індекса i ми матимемо кожного разу ансамбль трьох точок, наприклад, P_{1i}, Q_{1i}, Q_{2i}, що належать відповідно до точкових множин G_i та H_i і лежать на паралельних прямих U_i' та U_i'' відповідно, які обмежують нескінченно тонку смугу навколо „оптимальної" прямої U_i[3]). Неодноразово застосований в попередньому викладі тип міркування дозволяє встановити тепер існування принаймні одного граничного лінійного виразу

$$\varphi_0 = a_0 \xi + b_0 \eta + c_0 \tag{432}$$

[1]) Якщо ніякі три точки множини E не лежать на одній прямій, ми матимемо $\rho_1 = 0$.

[2]) Пор. відповідні міркування в §§ 18, 19, 22, а також виноску вище щодо розуміння терміну „неперервність" у випадку замкненої точкової множини, яка може містити й точки ізольовані.

[3]) Індексом i позначаємо тут скрізь залежність від лінійного виразу φ_i та числа R_i. Ми не маємо потреби окремо розглядати той випадок, коли кожному значенню індекса i відповідає ансамбль трьох точок P_{1i}, P_{2i}, Q_{1i}: для дальших висновків досить знати, що завжди можна виділити частинні послідовності (429), (430) так, щоб мати кожного разу одного й того самого типу ансамбль трьох точок, напр., P_{1i}, Q_{1i}, Q_{2i}, що ми й припускаємо для конкретності.

і відповідно до нього граничної прямої [1]) U_0, які мають таку властивість: різниця $F-\varphi_0$ набирає з альтернуючим знаком значень не менших за $\rho_1+\varepsilon$ абсолютною величиною в трьох послідовних точках прямої U_0, — які є відмінні між собою і належать до замкненої точкової множини E. Але це призводить до очевидної суперечності, коли зважити на прийняту вище умову щодо значення числа ρ_1.

Розуміючи далі під φ якесь вихідне наближене представлення типу (423) для функції F на точковій множині E, підпорядковане умові (424), задамо два числа R_1 та R_2 за умовами (298). Позначивши через G_1, H_1; G_2, H_2 відповідні точкові множини, нехай буде

$$\xi\cos\alpha+\eta\sin\alpha-p=0 \tag{433}$$

якась пряма, що розділяє точкові множини G_1 та H_1. Попереднє дослідження показує, що мінімальну віддаль d від точкової множини G_1+H_1 до прямої (433) можна припустити більшою за певну додатну величину δ, якщо різниця $R_1-\rho_1$ (де ρ_1 має з'ясоване вище значення) залишається вище якоїсь відповідної додатної границі. Віддаль від точкової множини $G_2+H_2\subset G_1+H_1$ до прямої (433) й поготів залишатиметься при тій самій умові більша, ніж δ. Позначаючи, з другого боку, через D максимальну віддаль двох точок на точковій множині E, легко бачити, що, додавши до виразу φ деяку поправку виду

$$\omega(\xi,\eta)=\lambda(\xi\cos\alpha+\eta\sin\alpha-p) \tag{434}$$

при

$$|\lambda|\geqslant\frac{(\vartheta_2-\vartheta_1)(L-A)}{D} \tag{435}$$

можна понизити величину наближення L щонайменше на величину, рівну меншій із двох таких:

$$\left.\begin{aligned}(1-\vartheta_2)(L-A)&>(1-\vartheta_2)(L-\rho),\\ \frac{\delta}{D}(\vartheta_2-\vartheta_1)(L-A)&>\frac{\delta}{D}(\vartheta_2-\vartheta_1)(L-\rho)\end{aligned}\right\} \tag{436}$$

Якщо покладемо означений тут алгорифм в основу нескінченного процесу послідовного наближення, добираючи числа ϑ_1 та ϑ_2 так, щоб різниці $1-\vartheta_2$, $\vartheta_2-\vartheta_1$ залишалися вище якоїсь додатної фіксованої границі, — міркування, цілком аналогічне до застосованого в § 22 n° 1, дозволяє довести, що різниця $L-\rho$ наближатиметься напевне до нуля при умові, що відношення $\frac{R_1-\rho_1}{L-\rho_1}$ залишатиметься також вище якоїсь фіксованої додатної границі при повторному застосуванні алгорифма, тобто коли

[1]) U_0 позначає не „оптимальну" пряму, а граничне положення „оптимальної" прямої.

брати кожного разу число R_1 не надто мале[1]). Виключаючи випадок $\rho = \rho_1$, можна буде при тій самій умові твердити, що $L - \rho$ наближатиметься до нуля принаймні як член спадної геометричної прогресії.

Другий алгорифм, означений в § 20, тут також допускає узагальнення — насамперед для того випадку, коли попереду будьяким способом уже одержано лінійний вираз φ, для якого

$$A > \rho_1, \tag{437}$$

де ρ_1 має з'ясоване вище значення.

В узагальненому алгорифмі за поправку $\Omega(\xi, \eta)$ береться найкраще наближене представлення функції $\Delta(\xi, \eta) = F - \varphi$ на ансамблі чотирьох точок області E, що з них ніякі три не лежать на одній прямій. При доборі чотирьох точок застосовуємо mutatis mutandis ті самі принципи, які було сформульовано в § 20 щодо добору ансамблю $n+2$ точок. Саму задачу щодо визначення поправки Ω розв'язуємо згідно з § 23 n° 5. Збіжність процесу послідовних наближень доводимо, базуючись на відповідно узагальнених лемі I та теоремах I, II параграфа 17.

При узагальненні леми I, замінюючи сегмент (a, b) на область квадрата чи то круга, замість „обмеженого числа інтервалів загальною довжиною $\leqslant \sigma$", ми тут матимемо смужку між двома скільки завгодно близькими паралельними прямими.

Узагальнюючи далі теорему I § 17 та замінюючи при цьому $k + \varepsilon$ на $\rho_1 + \varepsilon$, ми тут, при аналогічному (до певної міри) доводженні, прийдемо до якогось граничного ансамблю чотирьох точок $K_{\nu 0}(\xi_{\nu 0}, \eta_{\nu 0})$, $\nu = 1, 2, 3, 4$, і до граничних відхилів $\Delta^{(1)}$, $\Delta^{(2)}$, $\Delta^{(3)}$, $\Delta^{(4)}$, абсолютною величиною не менших ніж $\rho_1 + \varepsilon$. Відповідна до (260′) послідовність лінійних виразів нехай тут буде $\varphi_1, \varphi_2, \varphi_3, \ldots$, при чому

$$\max_{(\xi, \eta) \subset E} |\varphi_i(\xi, \eta)| = M_i, \quad \lim_{i \to \infty} M_i = \infty$$

Ансамбль чотирьох точок, що за його допомогою встановлюється число $A = A_i$ для виразу φ_i, нехай буде K_{1i}, K_{2i}, K_{3i}, K_{4i}. Різниця $F - \varphi_i$ нехай буде $\Delta_i(\xi, \eta)$. Якщо із точок $K_{\nu 0}$, $\nu = 1, 2, 3, 4$, ніякі три не лежать на одній прямій, то, як не важко пересвідчитись, при досить великому значенні індекса i досить було б скільки завгодно малої варіації координат щонайбільше двох точок[2]) з ансамблю чотирьох точок $K_{\nu i}$, $\nu = 1, \ldots, 4$, — варіації, яка не міняє типу розміщення чотирьох точок $K_{\nu i}$ — аби дістати

[1]) Як уже було зауважено, ми матимемо $\rho_1 = 0$, якщо ніякі три точки множини E не лежать на одній прямій. Такий є взагалі практично важливий випадок системи емпіричних рівнянь. Щождо інших випадків, коли $\rho_1 > 0$, оцінювати попереду величину ρ_1 доконечної потреби у всякому разі не становить.

[2]) Саме тих точок, які, можливо, попадають у нескінченно вузьку смугу, в якій $|\varphi_i| \leqslant M$, де M має значення (425). При цьому нас мало обходить, чи залишаться після розглядуваної варіації зміщені точки на точковій множині E.

такі чотири точки K'_{ν_i}, щоб знаки чотирьох величин $-\varphi_i(K'_{\nu_i}) = -\varphi_i(\xi'_{\nu_i}, \eta'_{\nu_i})$ збігалися з відповідними знаками чотирьох відхилів $\Delta_i(\xi_{\nu_i}, \eta_{\nu_i}) = \Delta_i(K_{\nu_i})$. Але це містить в собі очевидну суперечність, бо знаки відхилів $\Delta_i(K_{\nu_i})$ є якраз такі, які не можуть збігатися із знаками числових значень, що їх набирає будьякий лінійний вираз $a\xi + b\eta + c$ в тих самих чотирьох точках, або в нескінченно близьких чотирьох точках, які являють такий самий тип розміщення. В іншому можливому випадку, коли три якісь точки з числа чотирьох точок K_{ν_0} (або навіть усі чотири) лежать на одній прямій, завжди знайдеться якийсь лінійний вираз $\bar{\varphi}(\xi, \eta)$, відхили якого $F(\xi, \eta)$ в чотирьох точках K_{ν_0} абсолютною величиною не перевищують ρ_1. А в такому разі, при досить великому значенні індекса i, різниця $\bar{\varphi} - \varphi_i$ набиратиме в чотирьох точках K_{ν_i} числових значень з такими самими знаками, як Δ_i, а це знову приводить до аналогічної суперечності.

Щождо узагальнення теореми II § 17, то воно вже не становить ніякої трудності. Нарешті, встановляючи нерівність аналогічну до (316), ми візьмемо до уваги геометричне значення множників C_1, C_2, C_3, C_4 (задача § 23 п° 5). Тут відповідні площі трикутників залишатимуться більші за якусь додатну фіксовану величину, зважаючи на умову (437): справді, при супротивному припущенні знайомий тип міркування дозволяє тут встановити існування такого граничного лінійного виразу φ та відповідної граничної прямої, що різниця $F - \varphi$ набирала б числових значень, більших за ρ_1 абсолютною величиною, з альтернуючим знаком у трьох послідовних точках граничної прямої, а це призводить до очевидної суперечності.

Уваги параграфа 20 щодо можливого поєднання істотних моментів обох алгорифмів мають і тут силу. Метод залишається прикладальний, з деякими можливими модифікаціями, також і в таких випадках, де умова (437) не має місця.

Геометричний характер попередніх міркувань підказує застосування відповідних графічних засобів при реалізації розглядуваного методу послідовних наближень. В разі неперервної області E графічне представлення ліній рівня функції $\Delta(\xi, \eta)$ часто буде корисне.

Приклад I. Визначити найкраще наближене представлення функції $f(x, y) = x^2 - y^2$ з допомогою лінійного виразу $\Phi(x, y; \alpha, \beta, \gamma) = \alpha x + \beta y + \gamma$ в області трикутника $1 \geqslant x \geqslant y \geqslant 0$.

За вихідне наближене представлення беремо вираз $\varphi(x, y) = x - y$, який збігається з $f(x, y)$ у вершках даного трикутника. Різниця $\Delta(x, y)$ має дві екстремальні точки:

$$\Delta\left(\frac{1}{2}, 0\right) = -\frac{1}{4}, \quad \Delta\left(1, \frac{1}{2}\right) = +\frac{1}{4}$$

На підставі критерія, одержаного в § 24, цей вираз $\varphi(x, y)$ не являє собою найкращого наближення. Першу поправку $\omega_1(x, y)$ визначаємо так, щоб зрівняти абсолютною величиною від'ємні значення $\Delta(x, y)$ у точках

$\left(\frac{1}{2}, 0\right)$ та (1, 1) і додатні значення в точках $\left(1, \frac{1}{2}\right)$ та (0, 0). Ця поправка буде

$$\omega_I(x, y) = \frac{1}{6}(x + y - 1),$$

після чого маємо:

$$\varphi_I(x, y) = \varphi + \omega_I = \frac{7}{6}x - \frac{5}{6}y - \frac{1}{6} \tag{438}$$

Екстремальні значення для $\Delta_I = f - \varphi_I$:

$$\Delta_I\left(\frac{7}{12}, 0\right) = -\frac{25}{144}, \quad \Delta_I\left(1, \frac{5}{12}\right) = \frac{25}{144}, \quad \Delta_I(0, 0) = \frac{1}{6}, \quad \Delta_I(1, 1) = -\frac{1}{6}$$

Тут, очевидно, $L = \frac{25}{144}, A = \frac{1}{6}, \; L - \rho < L - A = \frac{1}{144}$.

Складаючи дальшу поправку ω_{II}, зрівнюємо абсолютною величиною значення Δ_{II} в точках $\left(\frac{7}{12}, 0\right)$, $\left(1, \frac{5}{12}\right)$, (0, 0), (1, 1):

$$\omega_{II}(x, y) = \frac{1}{204}(x + y - 1)$$

Вона дає нам:

$$\varphi_{II}(x, y) = \frac{239}{204}x - \frac{169}{204}y - \frac{35}{204} \tag{439}$$

$$\Delta_{II}(0, 0) = -\Delta_{II}(1, 1) = \frac{35}{204} \approx 0{,}171569$$

$$\Delta_{II}\left(\frac{239}{408}, 0\right) = -\Delta_{II}\left(1, \frac{169}{408}\right) = -\left(\frac{169}{408}\right)^2 = -0{,}171575$$

$$L = 0{,}171575; \; A = 0{,}171569; \; L - \rho < L - A = 0{,}000006$$

Симетричний характер послідовних поправок виявляє можливість знайти точний розв'язок задачі у виді виразу

$$\varphi(x, y; \lambda) = x - y + \lambda(x + y - 1) \tag{440}$$

Відповідні екстремальні значення різниці $\Delta = f - \varphi$ тут будуть:

$$\Delta(0, 0; \lambda) = -\Delta(1, 1; \lambda) = \lambda; \; \Delta\left(\frac{1+\lambda}{2}, 0; \lambda\right) = -\Delta\left(1, \frac{1-\lambda}{2}; \lambda\right) = -\left(\frac{1-\lambda}{2}\right)^2$$

Шуканий точний розв'язок дістаємо, взявши λ за умовою $\lambda = \left(\frac{1-\lambda}{2}\right)^2$, звідки

$$\lambda = 3 - \sqrt{8} \approx 0{,}171573 = \rho$$

Приклад II. Знайти найкраще лінійне наближене представлення функції $f(x, y) = e^{\frac{1}{2}(x+y)}$ в області трикутника $1 \geqslant x \geqslant y \geqslant 0$, узявши за вихідне наближене представлення вираз

$$\varphi(x, y) = 0{,}908 + 0{,}638x + 1{,}050y,$$

одержаний у § 14 n° 3, прикл. II.

Екстремальні значення різниці $\Delta = f - \varphi$ тут такі:

$$\Delta(0{,}524;\ 0{,}524) = -0{,}104;\ \Delta(0{,}487;\ 0) = 0{,}057;\ \Delta(1;\ 0{,}484) = 0{,}046;$$

$$\Delta(0{,}0) = 0{,}092;\ \Delta(1,\ 0) = 0{,}103;\ \Delta(1,\ 1) = 0{,}122$$

Скорегувавши $\varphi(x, y)$ з допомогою аддитивної сталої $\frac{0{,}122 - 0{,}104}{2} = 0{,}009$ (пор. § 21), матимемо поліпшений вираз

$$\varphi(x, y) = 0{,}917 + 0{,}638x + 1{,}050y, \tag{441}$$

для якого різниця $\Delta = f - \varphi$ матиме такі значення в екстремальних точках:

$$\Delta(0{,}524;\ 0{,}524) = -0{,}113;\ \Delta(0{,}487;\ 0) = 0{,}048;\ \Delta(1;\ 0{,}484) = 0{,}037;$$

$$\Delta(0, 0) = 0{,}083;\ \Delta(1, 0) = 0{,}094;\ \Delta(1, 1) = 0{,}113 \text{ (див. рис. 8).}$$

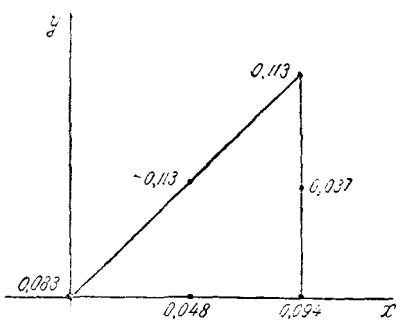

Рис. 8.

Звертаючи увагу на три екстремальні точки на гіпотенузі, маємо:

$$L = 0{,}113;\ A = 0{,}083$$

Визначаємо поправку $\Omega_1(x, y)$ з таких умов:

1) $\Omega_1(x, y)$ дає найкраще лінійне наближене представлення функції $\Delta(x, y)$ на ансамблі трьох екстремальних точок, що знаходяться на гіпотенузі трикутника;

2) $\Omega_1(1, 0) = 0$.

Така поправка буде [1]):

$$\Omega_1(x, y) = -0{,}023 + 0{,}023x + 0{,}007y \tag{442}$$

Це дає нам

$$\varphi_1(x, y) = 0{,}894 + 0{,}661x + 1{,}057y, \tag{443}$$

[1]) Поклавши $x = y = t$, знаходимо безпосередньо, за способом розділу I цієї роботи, відповідно пристосованим до випадку дискретної точкової множини, найкраще лінійне наближене представлення функції одного змінного $\Delta(t, t)$ на ансамблі трьох точок (0,0) (0,524; 0,524), (1,1) у виді $-0{,}023 + 0{,}03\ t$ або, повернувшись до змінних x, y, у виді

$$-0{,}023 + \mu x + (0{,}03 - \mu)\, y,$$

де $\mu = 0{,}023$ на підставі умови 2).

при чому різниця $\Delta_I = f - \varphi_I$ має такі екстремальні значення:

$$\Delta_I(0,\ 0) = 0{,}106;\ \Delta_I(0{,}541;\ 0{,}541) = -0{,}106;\ \Delta_I(1,\ 1) = 0{,}106$$

$$\Delta_I(1,\ 0) = 0{,}094;\ \Delta_I(0{,}558;\ 0) = 0{,}059;\quad \Delta_I(1;\ 0{,}497) = 0{,}034$$

Отже, з точністю до десятитисячних, маємо тепер [1])

$$L = A = \rho = 0{,}106$$

Приклад III [2]). Емпіричну функцію $y = f(x)$ задано 13 точками:

$$f(0) = 0{,}0;\quad f(5) = 25{,}1;\quad f(10) = 59{,}2;\quad f(15) = 92{,}9;$$
$$f(20) = 119{,}6;\ f(25) = 135{,}6;\ f(30) = 140{,}2;\ f(35) = 135{,}4;$$
$$f(40) = 124{,}6;\ f(45) = 111{,}4;\ f(50) = 99{,}2;\ f(55) = 90{,}2;$$
$$f(60) = 85{,}4$$

Треба визначити найкраще наближене представлення функції $f(x)$ на точковій множині $x = 0,\ 5,\ 10, \ldots,\ 60$ за допомогою виразу виду

$$\left.\begin{aligned}\varphi(x;\ a_1,\ a_2,\ a_3) = a_1 + a_2 e^{-0{,}0296x}\cos 5{,}36^\circ x + \\ + a_3 e^{-0{,}0296x}\sin 5{,}36^\circ x\end{aligned}\right\} \tag{444}$$

Позначивши

$$e^{-0{,}0296x}\cos 5{,}36^\circ x = \xi,\ e^{-0{,}0296x}\sin 5{,}36^\circ x = \eta, \tag{445}$$

ми можемо трактувати дану задачу як задачу найкращого наближеного представлення функції двох змінних $f(x) = F(\xi,\ \eta)$ з допомогою лінійного виразу

$$\varphi(x;\ a_1,\ a_2,\ a_3) = \Phi(\xi,\ \eta;\ a_1,\ a_2,\ a_3) = a_1 + a_2\xi + a_3\eta \tag{446}$$

на множині 13 точок, що лежать на логарифмічній спіралі (див. рис. 9). Відповідні значення ξ, η та $F(\xi,\ \eta)$ вміщено в другому, третьому та четвертому стовпцях таблиці (див. на с. 156).

З другого боку, ми можемо сформулювати дану задачу як задачу мінімального наближення для несумісної системи 13-тьох лінійних рівнянь, з яких останнє, наприклад, є

$$a_1 + 0{,}13269\,a_2 - 0{,}10517\,a_3 = 85{,}4 \tag{447_{13}}$$

Ми в дальшому користуватимемось паралельно одним і другим сформулюванням задачі, маючи на увазі їх повну еквівалентність.

Як вихідне наближене представлення функції $F(\xi,\ \eta)$ ми візьмемо вираз

$$\Phi(\xi, \eta) = 100{,}0 - 100{,}5\xi + 6{,}105\eta, \tag{448}$$

знайдений за способом найменших квадратів [3]).

[1]) Беручи знову таки на увагу, що перші три екстремальні точки лежать на одній прямій.

[2]) C. Runge und H. König. Numerisches, Rechnen. Berlin, 1924, сс. 246—249. Автори трактують цей приклад за способом найменших квадратів.

[3]) Пор. C. Runge und H. König, ibidem.

№ точки	ξ	η	$F(\xi, \eta)$	$\Delta(\xi, \eta)$	$\Delta_I(\xi, \eta)$
1	1,000 00	0,000 00	0,0	0,500	0,488
2	0,769 78	0,388 85	25,1	0,089	— 0,114
3	0,441 38	0,598 67	59,2	— 0,096	— 0,381
4	0,106 98	0,632 49	92,9	— 0,210	— 0,479
5	— 0,163 59	0,528 49	119,6	— 0,067	— 0,251
6	— 0,331 43	0,343 20	135,6	0,196	0,131
7	— 0,388 60	0,135 32	140,2	0,320	0,375
8	— 0,351 75	— 0,046 93	135,4	0,336	0,488
9	— 0,252 53	— 0,172 91	124,6	0,277	0,488
10	— 0,127 16	— 0 231 30	111,4	0,032	0,264
11	— 0,007 94	— 0,227 50	99,2	— 0,209	0,007
12	0,082 35	— 0,178 22	90,2	— 0,436	— 0,256
13	0,132 69	— 0,105 17	85,4	— 0,623	— 0,488

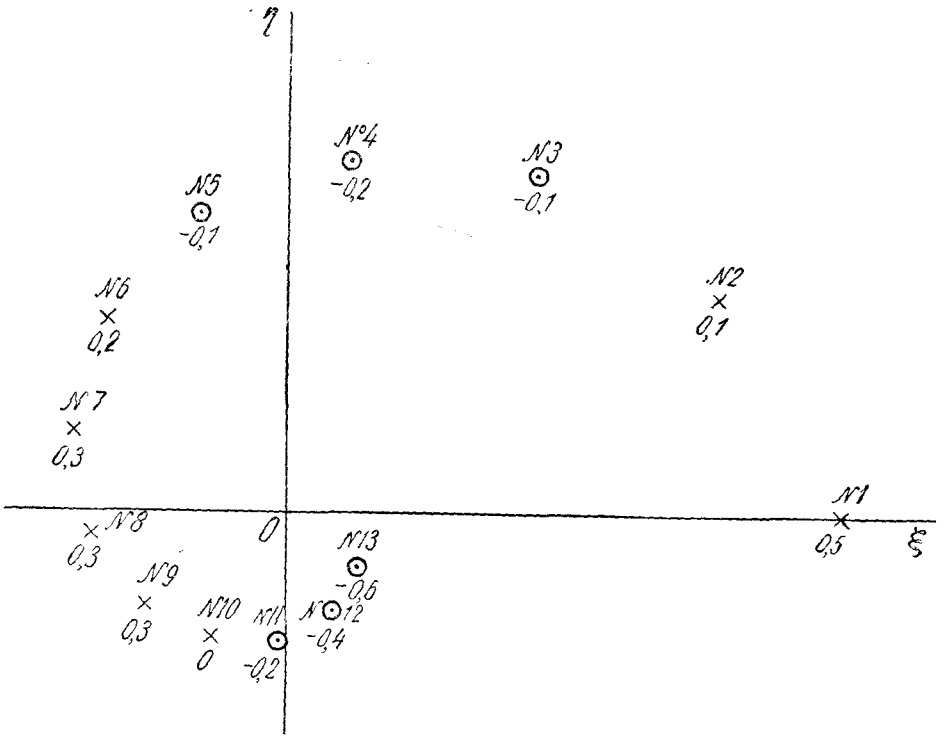

Рис. 9.

Значення різниці $F - \Phi = \Delta$ вміщене в 5-му стовпці таблиці. Їх зазначено, крім того, на рис. 9 біля відповідних точок (ξ, η).

Згідно з висновками § 23 n° 8, ми могли б безпосередньо добути точне розв'язання задачі, коли б нам вдалося дібрати ту комбінацію

чотирьох рівнянь із загального числа 13-тьох рівнянь, для якої мінімальне наближення має найбільше значення. При існуванні кількох таких комбінацій нам досить було б знайти одну з них — на підставі прямого розуміння результатів того самого § 23 n° 8, оскільки ніякі три з розглядуваних 13-тьох точок не лежать на одній прямій. Відповідно до основних принципів, на яких базується загальний метод послідовного наближення §§ 20, 22, 25, нам буде далеко зручніше шукати потрібну оптимальну комбінацію чотирьох рівнянь, визначаючи найкраще наближене представлення для різниці $\Delta = F - \Phi$ у виді лінійної поправки

$$\Omega_I(\xi, \eta) = \alpha_1 + \alpha_2 \xi + \alpha_3 \eta, \tag{449}$$

яку маємо додати до виразу (448). При цьому ми користуватимемось перетвореною системою 13-тьох рівнянь, де на лівих сторонах, замість коефіцієнтів a_1 a_2, a_3 виразу $\Phi(\xi, \eta; a_1, a_2, a_3)$ фігурують коефіцієнти α_1, α_2, α_3 шуканої поправки Ω_I, а на правих сторонах, замість значень $F(\xi, \eta)$, значення $\Delta(\xi, \eta)$. Напишімо для прикладу останнє з цих рівнянь:

$$\alpha_1 + 0{,}132\,69\alpha_2 - 0{,}105\,17\alpha_3 = -0{,}623 \tag{$447'_{13}$}$$

Користуючись формулою (377) та добираючи комбінації чотирьох точок за рисунком 9 так, щоб усі члени в чисельнику формули мали однаковий знак [1]), ми маємо п'ять комбінацій по чотири точки, вартих нашої уваги [2]): 1) 1, 9, 12, 13; 2) 1, 7, 9, 13; 3) 1, 6, 9, 13; 4) 1, 8, 9, 13; 5) 1, 4, 8, 13. Обчислюючи (з трьома-чотирма значущими цифрами) для кожної із цих п'ятьох комбінацій число ρ' за формулою (377) [3]), дістанемо для ρ такі значення: 1) 0,457; 2) 0,481; 3) 0,480; 4) 0,4885; 5) 0,4878 [4]).

Отже, можна думати, що комбінація рівнянь (або точок) №№ 1, 8, 9, 13 дає найбільше значення для ρ'. За поправку Ω_I ми й візьмемо найкраще лінійне наближене представлення функції $\Delta(\xi, \eta)$ на ансамблі чотирьох точок №№ 1, 8, 9, 13. Обчисливши повторно з більшою точністю множники

[1]) Це будуть, очевидно, якраз ті комбінації точок, які допускають застосування теореми § 24 n° 4. Одна з найістотніших переваг даного методу, який виходить із розшукання найкращого наближеного представлення для різниці $\Delta = F - \Phi$, замість аналогічного питання щодо самої функції F, полягає якраз, поруч мализни шуканих параметрів у перетвореній задачі, в порівняльній легкості відшукання згаданих комбінацій рівнянь, при чому перший добір „вартих уваги" комбінацій можна робити „на око" безпосередньо з рисунка.

[2]) Числа позначають №№ точок або відповідних лінійних рівнянь.

[3]) Множники C_1, C_2, C_3, C_4, для кожної комбінації чотирьох рівнянь визначаємо, розв'язуючи безпосередньо (без застосування детермінантів) відповідну систему трьох лінійних однорідних рівнянь — напр., для комбінації 4) маємо систему рівнянь:

$$\left.\begin{aligned} C_1 + \quad C_2 + \quad C_3 + \quad C_4 &= 0 \\ C_1 - 0{,}352\, C_2 - 0{,}253\, C_3 + 0{,}133\, C_4 &= 0 \\ -0{,}0469\, C_2 - 0{,}173\, C_3 - 0{,}105\, C_4 &= 0 \end{aligned}\right\} \tag{450}$$

Провадячи це обчислення, користуємося логарифмічною 25-сантиметровою лінійкою.

[4]) Цей попередній підрахунок провадимо, згідно з наміченим тільки но загальним планом розв'язання задачі, власне для того, щоб реалізувати викладений вище метод послідовних наближень при можливо меншому числі кроків.

$$C_1 : C_2 : C_3 : C_4 = -0{,}315\,91 : -0{,}105\,48 : -0{,}578\,61 : 1 \qquad (451)$$

та значення $\rho' = 0{,}4879$ для цієї комбінації лінійних рівнянь, маємо далі [1]) таку систему рівнянь, щоб визначити коефіцієнти $\alpha_1, \alpha_2, \alpha_3$ поправки Ω_l:

$$\left.\begin{aligned} \alpha_1 + \alpha_2 &= 0\,0121 \\ \alpha_1 - 0{,}35175\,\alpha_2 - 0{,}04693\,\alpha_3 &= -0{,}1523 \\ \alpha_1 - 0{,}25253\,\alpha_2 - 0{,}17291\,\alpha_3 &= -0{,}2116 \end{aligned}\right\} \qquad (452)$$

Звідкіля [2])

$$\alpha_1 = -0{,}0904;\ \alpha_2 = 0{,}1025;\ \alpha_3 = 0{,}5514$$

$$\Omega_l(\xi, \eta) = -0{,}0904 + 0{,}1025\,\xi + 0{,}5514\,\eta$$

$$\Phi_l(\xi, \eta) = 99{,}9096 - 100{,}3975\,\xi + 6{,}6564\,\eta \qquad (453)$$

Відповідні значення різниці $\Delta_l = F - \Phi_l$ уміщено в 6-му стовпці таблиці. Вони показують, що одержане наближене представлення є найкраще для системи тринадцятьох рівнянь в цілому, бо відхили для останніх дев'ятьох рівнянь є менші за ρ'. Отже матимемо остаточно:

$$\left.\begin{aligned} f(x) \approx 99{,}91 - 100{,}4\,e^{-0{,}0296x} \cos 5{,}36^\circ x \\ + 6{,}656\,e^{-0{,}0296x} \sin 5{,}36^\circ x \end{aligned}\right\} \qquad (454)$$

при максимальному відхилі $\rho = 0{,}49$.

Приклад IV. Визначити оптимальні, за критерієм мінімізації найбільшої відносної похибки, значення констант F та k в емпіричній лінійній формулі $D = F + kH$, якщо з вимірянь знайдено [3]) десять пар відповідних значень величин H та D, що їх вміщено у стовпцях 2-му та 3-му таблиці:

№	H	D	$R = \frac{F + kH}{D}$
1	0,200	20	1,000 00
2	0,400	40	0,989 40
3	0,600	60	0,985 86
4	0,800	80	0 984 09
5	1,008	100	0,990 86
6	1,218	120	0,997 01
7	1,426	140	1,000 00
8	1,628	160	0,998 57
9	1,830	180	0,997 46
10	2,036	200	0,998 53

[1]) Згідно з (376).

[2]) Тут взято деякі цифри „в запас", маючи на увазі демонстративне призначення поданих тут викладок.

[3]) E. Goedseels... Etalonnage des lunettes stadimetriques (Ann. de la Soc. Sc. de Bruxelles, 1910–1911, p. 1—16).

Інтерпретуючи (H, D) як точку на площині, при графічному зображенні, а то й просто при уважному розгляді таблиці — не важко помітити, що пряма, яка переходить через точки №№ 1 та 7, залишає по одну сторону від себе всі інші вісім точок; до того ж точки №№ 8, 9, 10 лежать досить близько до цієї прямої. Це безпосередньо підказує наступний шлях розв'язання, виходячи з відповідно узагальнених (для дискретної області) положень, розвинутих у розділі I цієї роботи: знайшовши рівняння згаданої прямої у виді

$$D = 0{,}4242 + 97{,}879H, \tag{455}$$

визначаємо для кожної точки відношення [1])

$$R = \frac{0{,}4242 + 97{,}879H}{D}. \tag{456}$$

4-ий стовпець таблиці показує, що $R \leqslant 1$, $R_{\min} = 0{,}98409$ для точки № 4, яка лежить між точками № 1 та № 7. Звідси, відповідно до (32) — (34) § 2 [2]), множачи праву сторону (455) на множник $\sigma = \frac{2}{1 + R_{\min}} = 1{,}00802$, дістаємо (єдину) шукану формулу найкращого наближення у виді

$$D = 0{,}4276 + 98{,}664H, \tag{457}$$

при чому максимальний відносний відхил є $0{,}00802 \approx 8^0/_{00}$ [3]).

[1]) Пор. (3) в § 1.

[2]) Зважаємо також на увагу наприкінці § 1 щодо продовження області, для якої визначено формулу найкращого наближення.

[3]) У Goedseels-а (op. cit.) розв'язання цієї задачі займає понад вісім сторінок. Його спосіб полягає в розв'язанні системи нерівностей, кількість яких дуже значна, користуючись деякими спеціальними правилами, що мають полегшити розв'язання такої системи.

ЗМІСТ

ЧАСТИНА ДРУГА

Методи більш загального характеру

Розділ VI

Розділ VII

Розділ VIII

Уповнов. Київськ. Облліту № 88. Вид. № 39. Зам. № 375. Тир. 1000. Ф. пап. 72×108 *см*. Вага 50,5 *кг*. Папер. арк. $5^1/_{16}$. Друк. зн. 112 тис. Здано до друкарні 14/IV 1935 р. Підписано до друку 22/VI 1935 р.