MIHAI TURINICI

MATEMATICI SPECIALE APLICATE ÎN ECONOMIE

Partea II: Programare neliniară şi dinamică

Casa de Editură VENUS
Iaşi 1999

Cuprins

Introducere

Lucrarea de faţă reprezintă cea de-a doua parte a cursului (cu acelaşi titlu) pe care autorul îl predă, de mai mulţi ani, la secţia *Economia şi Gestiunea Produselor Agroalimentare* de la Facultatea de Ştiinţe Economice a Universităţii "Al.I.Cuza" Iaşi. Materialul prezentat este împărţit în trei capitole (numerotate în continuarea celorlalte trei ale primei părţi). Capitolul al patrulea este consacrat Analizei. Se tratează astfel probleme legate de şiruri şi serii, limite şi continuitate, derivate parţiale, formule de medie, serii de puteri,integrale,funcţii speciale, şi alte aspecte conexe. Capitolul al cincilea are ca domenii de preocupări funcţiile convexe/concave, extremele libere, funcţiile implicite şi extremele condiţionate. In fine, capitolul şase tratează unele chestiuni de teoria sistemelor dinamice: şiruri recurente, programare dinamică, ecuaţii diferenţiale şi gestiunea stocurilor.

Concepţia generală a lucrării de faţă este aceeaşi ca şi la prima parte. Astfel, în funcţie de importanţa temei, unele demonstraţii au fost prezentate complet, iar altele doar schiţat; dar, în orice caz, toate acestea au fost ilustrate cu exemple numerice concrete – şi, la nevoie, cu interpretări economice – care să ajute la o mai bună înţelegere a temei tratate. Tot în acest scop au fost redactate şi exerciţiile de la sfârşitul fiecărui paragraf. Relativ la limbajul şi notaţiile folosite, s-a urmărit ca acestea să fie cât mai apropiate de cele standard. In fine, din bogata bibliografie existentă, s-au ales doar titlurile cele mai accesibile, pentru un cititor de specialitate economică.

Autorul

Capitolul 4

Complemente de analiză

4.1 Structuri de convergenţă pe spaţii finit dimensionale

4.1.1 Şiruri numerice convergente/divergente

Fie x un număr real arbitrar. Numim *vecinătate* a acestui număr

(D1) orice interval deschis care conţine punctul x,

(D2) orice submulţime care include un interval de forma descrisă.

Familia tuturor vecinătăţilor numărului x se va nota $\mathcal{V}(x)$. De exemplu, dacă $x = 2$, mulţimea $A = (-1, 5)$ este o vecinătate a acestui punct. De asemenea, $B = [0, \infty)$ este tot o vecinătate a punctului considerat; deoarece, de pildă,

$$B \supseteq (0, 3) = \text{ interval deschis care conţine punctul 2.}$$

În particular, toate intervalele centrate în x de forma

$$S(x, \varepsilon) = (x - \varepsilon, x + \varepsilon), \quad \varepsilon > 0$$

sunt vecinătăţi ale lui x. Menţionăm cu această ocazie proprietatea:

$$\text{(P1)} \begin{cases} \text{oricare ar fi } V \in \mathcal{V}(x) \text{ există } \varepsilon > 0 \text{ cu } V \supseteq S(x, \varepsilon) \\ \text{(orice vecinătate include o asemenea vecinătate simetrică).} \end{cases}$$

Fie (x_n) un şir de numere reale. Zicem că acesta *converge* către limita x (şi scriem $x_n \longrightarrow x$ sau $\lim\limits_{n \to \infty} x_n = x$) dacă

$$\text{(D3)} \left\{ \begin{array}{l} \text{pentru orice vecinătate } V \in \mathcal{V}(x) \text{ există un rang } n(V) \\ \text{aşa ca } x_n \in V, \text{ oricare ar fi rangul } n \geq n(V). \end{array} \right.$$

Adică: înafara oricărei vecinătăţi se află doar un număr finit de termeni ai şirului. O formă echivalentă de scriere a definiţiei este prin utilizarea vecinătăţilor simetrice. Anume, zicem că (x_n) *converge către limita* x dacă pentru orice $\varepsilon > 0$ există un rang $n(\varepsilon)$ şa ca

$$n \geq n(\varepsilon) \Longrightarrow x_n \in (x - \varepsilon, x + \varepsilon) \text{ (adică } |x_n - x| < \varepsilon).$$

Prin definiţie şirul (x_n) se va numi *convergent* dacă are o limită. (Să notăm că limita respectivă este neapărat unică.) În caz contrar, şirul (x_n) se va zice *divergent*. Cu privire la acest aspect, vom mai face o completare utilă. Introducem două elemente,$(+\infty)$ şi $(-\infty)$, exterioare mulţimii R, prin intermediul vecinătăţilor acestora.Mai precis, să numim vecinătate a lui $(+\infty)$

(i) orice interval de forma $(a, +\infty)$ cu $a \in R$,

(ii) orice submulţime a lui R ce include un asemenea interval.

Mulţimea tuturor vecinătăţilor lui $(+\infty)$ se va nota $\mathcal{V}(+\infty)$. Analog, se introduc vecinătăţile lui $(-\infty)$; mulţimea lor o vom nota corespunzător cu $\mathcal{V}(-\infty)$. Vom zice acum că şirul (x_n) are ca *limită* $(+\infty)$ (şi scriem $x_n \longrightarrow +\infty$ sau $\lim\limits_{n\to\infty} x_n = +\infty$) dacă condiţia (D3) are loc cu $x = +\infty$; analog se prezintă cazul limitei $(-\infty)$. Numim atunci *şir generalizat convergent* orice şir care este fie convergent (propriu), fie divergent dar cu limita unul din simbolurile $(+\infty)$ sau $(-\infty)$.

Zicem că şirul (x_n) este *de tip Cauchy* (sau *fundamental*) dacă

$$\text{(D4)} \left\{ \begin{array}{l} \text{oricare ar fi } \varepsilon > 0 \text{ există un rang } n(\varepsilon) \text{ aşa încât} \\ \text{oricare ar fi } n \geq n(\varepsilon) \text{ şi } p \geq 1 \text{ să avem } |x_{n+p} - x_n| < \varepsilon. \end{array} \right.$$

Adică: de la un rang suficient de mare încolo, distanţa dintre doi termeni ai şirului devine oricât de mică. Un rezultat important este aici

Teorema 4.1.1 *Un şir de numere reale este convergent dacă şi numai dacă are proprietatea Cauchy* (*sau, altfel spus, este fundamental*).

Nu dăm demonstraţia acestui fapt; ne mulţumim doar să semnalăm că un mod echivalent de scriere a concluziei este

(P2) mulţimea numerelor reale R este completă.

Fie (x_n) un şir de numere reale. Dat şirul de ranguri

$$k_1 < k_2 < \cdots < k_n < \cdots \quad (\text{deci} \quad k_n \longrightarrow \infty)$$

putem construi un nou şir de numere reale (y_n) prin convenţia

$$y_1 = x_{k_1}, y_2 = x_{k_2}, ..., y_n = x_{k_n}, ... \, .$$

Numim acesta un *subşir* al şirului iniţial (x_n). Avem acum

(P3) $x_n \longrightarrow x \Longrightarrow y_n \longrightarrow x$ pentru orice subşir (y_n) al lui (x_n).

Deci, în particular, orice subşir al unui şir generalizat convergent este şi el generalizat convergent, cu aceeaşi limită. Reciproca acestei afirmaţii este, de asemenea, valabilă. Deci, un şir cu proprietatea că două subşiruri ale sale converg la limite distincte este în mod necesar divergent.

Exemplu. Fie dat şirul de numere reale

$$x_n = (-1)^n + \frac{n}{n+1}, \quad n = 1, 2, ...$$

Să analizăm subşirul termenilor pari şi, respectiv, impari

$$\begin{array}{l} x_{2n} = 1 + \frac{2n}{2n+1}; \quad \text{deci} \lim\limits_{n\to\infty} x_{2n} = 1 + 1 = 2, \\ x_{2n+1} = -1 + \frac{2n+1}{2n+2}; \quad \text{deci} \lim\limits_{n\to\infty} x_{2n+1} = -1 + 1 = 0. \end{array}$$

Cum limitele obţinute sunt distincte, şirul (x_n) diverge.

Calculele cu şirurile generalizat convergente sunt cuprinse în următorul rezultat (dat fără demonstraţie).

Teorema 4.1.2 *Fie (x_n) şi (y_n) două şiruri generalizat convergente şi fie $x = \lim\limits_{n\to\infty} x_n$, $y = \lim\limits_{n\to\infty} y_n$. Atunci,*

(P4) $\lim\limits_{n\to\infty}(x_n + y_n) = x + y$, *cu excepţia cazului* $\infty - \infty$

(P5) $\lim\limits_{n\to\infty}(x_n y_n) = xy$, *cu excepţia cazului* $0 \cdot \infty$

(P6) $\lim\limits_{n\to\infty}\left(\dfrac{x_n}{y_n}\right) = \dfrac{x}{y}$, *cu excepţia cazurilor* $\dfrac{0}{0}, \ \dfrac{\infty}{\infty}$

(P7) $\lim\limits_{n\to\infty}(x_n)^{y_n} = x^y$, *cu excepţia cazurilor* $\infty^\circ, 0^\circ, 1^\infty$.

De asemenea, mai avem valabilă implicaţia

(P8) $x_n \le y_n,\ \forall n \Longrightarrow x \le y$.

O clasă importantă de şiruri generalizat convergente este dată de

Teorema 4.1.3 *Orice şir monoton este generalizat convergent. Mai exact, un astfel de şir este*

(a) *convergent (propriu) în caz de mărginire,*

(b) *cu limita* $(+\infty)$ *(respectiv,* $(-\infty)$*) în caz de nemărginire superioară (respectiv, inferioară).*

Un exemplu standard de acest tip este furnizat de şirul

$$x_n = \left(1 + \frac{1}{n}\right)^n,\ n = 1, 2, ...$$

Se arată relativ simplu că (x_n) este monoton crescător şi mărginit. Deci, obligatoriu, (x_n) converge. Notăm

(D5) $\lim\limits_{n\to\infty}\left(1 + \dfrac{1}{n}\right)^n = e$ (numărul lui Euler).

Numărul e apare ca iraţional, cu valoarea cuprinsă în intervalul $(2, 3)$; mai exact, $e \approx 2,71$. Se poate demonstra implicaţia (utilă în practică)

$$\lim_{n\to\infty} u_n = 0 \Longrightarrow \lim_{n\to\infty}(1 + u_n)^{1/u_n} = e.$$

Mai adăugăm, la cele de mai sus, următoarele strategii de lucru pentru stabilirea convergenţei şi limitei unui şir:

(A) Principiul încadrării: dacă (x_n) şi (y_n) sunt (generalizat) convergente cu aceeaşi limită ℓ, atunci orice şir (z_n) cu $x_n \le z_n \le y_n$, $\forall n$, este de asemenea (generalizat) convergent cu $\lim\limits_{n\to\infty} z_n = \ell$.

În particular, de aici rezultă implicaţia

(P9) $x_n \longrightarrow 0,\ (y_n) = \text{mărginit} \Longrightarrow x_n y_n \longrightarrow 0$.

Într-adevăr, nu avem decât să ţinem cont de

$$-\mu|x_n| \leq x_n y_n \leq \mu|x_n|, \ n = 1, 2, ...,$$

unde μ este o barieră superioară a şirului $(|y_n|)$.

(B) Principiul lui Stolz: Dacă (y_n) este un şir strict crescător cu limita $(+\infty)$, atunci (dat încă un şir (x_n))

$$\lim_{n\to\infty} \frac{x_n}{y_n} = \lim_{n\to\infty} \frac{x_{n+1} \to x_n}{y_{n+1} - y_n}, \text{ dacă limita dreaptă există.}$$

În particular, de aici rezultă implicaţia

(P10) $\displaystyle\lim_{n\to\infty} \frac{x_1 + \cdots + x_n}{n} = \lim_{n\to\infty} x_n$, dacă limita dreaptă există

şi, de asemenea, ((x_n) fiind un şir strict pozitiv)

(P11) $\displaystyle\lim_{n\to\infty} \sqrt[n]{x_n} = \lim_{n\to\infty} \frac{x_{n+1}}{x_n}$, dacă limita dreaptă există.

Pentru verificare, se scrie $\sqrt[n]{x_n} = e^{\frac{1}{n}\ln x_n}$ şi se aplică (B). Un exemplu de aplicare a acestei reguli este

$$\lim_{n\to\infty} \sqrt[n]{n} = \lim_{n\to\infty} \frac{n+1}{n} = 1.$$

(C) Principiul lui Töplitz: Fie dată matricea triunghiulară infinită

$$\begin{bmatrix} t_{11} \\ t_{21}, t_{22} \\ \dots\dots\dots\dots \\ t_{n1}, t_{n2}, ..., t_{nn} \\ \dots\dots\dots\dots \end{bmatrix}$$

ale cărei elemente satisfac la condiţiile

(C1) $t_{nm} \longrightarrow 0$ când $n \longrightarrow \infty$, pentru orice m
(şirurile formate pe coloane au limita zero)

(C2) $|t_{n1}| + \cdots + |t_{nn}| \leq \mu$, pentru toţi n
(sumele modulelor elementelor de pe linii sunt mărginite)

(C3) $t_{n1} + \cdots + t_{nn} \longrightarrow 1$ pentru $n \longrightarrow \infty$.

Atunci (pentru şirul (x_n) dat)

$$\begin{cases} \lim\limits_{n\to\infty}(t_{n1}x_1 + \cdots + t_{nm}x_n) = \lim\limits_{n\to\infty} x_n, \\ \text{dacă limita din dreapta există şi este finită.} \end{cases}$$

În particular, de aici rezultă implicaţia

$$\text{(P12)} \quad \begin{cases} x_n \to 0,\ y_n \to 0,\ |y_1| + \cdots + |y_n| \le \mu,\ \forall n, \\ \Longrightarrow \lim\limits_{n\to\infty}(x_1y_n + \cdots + x_ny_1) = 0. \end{cases}$$

Pentru aceasta va fi suficient să punem

$$t_{nm} = y_{n-m+1},\ n = 1, 2, ...,\ 1 \le m \le n.$$

4.1.2 Serii numerice. Concepte şi rezultate de bază

Fie (u_n) un şir de numere reale. Numim *serie numerică* de *termen general* u_n, simbolul

$$\text{(D1)} \quad u_1 + u_2 + \cdots + u_n + \cdots \ \text{(notat şi prin } \sum_{n=1}^{\infty} u_n \text{ sau } \sum_n u_n).$$

Semnificaţia acestuia este legată de conceptele anterior introduse pentru şiruri. Mai exact, să construim şirul de semne parţiale

$$s_1 = u_1, s_2 = u_1 + u_2, ..., s_n = u_1 + \cdots + u_n, ...$$

Dacă (s_n) este convergent,atunci prin definiţie spunem că seria $\sum_n u_n$ *converge*; iar $s = \lim_{n\to\infty} s_n$ se va numi *suma seriei* şi scriem $s = \sum_{n=1}^{\infty} u_n$. (Aceasta înseamnă că mulţimea termenilor şirului (u_n) poate fi adunată). Dacă (s_n) este divergent,atunci prin definiţie seria $\sum_n u_n$ *diverge*; adică,nu putem aduna mulţimea de termeni ai şirului (u_n).

Din aceste consideraţii rezultă următoarele probleme de bază care se pun în studiul seriilor

(a) stabilirea naturii seriei (convergenţă sau divergenţă),

(b) stabilirea sumei seriei (în caz de convergenţă).

În ce priveşte prima problemă, avem următorul rezultat de principiu (numit "Criteriul de convergenţă Cauchy").

Teorema 4.1.4 *Seria* $\sum\limits_n u_n$ *este convergentă dacă şi numai dacă pentru orice* $\varepsilon > 0$ *există un rang* $n(\varepsilon)$ *cu proprietatea*

(PC) $n \geq n(\varepsilon),\ p \in N \Longrightarrow |u_{n+1} + \cdots + u_{n+p}| < \varepsilon.$

Desigur, aceasta nu înseamnă nimic altceva decât caracterizarea de tip Cauchy a şirului sumelor parţiale (s_n) (vezi Teorema 4.1.1). În particular, pentru $p = 1$ deducem implicaţia

(P1) $\sum\limits_n u_n = \text{convergentă} \Longrightarrow \lim\limits_{n\to\infty} u_n = 0.$

(Termenul general al unei serii convergente tinde la zero.)

Criteriul de convergenţă stabilit anterior este în general, greu aplicabil direct; dar, metodologic, el are o importanţă deosebită în stabilirea unor criterii de convergenţă cu valoare practică mai mare. Vom exemplifica aceasta mai întâi pe seriile cu termeni pozitivi; iar apoi vom trece la cazul general.

Relativ la cea de-a doua problemă, se cunosc puţine clase de serii la care se poate calcula suma. Una dintre acestea este cuprinsă în următorul

Exemplu. Să se studieze seria (geometrică)

$$\sum_{n=0}^{\infty} q^n = 1 + q + q^2 + \cdots + q^n + \cdots$$

Aici q este un număr real fixat. Dăm deoparte cazul $q = 1$ când

$$s_n = n + 1,\ \forall n \text{ (şi deci seria diverge).}$$

Avem atunci (pentru $q \neq 1$)

$$s_0 = 1, s_1 = 1 + q = \frac{1 - q^2}{1 - q}, ..., s_n = 1 + q + \cdots + q^n = \frac{1 - q^{n+1}}{1 - q}, ...$$

Din studiul convergenţei şirului (q^n) rezultă discuţia:

(i) dacă $|q| < 1$, şirul (s_n) converge, cu $\lim\limits_{n\to\infty} s_n = \frac{1}{1-q}$; şi, deci, seria geometrică converge, cu $\sum\limits_{n=0}^{\infty} q^n = \frac{1}{1-q}$.

(ii) dacă $|q| \geq 1$, şirul (s_n) este divergent; ca atare, seria geometrică este şi ea divergentă.

Vom mai avea ocazia să indicăm şi alte situaţii de acest fel; dar, după cum s-a spus deja, ele sunt relativ restrânse.

(A) Să ne ocupăm în continuare de seriile $\sum_n u_n$ cu

(C1) $u_n \geq 0, \quad$ pentru $n = 1, 2, \ldots$.

Începem prin a observa că şirul (s_n) al sumelor parţiale este monoton crescător. Deci, o asemenea serie este

– convergentă dacă (s_n) este mărginit,

– divergentă dacă (s_n) este nemărginit.

Această simplă caracterizare a naturii seriilor cu termeni pozitivi este de efect în unele cazuri concrete, după cum arată şi următorul

Exemplu. Să studiem seria (armonică)

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} = 1 + \frac{1}{2^\alpha} + \cdots + \frac{1}{n^\alpha} + \cdots$$

Aici, α este un număr real arbitrar fixat. Pentru analiza acesteia, vom porni de la anumite subşiruri ale şirului sumelor parţiale

$$s_1 = 1, s_2 = 1 + \frac{1}{2^\alpha}, \ldots, s_n = 1 + \frac{1}{2^\alpha} + \cdots + \frac{1}{n^\alpha}, \ldots$$

Să admitem că $\alpha > 1$. Notăm $q = \frac{1}{2^{\alpha-1}}$; deci $0 < q < 1$. Avem

$$\frac{1}{2^\alpha} + \frac{1}{3^\alpha} \leq \frac{2}{2^\alpha} = \frac{1}{2^{\alpha-1}} = q$$

$$\frac{1}{4^\alpha} + \frac{1}{5^\alpha} + \frac{1}{6^\alpha} + \frac{1}{7^\alpha} \leq \frac{4}{4^\alpha} = \left(\frac{1}{2^{\alpha-1}}\right)^2 = q^2$$

....

$$\frac{1}{(2^{n-1})^\alpha} + \frac{1}{(2^{n-1}+1)^\alpha} + \cdots + \frac{1}{(2^n-1)^\alpha} \leq \frac{2^{n-1}}{(2^{n-1})^\alpha} = q^{n-1}.$$

De aici deducem

$$s_{(2^n-1)} \leq 1 + q + \cdots + q^{n-1} = \frac{1-q^n}{1-q} \leq \frac{1}{1-q}, \; n = 1, 2, \ldots .$$

Ca atare, şirul sumelor parţiale al seriei este mărginit; şi, astfel, seria în chestiune converge. Să luăm în discuţie cazul $\alpha \leq 1$. Avem

$$\begin{array}{l} \frac{1}{3^\alpha} + \frac{1}{4^\alpha} \geq \frac{2}{4^\alpha} \geq \frac{2}{4} = \frac{1}{2} \\[2ex] \frac{1}{5^\alpha} + \frac{1}{6^\alpha} + \frac{1}{7^\alpha} + \frac{1}{8^\alpha} \geq \frac{4}{8^\alpha} \geq \frac{4}{8} = \frac{1}{2} \\ \dotfill \\ \frac{1}{(2^{n-1}+1)^\alpha} + \cdots + \frac{1}{(2^n)^\alpha} \geq \frac{2^{n-1}}{(2^n)^\alpha} \geq \frac{2^{n-1}}{2^n} = \frac{1}{2}. \end{array}$$

Se obţine de aici

$$s_{(2^n)} \geq 1 + \overbrace{\frac{1}{2} + \cdots + \frac{1}{2}}^{(n \text{ ori})} = \frac{n+2}{2}, \ n = 1, 2, \dots .$$

Cu alte cuvinte, şirul sumelor parţiale ale seriei este nemărginit; ceea ce înseamnă că seria în chestiune este divergentă.

Concluziile obţinute se dovedesc utile pentru criterii de convergenţă *relative* (la care, cunoaşterea naturii unei serii se obţine prin *comparaţie* cu o altă serie, cunoscută). Un asemenea criteriu este conţinut în

Teorema 4.1.5 *Fie* $\sum_n u_n$ *şi* $\sum_n v_n$ *două serii cu termeni pozitivi, pentru care (cu* $\alpha, \beta > 0$*)*

(C2) *există un rang* k*, aşa încât* $\alpha u_n \leq \beta v_n, \ \forall n \geq k$.

Au loc atunci implicaţiile

(i) $\sum_n v_n$ *convergentă* $\Longrightarrow \sum_n u_n$ *convergentă*

(ii) $\sum_n u_n$ *divergentă* $\Longrightarrow \sum_n v_n$ *divergentă.*

În particular, condiţia (C2) *are loc dacă*

(C3) *există un rang* k*, aşa încât* $\dfrac{u_{n+1}}{u_n} \leq \dfrac{v_{n+1}}{v_n}, \ \forall n \geq k$.

Demonstraţie. Partea întâia este clară, din compararea şirurilor sumelor parţiale asociate. Pentru partea a doua, nu avem decât să observăm că, prin înmulţiri succesive ale relaţiilor din (C2), rezultă

$$\frac{u_n}{u_k} \leq \frac{v_n}{v_k}, \quad \text{pentru orice} \ \ n \geq k;$$

iar,de aici, concluzia este clară. ∎

Să notăm că de aici mai avem şi implicaţia

(P2) $0 < \lim\limits_{n\to\infty} \dfrac{u_n}{v_n} < \infty \Longrightarrow \sum\limits_n u_n$ şi $\sum\limits_n v_n$ au aceeaşi natură.

Exemplu. Să se studieze natura seriei $\sum\limits_n \dfrac{1}{n\sqrt{n+1}}$. Vom compara aceasta cu seria $\sum\limits_n \dfrac{1}{n\sqrt{n}} = \sum\limits_n \dfrac{1}{n^{3/2}}$ care este convergentă ($\alpha = \frac{3}{2} > 1$). Facem limita raportului termenilor generali

$$\lim_{n\to\infty} \frac{u_n}{v_n} = \lim_{n\to\infty} \frac{\frac{1}{n\sqrt{n+1}}}{\frac{1}{n\sqrt{n}}} = \lim_{n\to\infty} \sqrt{\frac{n}{n+1}} = 1 \in (0, \infty).$$

Cu cele spuse anterior, cele două serii au aceeaşi natură; deci, seria în chestiune este convergentă.

Vom completa aceste consideraţii cu criterii de convergenţă *absolute* (care, din analiza doar a termenilor seriei să dea răspunsul privind natura acesteia). Avem în acest sens

Teorema 4.1.6 *Fie* $\sum\limits_n u_n$ *o serie cu termeni pozitivi pentru care,fie*

(C4) *există* $\lambda = \lim\limits_{n\to\infty} \dfrac{u_{n+1}}{u_n} \quad (0 \leq \lambda \leq \infty)$

fie că (independent de aceasta)

(C5) *există* $\lambda = \lim\limits_{n\to\infty} \sqrt[n]{u_n} \quad (0 \leq \lambda \leq \infty)$.

Atunci,dacă $\lambda < 1$, *seria converge ,iar dacă* $\lambda > 1$, *seria diverge.*

Demonstraţie. Vom studia numai cazul descris de (C4). Fie $\lambda < 1$; există $\varepsilon > 0$ cu $q = \lambda + \varepsilon < 1$. Conform definiţiei limitei, se poate găsi rangul $k = n(\varepsilon)$, aşa ca

$$\frac{u_{n+1}}{u_n} \leq q = \frac{q^{n+1}}{q^n}, \quad \forall n \geq k;$$

şi cum $\sum\limits_n q^n$ converge, seria $\sum\limits_n u_n$ de asemenea converge, în baza rezultatului precedent. Cazul $\lambda > 1$ se analizează în acelaşi mod. ■

Prin convenţie, partea întâia a acestei teoreme alcătuieşte ceea ce se cheamă *criteriul raportului (d'Alembert)*, iar partea a doua, *criteriul rădăcinii (Cauchy)*.

Exemplu. Să se studieze natura seriei $\sum\limits_n \frac{x^n}{n!}$, unde $x \geq 0$ este un parametru. Aplicăm criteriul raportului

$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = \lim_{n\to\infty} \frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}} = \lim_{n\to\infty} \frac{x}{n+1} = 0 < 1.$$

Seria în discuţie este deci convergentă, pentru oricare $x \geq 0$.

Să observăm că, pentru $\lambda = 1$, teorema precedentă nu ne arată ce se întâmplă cu seria. Spunem că suntem într-un *caz de nedeterminare.* Într-o asemenea situaţie, pentru a decide asupra naturii seriei, este nevoie de criterii mai fine. Unul dintre acestea este aşa–numitul *criteriu Raabe–Duhamel*:

Teorema 4.1.7 *Fie* $\sum\limits_n u_n$ *o serie cu termeni pozitivi, pentru care*

(C6) *există* $\mu = \lim\limits_{n\to\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) \quad (0 \leq \mu \leq \infty).$

Atunci,dacă $\mu > 1$, *seria converge, iar dacă* $\mu < 1$, *seria diverge.*

Demonstraţie. Să admitem $\mu > 1$; există $\varepsilon > 0$ cu $q = \mu - \varepsilon > 1$. Conform definiţiei limitei,se poate găsi rangul $k = n(\varepsilon)$ aşa încât

$$n\left(\frac{u_n}{u_{n+1}} - 1\right) \geq q, \quad \forall n \geq k;$$

sau, echivalent,

$$nu_n - (n+1)u_{n+1} \geq (q-1)u_{n+1},\ \forall n \geq k.$$

Fără a restrânge din generalitate, putem lua $k = 1$. Scriem succesiv primele n inegalităţi şi apoi le adunăm. Obţinem astfel relaţiile

$$u_1 \geq (q-1)(u_1 + \cdots + u_n),\ \forall n \geq 1.$$

Adică, şirul sumelor parţiale al seriei $\sum_n u_n$ este mărginit, ceea ce înseamnă că seria converge. Cazul $\mu < 1$ se studiază analog. ∎

Ca mai înainte, avem şi aici caz de nedeterminare ($\mu = 1$). Atunci, se pune problema obţinerii de alte criterii mai fine, etc. Cele spuse până acum sunt totuşi operante în destule situaţii concrete.

Exemplu. Să se studieze natura seriei $\sum_n \frac{1 \cdot 3 \cdots (2n-1)}{2 \cdot 4 \cdots (2n)} \frac{1}{2n+1}$. Evaluăm deocamdată raportul a doi termeni consecutivi

$$\frac{u_{n+1}}{u_n} = \frac{\frac{1\cdot 3\cdots(2n-1)(2n+1)}{2\cdot 4\cdots(2n)(2n+2)}\ \frac{1}{2n+3}}{\frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots(2n)}\ \frac{1}{2n+1}} = \frac{(2n+1)^2}{(2n+2)(2n+3)}.$$

În felul acesta, $\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = 1$; adică, criteriul raportului nu este operant (caz de nedeterminare). Să aplicăm criteriul Raabe–Duhamel:

$$\begin{aligned} \lim_{n\to\infty} n\left(\tfrac{u_n}{u_{n+1}} - 1\right) &= \lim_{n\to\infty} n\left(\tfrac{(2n+2)(2n+3)}{(2n+1)^2} - 1\right) = \\ &= \lim_{n\to\infty} \tfrac{6n^2+5n}{(2n+1)^2} = \tfrac{3}{2} > 1. \end{aligned}$$

Seria noastră este deci convergentă.

(B) Trecem acum la studiul seriilor cu termeni oarecare. Începem cu următorul *criteriu de comparaţie* (datorat lui Weierstrass).

Teorema 4.1.8 *Fie* $\sum_n u_n$ *o serie cu termenii oarecare pentru care*

(C7) $|u_n| \leq \beta_n$, *pentru orice* n.

Dacă seria (*cu termeni pozitivi*) $\sum_n \beta_n$ *este convergentă, atunci şi seria* $\sum_n u_n$ *este convergentă.*

Demonstraţie. Se foloseşte criteriul de convergenţă Cauchy. ■

Exemplu. Să se studieze seria $\sum_n \frac{x^n}{n!}$, unde $x \in R$. Avem

$$\left|\frac{x^n}{n!}\right| \le \frac{|x|^n}{n!},\ n = 0, 1, ...$$

Cum seria cu termeni pozitivi $\sum_n \frac{|x|^n}{n!}$ converge (a se vedea un exemplu anterior), va rezulta că şi seria iniţială converge, oricare ar fi $x \in R$.

În particular, seria $\sum_n u_n$ converge dacă seria modulilor $\sum_n |u_n|$ converge. (Vom numi seria iniţială, *absolut convergentă* în acest caz.) Cu alte cuvinte,

(P3) orice serie absolut convergentă este convergentă.

Menţionăm că, astfel de serii au proprietăţile importante:

(i) *asociativitatea*: în orice mod am grupa termenii, rezultatul sumării acestor fragmente de serie este acelaşi;

(ii) *comutativitatea*: în orice ordine am suma termenii, rezultatul este acelaşi.

Ne interesează în continuare studiul acelor serii $\sum_n u_n$ pentru care tehnicile anterioare nu sunt aplicabile. O modalitate utilă de abordare a lor este furnizată de aşa–numitul *criteriu al lui Abel*:

Teorema 4.1.9 *Fie dată seria* $\sum_n \lambda_n v_n$ *în care*

(C8) *şirul sumelor parţiale ale seriei* $\sum_n \lambda_n$ *este mărginit*

(C9) (v_n) *este un şir monoton convergent la zero.*

Atunci, obligatoriu $\sum_n \lambda_n v_n$ *converge.*

Demonstraţie. Să notăm

$$\sigma_n = \lambda_1 + \cdots + \lambda_n,\ n \ge 1 \text{ (deci } \lambda_n = \sigma_n - \sigma_{n-1},\ \forall n \ge 2).$$

Ipoteza (C8) ne spune că (σ_n) este mărginit; deci

$$\text{există } M > 0 \text{ aşa ca } |\sigma_n| \leq M, \text{ pentru toţi } n.$$

Fie $n \geq 1$, $p \geq 1$ numere naturale arbitrare deocamdată. Avem evaluările (cu (v_n) descrescător, de pildă)

$$\begin{aligned}
&|\sum_{h=1}^{p} \lambda_{n+h} v_{n+h}| = |\sum_{h=1}^{p} (\sigma_{n+h} - \sigma_{n+h-1}) v_{n+h}| = \\
&= | - \sigma_n v_{n+1} + \sigma_{n+p} v_{n+p} + \sum_{k=1}^{p-1} \sigma_{n+k}(v_{n+k} - v_{n+k+1})| \leq \\
&\leq |\sigma_n||v_{n+1}| + |\sigma_{n+p}|v_{n+p}| + \sum_{k=1}^{p-1} |\sigma_{n+k}|(v_{n+k} - v_{n+k+1}) \leq \\
&\leq M(|v_{n+1}| + |v_{n+p}| + \sum_{k=1}^{p-1} (v_{n+k} - v_{n+k+1})) \leq 2M(|v_{n+1}| + |v_{n+p}|).
\end{aligned}$$

Conform ipotezei (C9), rezultă că, dat $\varepsilon > 0$, există un rang $n(\varepsilon)$ cu

$$n \geq n(\varepsilon) \Longrightarrow |v_n| \leq \varepsilon/2M \text{ (deci } \ |v_{n+p}| \leq \varepsilon/2M).$$

Ca atare, criteriul Cauchy se verifică pentru seria noastră şi concluzia reiese (cu Teorema 4.1.4). ■

Exemplu. Să se studieze seria $\sum_n \frac{1}{n} \sin(nx)$ unde x este un parametru. Criteriul comparaţiei nu se poate aplica deoarece acesta revine la

$$|\frac{1}{n} \sin(nx)| \leq \frac{1}{n}, \ n = 1, 2,$$

cu seria de comparaţie $\sum_n \frac{1}{n}$, divergentă. Aplicăm criteriul Abel cu

$$\lambda_n = \sin(nx), \ v_n = \frac{1}{n}, \ n = 1, 2, ...$$

Avem (conform unor formule cunoscute)

$$|\lambda_1 + \cdots + \lambda_n| = \left| \frac{\cos \frac{x}{2} - \cos(nx + \frac{x}{2})}{2 \sin \frac{x}{2}} \right| \leq \frac{1}{|\sin \frac{x}{2}|}, \ n = 1, ...$$

(Desigur, aceasta cere $x \neq 2k\pi$, $k \in Z$, ceea ce şi presupunem.) Pe de altă parte, (v_n) este monoton descrescător şi convergent la zero. Pe baza criteriului menţionat se deduce atunci că seria noastră converge.

În particular, condiţia de mărginire (C8) are loc când

$$\lambda_n = (-1)^n,\ n = 1, 2, ...;$$

adică, atunci când seria în discuţie devine $\sum_n (-1)^n v_n$. (O asemenea serie se zice *alternantă.*) Ajungem atunci să formulăm aşa–numitul *criteriu de convergenţă Leibniz* pentru aceste serii:

$$\text{(P4)}\quad \begin{cases} \text{seria alternantă } \sum_n (-1)^n v_n \text{ converge de fiecare dată când} \\ (v_n) \text{ este monoton şi convergent la zero.} \end{cases}$$

Exemplu. Seria alternantă $\sum_n (-1)^n \frac{1}{n}$ satisface condiţiile precedente; deci, ea este convergentă.

Să observăm că pentru seria din acest exemplu, seria modulilor (adică $\sum_n \frac{1}{n}$) diverge; o serie cu această proprietate se zice *semiconvergentă.* Notăm că în acest caz, nici una din proprietăţile menţionate la cazul seriilor absolut convergente nu se mai păstrează.

(C) În ce priveşte suma unei serii numerice următoarea observaţie se poate dovedi utilă.

Teorema 4.1.10 *Fie* $\sum_n u_n$ *o serie numerică la care termenul general al ei admite reprezentarea*

(C10) $u_n = \alpha_n - \alpha_{n+1},\ n = 1, 2, ...\ .$

Atunci, seria considerată converge sau diverge odată cu şirul (α_n)*; iar, în caz de convergenţă, suma ei are expresia*

$$\sum_{n=1}^{\infty} u_n = \alpha_1 - \lim_{n\to\infty} \alpha_n.$$

Demonstraţie. Să evaluăm şirul sumelor parţiale ale seriei. Avem

$$\begin{cases} s_1 = u_1 = \alpha_1 - \alpha_2 \\ s_2 = u_1 + u_2 = \alpha_1 - \alpha_2 + \alpha_2 - \alpha_3 = \alpha_1 - \alpha_3 \\ \vdots \\ s_n = u_1 + \cdots + u_n = \alpha_1 - \alpha_2 + \cdots + \alpha_n - \alpha_{n+1} = \alpha_1 - \alpha_{n+1} \end{cases}$$

De aici, urmează că (s_n) converge dacă şi numai dacă (α_n) converge. În plus, trecând la limită în relaţia anterioară avem, în caz de convergenţă

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty}(\alpha_1 - \alpha_{n+1}) = \alpha_1 - \lim_{n\to\infty}\alpha_n.$$

Teorema este astfel dovedită. $\blacksquare$

Prin definiţie, vom numi *telescopică* orice serie numerică la care termenul general verifică (C10).

Exemplu. Să se studieze convergenţa seriei $\sum\limits_n \frac{1}{n(n+1)}$; iar, în caz afirmativ, să se calculeze şi suma acesteia. Avem, evident (prin descompunere în fracţii simple),

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1},\ n = 1, 2, \dots\ .$$

Suntem deci în cazul rezultatului anterior, cu $\alpha_n = \frac{1}{n}$, $n = 1, 2, \dots$. Cum (α_n) converge, cu $\lim\limits_{n\to\infty}\alpha_n = 0$, urmează că seria noastră este convergentă, iar suma ei este

$$\sum_{n=1}^{\infty}\frac{1}{n(n+1)} = \alpha_1 - \lim_{n\to\infty}\alpha_n = 1.$$

(D) Ca o completare a celor spuse anterior, să mai facem câteva observaţii privind operaţiile cu seriile. Fie $\sum\limits_n u_n$ şi $\sum\limits_n v_n$ două serii numerice. Numim seria $\sum\limits_n z_n$, unde

$$z_n = u_n + v_n,\ n = 1, 2, \dots$$

suma celor două serii; o vom mai nota şi prin $\sum\limits_n (u_n + v_n)$. Este o consecinţă imediată a definiţiei implicaţia

(P5) $\sum\limits_n u_n = conv.,\ \sum\limits_n v_n = conv. \Longrightarrow \sum\limits_n (u_n + v_n) = conv.$

În plus, are loc şi formula (în caz de convergenţă desigur)

(P6) $\sum\limits_{n=1}^{\infty}(u_n + v_n) = \sum\limits_{n=1}^{\infty} u_n + \sum\limits_{n=1}^{\infty} v_n.$

Mai departe, să numim seria $\sum\limits_n t_n$, unde

$$t_n = u_1 v_n + u_2 v_{n-1} + \cdots + u_{n-1} v_2 + u_n v_1, \; n = 1, 2, ...,$$

produsul celor două serii. Analiza convergenţei acestei serii nu mai este la fel de simplă ca precedenta. Dăm fără demonstraţie următorul rezultat (numit " Teorema lui Cauchy"):

Teorema 4.1.11 *Dacă ambele serii $\sum\limits_n u_n$ şi $\sum\limits_n v_n$ sunt absolut convergente, atunci seria produs $\sum\limits_n t_n$ este,şi ea, absolut convergentă , iar suma ei este produsul sumelor celor două serii*

$$\sum_{n=1}^{\infty} t_n = \left(\sum_{n=1}^{\infty} u_n \right) \left(\sum_{n=1}^{\infty} v_n \right).$$

Exemplu. După cum s-a arătat deja într-un loc anterior, seria $\sum\limits_n \frac{x^n}{n!}$ converge absolut, pentru orice valoare a parametrului real x. Să notăm

$$\varphi(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in R.$$

Fie x şi y valori reale arbitrare. Produsul seriei $\sum\limits_n \frac{x^n}{n!}$ cu seria $\sum\limits_n \frac{y^n}{n!}$ va fi, conform definiţiei, seria $\sum\limits_n t_n$ unde

$$\begin{cases} t_n = \frac{x^0}{0!}\frac{y^n}{n!} + \frac{x^1}{1!}\frac{y^{n-1}}{(n-1)!} + \cdots + \frac{x^{n-1}}{(n-1)!}\frac{y^1}{1!} + \frac{x^n}{n!}\frac{y^0}{0!} = \\ = \sum\limits_{k=0}^{n} \frac{1}{k!(n-k)!} x^k y^{n-k} = \frac{1}{n!} \sum\limits_{k=0}^{n} C_n^k x^k y^{n-k} = \frac{1}{n!}(x+y)^n, n=0, 1, ... \end{cases}$$

Cum ambele serii sunt absolut convergente, rezultă că (din teorema anterioară) şi seria $\sum\limits_n t_n$ este absolut convergentă; în plus,

$$\sum_{n=0}^{\infty} t_n = \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right) \left(\sum_{n=0}^{\infty} \frac{y^n}{n!} \right) = \varphi(x)\varphi(y).$$

Dar, evident (prin calculele anterioare),

$$\sum_{n=0}^{\infty} t_n = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = \varphi(x+y).$$

Am ajuns deci la concluzia că

(R1) $\varphi(x+y) = \varphi(x)\varphi(y),\ \forall x, y \in R.$

Să completăm informaţiile anterioare cu încă una:

(R2) $(\varphi(1) =) \sum_{n=0}^{\infty} \dfrac{1}{n!} = e$ (= numărul lui Euler).

Pentru demonstrarea acestei formule, să notăm

$$x_n = 1 + \frac{1}{1!} + \cdots + \frac{1}{n!},\ y_n = \left(1 + \frac{1}{n}\right)^n,\ n \geq 1.$$

Din formula binomului a lui Newton avem

$$\begin{aligned} y_n &= 1 + \tfrac{n}{1!}\tfrac{1}{n} + \tfrac{n(n-1)}{2!}\tfrac{1}{n^2} + \cdots + \tfrac{n(n-1)\cdots 1}{n!}\tfrac{1}{n^n} = \\ &= 1 + \tfrac{1}{1!} + \tfrac{1}{2!}(1 - \tfrac{1}{n}) + \cdots + \tfrac{1}{n!}(1 - \tfrac{1}{n})\cdots(1 - \tfrac{n-1}{n}). \end{aligned}$$

Şi, cum fiecare paranteză poate fi majorată prin 1, avem

$$y_n \leq x_n,\ n = 1, 2, ...$$

iar de aici, prin trecere la limită, $e \leq \lim\limits_{n\to\infty} x_n$. Pe de altă parte, oricare ar fi $k \geq 2$ arbitrar fixat, putem scrie din evaluarea precedentă

$$y_n \geq 1 + \frac{1}{1!} + \frac{1}{2!}(1 - \frac{1}{n}) + \cdots + \frac{1}{k!}(1 - \frac{1}{n})\cdots(1 - \frac{k-1}{n}),\ \forall n > k.$$

Trecând la limită pentru $n \longrightarrow \infty$, găsim atunci

$$e \geq x_k,\ \text{oricare ar fi} \quad k \geq 2,$$

şi, deci, $e \geq \lim\limits_{k\to\infty} x_k$. Combinând informaţiile găsite, deducem $e = \lim\limits_{n\to\infty} x_n$ care înseamnă tocmai ceea ce trebuia dovedit.

Vom vedea ulterior că din (R1)+(R2) deducem forma funcţională a lui φ; anume

$$\varphi(x) = e^x,\ \forall x \in R.$$

Cu alte cuvinte, avem ca valabilă formula

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots,\ \forall x \in R.$$

4.1.3 Convergenţa pe spaţiile R^m

Dat numărul natural m, fie R^m spaţiul liniar m–dimensional (al vectorilor linie)

$$R^m = \{x = (x(1), ..., x(m));\ x(1), ..., x(m) \in R\}.$$

Cu alte cuvinte, vom identifica fiecare vector x din R^m cu funcţia $x : \{1, ..., m\} \longrightarrow R$ definită prin

$$x(i) = \text{coordonata de indice } i \text{ din vectorul } x.$$

Am introdus (Secţiunea 1.1.8) *norma* unui vector din R^m prin

$$\|x\| = [(x(1))^2 + \cdots + (x(m))^2]^{1/2},\ x = (x(1), ..., x(m)) \in R^m.$$

Proprietăţile mai importante ale acesteia sunt

$$\text{(P1)} \begin{cases} \|x + y\| \le \|x\| + \|y\|,\ \|\lambda x\| = |\lambda| \cdot \|x\|,\ \forall x, y \in R^m, \forall \lambda \in R; \\ \|x\| \ge 0,\ \forall x \in R^m;\ \|x\| = 0 \iff x = 0. \end{cases}$$

Cu alte cuvinte, norma unui vector apare ca o generalizare a modulului unui număr real. Ar fi de aşteptat astfel să putem construi o teorie a convergenţei în aceste spaţii, după modelul precedent; este tocmai ceea ce intenţionăm în continuare.

Pentru fiecare cuplu $x, y \in R^m$, numim *distanţa* dintre aceşti vectori, numărul $\|x - y\|$. Să facem notaţia

$$S(x, \varepsilon) = \{y \in R^m;\ \|x - y\| < \varepsilon\},\ x \in R^m,\ \varepsilon > 0.$$

Vom numi aceasta, *sferă deschisă* cu centrul x de rază ε. În continuare, să numim *vecinătate* a punctului $x \in R^m$,

(D1) orice asemenea sferă deschisă cu centrul în x,

(D2) orice parte din R^m care include o asemenea sferă.

Familia tuturor vecinătăţilor punctului x se va nota prin $\mathcal{V}(x)$.

Fie dat un şir de vectori din R^m

$$x_1 = (x_1(1), ..., (x_1(m)), ..., x_n = (x_n(1), ..., x_n(m)), ...;$$

şi, de asemenea, vectorul $x = (x(1), ..., x(m))$ din R^m. Zicem că (x_n) *converge* către x (şi scriem aceasta prin $x_n \longrightarrow x$ sau $\lim\limits_{n\to\infty} x_n = x$) dacă

(D3) oricare ar fi vecinătatea $V \in \mathcal{V}(x)$ există un rang $n(V)$ aşa ca $x_n \in V$, oricare ar fi rangul $n \geq n(V)$.

Definiţia este formal identică cu aceea din cazul $m = 1$. O formă echivalentă de scriere a ei este următoarea:

(D4) pentru orice $\varepsilon > 0$ există un rang $n(\varepsilon)$ aşa ca
$n \geq n(\varepsilon) \Longrightarrow \|x_n - x\| < \varepsilon$ (sau $x_n \in S(x, \varepsilon)$).

Următorul rezultat este important de semnalat.

Teorema 4.1.12 *Avem echivalenţa*

(P2) $x_n \longrightarrow x \Longleftrightarrow x_n(1) \longrightarrow x(1), ..., x_n(m) \longrightarrow x(m)$.

Adică: şirul de vectori (x_n) *converge către vectorul* x *dacă şi numai dacă şirul componentelor de un anume rang al lui* (x_n) *converge la componenta de acelaşi rang a lui* x.

Demonstraţia se bazează pe dubla inegalitate

$$\max_{1 \leq i \leq m} |x_n(i) - x(i)| \leq \|x_n - x\| \leq m \max_{1 \leq i \leq m} |x_n(i) - x(i)|, \ n = 1, 2, ... \ .$$

Urmează de aici o importantă concluzie de natură metrică

(P3) R^m este spaţiu complet, pentru orice m.

Exemplu. Fie dat şirul de vectori din R^3

$$x_n = \left(\frac{n+1}{3n+2} \ , 2^{-n} \ , \frac{n^3}{5^n} \right), \ n = 1, 2, ... \ .$$

Limita şirului respectiv este în aceste condiţii

$$\lim_{n \to \infty} x_n = \left(\lim_{n \to \infty} \frac{n+1}{3n+2} \ , \lim_{n \to \infty} 2^{-n} \ , \lim_{n \to \infty} \frac{n^3}{5^n} \right) = \left(\frac{1}{3}, 0, 0 \right).$$

Să descriem acum unele mulţimi din R^m asociate structurilor introduse. Fie D o mulţime nevidă de vectori din R^m.

(i) Vectorul $x \in R^m$ se va numi *punct interior* pentru D dacă există $\varepsilon > 0$ aşa încât $S(x, \varepsilon) \subseteq D$ (deci, în particular, $x \in D$). Mulţimea tuturor acestor puncte alcătuieşte *interiorul* mulţimii D şi se notează $\mathrm{int}\,(D)$. Mulţimea D se zice *deschisă* dacă toate punctele ei sunt interioare. Ca exemplu imediat de asemenea mulţimi avem orice sferă deschisă din spaţiul considerat; şi, de asemenea,

$$(R_+^0)^m = \{x = (x(1), ..., x(m));\ x(1) > 0, ..., x(m) > 0\}$$

(mulţimea vectorilor cu toate componentele strict pozitive.)

(ii) Vectorul $x \in R^m$ se va numi *punct aderent* pentru D dacă există un şir (x_n) din D cu $x_n \longrightarrow x$. Mulţimea tuturor acestor puncte alcătuieşte *aderenţa* (sau *închiderea*) mulţimii D şi se notează $\mathrm{ad}\,(D)$. Mulţimea D se zice *închisă* dacă este identică cu închiderea sa. Exemple de asemenea mulţimi sunt toate sferele închise

$$S[x, \varepsilon] = \{y \in R^m;\ \|x - y\| \le \varepsilon\},\ x \in R^m,\ \varepsilon > 0;$$

şi, de asemenea,

$$R_+^m = \{x = (x(1), ..., x(m));\ x(1) \ge 0, ..., x(m) \ge 0\}$$

(conul vectorilor de coordonate pozitive ale spaţiului R^m.) Să mai remarcăm cu această ocazie proprietatea

(P4) o mulţime a lui R^m este deschisă (respectiv închisă) dacă şi numai dacă complementara sa este închisă (respectiv deschisă).

(iii) Vectorul $x \in R^m$ se va numi *punct de acumulare* pentru D dacă există un şir (x_n) în D cu

$$x_n \ne x,\ \forall n;\ x_n \longrightarrow x.$$

Mulţimea tuturor acestor puncte alcătuieşte *mulţimea derivată* a lui D şi se notează $\mathrm{ac}\,(D)$. Orice punct de acumulare este, desigur, punct aderent dar nu şi reciproc. Orice punct din D care nu este punct de acumulare al lui D se zice *punct izolat* al lui D.

(iv) Vectorul $x \in R^m$ se va zice *punct frontieră* al lui D dacă acesta este atât punct aderent al lui D, cât şi punct aderent al complementarei $D^c = R^m \setminus D$. Mulţimea tuturor punctelor frontieră ale lui D alcătuieşte *frontiera* mulţimii D şi se notează $\mathrm{fr}\,(D)$. De pildă, frontiera oricărei sfere închise sau deschise este coaja acesteia:

$$\mathrm{fr}\,(S(x, \varepsilon)) = \mathrm{fr}\,(S[x, \varepsilon]) = \{y \in R^m;\ \|x - y\| = \varepsilon\}.$$

Mulţimea de vectori D se zice *mărginită* dacă poate fi închisă într-o sferă; sau, echivalent, într-o sferă cu centrul în origine. Ca exemplu în acest sens poate servi orice sferă închisă sau deschisă. Menţionăm cu această ocazie proprietatea (Bolzano–Weierstrass)

(P5) orice mulţime infinită şi mărginită are puncte de acumulare.

Mulţimea de vectori D se zice *compactă* dacă este simultan mărginită şi închisă. Ca exemplu de asemenea mulţime poate servi orice sferă închisă. Alte exemple pot fi deduse din observaţia că orice parte închisă a unei mulţimi compacte este compactă.

În fine, mulţimea de vectori D se va zice *conexă* dacă în orice descompunere de forma $D = D_1 \cup D_2$ cu D_1, D_2 nevide disjuncte, măcar una din cele două părţi are un punct de acumulare în cealaltă. Ca exemplu de asemenea mulţimi avem orice parte convexă a lui R^m. În particular, pentru $m = 1$ intervalele sunt conexe; reciproca este, de asemenea, adevărată. O mulţime conexă şi deschisă se numeşte *domeniu*; iar,o mulţime conexă şi compactă se va zice *continuum*.

4.1.4 Serii de vectori (în spaţiile R^m)

Fie

$$x_1 = (x_1(1), ..., x_1(m)), ..., x_n = (x_n(1), ..., x_n(m)), ...,$$

un şir de vectori din R^m. Vom numi şi aici simcolul

$$x_1 + x_2 + \cdots + x_n + \cdots \text{ (notat, de asemenea, } \sum_n x_n)$$

o *serie vectorială*.Semnificaţia acestuia este următoarea.Definim şirul de vectori (ai sumelor parţiale) din R^m

$$s_1 = x_1, s_2 = x_1 + x_2, ..., s_n = x_1 + \cdots + x_n, ... \ .$$

Prin definiţie,vom spune că seria $\sum_n x_n$ *converge* (în R^m) dacă şirul (s_n) converge (în R^m); iar,în acest caz, vectorul $s = \lim_{n\to\infty} s_n$ se va numi *suma seriei* în chestiune şi vom nota aceasta $s = \sum_{n=1}^{\infty} s_n$. În caz contrar, seria $\sum_n x_n$ se va zice *divergentă*.

Proprietăţile de bază ale seriilor vectoriale rezultă în esenţă din următoarea teoremă de tip relativ.

Teorema 4.1.13 *Seria vectorială* $\sum_n x_n$ *este convergentă dacă şi numai dacă fiecare din seriile numerice* $\sum_n x_n(1), ..., \sum_n x_n(m)$, *este convergentă. Iar, în acest caz, suma ei se obţine prin formula*

$$\sum_n x_n = \left(\sum_n x_n(1), ..., \sum_n x_n(m)\right)$$

(adică, prin sumarea pe componente).

Demonstraţie. Nu avem decât să ţinem cont de definiţia convergenţei unei serii vectoriale şi caracterizarea convergenţei unui şir de vectori prin intermediul componentelor (rezultatul precedent). ∎

În felul acesta, întreaga problematică a seriilor vectoriale se reduce la aceea a seriilor numerice.

Exemplu. Seria vectorială $\sum_n \left(\frac{1}{n^2}, 2^{-n}, (-1)^n \frac{1}{n}\right)$ este convergentă (în R^3), deoarece eiecare din seriile numerice componente $\sum_n \frac{1}{n^2}, \sum_n 2^{-n}, \sum_n (-1)^n \frac{1}{n}$ este convergentă. Dimpotrivă, seria vectorială $\sum_n \left(\frac{1}{n}, 3^{-n}\right)$ este divergentă (în R^2); căci una din seriile componente (anume $\sum_n \frac{1}{n}$) este divergentă (în R).

În particular, de aici rezultă versiunea vectorială a Criteriului de comparaţie Weierstrass (formulat pentru serii numerice).

Teorema 4.1.14 *Fie* $\sum_n x_n$ *o serie vectorială pentru care*

$$\|x_n\| \leq \beta_n, \text{ pentru orice } n.$$

Dacă seria (cu termeni pozitivi) $\sum_n \beta_n$ *este convergentă, atunci şi seria vectorială* $\sum_n x_n$ *converge.*

Demonstraţia se reduce în ultimă instanţă la aceea a teoremei corespunzătoare; nu mai dăm detalii.

4.1.5 Şiruri şi serii de matrici

Date numerele naturale p, q să notăm cu $\mathcal{M}(p,q)$ spaţiul liniar al tuturor matricilor de tip (p,q)

$$\mathcal{M}(p,q) = \{A = (a(i,j));\ a(i,j) \in R,\ 1 \le i \le p, 1 \le j \le q\}.$$

Şi aici vom identifica fiecare matrice A din $\mathcal{M}(p,q)$ cu funcţia $a : \{1,...,p\} \times \{1,...,q\} \longrightarrow R$ definită prin

$$a(i,j) = \text{elementul de pe linia } i \text{ şi coloana } j \text{ din matricea } A.$$

Introducem o normă pe acest spaţiu prin convenţia

$$\|A\| = \left(\sum_{i,j}(a(i,j))^2\right)^{1/2},\ A = (a(i,j)) \in \mathcal{M}(p,q).$$

Se constată valabilitatea proprietăţilor cunoscute; anume

$$\text{(P1)} \begin{cases} \|A+B\| \le \|A\| + \|B\|, \|\lambda A\| = |\lambda| \cdot \|A\|, \forall A, B \in \mathcal{M}(p,q), \forall \lambda \in R; \\ \|A\| \ge 0,\ \forall A \in \mathcal{M}(p,q);\ \|A\| = 0 \iff A = 0. \end{cases}$$

În plus, mai avem valabilă şi

(P2) $\|AC\| \le \|A\| \cdot \|C\|,\ A \in \mathcal{M}(p,q),\ C \in \mathcal{M}(q,r)$.

(Aici, r este un alt număr natural arbitrar fixat.)

Numim $\|A - B\|$ *distanţa* dintre matricile A şi B din $\mathcal{M}(p,q)$. Noţiunile de sferă şi vecinătate se introduc în acelaşi mod ca la vectori. Acestea, la rândul lor, permit introducerea noţiunii de convergenţă a unui şir de matrici

$$A_1 = (a_1(i,j)), A_2 = (a_2(i,j)), ..., A_n = (a_n(i,j)), ...$$

către o matrice limită $A = (a(i,j))$ (toate acestea luate ca elemente din $\mathcal{M}(p,q)$). Rezultatul anterior devine în acest context

Teorema 4.1.15 *Avem echivalenţa*

$$A_n \longrightarrow A \iff a_n(i,j) \longrightarrow a(i,j),\ 1 \le i \le p,\ 1 \le j \le q.$$

Adică: şirul de matrici (A_n) *converge la matricea* A *dacă şi numai dacă şirul elementelor de pe acelaşi loc din* (A_n) *converge la elementul de pe locul respectiv din* A.

Fie (A_n) un şir de matrici din $\mathcal{M}(p,q)$. Zicem că seria matricială $\sum\limits_n A_n$ este *convergentă* dacă şirul matricial (S_n) al sumelor parţiale

$$S_1 = A_1, S_2 = A_1 + A_2, ..., S_n = A_1 + \cdots + A_n, ...$$

converge (în $\mathcal{M}(p,q)$); iar, în acest caz, matricea limită $S = \lim\limits_{n\to\infty} S_n$ se va numi *suma seriei*, şi o notăm $S = \sum\limits_{n=1}^{\infty} A_n$. În caz contrar, seria se zice *divergentă*. Proprietăţile de bază ale seriilor matriciale rezultă din Teorema 4.1.13 (corespunzător adaptată). În particular, putem formula şi aici un rezultat corespunzător Teoremei 4.1.14 ; nu mai dăm alte detalii.

Vom da acum două aplicaţii importante ale acestor dezvoltări. Fie p număr natural dat şi $\mathcal{M}(p)$ spaţiul tuturor matricilor pătrate cu p linii şi p coloane. Să notăm că din (P2) avem

(P3) $\|C^n\| \leq \|C\|^n,\ n = 0, 1, ..., C \in \mathcal{M}(p)$.

(i) Oricare ar fi matricea $A = (a_{ij})$ din $\mathcal{M}(p)$ cu $\rho = \|A\| < 1$, seria (geometrică) matricială $\sum\limits_n A^n$ converge. Aceasta rezultă din Criteriul de comparaţie Weierstrass pentru serii matriciale, dacă ţinem cont de evaluările (deduse din (P3))

$$\|A^n\| \leq \rho^n,\ n = 0, 1, 2, ...$$

În ce priveşte suma acestei serii, avem formula

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n \text{ (pentru asemenea matrici } A).$$

Adică, mai exact, oricare ar fi matricea A din $\mathcal{M}(p)$ cu $\|A\| < 1$, $(I - A)^{-1}$ există şi apare drept sumă a seriei considerate. Pentru aceasta, nu avem decât să pornim de la identitatea (matricială)

$$(I - A)(I + A + \cdots + A^n) = I - A^{n+1},\ n = 1, 2, ...,$$

şi să facem apoi o trecere la limită pentru $n \to \infty$. (Notăm că, din $\|A\| < 1$, avem $\lim\limits_{n\to\infty} A^n = 0$.)

(ii) Oricare ar fi matricea $A = (a_{ij})$ din $\mathcal{M}(p)$, seria matricială $\sum\limits_n \frac{1}{n!} A^n$ converge. Aceasta rezultă din nou cu Criteriul de comparaţie Weierstrass (şi evaluările anterioare) dacă ne aducem aminte că

$$\text{seria (numerică) } \sum_n \frac{1}{n!}\rho^n \text{ converge}, \quad \forall \rho \geq 0.$$

Să notăm cu $\varphi(A)$ suma acesteia; adică

$$\varphi(A) = I + \frac{1}{1!}A + \cdots + \frac{1}{n!}A^n + \cdots, \ A \in \mathcal{M}(p).$$

În primul rând, evident,

(R1) $\varphi(xI) = \left(1 + \frac{x}{1!} + \cdots + \frac{x^n}{n!} + \cdots\right) I = e^x I, \ x \in R.$

În al doilea rând, prin raţionamente similare cu acelea din cazul numeric (Secţiunea 4.1.2), deducem

(R2) $\varphi(A+B) = \varphi(A) \cdot \varphi(B), \ \text{dacă } AB = BA.$

Vom folosi faptul că, oricare ar fi cuplul de matrici A, B care permută între ele ($AB = BA$), este valabilă formula binomului (Newton)

$$(A+B)^n = \sum_{n=0}^{n} C_n^k A^{n-k} B^k, \ n = 1, 2, \dots \ .$$

Din aceste două proprietăţi urmează că $A \vdash \varphi(A)$ acţionează ca o funcţie exponenţială de matrici. Vom nota aceasta în forma

$$\varphi(A) = e^A, \quad A \in \mathcal{M}(p).$$

În acest caz, se poate scrie formula

$$e^A = I + \frac{1}{1!}A + \cdots + \frac{1}{n!}A^n + \cdots, \ \forall A \in \mathcal{M}(p).$$

Probleme propuse la §4.1

1. Să se calculeze limitele şirurilor

a) $x_n = n(1 - \sqrt[n]{a})$, (a=număr pozitiv arbitrar)

b) $x_n = \dfrac{1 \cdot 3 \cdots (2n-1)}{2 \cdot 4 \cdots (2n)}$ (Indicaţie: $x_n \leq \dfrac{1}{\sqrt{2n+1}}$, $\forall n$.)

2. Să se studieze şirurile recurente (r=parametru pozitiv)

a) $x_{n+1} = \sqrt{r + x_n}, \forall n \geq 1$; $x_1 = 0$

b) $x_{n+1} = \dfrac{1}{2}\left(x_n + \dfrac{r}{x_n}\right), \forall n \geq 1$; $x_1 > 0$.

3. Să se studieze seriile cu termeni pozitivi

a) $\sum\limits_n \left(r\dfrac{n^2+n+1}{n^2}\right)^n$, $r \geq 0$ (Criteriul rădăcinii)

b) $\sum\limits_n nr^n$, $r \geq 0$ (Criteriul raportului)

c) $\sum\limits_n \dfrac{(2n)!}{4^n (n!)^2}$ (Criteriul Raabe-Duhamel)

4. Să se precizeze care din seriile următoare este absolut convergentă sau semi–convergentă

(a): $\sum\limits_n (-1)^{n+1} \dfrac{2^n \sin^{2n} x}{n+1}$; (b): $\sum\limits_n (-1)^{n+1} \dfrac{1}{\sqrt{n(n+1)}}$.

5. Să se demonstreze egalitatea

$$\left(\sum_{n=0}^{\infty} q^n\right)^2 = \sum_{n=0}^{\infty} (n+1)q^n, \text{ dacă } |q| < 1.$$

6. Să se demonstreze convergenţa şi să se calculeze suma seriilor

a) $\sum\limits_{n=1}^{\infty} \dfrac{1}{n(2n+1)}$; b) $\sum\limits_{n=0}^{\infty} na^n$, $|a| < 1$; c) $\sum\limits_{n=1}^{\infty} \dfrac{n^2 - 3n + 2}{n!}$.

4.2 Funcţii de mai multe variabile. Limită şi continuitate

4.2.1 Limita unei funcţii de o variabilă

Fie $f : D \longrightarrow R$ o funcţie definită pe mulţimea $D \subseteq R$ şi a = punct de acumulare al lui D. Zicem că mărimea ℓ este *limita funcţiei f în punctul a* (şi scriem $\ell = \lim\limits_{x \to a} f(x)$) dacă

$$\text{(D1)}\quad \begin{cases} \text{oricare ar fi şirul } (x_n) \text{ din } D\setminus\{a\} \text{ cu } x_n \longrightarrow a, \\ \text{avem } f(x_n) \longrightarrow \ell \text{ (şirul imaginilor converge către } \ell). \end{cases}$$

Formal, definiţia rămâne aceeaşi chiar în cazul când a sau ℓ este unul din simbolurile $(+\infty)$ sau $(-\infty)$. Un mod echivalent de a scrie definiţia este următorul: $\ell = \lim\limits_{x\to a} f(x)$ dacă

$$\text{(D2)}\quad \begin{cases} \text{oricare ar fi vecinătatea } V\in\mathcal{V}(\ell), \text{ există vecinătatea } U\in\mathcal{V}(a) \\ \text{aşa încât } x\in U\cap(D\setminus\{a\}) \Longrightarrow f(x)\in V. \end{cases}$$

În legătură cu definiţia dată este important de subliniat următoarele aspecte de principiu:

(i) punctul (de acumulare) a nu este necesar să fie în mulţimea de definiţie D (adică, $f(a)$ poate să nu fie definit). Dacă a este punct izolat al mulţimii D, problema limitei în acest punct nu se pune;

(ii) mărimea ℓ nu este legată de valoarea funcţiei în punctul a (dacă există). Aceasta deoarece, în procesul trecerii la limită în punctul a, nu interesează comportarea funcţiei în punctul a, ci în punctele vecine acestuia (din mulţimea de definiţie D).

Rezultă din definiţia (D1) că, în oricare din următoarele situaţii, nu există limita funcţiei f în punctul de acumulare a:

(a) pentru cel puţin un şir (x_n) din $D\setminus\{a\}$ cu $x_n \longrightarrow a$, şirul $(f(x_n))$ nu are limită,

(b) pentru două şiruri (x'_n) şi (x''_n) din $D\setminus\{a\}$ cu $x'_n \longrightarrow a$, $x''_n \longrightarrow a$, limitele şirurilor $(f(x'_n))$ şi $(f(x''_n))$ există, dar sunt distincte.

Exemplu. Fie dată funcţia

$$f(x) = \cos\frac{1}{x},\ x\neq 0.$$

Mulţimea de definiţie este $D = R\setminus\{0\}$. Punctul $a=0$ este evident punct de acumulare al acesteia. Să studiem existenţa limitei în acest punct. Considerăm şirurile (convergente spre $a=0$)

$$\begin{array}{lll} x'_n = \frac{1}{2n\pi}, n\in N & \Longrightarrow f(x'_n) = 1, \forall n \Longrightarrow & \lim\limits_{n\to\infty} f(x'_n) = 1 \\ x,'_n = \frac{1}{\frac{\pi}{2}+2n\pi}, n\in N & \Longrightarrow f(x''_n) = 1, \forall n \Longrightarrow & \lim\limits_{n\to\infty} f(x''_n) = 0 \end{array}$$

Cum şirurile valorilor funcţiei asociate acestora dau limite distincte, urmează că nu există $\lim\limits_{x\to 0} f(x)$.

Ca o ultimă observaţie cu caracter general mai menţionăm şi faptul că, în analogie cu ceea ce se întîmplă la şiruri, există şi aici posibilitatea de a caracteriza existenţa limitei într-un punct fără a face să intervină explicit valoarea acesteia. Mai precis, avem (criteriul Cauchy–Bolzano):

Teorema 4.2.1 *Funcţia $f : D \longrightarrow R$ are limită finită în punctul de acumulare a, dacă şi numai dacă*

$$\text{(D3)} \quad \begin{cases} \textit{pentru orice } \varepsilon > 0 \textit{ există o vecinătate } V \in \mathcal{V}(a) \textit{ aşa încât} \\ x', x'' \in V \cap (D \setminus \{a\}) \Longrightarrow |f(x') - f(x'')| < \varepsilon. \end{cases}$$

Cu alte cuvinte, există limită în punct dacă, plasându-ne suficient de aproape de acesta, distanţele dintre valorile funcţiei în aceste puncte sunt oricât de mici dorim. Desigur, cazul $a = +\infty$ sau $a = -\infty$ nu este exclus în enunţul anterior.

Datorită structurii particulare a mulţimii numerelor reale, noţiunea de limită are două determinări posibile. Mai precis, fie *a punct de acumulare înspre dreapta* al lui D (adică există şiruri (x_n) din $D \cap (-\infty, a)$ cu $x_n \longrightarrow a$.) Zicem că mărimea ℓ_s este *limita la stânga* a funcţiei f în punctul a, şi scriem aceasta

$$\ell_s = \lim_{x \to a-} f(x) = \lim_{\substack{x \to a \\ x < a}} f(x) = f(a - 0)$$

dacă

$$\text{(D4)} \quad \begin{cases} \text{oricare ar fi şirul } (x_n) \text{ din } D \setminus \{a\} \\ \text{cu } x_n < a, \forall n, \text{ şi } x_n \longrightarrow a, \text{ avem } f(x_n) \longrightarrow \ell_s. \end{cases}$$

(Desigur, ℓ_s poate fi şi unul din simbolurile $(+\infty)$ sau $(-\infty)$.) Analog, fie *a punct de acumulare înspre stânga* al lui D (adică există şiruri (x_n) din $D \cap (a, \infty)$ cu $x_n \longrightarrow a$). Zicem că mărimea ℓ_d este *limita la dreapta* a funcţiei f în punctul a, şi scriem aceasta

$$\ell_d = \lim_{x \to a+} f(x) = \lim_{\substack{x \to a \\ x > a}} f(x) = f(a + 0)$$

dacă

$$\text{(D5)} \quad \begin{cases} \text{oricare ar fi şirul } (x_n) \text{ din } D \setminus \{a\} \\ \text{cu } x_n > a, \forall n, \text{ şi } x_n \longrightarrow a, \text{ avem } f(x_n) \longrightarrow \ell_d. \end{cases}$$

(Ca şi mai sus, $\ell_d = +\infty$ sau $\ell_d = -\infty$ sunt cazuri permise.) Relaţia acestor limite laterale cu limita (bilaterală) este dată de formula

(RL) $\lim\limits_{x \to a} f(x) = f(a-0) = f(a+0)$.

Adică: dacă există limită, atunci există şi limite laterale egale cu aceasta; şi, reciproc, dacă limitele laterale sunt egale, valoarea lor comună reprezintă tocmai limita în acel punct (a funcţiei considerate).

În ce priveşte calculele cu limite de funcţii, acestea rezultă în esenţă din cele spuse la şiruri. Mai precis, avem următorul rezultat sintetic:

Teorema 4.2.2 *Fie $f : D \longrightarrow R$, $g : D \longrightarrow R$ două funcţii pentru care există $\lambda = \lim\limits_{x \to a} f(x)$, $\mu = \lim\limits_{x \to a} g(x)$. Atunci*

(P1) $\lim\limits_{x \to a}(f(x) + g(x)) = \lambda + \mu$ *(caz exceptat: $\infty - \infty$);*

(P2) $\lim\limits_{x \to a}(f(x)g(x)) = \lambda\mu$ *(caz exceptat: $0 \cdot \infty$);*

(P3) $\lim\limits_{x \to a}(f(x)/g(x)) = \lambda/\mu$ *(cazuri exceptate: $0/0$, ∞/∞);*

(P4) $\lim\limits_{x \to a}(f(x))^{g(x)} = \lambda^{\mu}$ *(cazuri exceptate: ∞°, 0°, 1^{∞});*

(P5) $f(x) \leq g(x)$, $\forall x \in D \Longrightarrow \lambda \leq \mu$;

(P6) $f(x) \leq h(x) \leq g(x)$, $\forall x \in D$, $\lambda = \mu \Longrightarrow \lim\limits_{x \to a} h(x) = \lambda = \mu$.

Toate formulele de mai sus sunt valabile şi pentru limite laterale. O situaţie importantă în care există asemenea limite este descrisă de

Teorema 4.2.3 *Orice funcţie monotonă admite limite laterale finite în orice punct de acumulare al mulţimii de definiţie.*

În ce priveşte unele cazuri de excepţie, menţionăm formulele

$$\lim_{x \to a-} \frac{1}{x-a} = -\infty, \quad \lim_{x \to a+} \frac{1}{x-a} = +\infty,$$

precum şi limita remarcabilă

$$\lim_{x \to 0}(1+x)^{1/x} = e.$$

Mai adăugăm la acestea şi următoarea egalitate utilă în practică (formula schimbării de variabilă)

$$\text{(P7)} \quad \begin{cases} \lim\limits_{x \to a} f(u(x)) = \lim\limits_{y \to b} f(y) \text{ dacă această ultimă limită există} \\ \text{şi dacă } \lim\limits_{x \to a} u(x) = b. \end{cases}$$

Exemplu. Să se calculeze $\lim\limits_{x \to a}(1 + u(x))^{1/u(x)}$, dacă $\lim\limits_{x \to a} u(x) = 0$. Făcând $y = u(x)$ şi folosind formula (P7), avem

$$\lim_{x \to a}(1 + u(x))^{1/u(x)} = \lim_{y \to 0}(1 + y)^{1/y} = e.$$

Consideraţiile dezvoltate până aici pot fi extinse la funcţii de o variabilă cu valori vectoriale. Anume, fie $f : D \longrightarrow R^p$, o asemenea funcţie, unde p este un număr natural dat. Putem scrie

$$f(x) = (f_1(x), ..., f_p(x)),\ x \in D$$

unde $f_1, ..., f_p$ sunt funcţii de o variabilă (definite pe D) cu valori reale. Dacă a este punct de acumulare al lui D, atunci prin definiţie

(D6) $\lim\limits_{x \to a} f(x) = (\lim\limits_{x \to a} f_1(x), ..., \lim\limits_{x \to a} f_p(x))$.

Calculele cu limite de acest fel sunt reductibile la acelea (precizate anterior) pentru funcţii reale.

Exemplu. Să se studieze limita în origine a funcţiei

$$f(x) = \left(\frac{\sin x}{x}\ ,\ \frac{3 - x^2}{1 + x^2}\right),\ x \neq 0.$$

Avem, evident,

$$\lim_{x \to 0} f(x) = \left(\lim_{x \to 0}\frac{\sin x}{x}\ ,\ \lim_{x \to 0}\frac{3 - x^2}{1 + x^2}\right) = (1, 3).$$

4.2.2 Funcţii de mai multe variabile. Limită într-un punct

Fie m număr natural dat, şi D o anumită submulţime din R^m. Presupunem că s-a indicat o lege prin care fiecărui $x = (x(1), ..., x(m))$ din D îi corespunde un număr real $f(x) = f(x(1), ..., x(m))$. Zicem atunci că am definit o funcţie $f : D \longrightarrow R$; iar D se va numi *mulţimea de definiţie* a acelei funcţii. În cazul unui număr mic de dimensiuni ($m = 2$ sau $m = 3$ fiind cazurile standard pentru noi), notaţia indicială nu este indicată; astfel că – pentru $m = 2$, de pildă – vom nota punctul curent al mulţimii $D \subseteq R^2$ prin (x, y), iar valoarea funcţiei în acest punct prin $f(x, y)$.

Să considerăm dată o funcţie $f : D \longrightarrow R$ şi fie $a = (a(1), ..., a(m))$ un punct de acumulare al lui D. Zicem că mărimea $\ell \in R$ este *limita funcţiei f în punctul a* şi scriem $\ell = \lim_{x \to a} f(x)$, dacă

$$\text{(D1)} \quad \begin{cases} \text{oricare ar fi şirul vectorial } (x_n = (x_n(1), ..., x_n(m))) \text{ din} \\ D \setminus \{a\} \text{ cu } x_n \longrightarrow a, \text{ avem } f(x_n) \longrightarrow \ell. \end{cases}$$

De fapt, ℓ poate fi aici chiar $(+\infty)$ sau $(-\infty)$. Definiţia este formal identică cu aceea din cazul $m = 1$. Pentru a sesiza unele detalii tehnice ale acesteia să studiem cazul $m = 2$. Fie deci $f : D \longrightarrow R$ o funcţie definită pe mulţimea $D \subseteq R^2$ şi (a, b) un punct de acumulare al lui D. Zicem că mărimea ℓ este limita funcţiei f în punctul (a, b) şi scriem $\ell = \lim_{\substack{x \to a \\ y \to b}} f(x, y)$, dacă

$$\text{(D2)} \quad \begin{cases} \text{oricare ar fi şirul vectorial } ((x_n, y_n)) \text{ din } D \setminus \{(a, b)\} \\ \text{cu } x_n \longrightarrow a, \ y_n \longrightarrow b, \text{ avem } f(x_n, y_n) \longrightarrow \ell. \end{cases}$$

Aspectele de principiu discutate în cazul $m = 1$ rămân valabile şi în acest cadru vectorial (deci şi pentru $m = 2$). Precizăm doar (din motive practice) situaţia în care funcţia de două variabile f nu are limită în punctul de acumulare (a, b) :

$$\text{(D3)} \quad \begin{cases} \text{există două şiruri vectoriale } ((x'_n, y'_n)),\ ((x"_n, y"_n)) \text{ din} \\ D \setminus \{(a, b)\}, \text{cu } x'_n, x"_n \longrightarrow a,\ y'_n, y"_n \longrightarrow b, \\ \text{pentru care } \lim_{n \to \infty} f(x'_n, y'_n) \neq \lim_{n \to \infty} f(x"_n, y"_n). \end{cases}$$

Exemplu. Fie dată funcţia

$$f(x, y) = \frac{xy}{x^2 + y^2}, \quad x \neq 0 \ \text{ sau } \ y \neq 0.$$

Mulţimea maximală de definiţie este $D = R^2 \setminus \{(0, 0)\}$. Să studiem existenţa limitei în $(0, 0)$. Fie (r_n) un şir de numere reale cu proprietăţile

$r_n \neq 0$, pentru orice n; $\quad r_n \longrightarrow 0$ pentru $n \longrightarrow \infty$.

Definind şirul vectorial

$(x'_n, y'_n) = (r_n, r_n)$, $\forall n$ (deci $(x'_n, y'_n) \longrightarrow (0, 0)$), avem

$f(x'_n, y'_n) = \frac{r_n^2}{2r_n^2} = \frac{1}{2}, \ \forall n \ (\Longrightarrow \lim\limits_{n\to\infty} f(x'_n, y'_n) = \frac{1}{2})$.

De asemenea, pentru şirul vectorial

$(x''_n, y''_n) = (r_n, -r_n), \ \forall n$ (deci $(x''_n, y''_n) \longrightarrow (0,0)$)

avem

$f(x''_n, y''_n) = \frac{-r_n^2}{2r_n^2} = -\frac{1}{2}, \ \forall n \ (\Longrightarrow \lim\limits_{n\to\infty} f(x''_n, y''_n) = -\frac{1}{2})$.

Deci,cu (D3), f nu are limită în punctul $(0,0)$.

Toate formulele de calcul stabilite pentru cazul $m = 1$ sunt valabile şi în acest cadru vectorial; nu mai dăm alte amănunte.

În încheiere, menţionăm că teoria expusă anterior poate fi extinsă la funcţiile de mai multe variabile cu valori vectoriale $f : D \longrightarrow R^p$, unde $D \subseteq R^m$; modul de lucru este, desigur, cel indicat de cazul $m = 1$.

4.2.3 Funcţii continue de o variabilă

Fie $f : D \longrightarrow R$ o funcţie definită pe mulţimea (nevidă) fără puncte izolate, $D \subseteq R$; să mai ne dăm un punct $a \in D$. Zicem că f este *continuă în punctul* a dacă

(D1) $\begin{cases} \text{există limita funcţiei } f \text{ în punctul } a \\ \text{şi aceasta egalează valoarea } f(a) \text{ (adică } \lim\limits_{x\to a} f(x) = f(a)). \end{cases}$

Definiţia este consistentă; deoarece din ipoteza asupra mulţimii D, orice punct al ei este de acumulare. Intuitiv, continuitatea în acest punct se exprimă astfel: la valori ale variabilei apropiate de a, valorile funcţiei sunt apropiate de $f(a)$. Din posibilitatea de a descrie în mai multe moduri limita rezultă şi aici mai multe modalităţi de a exprima continuitatea în punct; nu mai dăm amănunte. Următoarele observaţii se impun în acest context

(i) dacă a este punct de acumulare bilateral al lui D, egalitatea din definiţia continuităţii apare ca o suprapunere a egalităţilor

(D2) $f(a - 0) = f(a); \ f(a + 0) = f(a)$.

Prima dintre ele înseamnă *continuitatea la stânga* a funcţiei f în punctul $a \in D$; iar a doua, *continuitatea la dreapta* a funcţiei f în punctul $a \in D$. Cu alte cuvinte, f este continuă într-un astfel de punct dacă şi numai dacă este simultan continuă la stânga şi la dreapta în acest punct.

(ii) dacă a este doar punct de acumulare unilateral (spre dreapta sau spre stânga) al lui D, atunci egalitatea din definiţia continuităţii înseamnă doar una (prima sau a doua) din relaţiile (D2). Deci, într-un asemenea caz, "continuă" înseamnă doar "continuă la stânga" (respectiv, "continuă la dreapta"); deoarece proprietatea de continuitate simetrică acesteia nu mai poate fi discutată (fiind fără obiect).

Fie din nou $f : D \longrightarrow R$ o funcţie definită pe o mulţime D fără puncte izolate. Zicem că aceasta este *discontinuă* în punctul $a \in D$ dacă nu sunt întrunite condiţiile din definiţia (D1) (a continuităţii în acest punct). Mai precis, aceasta înseamnă fie afirmaţia

(a) există limita funcţiei în acest punct dar este distinctă de valoarea funcţiei în punctul respectiv,

fie afirmaţia

(b) nu există limita funcţiei f în acest punct.

Fiecare din aceste posibilităţi descrie o clasă (speţă) de discontinuităţi. Anume, vom zice că punctul de discontinuitate $a \in D$ este

(c1) *de prima speţă* (clasă) dacă limitele laterale în acest punct există şi sunt finite,

(c2) *de a doua speţă* (clasă) dacă nu este de prima speţă; adică, cel puţin una din limitele laterale este infinită sau nu există.

Un exemplu important în acest sens care completează unul anterior este furnizat prin

Teorema 4.2.4 *Orice funcţie monotonă (definită pe astfel de mulţimi D) nu are decât puncte de discontinuitate de prima speţă. În plus, mulţimea acestora este cel mult numărabilă.*

Să ne întoarcem la funcţiile continue. În ce priveşte calculele cu acestea, avem următorul rezultat sintetic.

Teorema 4.2.5 *Dacă $f : D \longrightarrow R$ şi $g : D \longrightarrow R$ şi sunt funcţii continue în punctul a, atunci*

(P1) *$f + g$ şi fg sunt continue în punctul a;*

(P2) *f/g este continuă în punctul a, dacă $g(a) \neq 0$;*

(P3) *f^g este continuă în punctul a, dacă simultan $f(a) > 0$, $g(a) > 0$.*

De asemenea, dacă $f : R \longrightarrow R$ *este continuă în punctul* $b = g(a)$, *iar* $g : D \longrightarrow R$ *este continuă în punctul* a, *atunci funcţia compusă* $h : D \longrightarrow R$ *definită de*

$$h = f \circ g \ (\textit{adică } h(x) = f(g(x)),\ x \in D)$$

este continuă în punctul a.

Exemple concrete de astfel de funcţii vor putea fi extrase din

Teorema 4.2.6 *Următoarele funcţii elementare sunt continue în fiecare punct al mulţimii lor* (*maximale*) *de definiţie*

(e1) *funcţiile putere şi radical;*

(e2) *funcţiile exponenţiale şi logaritmice;*

(e3) *funcţiile trigonometrice şi inversele lor.*

Trecem acum la unele proprietăţi speciale ale funcţiilor continue. Începem întâi cu următoarea proprietate locală. Fie $f : D \longrightarrow R$ o funcţie definită pe mulţimea D (fără puncte izolate).

Teorema 4.2.7 *Dacă* f *este continuă în* $a \in D$, *atunci există* α, β *şi vecinătatea* $V \in \mathcal{V}(a)$ *încât* $\alpha \leq f(x) \leq \beta$, $\forall x \in D \cap V$. (*Adică,* f *este local mărginită în* a). *Iar dacă în plus* $f(a) > 0$ (*respectiv* $f(a) < 0$), *atunci* $\alpha > 0$ (*respectiv* $\beta < 0$); *deci, semnul valorilor funcţiei pe* $D \cap V$ *coincide cu semnul lui* $f(a)$.

Trecând la proprietăţi globale, fie $f : D \longrightarrow R$ funcţie dată unde D este fără puncte izolate. Zicem că f este *continuă pe submulţimea* $D_1 \subseteq D$ dacă ea este continuă în fiecare punct al lui D_1. Vom lua de obicei $D_1 = D$.

Teorema 4.2.8 *Fie* $f : D \longrightarrow R$ *continuă pe* D. *Atunci*

(A) (Principiul de conexiune Cauchy) *Dacă* D *este conexă* (*deci interval*), *atunci* $E = f(D)$ *este şi ea conexă* (*deci interval*). *În particular,* f *are proprietatea Darboux pe* D, *în sensul*

(P4) $x_1, x_2 \in D,\ \lambda \in (f(x_1), f(x_2)) \Longrightarrow$ *există* $x_3 \in (x_1, x_2)$ *cu* $f(x_3) = \lambda$.

Iar dacă, în plus, f este şi strict monotonă pe D, atunci inversa $f^{-1} : E \longrightarrow D$ *este continuă şi strict monotonă (de acelaşi sens cu f).*

(B) (Principiul de compacitate Weierstrass) *Dacă D este compactă, atunci* $E = f(D)$ *este, de asemenea, compactă. Deci, în particular, combinând şi cu cele anterioare*

$$\text{(P5)} \quad \begin{cases} D = [a, b] \Longrightarrow f(D) = [m, M] \text{ unde } m = \inf\{f(x), x \in D\}, \\ M = \sup\{f(x); x \in D\} \text{ sunt efectiv atinse.} \end{cases}$$

În acelaşi context, f este chiar uniform continuă pe D, în sensul

$$\text{(P6)} \quad \begin{cases} \text{pentru orice } \varepsilon > 0, \text{ există } \delta > 0 \text{ aşa ca } x', x'' \in D, \\ |x' - x''| < \delta \Longrightarrow |f(x') - f(x'')| < \varepsilon. \end{cases}$$

Adică: în definiţia continuităţii de tip (ε, δ)*, numărul* δ *(care depinde în general de punctul în care se scrie continuitatea) poate fi făcut dependent doar de numărul* ε*.*

În încheiere, menţionăm că studiul continuităţii funcţiilor cu valori vectoriale se reduce, prin schemele deja expuse, la acela al funcţiilor cu valori reale.

4.2.4 Funcţii continue de mai multe variabile

Fie m un număr natural dat şi D o submulţime nevidă a lui R^m, fără puncte izolate. Fie, de asemenea, a punct fixat în D. Zicem că funcţia f este *continuă în punctul* a dacă condiţia (D1) (Secţiunea 4.2.3) este (formal) valabilă; sau, echivalent,

$$\text{(D1)} \quad \begin{cases} \text{pentru orice vecinătate } V \in \mathcal{V}(f(a)) \text{ există} \\ \text{vecinătatea } U \in \mathcal{V}(a) \text{ aşa ca } x \in D \cap U \Longrightarrow f(x) \in V. \end{cases}$$

Menţionăm că, formal, definiţia se păstrează şi pentru funcţiile $f : D \longrightarrow R^p$, unde p este un număr natural. De fapt, o astfel de funcţie se scrie sub forma $f = (f_1, ..., f_p)$, adică

$$\text{(D2)} \quad f(x) = (f_1(x), ..., f_p(x)), \; x \in D.$$

unde pentru fiecare $i \in \{1, ..., p\}$, f_i este o funcţie definită pe mulţimea $D \subseteq R^m$ cu valori reale. Mai mult, avem echivalenţa

(P) $\quad f$ continuă în $a \Longleftrightarrow f_i$ continuă în $a,\ 1 \leq i \leq p.$

Cu alte cuvinte, studiul continuităţii funcţiilor vectoriale se reduce la acela al funcţiilor cu valori reale.

În ce priveşte calculele cu funcţii continue, avem

Teorema 4.2.9 *Dacă $f : D \longrightarrow R$ şi $g : D \longrightarrow R$ sunt funcţii continue în punctul a, atunci concluziile* (P1)–(P3) *din Teorema* 4.2.5*, pot fi reţinute (formal) şi în acest cadru. În plus, dacă $f : R^p \longrightarrow R$ este continuă în punctul $b = f(a)$, iar $g : D \longrightarrow R^p$ este continuă în punctul a, atunci funcţia compusă $h : D \longrightarrow R$ definită prin convenţia*

$$h = f \circ g \ \textit{(adică } h(x) = f(g(x)),\ x \in D\textit{)}$$

este continuă în punctul a.

Combinând acest rezultat cu Teorema 4.2.6, rezultă o clasă largă de funcţii continue care acoperă, în principiu, toate aplicaţiile uzuale. Pe de altă parte, proprietăţile locale exprimate prin Teorema 4.2.7 se păstrează şi în acest caz. În fine, proprietăţile globale date în Teorema 4.2.8 rămân valabile cu modificări minore.

Exemplu. Fie dată funcţia

$$f(x, y) = x^2 + y^2 + 3x + 2y,\ (x, y) \in R^2.$$

Evident, f este continuă pe $D = R^2$. Considerăm mulţimea

$$K = [0, 2] \times [1, 5] \quad \text{(dreptunghi în plan).}$$

Această mulţime este conexă, deoarece este convexă. Deci, conform celor spuse anterior

$$f(K) = \text{mulţime conexă (deci interval) în } R.$$

Pe de altă parte, mulţimea K este şi compactă, deoarece este mărginită şi închisă; deci

$$f(K) = [m, M] \ \text{(interval compact din } R\text{).}$$

Se observă că $m = f(0, 1) = 3$, $M = f(2, 5) = 45$. Adică, $f(K) = [3, 45]$; iar valorile extreme sunt efectiv atinse, după cum s-a arătat.

În încheiere, subliniem că toate cele spuse pot fi extinse la funcţii vectoriale de mai multe variabile; nu mai dăm amănunte.

Probleme propuse la §4.2

1. Să se arate că următoarele limite nu există

$$\lim_{x\to 0}\frac{1}{x}\sin\frac{1}{x};\quad \lim_{x\to\infty}(1+\cos x)e^{2x};\quad \lim_{x\to 1}e^{\frac{1}{x-1}}.$$

2. Să se calculeze limitele

$$\lim_{x\to 0}\frac{1-\cos x}{x\sin x};\quad \lim_{x\to 0}x\sin\frac{1}{x};\quad \lim_{x\to\infty}\left(\frac{x^2+1}{x^2-1}\right)^{x^2}.$$

3. Să se studieze existenţa limitei în origine pentru funcţia

$$f(x,y)=(x+y)\sin\frac{1}{x}\sin\frac{1}{y}\cdot$$

4. Să se calculeze limitele

$$\lim_{\substack{x\to 0\\ y\to 0}}\frac{xy}{\sqrt{1+xy}-1};\quad \lim_{\substack{x\to 0\\ y\to 2}}\frac{\sin xy}{x};\quad \lim_{\substack{x\to 0\\ y\to 0}}\frac{x^2y}{x^2+y^2}\cdot$$

5. Să se studieze continuitatea funcţiei

$$f(x)=\begin{cases}2x-1, & x=\text{raţional}\\ 3x, & x=\text{iraţional.}\end{cases}$$

(Indicaţie: Punctul de continuitate este acela în care $2x-1=3x$, deci $x=-1$; celelalte sunt puncte de discontinuitate.)

6. Să se arate că funcţia

$$f(x)=\begin{cases}2x, & x\le 0\\ x-1, & x>0\end{cases}$$

are proprietatea Darboux dar nu este continuă (în $x = 0$).

7. Să se studieze continuitatea uniformă a funcţiei (de o variabilă)

$$f(x) = \operatorname{arctg}\frac{1+x}{1-x},\ x \in (1, \infty).$$

8. Să se arate că funcţia

$$f(x) = x - 2\sqrt{x+3},\quad x \in [-3, \infty)$$

este mărginită pe intervalul $[0, 7]$ şi să se calculeze efectiv marginile inferioară şi superioară ale valorilor sale pe acest interval.

9. Să se studieze continuitatea funcţiei

$$f(x) = \begin{cases} \sqrt{1 - x^2 - y^2}, & x^2 + y^2 \leq 1, \\ 0, & x^2 + y^2 > 1. \end{cases}$$

10. Să se calculeze marginea inferioară şi superioară ale valorilor funcţiei (de două variabile)

$$f(x, y) = x^2 + y^2 + x + 3y$$

pe compactul $D = [0, 1] \times [0, 3]$.

11. Să se studieze continuitatea uniformă a funcţiei

$$f(x, y) = \frac{x - y}{x + y},\quad (x, y) \in (1, 2) \times (1, 2).$$

4.3 Teoria diferenţială a funcţiilor

4.3.1 Derivata unei funcţii de o variabilă

Fie $f : I \longrightarrow R$ o funcţie definită pe intervalul I şi $a \in I$ un punct arbitrar fixat. Numim *derivată* a funcţiei f în punctul a mărimea

(D1) $$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \quad \left(= \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \right).$$

Desigur, aceasta poate fi şi unul din simbolurile $(+\infty)$ sau $(-\infty)$. Dacă $f'(a)$ este finită, vom zice că f este *derivabilă* în punctul a. Următoarele observaţii se impun aici:

(i) Dacă a este punct interior al intervalului I, atunci derivata funcţiei f în acest punct apare ca valoare comună a mărimilor

$$f'_s(a) = \lim_{\substack{x\to a\\ x<a}} \frac{f(x)-f(a)}{x-a} \quad (\textit{derivata la stânga} \text{ în } a)$$
$$f'_d(a) = \lim_{\substack{x\to a\\ x>a}} \frac{f(x)-f(a)}{x-a} \quad (\textit{derivata la dreapta} \text{ în } a).$$

În mod corespunzător, f este derivabilă în punctul a dacă şi numai dacă ea este simultan *derivabilă* la *stânga* şi la *dreapta* în acel punct, iar derivatele laterale obţinute sunt (finite şi) egale.

(ii) Dacă a este extremitate (stângă, respectiv dreaptă) a intervalului I, atunci derivata funcţiei f în acest punct apare ca derivată la dreapta (respectiv stânga); deoarece derivata simetrică acesteia nu poate fi definită într-un asemenea caz.

Legătura dintre derivabilitate şi continuitate este precizată de

Teorema 4.3.1 *Orice funcţie derivabilă într-un punct este continuă în acel punct. Reciproca nu este în general valabilă.*

În ce priveşte calculele cu derivatele, avem următorul rezultat sintetic (cu caracter local).

Teorema 4.3.2 *Dacă $f : I \longrightarrow R$ şi $g : I \longrightarrow R$ sunt derivabile în punctul a, atunci*

(P1) *$f+g$ şi fg sunt, de asemenea, derivabile în acest punct cu*

$$(f+g)'(a) = f'(a) + g'(a), (fg)'(a) = f'(a)g(a) + f(a)g'(a);$$

(P2) *f/g este derivabilă în punctul a dacă $g(a) \neq 0$ şi*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)};$$

(P3) f^g *este derivabilă în punctul* a (*dacă* $f(a) > 0, g(a) > 0$) *şi*

$$(f^g)'(a) = (f(a))^{g(a)} \left(g'(a) \ln f(a) + g(a) \frac{f'(a)}{f(a)} \right).$$

De asemenea, dacă $f : R \longrightarrow R$ *este derivabilă în punctul* $a = u(b)$ *unde* $u : I \longrightarrow R$ *este derivabilă în punctul* b*, funcţia compusă* $h : I \longrightarrow R$ *dată de*

$$h(x) = f(u(x)),\ x \in D \quad (\textit{adică } h = f \circ u)$$

este derivabilă în punctul b *cu* $h'(b) = f'(u(b)) \cdot u'(b)$.

Mai menţionăm aici şi următorul rezultat (de tip local–global) privind derivarea funcţiilor inverse:

Teorema 4.3.3 *Fie* $f : I \longrightarrow R$ *continuă şi strict monotonă pe intervalul* I *şi fie* $f^{-1} : J \longrightarrow R$ (*unde* $J = f(I)$ *este intervalul imagine*) *funcţia sa inversă. Dacă* f *este derivabilă în* $a \in I$*, cu* $f'(a) \neq 0$*, atunci* f^{-1} *este derivabilă în* $b = f(a) \in J$ *cu* $(f^{-1})'(b) = 1/f'(a)\cdot$

Exemple concrete de astfel de funcţii derivabile se deduc acum din

Teorema 4.3.4 *Orice funcţie elementară este derivabilă în interiorul domeniului său maximal de definiţie.*

Fie $f : I \longrightarrow R$ o funcţie definită pe intervalul I şi a un punct al acestui interval. Zicem că f este *diferenţiabilă în sens Fréchet* în punctul a, dacă

(D2) f este derivabilă în punctul a,

(D3) $\lim\limits_{x \to a} \dfrac{1}{x-a} \left(f(x) - f(a) - f'(a)(x-a)\right) = 0.$

Desigur, condiţia (D3) se poate scrie echivalent astfel

(D4) $f(x) - f(a) = f'(a)(x-a) + \varphi(x)(x-a), \forall x \in I,$

unde funcţia $\varphi : I \longrightarrow R$ are proprietăţile

$$\varphi(a) = 0,\ \lim_{x \to a} \varphi(x) = 0.$$

Funcţia liniară $df(a) : R \longrightarrow R$ definită de

(D5) $df(a)(h) = f'(a) \cdot h, \quad h \in R$

se va numi *diferenţiala* funcţiei f în punctul curent a din I. Ea reprezintă *partea principală* a creşterii funcţiei f în jurul punctului a:

$$f(a+h) - f(a) \approx f'(a) \cdot h, \quad \text{dacă } |h| \text{ este mic.}$$

Din însăşi această definiţie, rezultă că o funcţie diferenţiabilă (Fréchet) în $a \in I$ este derivabilă în I. Reciproca este, de asemenea, adevărată din (D4); deci avem

Teorema 4.3.5 *Funcţia f este diferenţiabilă (Fréchet) în a dacă şi numai dacă este derivabilă în a.*

Să observăm acum că pentru funcţia identică

$$i(x) = x, \; x \in R$$

diferenţiala acesteia în orice punct este

$$di(a)(h) = h, \; h \in R, \; a \in R.$$

Vom face convenţia de notaţie

$$di(a) = dx, \; a \in R.$$

În acest caz, diferenţiala unei funcţii are exprimarea (notând punctul generic a prin x)

$$df(x) = f'(x) \cdot dx \quad \left(\Longrightarrow f'(x) = \frac{df(x)}{dx} \right).$$

De aici provine şi scrierea derivatei într-un punct ca un cât de diferenţiale. Regulile de calcul cu diferenţiale se reduc imediat la acelea pentru derivate; nu mai dăm alte detalii.

Fie din nou $f : I \longrightarrow R$ o funcţie dată. Zicem că aceasta este *derivabilă pe submulţimea* $J \subseteq I$ dacă este derivabilă în fiecare punct $a \in J$. În particular, dacă avem proprietatea (globală)

(C1) f este derivabilă pe intervalul I, valorile găsite ale derivatelor definesc ceea ce se cheamă *funcţia derivată* $f' : I \longrightarrow R$. Pentru aceasta se poate pune din nou problema derivării. Să notăm pentru $a \in I$,

$$f''(a) = (f')'(a) = \lim_{x \to a} \frac{f'(x) - f'(a)}{x - a}\cdot$$

Prin definiţie vom numi aceasta *derivata a doua* a funcţiei f în $a \in I$; dacă este finită, vom zice că f este *derivabilă de două ori* în acest punct. Într-un asemenea caz, funcţia pătratică $d^2f(a) : R \longrightarrow R$ definită prin

$$d^2f(a)(h) = f''(a) \cdot h^2, \ h \in R$$

se zice *diferenţiala a doua* din funcţia f în punctul $a \in I$. Cu notaţiile acceptate anterior avem şi aici (notând a prin x)

$$d^2f(x) = f''(x)dx^2 \quad \left(\Longrightarrow f''(x) = \frac{d^2f(x)}{dx^2} \right)$$

cu interpretarea corespunzătoare pentru derivata a doua.

Acum, dacă se admite o ipoteză de forma

(C2) f este derivabilă de două ori pe intervalul I,

valorile găsite ale derivatelor definesc funcţia derivată dublă $f'' : I \to R$. Pentru aceasta, se poate pune din nou problema derivării şi diferenţierii, etc. În general, pentru numărul natural n fixat, vom pune

$$f^{(n)}(a) = \text{derivata de ordinul } n \text{ din funcţia } f \text{ în } a \in I.$$

Funcţia f se va zice *derivabilă de* n *ori* în $a \in I$ dacă $f^{(n)}(a)$ este finit. În acest caz, funcţia $d^nf(a) : R \longrightarrow R$ dată de formula

$$d^nf(a)(h) = f^{(n)}(a)h^n, \quad h \in R$$

se va numi *diferenţiala de ordinul* n a lui f în $a \in D$. Evident, avem şi aici (notând punctul curent a prin x)

$$d^nf(x) = f^{(n)}(x)dx^n \quad \left(\Longrightarrow f^{(n)}(x) = \frac{d^nf(x)}{dx^n} \right).$$

Regulile de calcul cu derivatele de ordin superior se reduc desigur la acelea pentru derivata de ordinul întâi (cu anumite accente specifice). Mai exact, avem

Teorema 4.3.6 *Dacă $f : I \longrightarrow R$, $g : I \longrightarrow R$ sunt derivabile de n ori pe I atunci*

(P3) *$f + g$ şi fg sunt derivabile de n ori pe I cu derivatele exprimate prin formulele (de tip Leibniz)*

(R1) $(f + g)^{(n)}(x) = f^{(n)}(x) + g^{(n)}(x),\ \forall x \in I$

(R2) $(fg)^{(n)}(x) = \sum_{k=0}^{n} C_n^k f^{(n-k)}(x) g^{(k)}(x),\ \forall x \in I$

(P4) *f/g este derivabilă de n ori pe I,dacă $g(x) \neq 0$, $\forall x \in I$*

(P5) *f^g este derivabilă de n ori pe I dacă $f(x) > 0,\ g(x) > 0$, oricare ar fi $x \in I$.*

În plus, dacă $f : R \longrightarrow R$ este derivabilă de n ori pe R, iar $g : R \longrightarrow R$ este derivabilă de n ori pe I, atunci funcţia compusă

$$h(x) = f(g(x)),\ x \in I \quad (\textit{adică } h = f \circ g)$$

este derivabilă de n ori pe intervalul I.

Consideraţiile dezvoltate anterior se pot aplica şi la funcţii de o singură variabilă cu valori vectoriale. Fie $f : I \longrightarrow R^p$ o asemenea funcţie unde I este un interval iar p un număr natural. După cum se ştie, putem reprezenta f sub forma

$$f(x) = (f_1(x), ..., f_p(x)),\ x \in I$$

unde $f_1, ..., f_p$ (componentele lui f) sunt funcţii de o variabilă (definite pe D) cu valori reale. Dacă $a \in I$ este arbitrar fixat, atunci punem

(D6) $f'(a) = (f_1'(a), ..., f_p'(a))$,

iar f se va zice *derivabilă în punctul a* dacă toate componentele sale sunt derivabile în punctul considerat. Calculul cu asemenea derivate se face după regulile anterioare. În particular, se pot considera şi funcţii cu valori matriciale, $F : I \longrightarrow \mathcal{M}(p, q)$ de forma

$$F(x) = (a_{ij}(x)),\ x \in I,\ 1 \leq i \leq p,\ 1 \leq j \leq q.$$

Derivata fiind introdusă ca mai sus, remarcăm, de exemplu, formula derivării unui produs matricial

$$(FG)'(x) = F'(x)G(x) + F(x)G'(x).$$

Aici, $F : I \longrightarrow \mathcal{M}(p, q)$ este ca mai sus, iar $G : I \longrightarrow \mathcal{M}(q, r)$ este o altă funcţie cu valori matriciale.

Funcţia $f : I \longrightarrow R$ se zice *indefinit derivabilă în punctul* a dacă $f^{(n)}(a)$ există, pentru orice n; şi, *indefinit derivabilă pe submulţimea* $J \subseteq I$, dacă aceasta se întâmplă în orice punct $a \in J$. Exemple imediate de asemenea funcţii sunt cele elementare (luate pe interiorul domeniilor maximale de definiţie). Definiţia se transferă şi la funcţii vectoriale de o variabilă în modul descris deja.

4.3.2 Funcţii de mai multe variabile. Derivate parţiale

Fie m număr natural dat şi D un domeniu (mulţime deschisă şi conexă) din R^m. (De exemplu, D poate fi o sferă deschisă din R^m.) Considerăm dată o funcţie $f : D \longrightarrow R$ şi un punct $a = (a_1, ..., a_m)$ din D. Pentru fiecare $i \in \{1, ..., m\}$, numărul (finit sau infinit)

$$\text{(D1)} \quad \frac{\partial f}{\partial x_i}(a) = \lim_{h \to 0} \frac{f(a_1, ..., a_i + h, ..., a_m) - f(a_1, ..., a_i, ..., a_m)}{h}$$

se va numi *derivata parţială,* faţă de x_i, din funcţia f în punctul a. Dacă aceasta este finită, f se va zice *derivabilă parţial faţă de variabila* x_i *în punctul* a.

Să observăm că derivata faţă de variabila x_i înseamnă derivata în sens uzual a funcţiei parţiale (de o variabilă)

$$f_i(x_i) = f(a_1, ..., a_{i-1}, x_i, a_{i+1}, ..., a_m)$$

menţinând celelalte variabile constante. Funcţia f se zice *derivabilă parţial în punctul* a dacă există $\dfrac{\partial f}{\partial x_i}(a)$ pentru $i \in \{1, ..., m\}$. În acest caz, numim *gradient* al lui f în a vectorul (din R^m)

$$\text{(D2)} \quad \nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), ..., \frac{\partial f}{\partial x_m}(a) \right).$$

Aici, operatorul asociat

$$\nabla = \left(\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_m}\right)$$

se va zice *operatorul nabla* (al lui Hamilton). În fine, f se zice *derivabilă parţial pe submulţimea* $D_1 \subseteq D$ dacă este derivabilă parţial în orice punct al lui D_1.

Exemplu. Fie dată funcţia de două variabile

$$f(x, y) = x^3y^2 + \frac{1}{xy}, \ x > 0, \ y > 0.$$

Funcţia este derivabilă parţial pe întreg domeniul său de definiţie cu

$$\frac{\partial f}{\partial x}(x, y) = 3x^2y^2 - \frac{1}{x^2y}, \quad \frac{\partial f}{\partial y}(x, y) = 2x^3y - \frac{1}{xy^2},$$
$$\nabla f(x, y) = \left(3x^2y^2 - \frac{1}{x^2y}, 2x^3y - \frac{1}{xy^2}\right).$$

Calculul cu derivate parţiale se face în esenţă după modelul Teoremei 4.3.2 , cu excepţia cazului compunerilor de funcţii.

Fie m şi p două numere naturale. Fixăm un domeniu D în R^m şi considerăm funcţia vectorială $f : D \longrightarrow R^p$. Aceasta, după cum ştim, admite reprezentarea

$$f(x) = (f_1(x), ..., f_p(x))^\top, \ x \in D$$

unde $f_1, ..., f_p$ sunt funcţii de m variabile (definite pe D) cu valori reale. Pentru fiecare $a \in D$ şi fiecare $i \in \{1, ..., m\}$, notăm

$$\frac{\partial f}{\partial x_i}(a) = \left(\frac{\partial f_1}{\partial x_i}(a), ..., \frac{\partial f_p}{\partial x_i}(a)\right)^\top, \ \text{(vector din } R^p\text{)}$$

vom numi aceasta *derivata parţială faţă de* x_i din funcţia vectorială f *în punctul* a; iar f se va zice *derivabilă parţial faţă de* x_i *în* a, dacă fiecare din derivatele parţiale din dreapta este finită. De asemenea, f se va numi *derivabilă parţial în* $a \in D$ dacă $\frac{\partial f}{\partial x_i}(a)$ există pentru $i \in \{1, ..., m\}$. În acest caz, matricea

(D3) $\dfrac{Df}{Dx}(a) = \left(\dfrac{\partial f_i}{\partial x_j}(a)\right)$ (element din $\mathcal{M}(p, m)$), obţinută punând unul sub altul gradienţii componentelor, se va numi *matricea jacobiană* a funcţiei f în a.

Trecem în continuare la formule de derivare pentru funcţiile compuse. Acestea, după cum se va vedea, necesită un concept potrivit de diferenţiabilitate. Fie $f : D \longrightarrow R$, o funcţie definită pe domeniul $D \subseteq R^m$, şi a punct arbitrar fixat din D. Zicem că f este *diferenţiabilă în sens Fréchet în punctul* a, dacă

(D4) f este derivabilă parţial în a

(D5) $\lim\limits_{x \to a} \dfrac{1}{\|x-a\|}(f(x) - f(a) - \nabla f(a) \cdot (x-a)) = 0.$

(Să notăm că $\nabla f(a) \cdot (x - a)$ este produsul unei matrici linie cu o matrice coloană.) Echivalent, (D5) ne spune că avem reprezentarea

$$f(x) - f(a) = \nabla f(a)(x-a) + \varphi(x)\|x-a\|,\ x \in D$$

unde funcţia $\varphi : D \longrightarrow R$ satisface condiţiile

$$\varphi(a) = 0,\ \lim_{x \to a} \varphi(x) = 0.$$

Se observă că pentru $m = 1$ definiţia coincide cu aceea dată deja. Că această extensie este potrivită în contextul de faţă, o demonstrează

Teorema 4.3.7 *Orice funcţie diferenţiabilă Fréchet în punctul a este continuă în acel punct.* (*Reciproca nu este în general valabilă.*)

Din însăşi această definiţie,rezultă că o funcţie diferenţiabilă Fréchet în punctul a este derivabilă parţial în acel punct.Se pune problema dacă reciproca este adevărată. Răspunsul este negativ, după cum rezultă din următorul

Exemplu. Fie dată funcţia de două variabile

$$f(x, y) = \begin{cases} \frac{x^4y^2}{x^8+y^4}, & \text{dacă } (x, y) \neq (0, 0), \\ 0, & \text{dacă } (x, y) = (0, 0). \end{cases}$$

Se constată că f este derivabilă parţial în origine, cu

$$\frac{\partial f}{\partial x}(0, 0) = 0,\ \frac{\partial f}{\partial y}(0, 0) = 0.$$

Cu toate acestea, f nu este diferenţiabilă Fréchet în $(0,0)$ deoarece nu este continuă în acest punct.

Apare ca justificată următoarea întrebare: în ce condiţii suplimentare avem asigurată totuşi o asemenea proprietate? Pentru aceasta, vom introduce următoarea convenţie. Zicem că f este *de clasă* C^1 *pe* D şi scriem $f \in C^1(D)$ dacă

(D6) f este derivabilă parţial în fiecare punct din D;

(D7) aplicaţiile $y \vdash \dfrac{\partial f}{\partial x_i}(y)$ 3unt continue pe D, $\forall i \in \{1, ..., m\}$.

Avem următorul rezultat (pe care îl dăm fără demonstraţie):

Teorema 4.3.8 *Să admitem ipoteza*

(C1) *f este de clasă C^1 pe D.*

Atunci, f este diferenţiabilă Fréchet în orice punct din D.

Pentru o asemenea funcţie să numim forma liniară $df(a) : R^m \to R$

(D8) $df(a)(h_1, ..., h_m) = \displaystyle\sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(a)h_i, \quad (h_1, ..., h_m) \in R^m,$

diferenţiala Fréchet a funcţiei f *în punctul* $a \in D$.

Să observăm că pentru funcţiile de proiecţie

$$i_k(x) = x_k,\ x = (x_1, ..., x_k) \in R^m,\ 1 \leq k \leq m$$

diferenţialele acestora în punctul $a \in R^m$ sunt

$$di_k(a)(h) = h_k,\ h = (h_1, ..., h_m) \in R^m,\ 1 \leq k \leq m.$$

Cu convenţia de notaţie

$$di_k(a) = dx_k,\ 1 \leq k \leq m.$$

diferenţiala funcţiei f are forma (notând punctul generic a prin x)

$$df(x) = \frac{\partial f}{\partial x_1}(x)dx_1 + \cdots + \frac{\partial f}{\partial x_m}(x)dx_m$$

iar operatorul de diferenţiere asociat este

$$d = \frac{\partial}{\partial x_1} dx_1 + \cdots + \frac{\partial}{\partial x_m} dx_m.$$

Putem acum să trecem la formula de derivare a funcţiilor compuse. Fie dată funcţia vectorială de m variabile $f : R^m \longrightarrow R^p$, prin formula

$$f(x_1, ..., x_m) = (f_1(x_1, ..., x_m), ..., f_p(x_1, ..., x_m)),\ (x_1, ..., x_m) \in R^m.$$

Să efectuăm schimbarea de variabilă

$$x_1 = u_1(t_1, ..., t_q), ..., x_m = u_m(t_1, ..., t_q),\ (t_1, ..., t_q) \in B,$$

prin intermediul funcţiei vectoriale de q variabile $u : B \longrightarrow R^m$ ($B =$ domeniu din R^q) dată de formula

$$u(t_1, ..., t_q) = (u_1(t_1, ..., t_q), ..., u_m(t_1, ..., t_q)),\ (t_1, ..., t_q) \in B.$$

Făcând înlocuirile indicate, se obţine funcţia vectorială compusă $F : B \longrightarrow R^p$, dată de formula

$$F(t_1, ..., t_q) = (F_1(t_1, ..., t_q), ..., F_p(t_1, ..., t_q)),\ (t_1, ..., t_q) \in B$$

unde, pentru $i \in \{1, ..., p\}$

$$F_i(t_1, ..., t_q) = f_i(u_1(t_1, ..., t_q), ..., u_m(t_1, ..., t_q)),\ (t_1, ..., t_q) \in B.$$

Teorema 4.3.9 *Să admitem condiţiile*

(C2) $f_1, ..., f_p$ *sunt de clasă* C^1 *pe* R^m

(C3) $u_1, ..., u_m$ *sunt de clasă* C^1 *pe* B.

Atunci, funcţia vectorială compusă F *este de clasă* C^1 *pe* B *cu*

(R1) $\displaystyle \frac{\partial F_i}{\partial t_k}(b) = \sum_{j=1}^{m} \frac{\partial f_i}{\partial x_j}(u(b)) \frac{\partial x_j}{\partial t_k}(b),\ 1 \le i \le p,\ 1 \le k \le q,\ b \in B,$

sau, echivalent, (*prin matricile jacobiene*)

$$\frac{DF}{Dt}(b) = \frac{Df}{Dx}(u(b)) \frac{Dx}{Dt}(b),\ b \in B.$$

Demonstraţie. Va fi suficient să dovedim cazul $p = 1$, deoarece condiţiile asupra componentelor lui f sunt identice. De asemenea, putem să ne limităm la cazul $q = 1$ din cauză că, în procesul derivării parţiale în raport cu o variabilă, celelalte sunt menţinute constante. Avem deci de verificat că, pentru funcţia compusă

$$F(t) = f(u_1(t), ..., u_m(t)),\ t \in B$$

are loc formula de derivare

$$F'(b) = \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}(u(b))u_j'(b),\ b \in B.$$

În acest scop, vom pleca de la faptul că ipotezele puse asigură diferenţiabilitatea Fréchet pentru f în punctul $u(b) \in R^m$. Deci,$\forall t \in B$,

$$f(u(t)) - f(u(b)) = \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}(u(b))(u_j(t) - u_j(b)) + \|u(t) - u(b)\|\varphi(u(t))$$

unde funcţia $\varphi : R^m \longrightarrow R$ satisface

$$\varphi(u(b)) = 0, \quad \lim_{x \to u(b)} \varphi(x) = 0.$$

Nu avem decât să împărţim acum cu $t - b$ şi să luăm limita pentru $t \to b$, ca să obţinem rezultatul cerut. ■

Exemplu. Fie dată funcţia de trei variabile $f(x, y, z)$. Cu schimbările

$$x = t^2 + s^2, \quad y = ts, \quad z = \frac{t}{s},$$

obţinem funcţia compusă de două variabile

$$F(t, s) = f(t^2 + s^2,\ ts,\ \frac{t}{s}).$$

Să calculăm derivatele parţiale ale acesteia. Avem

$$\frac{\partial F}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t} = 2t\frac{\partial f}{\partial x} + s\frac{\partial f}{\partial y} + \frac{1}{s}\frac{\partial f}{\partial z}$$
$$\frac{\partial F}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial s} = 2s\frac{\partial f}{\partial x} + t\frac{\partial f}{\partial y} - \frac{t}{s^2}\frac{\partial f}{\partial z}\cdot$$

Fie din nou $f : D \longrightarrow R$, o funcţie de m variabile definită pe domeniul $D \subseteq R^m$. Să admitem că pentru un anume $i \in \{1, ..., m\}$

(C4) f este derivabilă parţial faţă de x_i pe D.

Valorile astfel găsite ale derivatelor parţiale definesc ceea ce se cheamă *funcţia derivată parţială* $\frac{\partial f}{\partial x_i} : D \longrightarrow R$; iar operatorul corespunzător $\frac{\partial}{\partial x_i}$,*operatorul de derivare parţială* faţă de x_i. Pentru aceasta, se poate pune din nou problema derivării în raport cu o variabilă, fie aceasta x_j, într-un punct din D. Să notăm

(D9) $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})(a),\ a \in D.$

Vom numi aceasta *derivata mixtă de ordin doi* a funcţiei f, faţă de grupul de variabile (x_i, x_j) (în această ordine) în punctul $a \in D$. În particular, dacă $i = j$, vom nota

$$\frac{\partial^2 f}{\partial x_i^2}(a) = \frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_i})(a),\ a \in D.$$

Aceasta este derivata de ordin doi a funcţiei f faţă de variabila x_i, în punctul $a \in D$.

Următorul rezultat este cunoscut sub numele de " Teorema lui Schwarz". Vom spune că funcţia $f : D \longrightarrow R$ este *de clasă* C^2 *pe* D şi scriem $f \in C^2(D)$, dacă ,pentru toţi $i, j \in \{1, ..., m\}$,

(D10) f este derivabilă parţial pe D faţă de grupul (x_i, x_j),

(D11) aplicaţiile $y \vdash \frac{\partial^2 f}{\partial x_i \partial x_j}(y)$ sunt continue pe D.

Teorema 4.3.10 *Să presupunem că f este de clasă C^2 pe D. Atunci*

(R2) $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a),\ i, j \in \{1, ..., m\},\ a \in D.$

Nu dăm demonstraţia acestui rezultat. Ne mulţumim doar să remarcăm faptul că rezultatul în cauză spune următoarele: indiferent de ordinea în care se aplică doi operatori de derivare parţială asupra unei funcţii (de clasă C^2), rezultatul este acelaşi.

Exemplu. Să considerăm funcţia de două variabile

$$f(x, y) = x^4 y^7 - \frac{x}{y},\ y \neq 0.$$

Vom calcula derivatele mixte de ordin doi. Avem

$$\frac{\partial f}{\partial x} = 4x^3y^7 - \frac{1}{y}, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y}(\frac{\partial f}{\partial x}) = 28x^3y^6 + \frac{1}{y^2}$$
$$\frac{\partial f}{\partial y} = 7x^4y^6 + \frac{x}{y^2}, \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial x}(\frac{\partial f}{\partial y}) = 28x^3y^6 + \frac{1}{y^2}$$

Se vede clar că derivatele mixte obţinute sunt egale.

Admitem în continuare ipoteza

(C5) f este de clasă C^2 pe D.

Pentru o asemenea funcţie, să numim forma pătratică (simetrică) $d^2f(a) : R^m \longrightarrow R$ dată de

(D12) $d^2f(a)(h_1, ..., h_m) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)h_ih_j, \; (h_1, ..., h_m) \in R^m,$

diferenţiala (Fréchet) de ordinul doi a funcţiei f *în punctul* $a \in D$. Matricea (simetrică) a coeficienţilor acestei forme

(D13) $Hf(a) = (\frac{\partial^2 f}{\partial x_i \partial x_j}(a))$ (element din $\mathcal{M}(m)$)

se va numi *matricea hessiană a funcţiei f în $a \in D$*. Cu notaţiile acceptate anterior, avem (notând punctul curent a prin x)

$$d^2f(x) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i dx_j$$

iar operatorul de diferenţiere asociat este

$$d^2 = \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} dx_i dx_j = \left(\sum_i \frac{\partial}{\partial x_i} dx_i \right)^{(2)}$$

(pătratul simbolic al operatorului de diferenţiere d). Acest raţionament poate fi continuat. În general,dat numărul natural r , să admitem

(C6) f este de clasă C^r pe D.

(Sensul acestei convenţii este clar.) Conform Teoremei lui Schwarz, indiferent de ordinea în care aplicăm r operatori de derivare asupra lui f, rezultatul este acelaşi. Pentru o astfel de funcţie, să numim forma omogenă de grad r în m variabile $d^r f(a) : R^m \longrightarrow R$, dată de

$$d^r f(a)(h_1, ..., h_m) = \sum_{i_1,...,i_r} \frac{\partial^r f}{\partial x_{i_1}...\partial x_{i_r}}(a) h_{i_1}...h_{i_r}, \ (h_1, ..., h_m) \in R^m$$

diferenţiala Fréchet, de ordin r, a lui f în $a \in D$. Cu aceleaşi notaţii ca înainte, avem exprimarea generică

$$d^r f(x) = \sum_{i_1,...,i_r} \frac{\partial^r f}{\partial x_{i_1}...\partial x_{i_r}} dx_{i_1}...dx_{i_r},$$

iar operatorul de diferenţiere asociat este

$$d^r = \sum_{i_1,...,i_r} \frac{\partial^r}{\partial x_{i_1}...\partial x_{i_r}} dx_{i_1}...dx_{i_r} = \left(\sum_i \frac{\partial}{\partial x_i} dx_i \right)^{(r)}$$

(puterea simbolică de ordin r,a operatorului de diferenţiere d). Această afirmaţie se demonstrează prin inducţie.

Calculele cu derivatele mixte urmează în general metodologia cunoscută, combinată cu rezultatul privind derivarea funcţiilor compuse. Un exemplu important de acest tip este oferit în continuare.

Fie $f : R^m \longrightarrow R$ o funcţie de m variabile (notate $x_1, ..., x_m$). Facem schimbarea de variabile

$$x_1 = \lambda_1 t + \mu_1, ..., x_m = \lambda_m t + \mu_m, \ t \in I.$$

Se obţine în felul acesta funcţia compusă (de o variabilă)

$$F(t) = f(\lambda_1 t + \mu_1, ..., \lambda_m t + \mu_m), \ t \in I.$$

Desigur, aceasta nu este decât un caz particular al compunerii din cadrul Teoremei 4.3.9, cu

$$u(t) = (\lambda_1 t + \mu_1, ..., \lambda_m t + \mu_m)^\top, \ t \in I.$$

Ca o completare a rezultatului citat, avem

Teorema 4.3.11 *Să admitem condiţia*

(C7) f *este de clasă* C^r *pe* R^m.

Atunci, funcţia compusă F este, de asemenea, de clasă C^r pe I, cu derivatele succesive date de formulele

$$\text{(R3)}\ \ F^{(k)}(t) = \left(\sum_i \lambda_i \frac{\partial}{\partial x_i}\right)^{(k)} f(u(t)), \quad t \in I,\ 1 \leq k \leq r.$$

Demonstraţie. (Schiţă) Cazul $k = 1$ a fost deja dovedit; deoarece (R3) revine atunci la

$$F'(t) = \sum_i \lambda_i \frac{\partial f}{\partial x_i}(u(t)),\ t \in I.$$

Să verificăm (R3) pentru $k = 2$. Notăm pentru $1 \leq i \leq m$,

$$G_i(t) = \frac{\partial f}{\partial x_i}(u(t)) = \frac{\partial f}{\partial x_i}(\lambda_1 t + \mu_1, ..., \lambda_m t + \mu_m),\ t \in I.$$

Aplicând formula de derivare compusă din Teorema 4.3.9, ajungem la

$$G_i'(t) = \sum_j \lambda_j \frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})(u(t)) = \sum_j \lambda_j \frac{\partial^2 f}{\partial x_i \partial x_j}(u(t)),\ t \in I.$$

Înlocuind în relaţia anterioară, găsim

$$\begin{aligned} F''(t) &= \sum_i \lambda_i G_i'(t) = \sum_i \lambda_i (\sum_j \lambda_j \frac{\partial^2 f}{\partial x_i \partial x_j}(u(t))) = \\ &= \sum_{i,j} \lambda_i \lambda_j \frac{\partial^2 f}{\partial x_i \partial x_j}(u(t)), \quad t \in I, \end{aligned}$$

care demonstrează (R3) pentru cazul avut în vedere. Să notăm cu această ocazie că relaţia obţinută se mai scrie (cu notaţiile făcute) astfel

$$F''(t) = (\lambda_1, ..., \lambda_m)^\top Hf(u(t))(\lambda_1, ..., \lambda_m),\ t \in I.$$

Cazul general se tratează acum prin inducţie. ■

Vom spune că funcţia $f : D \longrightarrow R$ este *de clasă C^∞ pe D* şi scriem $f \in C^\infty(D)$, dacă ipoteza (C6) are loc pentru orice r. Desigur, într-un asemenea caz, sunt valabile concluziile

(a) există derivate mixte de orice ordin (ce nu depind de ordinea de aplicare a derivărilor componente);

(b) există diferenţiale de orice ordin (calculabile printr-o formulă de ridicare la putere simbolică).
De asemenea, formula precedentă are loc pentru orice ordin de derivare; nu mai dăm detalii.
Toate aceste consideraţii se aplică, cu modificări corespunzătoare şi la funcţii cu valori vectoriale.

4.3.3 Formule de medie Lagrange–Taylor

Dintre proprietăţile globale ale funcţiilor derivabile, cele mai importante sunt acelea cunoscute sub denumirea " formule de medie".

Începem deocamdată cu funcţiile de o variabilă. Fie $f : I \longrightarrow R$, unde I=interval real. Zicem că f este *funcţie Lagrange* pe I dacă

(D1) f este continuă pe intervalul I

(D2) f este derivabilă măcar în interiorul intervalului I.

Teorema 4.3.12 *Fie f funcţie Lagrange pe I. Atunci, pentru orice $a, b \in I$ cu $a < b$, există punctul intermediar $c \in (a, b)$ cu*

(R1) $f(b) - f(a) = f'(c)(b - a)$.

(*Deci, în particular, dacă $f(a) = f(b)$, atunci $f'(c) = 0$.*)

Prima parte a teoremei alcătuieşte ceea ce se cheamă "Teorema de medie Lagrange"; iar a doua, " Teorema lui Rolle".

Să prezentăm câteva consecinţe mai importante ale acestor rezultate. Fie $f : I \longrightarrow R$ funcţie Lagrange pe I.
(P1) Dacă derivata funcţiei f este pozitivă (negativă) în interiorul lui I, atunci f este crescătoare (descrescătoare) pe I. Deci, în particular, dacă $f'(x) = 0, \forall x \in \operatorname{int}(I)$, atunci f este constantă pe I.
(P2) Dacă derivata funcţiei f este strict pozitivă (negativă) în interiorul lui I, atunci f este strict crescătoare (descrescătoare) pe I.
(P3) Dacă $|f'(x)| \leq \mu$ pentru toţi $x \in \operatorname{int}(I)$, atunci

$$|f(b) - f(a)| \leq \mu|b - a|, \quad \text{pentru orice } a, b \in I.$$

(Zicem atunci că f este *lipschitziană* (de constantă Lipschitz μ)*pe intervalul* I.) Deci, în particular, f este uniform continuă pe I.

(P4) Dacă $a \in I$ este punct de acumulare înspre stânga al lui I, atunci

$$f'_d(a) = \lim_{x \to a+} f'(x) \quad \text{(dacă limita dreaptă există).}$$

(Un rezultat analog se poate da pentru punctele de acumulare înspre dreapta.) Să notăm că o asemenea proprietate se dovedeşte utilă în studiul derivabilităţii funcţiilor de forma

$$f(x) = \begin{cases} \varphi(x), & \text{dacă } x \leq c \\ \psi(x), & \text{dacă } x > c. \end{cases}$$

În fine, ca o aplicaţie a teoremei lui Rolle, avem
(P5) Între două rădăcini consecutive ale ecuaţiei $f'(x) = 0$ există cel mult o rădăcină a ecuaţiei $f(x) = 0$. Aceasta permite o separare a rădăcinilor ecuaţiei $f(x) = 0$ prin intermediul variaţiilor de semn din aşa–numitul *şir al lui Rolle*

$$f(-\infty), f(c_1), ..., f(c_p), f(+\infty),$$

unde $c_1, ..., c_p$ sunt rădăcinile consecutive ale derivatei.

Următoarea variantă utilă a teoremei Lagrange este de remarcat (aşa–numita " Teoremă a lui Cauchy"):

Teorema 4.3.13 *Fie f,g două funcţii Lagrange pe intervalul I, cu*

(C1) $g'(x) \neq 0$, *pentru orice x interior lui I.*

Atunci, pentru orice două puncte $a, b \in I$ cu $a < b$, avem $g(a) \neq g(b)$ şi există punctul intermediar c din (a, b) cu

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

Acest rezultat reprezintă o bază metodologică pentru calculul limitelor prin aşa–numita *metodă a derivării (regula l'Hospital).* Fie I un interval şi b un punct de acumulare spre dreapta al acestuia. (Să notăm că este posibil ca b să nu fie în I; şi, de asemenea, cazul $b = \infty$ nu poate fi exclus). În plus, fie f, g două funcţii Lagrange pe I din care a doua mai satisface condiţia (C1).

$$\text{(P6)} \quad \begin{cases} \text{Cazul } \frac{0}{0}\text{: Dacă } \lim_{x \to b^-} f(x) = 0, \ \lim_{x \to b^-} g(x) = 0, \text{ atunci} \\ \lim_{x \to b^-} \dfrac{f(x)}{g(x)} = \lim_{x \to b^-} \dfrac{f'(x)}{g'(x)}, \text{ dacă limita din dreapta există,} \end{cases}$$

$$(\mathrm{P7})\quad \begin{cases} \text{Cazul } \dfrac{\infty}{\infty}\text{: Dacă } \lim\limits_{x\to b^-} |g(x)| = \infty, \text{ atunci,necesar} \\ \lim\limits_{x\to b^-} \dfrac{f(x)}{g(x)} = \lim\limits_{x\to b^-} \dfrac{f'(x)}{g'(x)}, \text{ dacă limita din dreapta există.} \end{cases}$$

Un rezultat analog are loc şi pentru cazul punctelor de acumulare spre stânga ale lui I.

Acum, toate celelalte nedeterminări se reduc la precedentele, prin schimbări de funcţie. De pildă, cazul $\infty - \infty$ se reduce la $\dfrac{0}{0}$ prin

$$f(x) + g(x) = \frac{1}{\varphi(x)} + \frac{1}{\psi(x)} = \frac{\varphi(x) + \psi(x)}{\varphi(x)\psi(x)}\cdot$$

Pe de altă parte, cazul $0 \cdot \infty$ se reduce la $\dfrac{0}{0}$ prin scrierea

$$f(x)g(x) = f(x) \cdot \frac{1}{\psi(x)} = \frac{f(x)}{\psi(x)}\cdot$$

Iar, în fine, cazurile $\infty^\circ, 0^\circ, 1^\infty$ se reduc la $0 \cdot \infty$ (şi ulterior la $\dfrac{0}{0}$) prin

$$f(x)^{g(x)} = e^{g(x)\ln f(x)}.$$

Dacă limitele din dreapta relaţiilor (P6), (P7) nu există, atunci nu se poate spune că nici limitele din stânga acestor relaţii nu există.

Exemplu. Fie $f(x) = x - \sin x$, $g(x) = 2x + 3 + \cos x$, $b = \infty$. Avem, evident, $\lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)} = \dfrac{1}{2}$ (se scoate x în factor). Pe de altă parte, $\lim\limits_{x\to\infty} \dfrac{f'(x)}{g'(x)} = \lim\limits_{x\to\infty} \dfrac{1 - \cos x}{2 - \sin x}$ nu există (se lucrează cu şiruri).

Trecem acum la o generalizare utilă a formulei de medie Lagrange. Fie $f : I \longrightarrow R$ o funcţie definită pe intervalul I şi n număr natural dat. Zicem că f este *funcţie Taylor de ordinul* n pe I dacă

(D3) $f, f', ..., f^{(n)}$ există şi sunt continue pe I.

(D4) $f^{(n+1)}$ există măcar pe interiorul lui I.

Teorema 4.3.14 *Fie f funcţie Taylor de ordin n pe I. Atunci,date $a, b \in I$, $a \neq b$, există punctul intermediar $\theta = \theta(a, b)$ din $(0, 1)$ cu*

$$\text{(R2)} \quad \begin{cases} f(b){=}f(a) + \dfrac{b-a}{1!}f'(a) + \cdots + \dfrac{(b-a)^n}{n!}f^{(n)}(a) + \\ +\dfrac{(b-a)^{n+1}}{(n+1)!}f^{(n+1)}(a+\theta(b-a)). \end{cases}$$

Demonstraţie. (Schiţă.) Considerăm numărul μ dat de

$$f(b) = f(a) + \frac{b-a}{1!}f'(a) + \cdots + \frac{(b-a)^n}{n!}f^{(n)}(a) + \frac{(b-a)^{n+1}}{(n+1)!}\mu$$

şi definim funcţia $F : I \longrightarrow R$ prin

$$F(x) = f(x) + \frac{b-x}{1!}f'(x) + \cdots + \frac{(b-x)^n}{n!}f^{(n)}(x) + \frac{(b-x)^{n+1}}{(n+1)!}\mu, \ x \in I.$$

Se constată uşor că F este funcţie Lagrange pe I şi, în plus, avem $F(a) = F(b)$ (din alegerea lui μ). Conform Teoremei lui Rolle,

$$F'(a + \theta(b-a)) = 0, \text{ pentru măcar un } \theta = \theta(a, b) \text{ din } (0, 1).$$

Dar, derivata funcţiei F are forma

$$F'(x) = \frac{(b-x)^n}{n!}(f^{(n+1)}(x) - \mu).$$

Egalitatea dinainte ne spune că, obligatoriu $\mu = f^{(n+1)}(a + \theta(b-a))$. Iar de aici concluzia este clară. ■

Prin definiţie, (R2) se va zice *formula de medie Taylor (de ordin n)* pentru funcţia f. Evident, dacă $n = 0$, aceasta nu este altceva decât formula de medie Lagrange. Polinomul de grad n (în variabila x)

$$T_n(x : a) = f(a) + \frac{x-a}{1!}f'(a) + \cdots + \frac{(x-a)^n}{n!}f^{(n)}(a)$$

se va numi *polinom Taylor (de grad n) asociat lui f*, iar expresia (în variabilele x, a)

$$R_n(x : a) = \frac{(x-a)^{n+1}}{(n+1)!}f^{(n+1)}(a + \theta(x-a))$$

restul de ordinul n din formula de medie Taylor. Dacă, în particular, $a = 0$, atunci, din (R2) deducem ceea ce se cheamă *formula de medie Mac–Laurin* (în orice $b \neq 0$)

(R3) $f(b) = f(0) + \frac{b}{1!}f'(0) + \cdots + \frac{b^n}{n!}f^{(n)}(0) + \frac{b^{n+1}}{(n+1)!}f^{(n+1)}(\theta b)$,
pentru un anumit $\theta = \theta(b)$ în $(0, 1)$.

Este important de subliniat cu această ocazie caracterul bilateral al formulei Taylor. Mai exact, dat $a \in I$, formula în cauză este valabilă atât pentru $b > a$, cât şi pentru $b < a$. În primul caz, punctul intermediar $c = a+\theta(b-a)$ este în intervalul (a, b); iar în al doilea, punctul respectiv este în intervalul (b, a).

În particular, dacă $I = R$ şi (pentru orice $x \in R$)

$$f(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n \text{ (polinom de grad cel mult } n),$$

atunci formula de medie Taylor devine

$$f(x) = f(a) + \frac{x-a}{1!}f'(a) + \cdots + \frac{(x-a)^n}{n!}f^{(n)}(a),\ \forall x \in R.$$

Adică, f este identic cu polinomul Taylor de ordin n asociat acestuia (în orice punct a). Această relaţie ne ajută de exemplu, să găsim condiţia ca punctul $x = a$ să fie rădăcină multiplă de ordin k pentru ecuaţia (algebrică) $f(x) = 0$, în sensul descompunerii

$$f(x) = (x-a)^k g(x),\ \forall x;\ g = \text{polinom cu } g(a) \neq 0.$$

Anume, condiţia respectivă se scrie

$$f(a) = 0,\ f'(a) = 0, ..., f^{(k-1)}(a) = 0,\ f^{(k)}(a) \neq 0.$$

Să trecem acum la cazul funcţiilor de mai multe variabile. Fie m număr natural dat şi D, o parte convexă şi deschisă a spaţiului R^m. Să notăm că, din proprietatea de convexitate pentru orice două puncte $a = (a_1, ..., a_m)$ şi $b = (b_1, ..., b_m)$ din D, segmentul de dreaptă

$$[a, b] = \{a + t(b-a) = (a_1 + t(b_1 - a_1), ..., a_m + t(b_m - a_m)); 0 \leq t \leq 1\}$$

rămâne în mulţimea D. Fie, de asemenea, $f : D \longrightarrow R$ o funcţie dată.

Teorema 4.3.15 *Să admitem că f este de clasă C^{n+1} pe D. Atunci, oricare ar fi punctele distincte $a = (a_1, ..., a_m)$, $b = (b_1, ..., b_m)$ din D, există un punct intermediar $\theta = \theta(a, b)$ din $(0, 1)$ aşa încât*

$$
\text{(R4)} \left\{ \begin{array}{l} f(b) = f(a) + \frac{1}{1!}(\sum\limits_{i=1}^{m}(b_i - a_i)\frac{\partial}{\partial x_i})f(a) + ... + \frac{1}{n!}(\sum\limits_{i=1}^{m}(b_i - a_i)\frac{\partial}{\partial x_i})^{(n)}f(a) \\ + \frac{1}{(n+1)!}(\sum\limits_{i=1}^{m}(b_i - a_i)\frac{\partial}{\partial x_i})^{(n+1)}f(a + \theta(b-a)). \end{array} \right.
$$

Demonstraţie. Să definim funcţia compusă

$$\varphi(t) = f(a + t(b-a)),\ 0 \le t \le 1.$$

Condiţiile problemei au drept consecinţă faptul că φ este funcţie Taylor de ordinul n pe intervalul $I = [0,1]$. Putem deci scrie această formulă în raport cu orice pereche de puncte t_1, t_2 din I. Să facem acest lucru pentru $t_1 = 0$, $t_2 = 1$. Avem

$$\varphi(1) = \varphi(0) + \frac{1}{1!}\varphi'(0) + \cdots + \frac{1}{n!}\varphi^{(n)}(0) + \frac{1}{(n+1)!}\varphi^{(n+1)}(\theta),$$

unde θ este un punct intermediar din $(0,1)$. Dar,cu Teorema 4.3.11,

$$\varphi^{(k)}(t) = (\sum_{i=1}^{m}(b_i - a_i)\frac{\partial}{\partial x_i})^{(k)}f(a + t(b-a)),\ 0 \le t \le 1,$$

oricare ar fi $k \in \{1, ..., n+1\}$. Iar de aici, concluzia este clară. ∎

În particular, pentru $n = 0$, rezultatul de mai sus ne spune că pentru funcţia $f \in C^1(D)$ are loc formula de medie (de tip Lagrange)

$$
\text{(R5)}\ \begin{array}{l} f(b_1, ..., b_m) - f(a_1, ..., a_m) = \\ \sum\limits_{i=1}^{m}(b_i - a_i)\dfrac{\partial}{\partial x_i}f(a_1 + \theta(b_1 - a_1), ..., a_m + \theta(b_m - a_m)) \end{array}
$$

unde θ este un punct intermediar în $(0,1)$ care depinde de cuplul $a = (a_1, ..., a_m)$, $b = (b_1, ..., b_m)$. Aceasta, scrisă sub forma vectorială

$$f(b) - f(a) = (\nabla f(a + \theta(b-a)))(b-a),\ \text{unde } \theta \in (0,1)$$

şi combinată cu inegalitatea Cauchy–Schwarz, conduce la concluzia

$$
\text{(P8)} \left\{ \begin{array}{l} \text{Dacă } \|\nabla f(x)\| \le \mu, \text{ pentru orice } x \in D, \text{ atunci} \\ \|f(b) - f(a)\| \le \mu\|b-a\|, \text{ oricare ar fi } a, b \in D. \end{array} \right.
$$

Adică, f este lipschitziană (de constantă Lipschitz μ) pe mulţimea convexă deschisă D. (Şi, deci, în particular, f este uniform continuă pe D.) Vom numi (R4) *formula de medie Taylor (de ordinul n)* pentru funcţia f. Polinomul de grad n (în variabila $x = (x_1, ..., x_m)$), notat

$$T_n(x : a) = f(a) + \sum_{k=1}^{n} \frac{1}{k!}((x_1 - a_1)\frac{\partial}{\partial x_1} + \cdots + (x_m - a_m)\frac{\partial}{\partial x_m})^{(k)} f(a)$$

se va numi *polinomul Taylor (de grad n)* asociat lui f; iar expresia (în variabilele $x = (x_1, ..., x_m)$, $a = (a_1, ..., a_m)$)

$$R_n(x : a) = \frac{1}{(n+1)!}(\sum_{i=1}^{m}(x_i - a_i)\frac{\partial}{\partial x_i})^{(n+1)} f(a + \theta(x - a))$$

restul de ordinul n din formula de medie Taylor. În particular, pentru $a = 0$, (R4) se va numi *formula de medie Mac–Laurin*; scrierea ei fiind imediată, nu mai insistăm . Menţionăm şi aici faptul că pentru

$$f(x_1, ..., x_m) = \text{polinom de grad cel mult } n \text{ (în variabilele sale)}$$

formula de medie Taylor devine exactă, în sensul că restul asociat acesteia este identic nul.

4.3.4 Calcul aproximativ al valorilor funcţionale

Printre aplicaţiile practice ale formulelor de medie anterioare, cea referitoare la calculul aproximativ este de importanţă centrală. În esenţă, ea se reduce la următoarele. Să presupunem dată mărimea

$$\eta = f(x_1, ..., x_m) \text{ (expresie analitică în } m \text{ variabile).}$$

Ne interesează valoarea acesteia pentru un sistem fixat de valori ale variabilelor: $x_1 = b_1, ..., x_m = b_m$; adică, mărimea $\eta_0 = f(b_1, ..., b_m)$. De obicei, aceste valori nu sunt efectiv accesibile. (De pildă, dacă sunt numere iraţionale, nu putem lua în considerare decât un număr finit de zecimale.) În acest caz, se vor lua valori vecine acestora $x_1 = a_1, ..., x_m = a_m$; adică, se va lua în discuţie mărimea $\eta_0' = f(a_1, ..., a_m)$. Două probleme se pun în acest context

(i) în ce măsură şi cu ce eroare se poate construi η_0';

(ii) odată construit η_0', care este eroarea pe care o facem dacă luăm pentru η_0 valoarea η_0'.

Prima problemă se poate rezolva prin intermediul formulei de medie Mac–Laurin. Anume, vom folosi în calculul lui $\eta_0' = f(a_1, ..., a_m)$, aproximarea definită de polinomul Taylor de ordinul n

$$f(a_1, ..., a_m) \approx f(0, ..., 0) + \sum_{k=1}^{n}(a_1\frac{\partial}{\partial x_1} + \cdots + a_m\frac{\partial}{\partial x_m})^{(k)} f(0, ..., 0)$$

sau, cu altă notaţie,

$$f(a) \approx T_n(a; 0), \text{ dacă } a \text{ este apropiat de } 0.$$

Eroarea evaluării se obţine din restul asociat formulei Taylor:

$$|f(a) - T_n(a; 0)| \leq \frac{(m\mu)^{n+1}}{(n+1)!}\varepsilon^{n+1}$$

unde $\varepsilon = \max(|a_1|, ..., |a_n|)$, iar

$$\mu = \text{bariera superioară a modulilor derivatelor de ordin } n+1.$$

Exemplu. Să se calculeze eroarea comisă în evaluarea

$$\ln(1+x) \approx x - \frac{x^2}{2}, \text{ pentru } |x| \leq \frac{1}{10}.$$

Folosim în acest sens formula de medie Mac–Laurin:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3}\frac{1}{(1+\theta x)^3}, \text{ unde } \theta = \theta(x) \in (0,1).$$

Avem atunci din $|\theta x| = |\theta| \cdot |x| \leq 1 \cdot \frac{1}{10} = \frac{1}{10}$, evaluarea

$$\left|\ln(1+x) - x + \frac{x^2}{2}\right| = \left|\frac{x^3}{3}\right| \frac{1}{|1+\theta x|^3} \leq \frac{1}{10^3} \cdot \frac{1}{3}\left(\frac{10}{9}\right)^3 = \frac{1}{2787}.$$

A doua problemă se poate rezolva prin intermediul formulei de medie Lagrange. Avem mai exact

$$\eta_0 - \eta_0' = f(b_1, ..., b_m) - f(a_1, ..., a_m) = \sum_{i=1}^{m}(b_1 - a_i)\frac{\partial}{\partial x_i}(a + \theta(b-a)).$$

Şi în acest caz găsim de pildă evaluarea

$$|\eta_0 - \eta_0'| \leq m\delta\nu$$

unde $\delta = \max(|b_1 - a_1|, ..., |b_m - a_m|)$, iar

$$\nu = \text{bariera superioară a modulilor derivatelor de ordin 1.}$$

Exemplu. Să presupunem că mărimea η este calculată în conformitate cu regula $\eta = x^3yz$. Să apreciem eroarea făcută în calculul acestei mărimi în punctul $b = \left(\frac{101}{100}, \frac{102}{100}, \frac{1}{100}\right)$ dacă se va identifica aceasta cu valoarea luată în $a = (1, 1, 0)$. Notând cu η_0, respectiv η_0', aceste valori putem scrie

$$\eta_0 - \eta_0' \approx \frac{1}{100}3x^2yz + \frac{2}{100}x^3z + \frac{1}{100}x^3y,$$

unde (x, y, z) este un punct pe segmentul care uneşte a cu b. Obţinem

$$|\eta_0 - \eta_0'| \leq \frac{1}{100}\left[3\left(\frac{102}{100}\right)^3 \cdot \frac{1}{100} + 2\left(\frac{101}{100}\right)^3 \cdot \frac{1}{100} + \left(\frac{101}{100}\right)^3 \cdot \frac{102}{100}\right]$$

care ne dă răspunsul cerut.

Probleme propuse la §4.3

1. Pornind de la definiţie, să se calculeze derivata funcţiei $f(x) = \sqrt{3x + 4}$ în punctul $a = 1$.

2. Să se cerceteze derivabilitatea funcţiilor

$$\begin{array}{lll} \text{a)} & f(x) & = \begin{cases} \ln(x^2 + 3x), & 0 < x < 1 \\ \frac{5}{4}(x - 1) + 2\ln 2, & x \geq 1 \end{cases} \\ \text{b)} & f(x) & = |x^2 + 3x + 2|, \qquad x \in R \\ \text{c)} & f(x) & = \max(x - 1, x^2 - x - 1), \quad x \in R. \end{array}$$

3. Folosind regulile stabilite, să se calculeze derivatele funcţiilor

a) $f(x) = \sqrt{\dfrac{x^2 - 1}{x^2 + 1}}$, b) $f(x) = x^3e^{\lambda x}$,

c)$f(x) = \ln(x + \sqrt{x^2 + k})$, d) $f(x) = \operatorname{arctg}(1 + x + x^2)$.

4. Să se calculeze derivatele de ordinul 1 şi 2 ale funcţiei compuse $F(x) = f(x^2+1)$ unde f este o funcţie de două ori derivabilă pe R.

5. Să se arate că funcţia $f(x) = (x + \sqrt{x^2 - 1})^n$ verifică ecuaţia $(x^2 - 1)f''(x) + xf'(x) - n^2 f(x) = 0$.

6. Să se calculeze derivatele de ordinul n ale funcţiilor

a) $f(x) = \ln(ax + b)$; b) $f(x) = \frac{1}{x^2-1}$,
c) $f(x) = \cos x$; d) $f(x) = x^p e^{qx}$.

7. Pornind de la definiţie, să se calculeze $\dfrac{\partial f}{\partial x}(1,1)$ şi $\dfrac{\partial f}{\partial y}(1,1)$, dacă $f(x,y) = \ln(1 + x + y^2) + \operatorname{arctg}(x^2 + y)$.

8. Să se arate că funcţia

$$f(x) = \begin{cases} \dfrac{2xy}{x^2 + y^2}, & \text{dacă } x \neq 0 \text{ sau } y \neq 0 \\ 0, & \text{dacă } x = y = 0. \end{cases}$$

are derivate parţiale în origine dar este discontinuă în acest punct.

9. Să se calculeze derivatele parţiale de ordin 1 şi 2 ale funcţiilor (iar apoi gradientul şi hessianul lor)

$$f(x,y) = \operatorname{arctg} \frac{x + y}{1 - xy}, \quad f(x,y,z) = e^{xy} \cdot \sin z.$$

10. Să se calculeze derivatele parţiale ale funcţiei compuse $\varphi(x,y,z) = f(x + y + z, x^2 + y^2 + z^2)$, unde $f(u,v)$ este o funcţie de două variabile de clasă C^1 pe R^2.

11. Să se demonstreze că funcţiile următoare sunt lipschitziene pe domeniile indicate
a)$f(x) = \sin x,\ x \in R$; b)$f(x,y,z) = \frac{x-y+2z}{1+x^2+y^2+z^2},\ 0 \leq x,y,z \leq 1$;
c)$f(x) = 2xe^x,\ 0 \leq x \leq 1$; d)$f(x,y) = e^{x-2y},\ 0 \leq x,y \leq 1$.
(Indicaţie. Se va utiliza teorema lui Lagrange.)

12. Să se scrie formula Taylor de ordinul doi pentru funcţiile de mai jos, în punctele indicate

a) $f(x) = \sqrt{1+x}$, $a = 0$;
b) $f(x,y) = e^{x-y}$, $a = (0,1)$;
c) $f(x,y,z) = \ln(1 + x + y - 3z)$, $a = (1,1,0)$.

13. Să se calculeze (cu ajutorul regulii l'Hospital) limitele

a) $\lim\limits_{x\to 0} \dfrac{e^x \sin x - x(1+x)}{x^3}$; b) $\lim\limits_{x\to 0} \left(\dfrac{1}{x^2} - \dfrac{1}{\mathrm{tg}^2 x}\right)$

c) $\lim\limits_{x\to 0} x \ln x$; d) $\lim\limits_{x\to 0} \left(\dfrac{x-1}{x+1}\right)^x$.

14. Să se evalueze eroarea comisă la aproximările
a) $\sin x \approx x - \frac{x^3}{6}$, $|x| \le \frac{\pi}{20}$; b) $\sqrt{1+x} \approx 1 + \frac{x}{2} - \frac{x^2}{8}$, $|x| \le \frac{1}{100}$.
(Indicaţie. Se va utiliza formula Mac–Laurin.)

15. Să se găsească o barieră superioară a erorii (absolute) pe care o facem în calculul expresiei $\eta = x^2 y^3 e^z$ relativ la punctul $\left(\frac{11}{10}, \frac{11}{10}, \frac{1}{100}\right)$ dacă se folosesc valorile apropiate $(1,1,0)$.

4.4 Structuri de convergenţă pe spaţii de funcţii

4.4.1 Norma unei funcţii. Convergenţă uniformă

Fie $I = [c,d]$ un interval compact al axei reale R. Convenim să notăm

$$F(I,R) = \text{mulţimea tuturor funcţiilor } f : I \longrightarrow R.$$

Se ştie că aceasta este spaţiu liniar în raport cu operaţiile uzuale

$$\begin{aligned} &(f+g)(x) = f(x) + g(x), \quad x \in I, \quad \text{(suma a două funcţii)} \\ &(\lambda f)(x) = \lambda f(x), \qquad\quad x \in I, \quad \text{(produsul unei funcţii cu un număr).} \end{aligned}$$

Definim acum o *normă (generalizată)* pe acest spaţiu prin convenţia

(D1) $\|f\| = \sup\limits_{x \in I} |f(x)|, \ f \in (I,R)$.

Proprietăţile mai importante ale acesteia sunt

(P1) $\|f+g\| \le \|f\| + \|g\|$ (inegalitatea triunghiului)

(P2) $\|\lambda f\| = |\lambda| \cdot \|f\|$, (omogenitatea)

(P3) $\|f\| = 0$ dacă şi numai dacă $f = 0$ (funcţia nulă).

Ele coincid – formal – cu acelea stabilite pentru norma unui vector. Ar fi de aşteptat astfel să putem construi o teorie a convergenţei în aceste spaţii, după modelul expus acolo; este ceea ce intenţionăm să facem.

Pentru fiecare pereche de funcţii $f, g \in F(I, R)$, numim *distanţa* dintre acestea, numărul $\|f - g\|$. (Este de remarcat aici faptul că, spre deosebire de cazul vectorilor, distanţa dintre două funcţii poate fi chiar ∞.) Să facem notaţia

(D2) $S(f, \varepsilon) = \{g \in F(I, R);\ \|f - g\| < \varepsilon\},\ f \in F(I, R),\ \varepsilon > 0.$

Vom numi aceasta *sferă deschisă* cu centrul f şi raza ε. În continuare, să numim *vecinătate* a unei funcţii $f \in F(I, R)$

(D3) orice asemenea sferă deschisă cu centrul în f

(D4) orice parte din $F(I, R)$ care include o asemenea sferă.

Mulţimea tututor vecinătăţilor funcţiei f o vom nota prin $\mathcal{V}(f)$. Fie acum dat un şir de funcţii din $F(I, R)$:

$$f_1 : I \longrightarrow R, f_2 : I \longrightarrow R, ..., f_n : I \longrightarrow R, ... ,$$

şi, de asemenea, funcţia $f : I \longrightarrow R$ (element din $F(I, R)$). Zicem că şirul (f_n) *converge* (în sensul normei) către f dacă

(D5) oricare ar fi vecinătatea $V \in \mathcal{V}(f)$ există un rang $n(V)$
aşa ca $f_n \in V$, pentru toţi $n \geq n(V)$.

sau,echivalent,

(D6) pentru orice $\varepsilon > 0$ există un rang $n(\varepsilon)$ aşa ca
$n \geq n(\varepsilon) \Longrightarrow \|f_n - f\| < \varepsilon$ (sau $f_n \in S(f, \varepsilon)$).

Vom mai exprima aceasta prin fraza: (f_n) este *uniform convergent* (pe intervalul I) către funcţia f (şi scriem $f_n \xrightarrow{u} f$) pe care o vom numi *limita uniformă* a şirului (f_n).Dacă această funcţie este generică,atunci spunem că (f_n) este *uniform convergent.*
Următorul rezultat este aproape evident

Teorema 4.4.1 *Şirul de funcţii (f_n) din $F(I, R)$ converge uniform către funcţia $f \in F(I, R)$ dacă şi numai dacă şirul numeric*

$$\lambda_n = \|f_n - f\|,\ n = 1, 2, ... ,$$

converge către zero.

Exemplu. Cu $I = [-a, a]$, să construim şirul de funcţii

$$f_n(x) = \frac{x^n}{n!},\ x \in I,\ n = 1, 2, \dots .$$

Avem, pentru orice n, evaluarea

$$|f_n(x)| \le \frac{a^n}{n!},\ x \in I \text{ (adică } \|f_n\| \le \frac{a^n}{n!}).$$

Şi, cum şirul numeric $\left(\frac{a^n}{n!}\right)$ converge către zero, urmează că $f_n \xrightarrow{u} 0$.

În analogie cu cazul spaţiilor finit dimensionale, există şi aici posibilitatea să caracterizăm convergenţa uniformă fără să utilizăm direct funcţia limită. Anume, fie (f_n) un şir de funcţii din $F(I, R)$. Zicem că acesta este un şir *uniform Cauchy* (sau *fundamental*) dacă

(D7) oricare ar fi $\varepsilon > 0$, există un rang $n(\varepsilon)$ aşa ca
$\|f_{n+p} - f_n\| \le \varepsilon$, pentru orice $n \ge n(\varepsilon)$ şi $p \ge 1$.

Definiţia este formal aceeaşi ca în cazul vectorilor.

Teorema 4.4.2 *Un şir de funcţii (f_n) din $F(I, R)$ este uniform convergent dacă şi numai dacă este şir uniform Cauchy.* (*Adică, $F(I, R)$ este spaţiu complet.*)

Demonstraţie. Necesitatea condiţiei este imediată. (Se ţine cont de inegalitatea triunghiului.) Pentru suficienţă vom pleca de la faptul că, dacă (f_n) este un şir uniform Cauchy, atunci

$$(f_n(x)) \text{ este şir (numeric) Cauchy, pentru orice } x \in I.$$

Cum R este complet, toate aceste şiruri converg. Să punem

$$f(x) = \lim_{n\to\infty} f_n(x),\ x \in I.$$

Din condiţia de şir uniform Cauchy

$$|f_{n+p}(x) - f_n(x)| \le \varepsilon,\ x \in I,\ n \ge n(\varepsilon),\ p \ge 1,$$

deducem, trecând la limită în raport cu p,

$$|f(x) - f_n(x)| \le \varepsilon,\ x \in I \text{ (deci } \|f - f_n\| \le \varepsilon),\ n \ge n(\varepsilon).$$

Aceasta arată că $f_n \xrightarrow{u} f$ şi încheie argumentul. ∎

Să descriem acum – în sensul analogiei discutate anterior – unele clase de mulţimi asociate structurii de spaţiu normat astfel intfoduse. Fie M mulţime nevidă de funcţii din $F(I, R)$.

(i) Funcţia $f \in F(I, R)$ se va numi *punct interior* al mulţimii M dacă există $\varepsilon > 0$ aşa ca $S(f, \varepsilon) \subseteq M$. (Deci, în particular, $f \in M$.) Mulţimea tuturor acestor puncte alcătuieşte *interiorul* mulţimii M şi se va nota $\text{int}\,(M)$. Vom zice că M este deschisă dacă coincide cu interiorul acesteia. Ca exemplu imediat de asemenea mulţimi avem orice sferă deschisă din $F(I, R)$.

(ii) Funcţia $f \in F(I, R)$ sa va numi *punct aderent* al mulţimii M dacă există un şir de funcţii (f_n) din M cu $f_n \xrightarrow{u} f$. Mulţimea tuturor acestor puncte aderente alcătuieşte *aderenţa* (sau *închiderea*) mulţimii M şi se notează $\text{ad}\,(M)$. Vom zice că M este *închisă* dacă este identică cu aderenţa sa. Ca un prim exemplu, avem toate sferele închise din spaţiul $F(I, R)$:

$$S[f, \varepsilon] = \{g \in F(I, R);\ \|f - g\| \leq \varepsilon\},\ f \in F(I, R),\ \varepsilon > 0.$$

Un alt exemplu important în acest sens este următorul. Să punem

$$C(I, R) = \{f \in F(I, R);\ f \text{ este continuă pe } I\}.$$

Este util de observat că, în baza teoremei lui Weierstrass,

$$\|f\| < \infty,\ \text{oricare ar fi } f \in C(I, R).$$

Teorema 4.4.3 *Subspaţiul* $C(I, R)$ *este închis în* $F(I, R)$; *şi deci* (*combinând cu rezultatul anterior*) $C(I, R)$ *este spaţiu complet.*

Demonstraţie. Fie (f_n) un şir de funcţii din $C(I, R)$ care converge uniform la o funcţie $f \in F(I, R)$. De asemenea, fie a un element arbitrar din I. Din definiţia convergenţei uniforme rezultă că, dat $\varepsilon > 0$, există un rang $n(\varepsilon)$ aşa încât

$$\|f_n - f\| \leq \varepsilon, \quad \text{pentru toţi } n \geq n(\varepsilon).$$

Punând pentru simplitate $k = n(\varepsilon)$ avem, în particular,

$$|f_k(x) - f(x)| \leq \varepsilon,\ \forall x \in I \text{ (şi deci } |f_k(a) - f(a)| \leq \varepsilon).$$

Pe de altă parte, din continuitatea lui f_k, există pentru acelaşi $\varepsilon > 0$, un $\delta > 0$ aşa încât

$$x \in I, |x - a| \leq \delta \Longrightarrow |f_k(x) - f_k(a)| \leq \varepsilon.$$

Combinând toate aceste inegalităţi, avem

$$\begin{cases} x \in I, |x-a| \leq \delta \Longrightarrow |f(x)-f(a)| \leq |f(x)-f_k(x)|+ \\ +|f_k(x)-f_k(a)|+|f_k(a)-f(a)| \leq \varepsilon+\varepsilon+\varepsilon=3\varepsilon. \end{cases}$$

Aceasta arată că f este continuă în a; şi, cum a este arbitrar în I, urmează $f \in C(I, R)$. ∎

În legătură cu concluzia de mai sus, se poate pune în mod natural problema dacă nu cumva aceeaşi proprietate o are şi familia de funcţii

$$D(I, R)=\{f \in F(I, R);\ f \text{ este derivabilă pe } I\}.$$

Răspunsul este negativ, în general. D2r, cu unele ipoteze suplimentare, concluzia poate fi (local) valabilă. Mai exact, dăm (fără demonstraţie)

Teorema 4.4.4 *Fie (f_n) un şir de funcţii din $D(I, R)$ care satisface*

(C1) (f_n) *converge uniform către funcţia $f \in F(I, R)$*

(C2) (f_n') *converge uniform către funcţia $g \in F(I, R)$*

Atunci, obligatoriu, $f \in D(I, R)$ şi $f'=g$.

Toate aceste consideraţii se pot uşor extinde la funcţii cu valori vectoriale. Mai exact, fie (pentru $p \geq 1$)

$$F(I, R^p)= \text{ mulţimea tuturor funcţiilor } f: I \longrightarrow R^p.$$

Fiecare funcţie f din această mulţime are forma

$$f(x)=(f_1(x), ..., f_p(x),\ x \in I$$

unde $f_1, ..., f_p$ (componentele lui f) sunt elemente din $F(I, R)$. Atunci, nu avem decât să punem

(D8) $\|f\|=\max(\|f_1\|, ..., \|f_p\|)$

pentru a traduce toate noţiunile anterioare în noul cadru. Menţionăm, în fine ,că o extensie ulterioară a acestor fapte se poate face prin înlociurea intervalului compact I din R cu un continuum D al unui spaţiu R^m (cu $m \geq 1$).

4.4.2 Convergenţă local–uniformă şi punctuală

Vom încerca acum să extindem rezultatele anterioare la cazul intervalelor oarecare. Fie J un asemenea interval; asta înseamnă că J poate fi chiar nemărginit (de pildă, $J = (e, \infty)$) sau mărginit, dar (eventual) deschis la unul din (sau ambele) capete. Să notăm şi aici

$$F(J, R) = \text{mulţimea tuturor funcţiilor } f : J \longrightarrow R.$$

Noţiunile de normă şi convergenţă uniformă se pot introduce exact ca la cazul compact; şi rezultatele corespunzătoare continuă să aibă loc şi în acest cadru. Avem însă şi unele aspecte noi. Anume, fie

$$f_1 : J \longrightarrow R, f_2 : J \longrightarrow R, ..., f_n : J \longrightarrow R, ...,$$

un şir de funcţii din $F(J, R)$ şi $f : J \longrightarrow R$ o funcţie din $F(J, R)$. Prin definiţie, vom spune că şirul (f_n) *converge uniform către f pe intervalul compact* $I = [c, d]$ din J dacă

(D1) restricţiile funcţiilor din şir la interval I converg uniform (în sensul normei din $F(I, R)$) către restricţia funcţiei f la intervalul I.

(Vom nota această proprietate prin $f_n \underset{I}{\overset{u}{\rightarrow}} f$. În particular, dacă I se reduceala un punct ($c = d$) vom zice că şirul de funcţii (f_n) *converge în punctul* c către funcţia f; adică,

(D2) $f_n(c) \longrightarrow f(c)$ pentru $n \longrightarrow \infty$.

Zicem că şirul (f_n) converge *local–uniform* către f pe intervalul (iniţial) J şi scriem aceasta $f_n \xrightarrow{lu} f$ dacă

(D3) $f_n \underset{I}{\overset{u}{\rightarrow}} f$, pentru orice interval compact $I \subset J$.

În fine, (f_n) se va numi *convergent punctual* către f pe intervalul (iniţial) J dacă

(D4) $f_n(x) \longrightarrow f(x)$, pentru toţi $x \in J$.

Vom desemna această proprietate prin $f_n \xrightarrow{p} f$. Avem, deci, în principiu, trei tipuri de convergenţă pe acest spaţiu. Relaţiile dintre acestea sunt precizate în

Teorema 4.4.5 *Avem implicaţiile*

(P1) $[f_n \xrightarrow{u} f] \Longrightarrow [f_n \xrightarrow{lu} f] \Longrightarrow [e_n \xrightarrow{p} f]$.

Adică: convergenţa uniformă o implică pe cea local–uniformă care, la rândul ei, are ca efect convergenţa punctuală.

Demonstraţia este imediată, din definiţie; de aceea nu mai insistăm asupra ei. Se pune însă problema în ce măsură pot fi inversate implicaţiile de mai sus. Răspunsul este următorul. Dacă J este interval compact, atunci implicaţia de la prima la a doua proprietate poate fi inversată; sau, mai precis,

(P2) $J = \text{compact} \Longrightarrow [\text{conv. uniformă} \equiv \text{conv. local-uniformă}]$

(Nu avem decât să luăm, printre intervalele compacte ale lui J pe J însuşi.) Dacă, însă, J nu este interval compact, implicaţia în chestiune nu mai poate fi inversată.

Exemplu. Cu $J = [0, 1)$, să construim şirul de funcţii

$$f_n(x) = x^n,\ x \in J,\ n = 1, 2, ...$$

Pe fiecare subinterval compact $I = [c, d]$, al intervalului (necompact) J,

$$|f_n(x)| \leq d^n,\ x \in I,\ n = 1, 2, ...$$

Deci, $f_n \underset{I}{\xrightarrow{u}} 0$; şi, cum I este arbitrar, $f_n \xrightarrow{lu} 0$. Pe de altă parte, luând în considerare şirul (din J)

$$x_n = 1/\sqrt[n]{2},\ n = 1, 2, ...$$

avem, evident, evaluările

$$f_n(x_n) = \frac{1}{2}\ (\text{deci, } \|f_n\| \geq \frac{1}{2}),\ n = 1, 2, ...$$

Aceasta, după un fapt anterior (Teorema 4.4.1) este în dezacord cu $f_n \xrightarrow{u} 0$ şi afirmaţia reiese.

Pe de altă parte, oricare ar fi intervalul J (compact sau nu) implicaţia între proprietatea a doua şi a treia din (P1) nu poate fi în general inversată. Aceasta se constată din următorul

Exemplu. Fie dat şirul de funcţii

$$g_n(x) = x^n(1 - x^n),\ 0 \leq x \leq 1,\ n = 1, 2, ...$$

Este uşor de văzut că $g_n \xrightarrow{p} 0$. În plus, o relaţie de forma $g_n \xrightarrow{lu} 0$ nu este valabilă deoarece, conform lui (P2) ar trebui să avem atunci $g_n \xrightarrow{u} 0$; iar aceasta nu se verifică. (Se utilizează o tehnică asemănătoare cu cea dinainte.) Iar de aici, concluzia este clară.

Trecând la proprietăţile acestor tipuri de convergenţe menţionăm, ca un fapt general, că rezultatele expuse anterior pe intervale compacte rămân valabile pe intervale necompacte şi pentru convergenţa uniformă înlocuită (eventual) cu convergenţa local–uniformă. Aceasta deoarece proprietatea de continuitate sau derivabilitate într-un punct angajează doar valorile funcţiei dintr-o vecinătate a acelui punct. Evident că toate aceste consideraţii se pot imediat extinde (prin tehnicile anterioare) la cazul funcţiilor cu valori vectoriale definite pe un interval (în general necompact) J al axei reale; sau, mai general, un domeniu al unui anumit spaţiu finit dimensional.

4.4.3 Serii de funcţii

În cele ce urmează vom aplica rezultatele descrise anterior în studiul seriilor cu elemente din spaţiile funcţionale deja introduse.

(A) Fie $I = [c, d]$ un interval compact al axei reale şi $F(I, R)$ clasa tuturor funcţiilor $f : I \longrightarrow R$. Fie

$$f_1 : I \longrightarrow R, f_2 : I \longrightarrow R, ..., f_n : I \longrightarrow R, ...,$$

un şir de funcţii din $F(I, R)$. Vom numi simbolul

(D1) $f_1 + f_2 + \cdots + f_n + \cdots$ (sau, echivalent, $\sum_n f_n$)

o *serie de funcţii.* Semnificaţia sa este legată de conceptele anterior introduse. Mai exact, să construim şirul de sume parţiale din $F(I, R)$

$$g_1 = f_1, g_2 = g_1 + f_2, ..., g_n = f_1 + \cdots + f_n, ...$$

Prin definiţie, vom spune că seria $\sum_n f_n$ *converge uniform* dacă şirul de funcţii (g_n) converge uniform în $F(I, R)$. În acest caz, funcţia $g \in F(I, R)$ cu

$g_n \xrightarrow{u} g$ se va numi *suma seriei* în chestiune şi vom nota aceasta prin $g = \sum\limits_{n=1}^{\infty} f_n$.

Următorul criteriu (Cauchy) de convergenţă uniformă rezultă cu uşurinţă din Teorema 4.4.2 anterioară.

Teorema 4.4.6 *Seria de funcţii* $\sum\limits_n f_n$ *este uniform convergentă dacă şi numai dacă oricare ar fi* $\varepsilon > 0$ *, există un rang* $n(\varepsilon)$ *aşa încât*

$$\|f_{n+1} + \cdots + f_{n+p}\| \leq \varepsilon, \text{ pentru orice } n \geq n(\varepsilon) \text{ şi } p \geq 1.$$

În particular, de aici găsim Criteriul de comparaţie Weierstrass pentru convergenţa uniformă.

Teorema 4.4.7 *Fie* $\sum\limits_n f_n$ *o serie de funcţii pentru care avem*

$$\|f_n\| \leq \beta_n, \text{ pentru toţi } n.$$

Dacă seria numerică (cu termeni pozitivi) $\sum\limits_n \beta_n$ *converge, atunci seria de funcţii* $\sum\limits_n f_n$ *este uniform convergentă.*

În ce priveşte suma unei serii de funcţii uniform convergente, rezultatele se deduc imediat din acelea privind limitele şirurilor uniform convergente. Astfel, cu Teorema 4.4.3 avem

Teorema 4.4.8 *Fie* $\sum\limits_n f_n$ *o serie uniform convergentă de funcţii şi* $f = f_1 + f_2 + \cdots$*, suma acesteia. Dacă toţi termenii seriei sunt funcţii continue pe* I*, atunci şi suma seriei este continuă pe* I*.*

Tot astfel, din Teorema 4.4.4, deducem

Teorema 4.4.9 *Fie* $\sum\limits_n f_n$ *o serie uniform convergentă de funcţii şi* $f = f_1 + f_2 + \cdots$*, suma acesteia. Dacă toţi termenii seriei sunt funcţii derivabile pe* I *şi seria de funcţii* $\sum\limits_n f_n'$ *este uniform convergentă, atunci* f *este derivabilă pe* I *şi* $f' = f_1' + f_2' + \cdots$ *.*

Exemplu. Fie $I = [-a, a]$ şi seria de funcţii $\sum_n f_n$, unde

$$f_n(x) = \frac{x^n}{n!} \cos nx,\ x \in I,\ n = 1, 2, ... \ .$$

Seria în chestiune este uniform convergentă, deoarece

$$|f_n(x)| = \left|\frac{x^n}{n!} \cos nx\right| \le \frac{a^n}{n!},\ \forall x \in I,\ \ (\text{deci}\quad \|f_n\| \le \frac{a^n}{n!}),\ n = 1, 2, ...,$$

iar seria numerică cu termeni pozitivi $\sum_n \frac{a^n}{n!}$ este convergentă. Să notăm cu f suma seriei

$$f(x) = \sum_{n=1}^{\infty} \frac{x^n}{n!} \cos nx,\quad x \in I.$$

Cum toţi termenii seriei sunt continui, f este de asemenea continuă. În plus, termenii respectivi sunt şi derivabili, cu

$$f_n'(x) = \frac{1}{(n-1)!}(x^{n-1} \cos nx - x^n \sin nx),\ x \in I, n = 1, 2, ... \ .$$

Deoarece, pentru orice $n \ge 1$,

$$|f_n'(x)| \le 2\frac{a^n}{(n-1)!},\ \forall x \in I \quad (\text{deci}\|f_n'\| \le 2\frac{a^n}{(n-1)!})$$

(am presupus aici $a \ge 1$), rezultă că seria derivatelor este şi ea uniform convergentă. Suma cestei serii este tocmai derivata sumei anterioare

$$f'(x) = \sum_{n=1}^{\infty} \frac{1}{(n-1)!}(x^{n-1} \cos nx - x^n \sin nx), x \in I.$$

Toate rezultatele obţinute pot fi uşor extinse la cazul funcţiilor cu valori vectoriale. Acestea, la rândul lor, se pot extinde la situaţia în care intervalul compact I din R se poate înlocui cu un continuum D dintr-un spaţiu R^m (cu $m \ge 1$).

(B) Fie acum J un interval arbitrar al axei reale şi $F(J, R)$ clasa tuturor funcţiilor $f : J \longrightarrow R$. Dat şirul de funcţii din $F(J, R)$

$$f_1 : J \longrightarrow R, f_2 : J \longrightarrow R, ..., f_n : J \longrightarrow R, ...$$

vom numi şi aici simbolul

(D2) $f_1 + f_2 + \cdots + f_n + \cdots$ (sau, echivalent, $\sum\limits_n f_n$)

serie de funcţii. Semnificaţia acestuia este legată de cele trei concepte de convergenţă care se pot introduce în acest context:

(a) convergenţa uniformă (pe intervalul J)

(b) convergenţa local–uniformă (care înseamnă o convergenţă uniformă pe fiecare interval compact al lui J)

(c) convergenţă punctuală (adică, o convergenţă la nivelul fiecărui punct din J).

Legătura dintre aceste tipuri de convergenţă este în fond aceea indicată de Teorema 4.4.5. Să notăm că pentru oricare tip de convergenţă există sumă a seriei respective ($f = \sum\limits_{n=1}^{\infty} f_n$ are sens ca element din $F(J, R)$).

În ce priveşte proprietăţile de bază menţionăm că rezultatele stabilite anterior rămân valabile chiar în condiţii de convergenţă local–uniformă a seriei considerate. Ilustrăm acestea prin următorul

Exemplu. Fie $J = (-1, 1)$ şi seria de funcţii $\sum\limits_n f_n$, cu

$$f_n(x) = x^n \cos(n\pi x),\ \ x \in J, n = 1, 2, ...$$

Dacă $I = [-a, a]$ este un interval compact al lui J, atunci, din evaluările

$$|f_n(x)| = |x^n \cos(n\pi x)| \leq a^n,\ \ x \in I, n = 1, 2, ...$$

şi din convergenţa seriei $\sum\limits_n a^n$ rezultă convergenţa uniformă a seriei noastre de funcţii pe I. Adică, $\sum\limits_n f_n$ este local–uniform convergentă pe (intervalul necompact) J. Conform unui rezultat precedent, suma seriei,

$$f(x) = \sum_{n=1}^{\infty} x^n \cos(n\pi x),\ \ x \in J$$

apare ca funcţie continuă pe J. În mod analog se lucrează şi în ce priveşte derivabilitatea.

Menţionăm, în final, că toate aceste dezvoltări sunt valabile şi pentru serii de funcţii vectoriale definite pe intervale oarecare (necompacte în general) ale axei reale. De asemenea, o extensie ulterioară este posibilă pentru funcţii definite pe domenii ale unui spaţiu finit dimensional; nu mai dăm alte detalii.

4.4.4 Serii de puteri. Dezvoltare în serie Taylor

Numim *serie de puteri* orice serie de funcţii $\sum_n f_n$ în care termenii acesteia au forma analitică

(D1) $f_n(x) = a_n x^n,\ x \in R,\ n = 0, 1, ...$.

(Aici, (a_n) este un şir dat de numere reale.) Desigur, toate rezultatele precedente sunt valabile şi pentru acest caz particular. Apar însă şi o serie de aspecte specifice pe care le vom evidenţia în continuare.

Începem prin a observa faptul că orice serie de puteri este convergentă (punctual) cel puţin în origine; aceasta deoarece toţi termenii seriei începând cu cel de rangul 1 se anulează în acest punct. Există însă şi serii de puteri ce converg doar în acest punct, ca, de pildă,

$$\sum_n n!x^n = 1 + 1!x + 2!x^2 + \cdots + n!x^n + \cdots \ .$$

Aceasta se constată imediat din faptul că

$$\lim_{n\to\infty} n!|x|^n = \infty, \quad \text{pentru toţi} \quad x \neq 0.$$

Pe de altă parte, există serii de puteri care converg în orice punct,după cum rezultă din exemplul

$$\sum_n \frac{x^n}{n!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots \ .$$

(A se vedea , în acest sens,discuţia făcută în Secţiunea 4.1.2.) În legătură cu mulţimea de convergenţă a unei serii de puteri, avem următorul rezultat important (numit "Teorema lui Abel"):

Teorema 4.4.10 *Pentru orice serie de puteri*

$$\sum_n a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots$$

există un număr $\rho \geq 0$ (finit sau infinit), cu proprietăţile

(i) *seria este convergentă pe intervalul $(-\rho, \rho)$*

(ii) *seria este divergentă pe $(-\infty, -\rho) \cup (\rho, +\infty)$.*

Demonstraţie. (Schiţă) Dacă seria de puteri converge doar în $x = 0$, atunci punem $\rho = 0$ şi teorema rezultă . În caz contrar, mulţimea

$$D = \{x \in R;\ \sum_n a_n x^n \ \text{ converge}\}$$

ar conţine şi puncte distincte de origine. Să notăm acum proprietatea

(P1) $z \neq 0,\ z \in D \Longrightarrow (-|z|, |z|) \subseteq D$.

(Adică, dacă z este un punct (nenul) de convergenţă al seriei, atunci orice w cu $|w| < |z|$ este, de asemenea, punct de convergenţă.) Căci, dacă z este astfel luat, avem, din $a_n z^n \longrightarrow 0$, că numărul

$$\mu = \sup\{|a_n z^n|;\ n = 0, 1, ...\}$$

este finit; şi atunci, cu evaluarea

$$|a_n w^n| = |a_n z^n| \cdot |\frac{w}{z}|^n \leq \mu|\frac{w}{z}|^n,\ n = 0, 1, ...\ ,$$

afirmaţia reiese, conform criteriului de comparaţie Weierstrass (Secţiunea 4.1.2). Este suficient acum să punem

(D2) $\rho =$ marginea superioară a mulţimii D

pentru a îndeplini toate cerinţele noastre. ∎

Numărul ρ definit de rezultatul anterior se numeşte *raza de convergenţă* a seriei de puteri $\sum_n a_n x^n$. Pentru calculul acesteia, este util de făcut observaţia că , în fond, din raţionamentul anterior avem chiar proprietatea mai tare

(P2) seria $\sum_n a_n x^n$ converge absolut pe $(-\rho, \rho)$.

Cu alte cuvinte, seria de funcţii (pozitive) $\sum_n |a_n| \cdot |x|^n$ converge când $|x| < \rho$ şi diverge când $|x| > \rho$. Acest fapt, combinat cu criteriul raportului sau rădăcinii aplicat seriei în cauză conduce la

Teorema 4.4.11 *Raza de convergenţă a seriei de puteri* $\sum_n a_n x^n$ *este egală cu oricare din numerele (dacă limita respectivă există)*

$$\rho = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}; \quad \rho = \lim_{n\to\infty} \sqrt[n]{1/|a_n|}$$

Fie acum $\sum_n a_n x^n$ o serie de puteri cu raza de convergenţă ρ. Următoarele două chestiuni ne vor interesa în continuare:

– precizarea tipului de convergenţă al seriei pe $(-\rho, \rho)$;

– stabilirea comportării seriei în punctele $-\rho$ şi ρ.

(A) În ce priveşte prima problemă, avem:

Teorema 4.4.12 *Seria de puteri* $\sum_n a_n x^n$ *este local-uniform convergentă pe intervalul (necompact)* $(-\rho, \rho)$.

Demonstraţie. Fie $[-r, r]$, cu $0 < r < \rho$, un interval compact din $(-\rho, \rho)$. Conform lui (P2), $\sum_n |a_n| \cdot r^n$ este o serie convergentă (cu termeni pozitivi). Aceasta, în combinaţie cu

$$|a_n x^n| \leq |a_n| r^n, \quad -r \leq x \leq r, \ n = 0, 1, ...,$$

ne duce, prin intermediul criteriului Weierstrass (Secţiunea 4.4.3) la concluzia cerută. ■

Fie $S : (-\rho, \rho) \longrightarrow R$ suma seriei de puteri

$$S(x) = \sum_n a_n x^n, \quad -\rho < x < \rho.$$

Din teorema anterioară rezultă că funcţia S este continuă pe $(-\rho, \rho)$. Să studiem şi derivabilitatea acesteia. Pentru aceasta, pornim de la

$$\text{(P3)} \quad \begin{cases} \text{seria de puteri derivată } \sum_{n=1}^{\infty} n a_n x^{n-1} \ (= \sum_n (n+1) a_{n+1} x^n) \\ \text{are aceeaşi rază de convergenţă ca seria iniţială.} \end{cases}$$

(Acest fapt este uşor de verificat,în cadrul oferit de Teorema 4.4.11. Căci, dacă σ ar fi raza seriei derivate, avem

$$\sigma = \lim_{n \to \infty} \frac{n+1}{n+2} \left| \frac{a_{n+1}}{a_{n+2}} \right| = 1 \cdot \rho = \rho,$$

dacă folosim formula de calcul dată acolo). Deci, seria derivată este, de asemenea, local–uniform convergentă pe $(-\rho, \rho)$. Şi atunci, pe baza unui fapt anterior (Teorema 4.4.9),

$$S'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}, \ -\rho < x < \rho.$$

Desigur, procedeul poate continua. Se ajunge astfel la proprietatea

$$\text{(P4)} \quad \begin{cases} S \text{ este indefinit derivabilă şi (pentru } k = 0, 1, ...) \\ S^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) a_n x^{n-k}, \ -\rho < x < \rho. \end{cases}$$

În particular, aceasta ne duce la evaluările

$$a_k = \frac{1}{k!} S^{(k)}(0), \quad k = 0, 1, ... \ ,$$

şi, implicit, la dezvoltarea (Mac –Laurin) a sumei

$$\text{(R1)} \quad S(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} S^{(n)}(0), \quad -\rho < x < \rho.$$

Informaţiile de acest tip se pot dovedi uneori hotărâtoare pentru determinarea expresiei analitice a sumei, după cum o arată şi următorul

Exemplu. Fie dată seria de puteri

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + (-1)^{n-1} \frac{x^n}{n} + \cdots \ .$$

Raza de convergenţă a acesteia este

$$\rho = \lim_{n\to\infty} \left| \frac{(-1)^{n-1}}{n} \cdot \frac{n+1}{(-1)^n} \right| = \lim_{n\to\infty} \frac{n+1}{n} = 1.$$

Seria este, deci, local–uniform convergentă pe $(-1, 1)$. Ca atare,suma

$$S(x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + (-1)^{n-1} \frac{x^n}{n} + \cdots, -1 < x < 1$$

este indefinit derivabilă pe $(-1, 1)$. Conform celor spuse anterior, avem pentru derivata sumei, dezvoltarea

$$S'(x) = 1 - x + x^2 + \cdots + (-1)^n x^n + \cdots = \frac{1}{1+x}, \ -1 < x < 1.$$

Cum, pe de altă parte, $S(0) = 0$, urmează

$$S(x) = \ln(1 + x), \ -1 < x < 1.$$

(Aceasta deoarece, două funcţii derivabile cu aceeaşi derivată şi aceeaşi valoare într-un punct coincid pe orice interval comun de definiţie care conţine acest punct în interior.)

Alte proprietăţi de acest tip rezultă şi din analiza operaţiilor cu seriile de puteri. Fie date două serii de puteri $\sum_n a_n x^n$ şi $\sum_n b_n x^n$, cu razele (nenule) de convergenţă ρ şi σ respectiv, şi cu sumele

$$A(x) = \sum_{n=0}^{\infty} a_n x^n, \ -\rho < x < \rho; \ B(x) = \sum_{n=0}^{\infty} b_n x^n, \ -\sigma < x < \sigma.$$

Vom numi seria de puteri $\sum_n c_n x^n$, cu coeficienţii daţi de

$$c_0 = a_0 + b_0, c_1 = a_1 + b_1, ..., c_n = a_n + b_n, ...,$$

suma celor două serii de puteri. Se constată că raza de convergenţă a acestei serii este superioară numărului $\gamma = \min(\rho, \sigma)$. În plus,

$$\sum_{n=0}^{\infty} c_n x^n = A(x) + B(x), \ \forall x \in (-\gamma, \gamma).$$

Apoi, să numim seria de puteri $\sum_n d_n x^n$, cu coeficienţii daţi de

$$d_0 = a_0 b_0, d_1 = a_0 b_1 + a_1 b_0, ..., d_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0, ...,$$

produsul celor două serii de puteri. Şi aici, raza de convergenţă a acestei serii este superioară numărului γ. (Aceasta rezultă din convergenţa absolută a seriilor de puteri componente şi din observaţia că seria $\sum_n d_n x^n$ este, în fond, produsul în sens Cauchy al seriilor $\sum_n a_n x^n$ şi $\sum_n b_n x^n$.) În plus, avem şi egalitatea

$$\sum_{n=0}^{\infty} d_n x^n = A(x) \cdot B(x), \quad -\gamma < x < \gamma.$$

(Se ţine cont de un fapt anterior (Teorema 4.1.11).)

Ca o ilustrare a acestor fapte vom trata următorul

Exemplu. Fie λ număr real dat. Definind combinările generalizate

$$C_\lambda^0 = 1, C_\lambda^1 = \lambda, ..., C_\lambda^n = \frac{\lambda(\lambda-1)\cdots(\lambda-n+1)}{n!}, ... ,$$

să considerăm seria de puteri

$$\sum_n C_\lambda^n x^n = C_\lambda^0 + C_\lambda^1 x + \cdots + C_\lambda^n x^n + \cdots .$$

Vom numi aceasta, *seria binomială.* Denumirea este legată de faptul că , pentru $\lambda =$ număr natural,

$$\sum_{n=0}^{\lambda} C_\lambda^n x^n = (1+x)^\lambda, \quad \text{pentru orice} \;\; x.$$

Adică, seria se reduce în acest caz la o dezvoltare de tip binomial. Raza de convergenţă a seriei este

$$\rho = \lim_{n\to\infty} \left| \frac{C_\lambda^n}{C_\lambda^{n+1}} \right| = \lim_{n\to\infty} \left| \frac{n+1}{\lambda - n} \right| = 1.$$

Să determinăm expresia analitică a sumei seriei

$$S(x) = C_\lambda^0 + C_\lambda^1 x + \cdots + C_\lambda^n x^n + \cdots, \; -1 < x < 1.$$

Avem, prin derivare

$$S'(x) = C_\lambda^1 + 2C_\lambda^2 x + \cdots + (n+1)C_\lambda^{n+1} x^n + \cdots, \; -1 < x < 1.$$

Cum, pe de altă parte,

$$xS'(x) = C_\lambda^1 + 2C_\lambda^2 x^2 + \cdots + nC_\lambda^n x^n + \cdots,$$

avem prin adunarea celor două serii,

$$\begin{cases} (1+x)S'(x) = S'(x) + xS'(x) = \sum\limits_{n=0}^{\infty} ((n+1)C_\lambda^{n+1} + nC_\lambda^n) x^n = \\ = \sum\limits_{n=0}^{\infty} \lambda C_\lambda^n x^n = \lambda S(x), \quad -1 < x < 1. \end{cases}$$

Deducem, astfel, că suma verifică condiţiile

$$(1+x)S'(x) - \lambda S(x) = 0, \; -1 < x < 1, \; S(0) = 1.$$

Dacă facem notaţia

$$\varphi(x) = S(x)/(1+x)^{\lambda}, \quad -1 < x < 1,$$

avem imediat din relaţiile anterioare

$$\varphi'(x) = 0, -1 < x < 1; \quad \varphi(0) = 1.$$

Deci, obligatoriu,

$$\varphi(x) = 1 \;\; (\text{adică} \;\; S(x) = (1+x)^{\lambda}), \quad -1 < x < 1,$$

şi astfel, forma analitică a sumei a fost determinată. Cu alte cuvinte, pentru orice λ real,

$$(1+x)^{\lambda} = 1 + \frac{\lambda}{1!}x + \cdots + \frac{\lambda(\lambda-1)\cdots(\lambda-n+1)}{n!}x^n + \cdots, -1 < x < 1.$$

(B) Să trecem acum la a doua problemă din cele formulate anterior. Fie, în general, $\sum_n a_n x^n$ o serie de puteri cu o rază de convergenţă ρ; acceptăm în continuare că $0 < \rho < \infty$ (deoarece altfel, problema nu are sens). Notând cu D mulţimea de convergenţă a seriei în chestiune, avem dubla incluziune

(P5) $(-\rho, \rho) \subseteq D \subseteq [-\rho, \rho]$.

Menţionăm în acest sens că, relativ la termenul din mijloc, toate cazurile sunt posibile. Adică, există serii care nu converg în nici unul din punctele $-\rho, \rho$; după cum există şi serii care converg în unul din punctele $-\rho$ sau ρ (sau în ambele). Presupunând că suntem în cel de-al doilea caz, să notăm suma seriei prin

$$S(x) = a_0 + a_1 x + \cdots + a_n x^n + \cdots, \quad x \in D.$$

Conform unui rezultat anterior, funcţia S este continuă pe intervalul $(-\rho, \rho)$. Se poate pune întrebarea dacă această proprietate se păstrează şi în punctul $-\rho$ sau ρ în care mai este definită această funcţie. Răspunsul este afirmativ şi este precizat în cadrul următorului rezultat (datorat, de asemenea, lui Abel) pe care îl dăm fără demonstraţie:

Teorema 4.4.13 *Dacă seria de puteri* $\sum_n a_n x^n$ *converge în punctul* $x = \rho$, *atunci ea converge local–uniform pe* $(-\rho, \rho]$; *şi, ca atare, funcţia sumă este continuă în* $x = \rho$.
(Un rezultat analog are loc şi în punctul $x = -\rho$.*)*

Putem astfel găsi suma seriei de puteri în capetele $x = \pm\rho$ (în cazul când există), prin intermediul expresiei analitice a acesteia pe intervalul $(-\rho, \rho)$.
Exemplu. Să considerăm seria de puteri

$$\sum_{n=1}(-1)^{n+1}\frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + (-1)^{n+1}\frac{x^n}{n} + \cdots \, .$$

Am stabilit deja că această serie are raza de convergenţă $\rho = 1$ şi că suma ei este dată de formula

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + (-1)^{n+1}\frac{x^n}{n} + \cdots, \ -1 < x < 1.$$

Pentru $x = -1$, seria de puteri devine

$$\sum_{n=1}^{\infty}(-1)^{n+1}\frac{(-1)^n}{n} = -(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots);$$

aceasta este divergentă (ca serie armonică de exponent $\alpha = 1$). Pentru $x = 1$, seria ce se obţine

$$\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} + \cdots + (-1)^{n+1}\frac{1}{n} + \cdots$$

este convergentă în baza criteriului Leibniz. Mulţimea de convergenţă a acestei serii este deci $D = (-1, 1]$. Pe baza teoremei anterioare avem

$$1 - \frac{1}{2} + \frac{1}{3} + \cdots + (-1)^{n+1}\frac{1}{n} + \cdots = \lim_{x \to 1-} \ln(1+x) = \ln 2.$$

(C) Să revenim la problemele discutate la cazul **(A)**. În legătură cu dezvoltarea Mac– Laurin găsită acolo, se ridică următoarea problemă. Fie I un interval (necompact în general) care conţine originea în interior, şi $f : I \longrightarrow R$, o funcţie care admite derivate de orice ordin în toate punctele lui $\operatorname{int}(I)$. Are sens astfel considerarea seriei de puteri $\sum_n \frac{f^{(n)}(0)}{n!}x^n$. În conformitate cu formula citată, ne putem întreba în ce măsură funcţia sumă a acestei serii coincide cu funcţia iniţială:

(R2) $f(x) = \sum_n \frac{f^{(n)}(0)}{n!} x^n$, pentru toţi $x \in I$.

(Vom zice atunci că f admite o *dezvoltare în serie* Mac– Laurin pe intervalul considerat.) Pentru a găsi condiţii care să asigure acest lucru să reamintim faptul că , în baza formulei de medie Mac– Laurin (Secţiunea 4.3.3) avem, pentru orice rang n

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \cdots + \frac{f^{(n)}(0)}{n!}x^n + R_n(x),\ x \in I,$$

unde funcţia rest R_n este dată de formula

$$R_n(x) = \frac{x^{n+1}}{(n+1)!} f^{(n+1)}(\theta x), \quad x \in I,$$

cu numărul θ (dependent de x şi n) situat în intervalul $(0,1)$. În aceste condiţii (combinând şi cu teorema a doua a lui Abel) următorul rezultat este aproape evident.

Teorema 4.4.14 *Avem dezvoltarea în serie Mac–Laurin* (R2) *dacă şi numai dacă şirul de funcţii rest* (R_n) *converge punctual pe mulţimea* I *către funcţia identic zero. Mai mult, în aceste condiţii, seria Mac–Laurin* $\sum_n \frac{f^{(n)}(0)}{n!}x^n$ *converge local–uniform pe* I *(şi, deci, şirul de funcţii* (R_n) *converge chiar local–uniform pe* I *către funcţia identic zero).*

Exemplu. Să studiem dezvoltarea în serie Mac–Laurin a funcţiei

$$f(x) = e^x, \quad x \in R.$$

Această funcţie este indefinit derivabilă pe R, cu

$$f^{(n)}(x) = e^x,\ x \in R \quad (\text{şi deci } f^{(n)}(0) = 1),\ n = 0, 1, \dots\ .$$

Formula de medie Mac–Laurin devine, pentru fiecare rang n

$$e^x = 1 + \frac{x}{1!} + \cdots + \frac{x^n}{n!} + R_n(x), \quad x \in R,$$

unde restul este dat prin

$$R_n(x) = \frac{x^{n+1}}{(n+1)!} e^{\theta x},\ \theta = \theta(x, n) \in (0, 1).$$

Avem în aceste condiţii, pentru x arbitrar fixat

$$|R_n(x)| = \left|\frac{x^{n+1}}{(n+1)!}e^{\theta x}\right| \leq \frac{|x|^{n+1}}{(n+1)!}e^{|x|},\ n = 0, 1, ... ;$$

şi, ca atare, pentru orice $x \in R$,

$$\lim_{n\to\infty} R_n(x) = 0 \quad (\text{deoarece } \lim_{n\to\infty} \frac{|x|^{n+1}}{(n+1)!}e^{|x|} = 0).$$

Ajungem astfel la dezvoltarea în serie Mac–Laurin

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots, \quad \forall x \in R.$$

Utilitatea unor astfel de dezvoltări în serie provine din faptul că acestea permit aproximarea, cu orice precizie dorim, a valorilor funcţiei sumă într-un punct dat.

Toate aceste consideraţii se pot extinde şi la cazul vectorial. Mai precis, fie m număr natural dat şi

$$a_0 = (a_0(1), ..., a_0(m)), ..., a_n = (a_n(1), ..., a_n(m)), ...$$

un şir de vectori din R^m. Vom numi şi aici simbolul $\sum_n a_n x^n$ o *serie de puteri* (cu coeficienţi vectoriali). Proprietăţile de bază ale acestora se deduc în esenţă din următoarea teoremă de caracterizare

Teorema 4.4.15 *Seria de puteri (cu coeficienţi vectoriali) $\sum_n a_n x^n$ este convergentă în punctul $x \in R$ dacă şi numai dacă fiecare din seriile de puteri componente (cu coeficienţi reali) $\sum_n a_n(1)x^n, ..., \sum_n a_n(m)x^n$ este convergentă în punctul considerat. Iar în acest caz, suma acesteia este vectorul din R^m dat de*

$$\sum_{n=0}^{\infty} a_n x^n = \left(\sum_{n=0}^{\infty} a_n(1)x^n, ..., \sum_{n=0}^{\infty} a_n(m)x^n\right).$$

(Demonstraţia revine la convergenţa unei serii vectoriale.)

Din rezultatul anterior deducem în primul rând expresia razei de convergenţă a unei asemenea serii de puteri

$$\begin{aligned}&\rho = \min(\rho_1, ..., \rho_m), \qquad \text{unde, prin definiţie}\\ &\rho_i = \text{raza de convergenţă a seriei } \sum_n a_n(i)x^n,\ 1 \leq i \leq m.\end{aligned}$$

Calculul efectiv al ei se face cu formulele (dacă limitele în cauză există)

$$\rho = \lim_{n\to\infty} \frac{\|a_n\|}{\|a_{n+1}\|}; \quad \rho = \lim_{n\to\infty} \sqrt[n]{1/\|a_n\|}.$$

(Aceasta rezultă din proprietăţile de absolută convergenţă a unei serii de puteri în interiorul mulţimii de convergenţă.) Proprietăţile de local–uniformă convergenţă stabilite în Teorema 4.4.12 rămân şi aici valabile; ca şi proprietăţile de continuitate şi derivabilitate ale sumei. În ce priveşte operaţile cu asemenea serii, avem posibilitatea de a aduna două serii de acest tip sau de a înmulţi o serie de puteri cu coeficienţi numerici cu o serie de puteri cu coeficienţi vectoriali; modul de operare va rezulta din cele spuse deja. În fine, comportarea unei serii de puteri cu coeficienţi vectoriali la capetele domeniului de convergenţă se reduce la aceea din Teorema 4.4.13.

Un caz particular important al acestor dezvoltări este acela referitor la serii de puteri cu coeficienţi matriciali. Anume, dată perechea (p, q) de numere naturale şi definind şirul de matrici din $\mathcal{M}(p, q)$

$$A_1 = (a_1(i, j)), A_2 = (a_2(i, j)), ..., A_n = (a_n(i, j)), ...$$

se poate considera seria de puteri $\sum_n A_n x^n$. În baza Teoremei 4.4.15 (scrisă în variantă matricială) toate formulele şi procedeele de lucru rămân valabile. Singura menţiune specială se referă la faptul că aici, în anume condiţii, putem efectua şi produsul (în sens Cauchy) a două serii de puteri matriciale.

Exemplu. Fie p număr natural dat şi $A = (a_{ij})$ o matrice din $\mathcal{M}(p)$. Să considerăm seria de puteri (cu coeficienţi din $\mathcal{M}(p)$)

$$\sum_n \frac{A^n}{n!} x^n = I + \frac{A}{1!} x + \frac{A^2}{2!} x^2 + \cdots + \frac{A^n}{n!} x^n + \cdots .$$

Raza de convergenţă a acestei serii este

$$\rho = \lim_{n\to\infty} \frac{\|A^n\|}{n!} \cdot \frac{(n+1)!}{\|A^{n+1}\|} \geq \lim_{n\to\infty} \frac{n+1}{\|A\|} = \infty.$$

Deci, seria este local–uniform convergentă pe toată axa reală (R). Conform unei convenţii anterioare, suma ei se va scrie sub forma:

(D3) $e^{Ax} = I + \frac{A}{1!} x + \frac{A^2}{2!} x^2 + \cdots + \frac{A^n}{n!} x^n + \cdots, \ x \in R$

Aceasta, din proprietatea anterioară, va avea derivate de orice ordin pe axa reală. Derivata acesteia se obţine prin însumarea seriei derivate

$$\begin{cases} \dfrac{d}{dx}(e^{Ax}) = A + \dfrac{A^2}{1!}x + \cdots + \dfrac{A^n}{(n-1)!}x^{n-1} + \cdots = \\ = A(I + \dfrac{A}{1!}x + \cdots + \dfrac{A^{n-1}}{(na)!}x^{n-1} + \cdots) = Ae^{Ax}, \ \forall x \in R. \end{cases}$$

Cum, pe de altă parte

$$e^{A0} = I + \frac{A}{1!}0 + \cdots \frac{A^n}{n!}0 + \cdots = I,$$

se poate spune că funcţia matricială

$$F : R \longrightarrow \mathcal{M}(p), \ F(x) = e^{Ax}, \quad x \in R$$

satisface condiţiile

$$F'(x) = AF(x), \ x \in R; \ F(0) = I.$$

Vom numi aceasta, *funcţia exponenţială de matrice* A. Ea joacă un rol important în teoria sistemelor diferenţiale, după cum vom vedea cu altă ocazie.

Încheiem aceste consideraţii asupra seriilor de puteri cu următoarea observaţie. Fie $x_0 \in R$ un număr finit. Numim *serie de puteri centrată în* x_0 orice serie de funcţii de forma

(D4) $\displaystyle\sum_n a_n(x - x_0)^n = a_0 + a_1(x - x_0) + a_n(x - x_0)^n + \cdots$,

unde (a_n) este un şir de numere reale. Toate proprietăţile acestor serii de puteri se reduc la acelea dinainte; deoarece, prin substituţia $x - x_0 = y$, o asemenea serie se reduce la

$$\sum_n a_n y^n = a_0 + a_1 y + \cdots + a_n y^n + \cdots \ .$$

De exemplu, dacă ρ este raza de convergenţă a acestei din urmă serii, atunci mulţimea de convergenţă a seriei iniţiale satisface

$$(x_0 - \rho, x_0 + \rho) \subseteq D \subseteq [x_0 - \rho, x_0 + \rho].$$

Proprietăţile de continuitate şi derivabilitate ale sumei seriei pe intervalul $(x_0-\rho, x_0+\rho)$ rămân aceleaşi.În particular,notând cu S funcţia sumă avem dezvoltarea Taylor (pe intervalul $(x_0-\rho, x_0+\rho)$)

$$S(x) = S(x_0) + \frac{S'(x_0)}{1!}(x-x_0) + \cdots + \frac{S^{(n)}(x_0)}{n!}(x-x_0)^n + \cdots.$$

Legătura cu dezvoltarea de tip Taylor a unei funcţii arbitrare se face în acelaşi mod ca mai înainte. De asemenea, comportarea în capetele mulţimii de convergenţă are acelaşi caracter ca în cazul $x_0 = 0$. În fine, este clar că toate aceste consideraţii se pot extinde la cazul coeficienţilor vectoriali (sau matriciali).

Probleme propuse la §4.4

1. Să se calculeze normele următoarelor funcţii din $F(R, R)$:

a) $f(x) = \dfrac{1}{1+|x|}$, b) $f(x) = \sin x$; *c*) $f(x) = e^x$.

2. Să se arate că şirul de funcţii (din $F((0,\infty), R)$), $f_n(x) = \frac{1}{n}e^{-nx}$, $x > 0$, $n = 1, 2, ...$, converge uniform către funcţia identic zero (din $F((0,\infty), R)$).

3. Se dă şirul de funcţii $f_n(x) = \dfrac{x}{n+x}$, $x \in (0,\infty), n = 1, 2, ...$ Să se precizeze tipul de convergenţă (uniformă, local–uniformă, punctuală) a acestui şir pe intervalul precizat.

4. Fie dată seria de funcţii (din $C([0,1], R)$),

$$\sum_{n=1}^{\infty}\left(\frac{nx}{1+n^2x^2} - \frac{(n-1)x}{1+(n-1)x^2}\right).$$

Să se arate că seria converge ne–uniform pe $[0,1]$ dar suma ei este totuşi o funcţie continuă pe $[0,1]$.

5. Să se arate că seriile următoare converg uniform pe mulţimile indicate, iar sumele lor sunt funcţii continue pe aceste mulţimi:

a) $\displaystyle\sum_n \frac{\sin nx}{\sqrt{n^n+x^2}}$, $x \in R$; b) $\displaystyle\sum_n \frac{a_n}{n^x}$, $0 \le x < \infty$.

(Aici, (a_n) este aşa încât seria $\displaystyle\sum_n |a_n|$ converge).

6. Să se determine mulţimea de convergenţă şi suma următoarelor serii de puteri:

a) $\sum_{n=0}^{\infty}(n+1)x^n$; b) $\sum_{n=0}^{\infty}\frac{2n+3}{n!}x^n$.

7. Să se arate că seria de puteri $\sum_{n=1}^{\infty}\frac{x^n}{n^2}$ este convergentă pe $[-1,1]$, iar suma acesteia, S, verifică ecuaţia

$$(1-x)S'(1-x)-xS'(x)=\ln\frac{1-x}{x},\quad 0<x\leq 1.$$

8. Date funcţiile de mai jos, se cere dezvoltarea în serie Mac Laurin a acestora, şi intervalul pe care este valabilă:

a) $f(x)=\frac{1}{2}\ln\frac{1-x}{1+x}, -1<x<1$; b) $f(x)=\frac{3x}{x^2+5x+6}, x>-2$;

c) $f(x)=\sin^3 x,\ x\in R$; d) $f(x)=e^{-x^2},\ x\in R$.

9. Să se găsească seria de puteri, convergentă pe R, a cărei sumă S, verifică ecuaţia $S''(x)-S(x)=0,\ \forall x\in R$.

10. Să se studieze convergenţa şi să se calculeze suma seriei de puteri vectoriale $\sum_n\left(\frac{1}{n+1},\frac{1}{n!}\right)x^n$.

11. Să se calculeze funcţia exponenţială matricială e^{Ax}, în fiecare din cazurile:

a) $A=\begin{pmatrix}3&0\\0&2\end{pmatrix}$, b) $A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$, c) $A=\begin{pmatrix}0&1&2\\0&0&1\\0&0&0\end{pmatrix}$.

12. Să se studieze seriile de puteri (centrate în jurul punctelor de mai jos):

a) $\sum_n n(x-1)^n$; b) $\sum_n\frac{2n+3}{n!}(x+1)^n$.

13. Să se dezvolte în serie Taylor în jurul punctului $x=\frac{\pi}{2}$, funcţia $f(x)=\sin x$.

14. Fie dată seria cu coeficienţi matriciali $\sum_n\begin{pmatrix}1&n\\-n&2\end{pmatrix}(x-2)^n$. Să se studieze convergenţa şi să se calculeze suma acesteia.

4.5 Integrarea funcţiilor

4.5.1 Primitiva (integrala nedefinită)

Fie I un interval din R şi $f : I \longrightarrow R$ funcţie dată. Spunem că funcţia $F : I \longrightarrow R$ este o *primitivă* a funcţiei f (pe intervalul I) dacă

(D1) F este derivabilă pe intervalul I,

(D2) $F'(x) = f(x)$, oricare ar fi $x \in I$.

Vom scrie acest lucru prin formula

$$F(x) = \int f(x)dx \ (\ F \text{ este integrala nedefinită a funcţiei } f).$$

Problema primitivelor comportă două aspecte:

– condiţii (necesare sau suficiente) de existenţă;

– formule/metode de calcul .

(A) În ce priveşte condiţiile necesare de existenţă a primitivei, următorul rezultat (dat fără demonstraţie) se va dovedi edificator.

Teorema 4.5.1 *Dacă funcţia $F : I \longrightarrow R$ este derivabilă pe I, atunci derivata sa, F', are proprietatea Darboux pe acel interval; adică*

(D3) $x_1, x_2 \in I,\ x_1 < x_2,\ \lambda \in (F'(x_1)), F'(x_2)) \Longrightarrow$
$\exists x_3 \in (x_1, x_2), F'(x_3) = \lambda.$

De aici urmează imediat implicaţia

(P1) f are primitive (pe I) $\Longrightarrow$ f are proprietatea Darboux (pe I).

Deci, dacă o asemenea proprietate nu are loc, funcţia în chestiune nu are primitive (pe intervalul considerat).

Exemplu. Fie dată funcţia

$$f(x) = \begin{cases} -1, & \text{dacă } x \leq 0 \\ \ \ 2, & \text{dacă } x > 0. \end{cases}$$

Funcţia f nu are proprietatea Darboux (deoarece, de pildă, valoarea intermediară $1 \in (-1, 2)$ nu este atinsă în nici un punct). Deci f nu are primitive pe R.

Relativ la condiţiile suficiente de existenţă, vom porni de la implicaţia (care va fi demonstrată mai târziu)

(P2) f este continuă (pe I) $\Longrightarrow$ f are primitive (pe I).

Acesta, de altfel, este cadrul natural al tuturor metodelor de calcul ce vor fi expuse în continuare. Trebuie însă de remarcat că există şi funcţii discontinue care admit primitive, după cum rezultă din următorul

Exemplu. Fie dată funcţia

$$f(x) = \begin{cases} 2x\sin\frac{1}{x} - \cos\frac{1}{x}, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

O primitivă a ei este, de exemplu, funcţia

$$F(x) = \begin{cases} x^2\sin\frac{1}{x}, & x \neq 0 \\ 0, & x = 0, \end{cases}$$

după cum se poate vedea imediat. Pe de altă parte, f nu este continuă în origine (deoarece $\lim\limits_{x\to 0}\cos\frac{1}{x}$ nu există); şi de aici, concluzia.

Un aspect important al chestiunilor dezbătute este acela privind structura mulţimii primitivelor unei funcţii. Avem în acest sens (din proprietăţile de medie ale funcţiilor derivabile)

(P3) $\begin{cases} F_1, F_2 & = & \text{primitive ale funcţiei } f \text{ (pe } I) \Longrightarrow \\ F_1 - F_2 & = & \text{constant (pe intervalul } I). \end{cases}$

Vom exprima aceasta prin formula

$$\int f(x)dx = F(x) + C, \quad C = \text{constantă}.$$

Pentru motive de simplitate, vom renunţa totuşi la scrierea acestei constante, în continuare.

(B) Metodele de calcul ale primitivelor se bazează în principiu pe următoarele proprietăţi ale acestora. (Toate funcţiile care apar sunt presupuse continue).

(P4) $\int(f(x) + g(x))dx = \int f(x)dx + \int g(x)dx$.

(Primitiva sumei este egală cu suma primitivelor.)

(P5) $\int \lambda f(x)dx = \lambda \int f(x)dx, \quad \lambda =$constantă.

(O constantă poate fi scoasă în faţa primitivei.)

(P6) $\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx.$

(Formula de integrare prin părţi.)
Mai adăugăm la acestea şi următorul rezultat important cunoscut sub numele de "Formula de Schimbare a Variabilei în Integrala Nedefinită":

Teorema 4.5.2 *Fie* $f : R \longrightarrow R$ *o funcţie continuă şi* $\varphi : I \longrightarrow J$ *o funcţie derivabilă (deci şi continuă), unde* I, J *sunt intervale. Au loc atunci concluziile*

(i) *dacă* φ' *este în plus continuă, atunci*

(R1) $\int f(x)dx = F(x) \Longrightarrow \int f(\varphi(t))\varphi'(t)dt = F(\varphi(t)).$

(ii) *dacă* φ' *este continuă şi nicăieri nulă, atunci funcţia inversă* $\varphi^{-1} : J \longrightarrow I$ *este derivabilă (cu derivata continuă) şi*

(R2) $$\int f(\varphi(t))\varphi'(t)dt = G(t) \Longrightarrow \int f(x)dx = G(\varphi^{-1}(x)).$$

(Demonstraţia se bazează pe formula de derivare a funcţiei compuse şi a funcţiei inverse.)

Toate aceste formule se pot acum suprapune în procesul calculului, peste diferite proceduri funcţionale (implicite sau explicite). De pildă, uneori primitiva poate fi găsită prin rezolvarea unei anumite ecuaţii funcţionale după cum o probează următorul

Exemplu. Să se calculeze $\int e^x \sin x\, dx$. Folosim în primul rând formula de integrare prin părţi, punând

$$f(x) = e^x \ (\Longrightarrow f'(x) = e^x);\ g'(x) = \sin x \ (\Longrightarrow g(x) = -\cos x).$$

Găsim deci egalitatea

$$\int e^x \sin x\, dx = -e^x \cos x + \int e^x \cos x\, dx.$$

Pentru aceasta din urmă folosim acelaşi procedeu, cu

$$f(x) = e^x \ (\Longrightarrow f'(x) = e^x);\ g'(x) = \cos x \ (\Longrightarrow g(x) = \sin x).$$

Obţinem, astfel

$$\int e^x \cos x\, dx = e^x \sin x - \int e^x \sin x\, dx.$$

Înlocuind în precedenta, ajungem la

$$\int e^x \sin x\, dx = e^x(\sin x - \cos x) - \int e^x \sin x\, dx.$$

Aceasta este o ecuaţie funcţională verificată de primitiva pe care o căutăm. Rezolvând aceasta (prin trecerea în partea stângă) găsim

$$\int e^x \sin x\, dx = \frac{1}{2} e^x(\sin x - \cos x).$$

În alte situaţii, primitiva poate fi determinată în urma unui proces iterativ funcţional. Pentru ilustrare, vom trata următorul

Exemplu. Să se calculeze primitivele

$$F_n(x) = \int \frac{dx}{(x^2+a^2)^n}, \quad n = 1, 2, \dots .$$

Folosim deocamdată scrierea (pentru $n \geq 2$)

$$F_n(x) = \frac{1}{a^2} \int \frac{x^2 + a^2 - x^2}{(x^2+1)^n} dx = \frac{1}{a^2} \left(F_{n-1}(x) - \int \frac{x^2 dx}{(x^2+a^2)^n} \right).$$

Ultima integrală ce apare o calculăm prin părţi, cu

$$f(x) = x, \quad g'(x) = \frac{x}{(x^2+a^2)^n}$$

Avem astfel (pentru $n \geq 2$)

$$\int \frac{x^2\, dx}{(x^2+a^2)^n} = \frac{-x}{2(n-1)(x^2+a^2)^{n-1}} + \frac{1}{2(n-1)} \int \frac{dx}{(x^2+a^2)^{n-1}}.$$

Şi atunci, înlocuind în relaţia dinainte, obţinem (pentru $n \geq 2$)

$$F_n(x) = \frac{1}{a^2} \left[F_{n-1}(x) - \left(\frac{-x}{2(n-1)(x^2+a^2)^{n-1}} + \frac{1}{2(n-1)} F_{n-1}(x) \right) \right],$$

adică

$$F_n(x) = \frac{1}{a^2} \left[\frac{2n-3}{2n-1} F_{n-1}(x) + \frac{x}{2(n-1)(x^2+a^2)^{n-1}} \right], n \geq 2.$$

Combinând aceasta cu $F_1(x) = \frac{1}{a}\text{arctg}\,\frac{x}{a}$ urmează că, din aproape în aproape, putem determina toate primitivele în cauză.

Putem trece acum la calculul primitivelor pentru anumite clase de funcţii. Începem cu cele raţionale. Să numim *fracţie simplă*, orice funcţie raţională de forma

$$\varphi_m(x) = \frac{\gamma}{(x-a)^m},\ m = 1, 2, ...,\quad (a \in R),$$

sau, respectiv,

$$\psi_n(x) = \frac{\lambda x + \mu}{(x^2 + px + q)^n},\ n = 1, 2, ...,\quad (p^2 - 4q < 0).$$

(Aici, γ, λ, μ sunt nişte constante.) Să notăm că, în baza exemplului precedent, este posibil să calculăm primitivele pentru toate aceste fracţii simple. Fie acum

$$f(x) = \frac{P(x)}{Q(x)},\quad P(x), Q(x) = \text{polinoame},$$

o fracţie raţională arbitrară. Presupunem cunoscute rădăcinile numitorului; şi, deci, factorizarea acestuia

$$Q(x) = \prod_{i=1}^{k}(x - a_i)^{m_i} \cdot \prod_{j=1}^{h}(x^2 + p_j x + q_j)^{n_j}.$$

(Să notăm că prima serie de factori corespunde rădăcinilor reale, iar a doua serie, rădăcinilor complex conjugate.)

Teorema 4.5.3 *Avem valabilă descompunerea*

$$\begin{aligned} &f(x) = R(x)+ \\ &+\sum_i \{\textit{fracţii simple ale factorului } (x-a_i)^{m_i}\}+ \\ &+\sum_j \{\textit{fracţii simple ale factorului } (x^2+p_jx+q_j)^{n_j}\} \end{aligned}$$

unde, prin definiţie,

$$R(x) = \textit{câtul împărţirii lui } P(x) \textit{ la } Q(x)$$

iar, în general,

$$\text{(D4)} \quad \begin{cases} \textit{fracţia simplă corespunzând factorului } (x-a)^m \textit{ este} \\ \varphi(x) = \dfrac{\gamma_1}{x-a} + \dfrac{\gamma_2}{(x-a)^2} + \cdots + \dfrac{\gamma_m}{(x-a)^m}, \end{cases}$$

$$\text{(D5)} \quad \begin{cases} \textit{fracţia simplă corespunzând factorului } (x^2+px+q)^n \textit{ este} \\ \psi(x) = \dfrac{\lambda_1 x + \mu_1}{x^2+px+q} + \cdots + \dfrac{\lambda_n x + \mu_n}{(x^2+px+q)^n}. \end{cases}$$

(Constantele care apar în toate aceste fracţii se obţin în mod univoc prin identificare.)

Cu acest rezultat, se poate spune că putem determina primitivele funcţiilor raţionale. Să trecem acum în revistă principalele metode de calcul ale primitivelor bazate pe reducerea la cazul raţional.

(i) O primă clasă este a primitivelor de forma $\displaystyle\int R(\varphi(x))dx$, în care

(C1) $R(x) = \dfrac{P(x)}{Q(x)}$ =funcţie raţională

iar funcţia φ este inversibilă cu inversa $\psi = \varphi^{-1}$ având proprietatea

(C2) $\psi'(t)$ =funcţie raţională de variabila t.

În acest caz, prin substituţia

$$t = \varphi(x), \quad \text{sau} \quad x = \psi(t) \quad (\Longrightarrow dx = \psi'(t)dt)$$

avem (cu Formula de Schimbare a Variabilei)

$$\int R(\varphi(x))dx = \int R(t)\psi'(t)dt,$$

iar primitiva din dreapta este dintr-o funcţie raţională.
Exemple de acest fel (cu substituţiile aferente):

$$\begin{array}{lll} \int R(e^{ax})dx; & t = e^{ax} & (\text{sau } x = \frac{1}{a}\ln(t)) \\ \int R(\operatorname{tg} bx)dx; & t = \operatorname{tg} bx & (\text{sau } x = \frac{1}{b}\operatorname{arctg}(t)). \end{array}$$

(ii) O a doua clasă este constituită de primitivele de forma $\int R(\sin x, \cos x)dx$, unde

(C3) $R(u,v) = \frac{P(u,v)}{Q(u,v)}$ (funcţie raţională de două variabile).

În acest caz, se face substituţia

$$\operatorname{tg}\frac{x}{2} = t \quad \text{sau} \quad x = 2\operatorname{arctg} t \ \ (\Longrightarrow dx = \frac{2\,dt}{1+t^2}).$$

Combinând cu formulele cunoscute

$$\sin x = \frac{2t}{1+t^2}, \quad \cos x = \frac{1-t^2}{1+t^2},$$

ajungem să expIimăm integrala sub forma (raţională)

$$\int R(\sin x, \cos x)dx = \int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)\frac{2}{1+t^2}dt.$$

În unele cazuri particulare se pot folosi substituţii mai avantajoase:

a) $R(-u, v) = -R(u, v) \Longrightarrow \cos x = t$

b) $R(u, -v) = -R(u, v) \Longrightarrow \sin x = t$

c) $R(-u, -v) = R(u, v) \Longrightarrow \operatorname{tg} x = t$

(iii) O altă clasă este dată de primitivele de forma

$$\int R(x, \sqrt[n]{\frac{\alpha x+\beta}{\gamma x+\delta}})dx \quad \text{unde} \quad R(u, v)\text{=funcţie raţională.}$$

Substituţia care se potriveşte aici este

$$t = \sqrt[n]{\frac{\alpha x+\beta}{\gamma x+\delta}} \ \left(\text{sau} \quad x = \frac{\delta t^n - \beta}{\alpha - \gamma t^n}\right).$$

Ca şi mai înainte, se obţine în final o reducere la primitive de funcţii raţionale; nu mai dăm alte detalii.

(iv) Urmează clasa de integrale de forma

$$\int R(x, \sqrt{ax^2+bx+c})dx \ \ , \ R(u, v) = \text{funcţie raţională.}$$

Substituţiile care se fac aici sunt următoarele

- în cazul $a > 0 : \sqrt{ax^2+bx+c} = x\sqrt{a} + t$

- în cazul $c > 0$: $\sqrt{ax^2 + bx + c} = tx + \sqrt{c}$

- în cazul $a < 0, c \leq 0$: $\sqrt{ax^2 + bx + c} = t(x - x_1)$ (sau $t(x - x_2)$).

(Aici, x_1, x_2 sunt rădăcinile ecuaţiei $ax^2 + bx + c = 0$.)

(v) În fine, o ultimă clasă este aceea a primitivelor de tip binom

$$\int x^m (ax^n + b)^p dx, \; m, n, p = \text{numere raţionale.}$$

Există doar trei cazuri în care se poate face reducerea la primitive de funcţii raţionale:

d) $p =$ număr întreg: se face substituţia $x = t^r$, unde $r =$ cel mai mic multiplu comun al numitorilor lui m şi n;

e) $\frac{m+1}{n} =$ număr întreg: se face substituţia $ax^n + b = t^s$, unde $s =$ numitorul lui p;

f) $\frac{m+1}{n} + p =$ număr întreg: se face substituţia $\frac{ax^n+b}{x^n} = t^s$, unde $s =$ numitorul lui p.

Exemplu. Să se calculeze $\int \frac{\sqrt[3]{1+\sqrt[4]{x}}}{\sqrt{x}} dx$. Este vorba de o primitivă de tip binom, cu $m = -\frac{1}{2}$, $n = \frac{1}{4}$, $p = \frac{1}{3}$· Deoarece $\frac{m+1}{n} = \frac{1}{2} \cdot \frac{4}{1} = 2 =$ număr întreg, se va face substituţia

$$1 + \sqrt[4]{x} = t^3 \;\text{ sau }\; x = (t^3 - 1)^4 \; (\Longrightarrow dx = 4(t^3 - 1)^3 \cdot 3t^2 dt).$$

Înlocuind în integrală, avem

$$\int \frac{\sqrt[3]{1 + \sqrt[4]{x}}}{\sqrt{x}} dx = \int \frac{t}{(t^3 - 1)^2} \cdot 4(t^3 - 1)^3 \cdot 3t^2 dt = 12 \int t^3 (t^3 - 1) dt =$$
$$= 12 \left(\frac{t^7}{7} - \frac{t^4}{4} \right) = \frac{12}{7} \sqrt[3]{(1 + \sqrt[4]{x})^7} - 3 \sqrt[3]{(1 + \sqrt[4]{x})^4}.$$

4.5.2 Integrala simplă (pentru funcţii de o variabilă)

Vom introduce în cele ce urmează noţiunea de integrală (simplă) pentru funcţii de o singură variabilă. Acest lucru va fi făcut în mai multe etape.

Începem discuţia prin precizarea convenţiilor şi noţiunilor de bază, folosite în continuare. Pentru fiecare interval (mărginit) I al axei reale, definim *lungimea* sa prin

(D1) $\text{lung}\,(I) = d - c$, dacă $I = [c, d), (c, d], [c, d]$ sau (c, d).

Vom zice acum că mulţimea A din R este de *lungime zero* dacă pentru orice $\varepsilon > 0$ există un şir (I_n) de intervale care acoperă A, cu suma lungimilor acestora mai mică decât ε :

(D2) $A \subseteq \bigcup\limits_{n=1}^{\infty} I_n, \quad \sum\limits_{n=1}^{\infty} \text{lung}\,(I_n) < \varepsilon.$

Vom scrie această constatare prin $\text{lung}\,(A) = 0$. Este uşor de arătat că implicaţiile următoare au loc:

(P1) $\text{lung}\,(A) = 0,\ B \subseteq A \Longrightarrow \text{lung}\,(B) = 0.$

(P2) $\text{lung}\,(A_n) = 0,\ n = 1, 2, ... \Longrightarrow \text{lung}\,(\bigcup\limits_{n=1}^{\infty} A_n) = 0.$

În particular, de aici rezultă constatarea:

(P3) orice mulţime numărabilă are lungime zero.

Prin definiţie, vom zice că proprietatea $P(x)$, de variabilă $x \in R$ are loc *aproape peste tot*, dacă

(D3) $\{x \in R;\ P(x) \text{ nu are loc}\}$ este de lungime zero;

sau, cu alte cuvinte: proprietatea $P(x)$ este verificată pentru toţi $x \in R$ cu excepţia unei mulţimi de lungime zero.

Definim în continuare o clasă de funcţii cu rol central în teoria integralei. Să notăm ca de obicei

$$F(R, R) = \text{mulţimea funcţiilor } f : R \longrightarrow R.$$

Zicem că funcţia g din această mulţime este *etajată* (sau *în scară*), dacă există o diviziune a axei reale

$$(\Delta) : -\infty < x_1 < x_2 < \cdots < x_n < \infty \quad (\text{unde } n \geq 2),$$

şi un sistem de constante $\{\lambda_1, ..., \lambda_{n-1}\}$, aşa încât

$$g(x) = \begin{cases} 0, & \text{dacă} \quad x \in (-\infty, x_1) \cup (x_n, \infty) \cup \{x_1, ..., x_n\} \\ \lambda_1, & \text{dacă} \quad x \in (x_1, x_2) \\ \cdots & \cdots \quad \cdots \\ \lambda_{n-1}, & \text{dacă} \quad x \in (x_{n-1}, x_n). \end{cases}$$

Să facem notaţia

$$E(R,R) = \{g : R \longrightarrow R; \quad g \text{ este etajată}\}.$$

Această mulţime apare ca subspaţiu liniar al lui $F(R,R)$, adică

(P4) $g_1, g_2 \in E(R,R) \Longrightarrow \alpha g_1 + \beta g_2 \in E(R,R),\ \forall \alpha, \beta \in R$

În plus, mai avem şi proprietatea

(P5) $g \in E(R,R) \Longrightarrow |g| \in E(R,R)$.

Să introducem acum aplicaţia de integrare Λ, pe clasa acestor funcţii etajate, prin convenţia

$$\text{(D4)} \begin{cases} \Lambda(g) = \lambda_1(x_2 - x_1) + \cdots + \lambda_{n-1}(x_n - x_{n-1}), \\ \text{dacă funcţia etajată } g \text{ este definită de diviziunea } \{x_1, ..., x_n\} \\ \text{şi de constantele } \{\lambda_1, ..., \lambda_{n-1}\}. \end{cases}$$

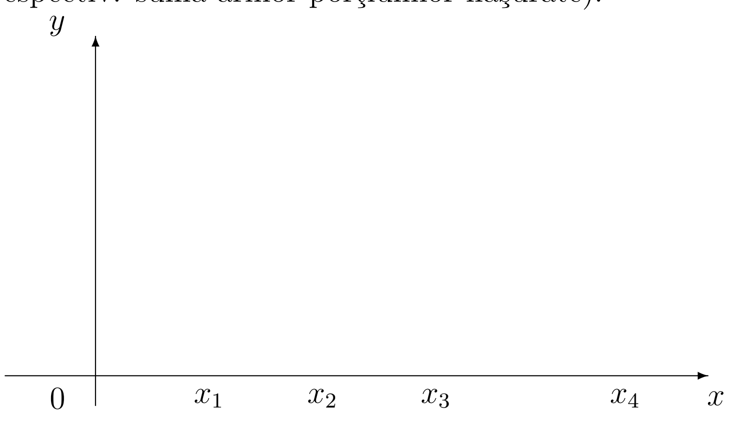
Astfel, de pildă, dacă funcţia etajată g are forma

$$g(x) = \begin{cases} 0, & \text{dacă} \quad x \in (-\infty, x_1) \cup (x_4, \infty) \cup \{x_1, x_2, x_3, , x_4\} \\ \lambda_1, & \text{pentru} \quad x \in (x_1, x_2) \\ \lambda_2, & \text{pentru} \quad x \in (x_2, x_3) \\ \lambda_3, & \text{pentru} \quad x \in (x_3, x_4), \end{cases}$$

atunci, conform definiţiei,

$$\Lambda(g) = \lambda_1(x_2 - x_1) + \lambda_2(x_3 - x_2) + \lambda_3(x_4 - x_3).$$

Semnificaţia geometrică a acestei mărimi este următoarea (în cazul pozitiv): $\Lambda(g)$ reprezintă aria porţiunii de plan delimitată de graficul funcţiei g şi axa $0x$ (respectiv: suma ariilor porţiunilor haşurate).



Integrala astfel definită are proprietatea de liniaritate

(P6) $\Lambda(\delta_1 g_1 + \delta_2 g_2) = \delta_1\Lambda(g_1) + \delta_2\Lambda(g_2),\ \delta_1, \delta_2 = \text{constante}$

şi proprietatea de monotonie

(P7) $g_2(x) \leq g_2(x),\ x \in R \Longrightarrow \Lambda(g_1) \leq \Lambda(g_2),$

O importantă consecinţă a acesteia este proprietatea de continuitate

(P8) $|\Lambda(g)| \leq \Lambda(|g|),\ g \in E(R, R).$

(A) Vom introduce acum noţiunea de integrală (simplă) a unei funcţii pe un interval (a, b), $-\infty \leq a < b \leq \infty$. Să notăm deocamdată

$$F((a, b), R) = \text{mulţimea tuturor funcţiilor } f : (a, b) \longrightarrow R.$$

O clasă importantă de asemenea funcţii este

$$E((a, b), R) = \{g \in E(R, R);\ g(x) = 0,\ \forall x \in R \setminus (a, b)\}$$

(mulţimea tuturor funcţiilor etajate care se anulează în afara intervalului (a, b)). Definim acum *integrala (Lebesgue)* a unei asemenea funcţii etajate prin convenţia

(D5) $\int_a^b g(x)dx = \Lambda(g)$, pentru $g \in E((a, b), R)$,

unde aplicaţia Λ este cea introdusă anterior.
Fie acum $f : (a, b) \longrightarrow [0, \infty)$ o funcţie cu valori pozitive. Vom zice că aceasta este *măsurabilă (Lebesgue)* dacă există un şir de funcţii etajate (pe (a, b)) cu valori pozitive, (g_n) care, aproape peste tot în (a, b), este crescător şi convergent către f, în sensul

(D6) $\begin{cases} (g_n(x)) \text{ este monoton crescător şi } g_n(x) \longrightarrow f(x) \text{ pentru toţi} \\ x \in (a, b) \text{ cu excepţia unei mulţimi de lungime zero.} \end{cases}$

În acest caz, vom pune prin definiţie

(D7) $\int_a^b f(x)dx = \lim\limits_{n\to\infty} \int_a^b g_n(x)dx,$

şi vom numi aceasta *integrala (Lebesgue)* a lui f pe intervalul (a, b). (Să notăm că limita din dreapta există, deoarece şirul $(\int_a^b g_n(x)dx)$ este crescător (conform proprietăţilor anterioare). În plus, definiţia nu depinde de şirul (g_n) folosit în caracterizarea lui f).

Desigur, această integrală poate fi şi $(+\infty)$; în cazul când ea este finită, vom zice că f este *sumabilă (în sens Lebesgue)* pe (a, b).

În fine, fie $f : (a, b) \longrightarrow R$ o funcţie cu valori oarecare. Construim următoarea pereche de funcţii cu valori pozitive

(D8) $\begin{cases} f_+(x) = \max(0, f(x)),\ x \in (a,b) \ \text{(partea pozitivă din } f) \\ f_-(x) = \max(0, -f(x)),\ x \in (a,b) \ \text{(partea negativă din } f). \end{cases}$

Constatăm imediat valabilitatea egalităţilor

(P9) $f(x) = f_+(x) - f_-(x),\ |f(x)| = f_+(x) + f_-(x),\ x \in (a,b).$

Vom zice că f este *măsurabilă (Lebesgue)* pe (a,b), dacă f_+ şi f_- sunt simultan măsurabile (în sensul anterior) pe (a,b). Dacă, în plus, acestea sunt sumabile pe (a,b), vom zice că f este *sumabilă* pe (a,b). Notăm, în acest caz

(D9) $\int_a^b f(x)dx = \int_a^b f_+(x)dx - \int_a^b f_-(x)dx$

şi vom numi aceasta *integrala (Lebesgue)* a lui f pe intervalul (a,b). Vom mai pune

(D10) $L((a,b),R) = \{f : (a,b) \longrightarrow R;\ f \text{ este sumabilă pe } (a,b)\}.$

În ce priveşte structura acestei clase şi proprietăţile integralei, menţionăm în primul rând (liniaritatea)

(P10) $\begin{cases} \delta_1,\delta_2 \in R, f_1, f_2 \in L((a,b),R) \Longrightarrow \delta_1 f_1 + \delta_2 f_2 \in L((a,b),R) \\ \text{şi } \int_a^b (\delta_1 f_1(x) + \delta_2 f_2(x))dx = \delta_1 \int_a^b f_1(x) + \delta_2 \int_a^b f_2(x)dx. \end{cases}$

(Deci, în particular, $L((a,b),R)$ apare ca subspaţiu liniar al lui $F((a,b),R)$). Avem apoi

(P11) $\begin{cases} \text{(monotonia): } f_1(x) \le f_2(x), \text{ pentru aproape toţi } x \in (a,b) \\ \Longrightarrow \int_a^b f_1(x)dx \le \int_a^b f_2(x)dx. \end{cases}$

Ca o importantă consecinţă a acestui fapt, avem de aici şi următoarea proprietate de continuitate

(P12) $\begin{cases} f \in L((a,b),R) \text{ dacă şi numai dacă } |f| \in L((a,b),R); \\ \text{şi, în acest caz, } \left|\int_a^b f(x)dx\right| \le \int_a^b |f(x)|dx. \end{cases}$

Mai adăugăm la aceasta proprietatea de identitate

(P13) $\begin{cases} f, g \in L((a,b),R), f(x) = g(x) \text{ pentru aproape toţi } x \in (a,b) \\ \Longrightarrow \int_a^b f(x)dx = \int_a^b g(x)dx. \end{cases}$

Menţionăm apoi şi proprietatea ereditar–aditivă

(P14) $\begin{cases} f \in L((a,b),R) \Longrightarrow f \in L((c,d),R), \\ \text{pentru orice subinterval } (c,d) \text{ din } (a,b); \text{ în plus,} \\ \int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx,\ \forall c \in (a,b). \end{cases}$

În fine, ca o ultimă proprietate de acest tip, dăm fără demonstraţie următorul rezultat (de tip Lebesgue).

Teorema 4.5.4 *Fie (f_n) un şir de funcţii sumabile pe (a,b) care converge aproape peste tot pe (a,b) către o funcţie $f:(a,b)\longrightarrow R$. Admitem, în plus, că una din condiţiile următoare are loc:*

(C1) *$(f_n(x))$ este crescător pentru aproape toţi $x\in(a,b)$ iar şirul numeric $(\int_a^b f_n(x)dx)$ este mărginit superior;*

(C2) *există o funcţie sumabilă $h:(a,b)\longrightarrow R$ cu $|f_n(x)|\le h(x)$, pentru aproape toţi $x\in(a,b)$.*

Atunci, obligatoriu, f este sumabilă pe (a,b), cu

$$\int_a^b f(x)dx = \lim_{n\to\infty}\int_a^b f_n(x)dx.$$

(B) Fie în continuare intervalul (a,b) mărginit $(-\infty<a<b<\infty)$. Ne interesează care sunt principalele clase de funcţii sumabile pe (a,b). Fie $f:(a,b)\longrightarrow R$ deocamdată arbitrară. Pentru orice pereche

$$\begin{array}{lll} (\Delta) & : \ a<x_1<x_2<\cdots<x_n<b & \text{(diviziune a lui } (a,b)) \\ (\xi) & : \ \xi_0\in(a,x_1),...,\xi_n\in(x_n,b) & \text{(sistem de puncte intermediare)} \end{array}$$

asociem funcţia etajată $g=g_{((\Delta),(\xi))}$ prin convenţia

$$g(x)=\begin{cases} 0, & \text{dacă} \quad x\in(\infty,a]\cup[b,\infty)\cup\{x_1,...,x_n\} \\ f(\xi_0), & \text{dacă} \quad x\in(a,x_1) \\ \cdots & \cdots \quad \cdots \\ f(\xi_n), & \text{dacă} \quad x\in(x_n,b). \end{cases}$$

Integrala acestei funcţii etajate

$$\int_a^b g(x)dx = f(\xi_0)(x_1-a)+f(\xi_1)(x_2-x_1)+\cdots+f(\xi_n)(b-x_n)$$

se va numi *suma Riemann* asociată funcţiei f diviziunii (Δ) şi punctelor intermediare (ξ). Să mai numim numărul

$$\text{norma } (\Delta)=\max(x_1-a,x_2-x_1,...,b-x_n)$$

norma diviziunii considerate. Să considerăm ipoteza

$$\text{(C3)} \left\{ \begin{array}{l} f \text{ este continuă aproape peste tot pe } (a,b) \\ \text{şi mărginită (peste tot) pe } (a,b). \end{array}\right.$$

Teorema 4.5.5 *În condiţiile date, f este sumabilă pe (a, b).*

Demonstraţie. (Schiţă) Fie

$$((\Delta_1), (\xi^{(1)})), ((\Delta_2), (\xi^{(2)})), ..., ((\Delta_n), (\xi^{(n)})), ... ,$$

un şir de diviziuni ale intervalului (a, b), împreună cu şirul corespunzător de puncte intermediare. Presupunem în plus că

$$\text{norma}\,(\Delta_n) \longrightarrow 0 \;\text{ pentru }\; n \to \infty.$$

Vom nota cu (g_n) şirul corespunzător de funcţii etajate asociate datelor de mai sus. Se poate arăta că

$$\lim_{n\to\infty} g_n(x) = f(x), \text{ aproape pentru toţi } \; x \in (a, b).$$

Cum funcţia f este în plus mărginită există $\mu \geq 0$ cu

$$|g_n(x)| \leq \mu, \; n \geq 1, \; x \in (a, b).$$

Şi atunci, din Teorema 4.5.4, f este sumabilă pe intervalul (a, b), cu

(R1) $\int_a^b f(x)dx = \lim\limits_{n\to\infty} \int_a^b g_n(x)dx.$

Teorema este dovedită. ∎

Prin definiţie, o funcţie mărginită $f : (a, b) \longrightarrow R$ cu proprietatea că limita din dreapta relaţiei (R1) există, se va zice *funcţie integrabilă Riemann* pe (a, b); iar limita însăşi se va numi *integrala Riemann* pe (a, b) din f. Faptul că nu este nevoie de o notaţie specială pentru această integrală rezultă din cele spuse mai înainte combinate cu următoarea proprietate structurală:

(P15) $\begin{cases} f = \text{mărginită şi integrabilă Riemann pe } (a, b) \Longrightarrow \\ f \text{ este continuă aproape peste tot pe } (a, b). \end{cases}$

În particular, de aici rezultă că , dacă

(C4) f este continuă şi mărginită pe (a, b)

atunci f este sumabilă pe (a, b). Concluziile obţinute permit acum şi abordarea unei chestiuni anterioare referitor la primitive.

Teorema 4.5.6 *În ipoteza de mai sus, funcţia*

$$F(x) = \int_a^x f(t)dt, \ a \leq x \leq b,$$

este funcţie Lagrange pe $[a,b]$ *şi* $F'(x) = f(x)$, $x \in (a,b)$. (*Adică* F *este o primitivă a funcţiei* f *pe* (a,b).)

Demonstraţie. Fie y arbitrar fixat în (a,b). Dat $\varepsilon > 0$, va exista $\delta > 0$ aşa încât (cum f este continuă)

$$t \in [a,b], |t-y| < \delta \Longrightarrow |f(t) - f(y)| < \varepsilon.$$

Avem, în acest fel, pentru orice $x \in (y, y+\delta) \cap [a,b]$,

$$\begin{aligned} &|F(x) - F(y) - (x-y)f(y)| = \left|\textstyle\int_y^x (f(t) - f(y))dt\right| \leq \\ &\leq \textstyle\int_y^x |f(t) - f(y)|dt \leq (x-y)\varepsilon. \end{aligned}$$

Iar, de aici, obţinem uşor evaluarea

$$\left|\frac{F(x) - F(y)}{x-y} - f(y)\right| \leq \varepsilon, \quad x \in (y, y+\delta) \cap (a,b),$$

ceea ce ne arată că $F'_d(y) = f(y)$. Analog, se arată că $F'_s(y) = f(y)$; şi, deci, $F'(y) = f(y)$. Continuitatea lui F pe $[a,b]$ este asigurată de mărginirea lui f pe (a,b), şi Teorema de medie a lui Lagrange. ■

Pentru asemenea funcţii are loc aşadar formula Newton–Leibniz

(R2) $\int_p^q f(x)dx = F(q) - F(p) \ \left(= F(x)|_p^q\right)$, $a \leq q < p \leq b$.

Rezultă de aici că fiecare metodă de calcul a primitivelor poate fi, în acelaşi timp, privită şi ca metodă de calcul a integralei (din asemenea funcţii). În particular, o condiţie suficientă pentru (C4) este

(C5) f este continuă pe $[a,b]$.

Într-un asemenea context, formula de integrare prin părţi (din teoria primitivei) are forma

(R3) $\int_a^b f(x)g'(x)dx = f(x)g(x)|_a^b - \int_a^b f'(x)g(x)dx.$

(Aici, funcţiile f şi g sunt de clasă C^1 pe $[a,b]$.)
De asemenea,mai trebuie menţionate formulele de medie

(R4) $\int_a^b f(x)dx = (b-a)f(c)$,pentru un $c \in (a,b)$,

(R5) $\int_a^b f(x)g(x)dx = g(a)\int_a^c f(x)dx + g(b)\int_c^b f(x)dx$,
pentru un $c \in [a,b]$.

(Funcţiile care apar aici sunt continue pe $[a,b]$; în plus,la a doua formulă, g este presupusă monotonă).

(C) Să ne întoarcem la cazul general. Admitem ipoteza

(C6) f este continuă pe (a,b).

Dacă a şi b sunt finite, iar f este mărginită pe (a,b), atunci, conform celor spuse anterior, f este sumabilă pe (a,b). Ne interesează ce se întâmplă când una din aceste două grupe de condiţii nu mai are loc; adică pentru capetele a şi b sunt deschise alternativele

(C7) $a = -\infty$; $\quad a > -\infty$, dar $\quad \lim\limits_{x \to a+} |f(x)| = \infty$.

(C8) $b = \infty$; $\quad b < \infty$, dar $\quad \lim\limits_{x \to b-} |f(x)| = \infty$.

Un răspuns util în acest sens este dat de

Teorema 4.5.7 *În condiţiile menţionate, f este sumabilă pe (a,b), dacă şi numai dacă*

(C9) $\lim\limits_{\substack{u \to a+ \\ v \to b-}} \int_u^v |f(x)|dx$ *este finită.*

Iar, în acest caz, integrala sa are evaluarea

(R6) $\displaystyle\int_a^b f(x)dx = \lim_{\substack{u \to a+ \\ v \to b-}} \int_u^v f(x)dx.$

(*Mai exact: limita din dreapta există şi egalează integrala în chestiune.*)

Demonstraţie. (Schiţă) Fie (u_n) şi (v_n) şiruri din (a,b), cu

$$u_1 > u_2 > \cdots,\ u_n \to a; \quad v_1 < v_2 < \cdots,\ v_n \to b.$$

Definim şirul de funcţii sumabile (f_n) prin

$$f_n(x) = \begin{cases} f(x), & x \in [u_n, v_n], \quad (n \geq 1) \\ 0 & \text{în rest.} \end{cases}$$

Se arată imediat că

$$|f_1(x)| \leq |f_2(x)| \leq \cdots;\ f_n(x) \to f(x),\ \forall x \in (a,b).$$

Cum încă mai avem şi evaluarea (pentru $n \geq 1$)

$$\int_a^b |f_n(x)|dx = \int_{u_n}^{v_n} |f(x)|dx \leq \lim_{n\to\infty} \int_{u_n}^{v_n} |f(x)|dx < \infty$$

urmează (cu Teorema 4.5.4) că $|f|$ este sumabilă pe (a,b) (deci şi f este sumabilă pe (a,b)). Ultima parte rezultă din criteriul Cauchy de existenţă a limitei (Teorema 4.2.1). ■

Să completăm aceasta cu o formulă de schimbare a variabilei, utilă în multe cazuri concrete:

Teorema 4.5.8 *Fie dată transformarea $x = \varphi(t)$, care aplică bijectiv intervalul (α,β) pe intervalul (a,b) unde φ este o aplicaţie de clasă C^1 pe (α,β) cu $\varphi(\alpha) = a, \varphi(\beta) = b$. Are loc atunci egalitatea*

(R7) $\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt.$

Mai precis: dacă funcţia $x \vdash f(x)$ este sumabilă pe (a,b) atunci funcţia $t \vdash f(\varphi(t))\varphi'(t)$ este sumabilă pe (α,β) şi are loc formula scrisă.

În legătură cu cele spuse anterior, următorul aspect important este de semnalat aici. Anume, ar fi posibil ca limita din partea dreaptă a relaţiei (R6) să existe fără ca funcţia f să fie sumabilă pe (a,b). Vom zice atunci că f este *sumabilă în sens generalizat* pe (a,b). (Notaţia pentru noua integrală va fi aceeaşi ca şi pentru integrala standard. Va trebui însă precizat în context cu ce fel de integrală se lucrează.) Să reţinem că, pe baza criteriului Cauchy de existenţă a limitei, f este sumabilă în sens generalizat pe (a,b) dacă şi numai dacă

$$\text{(C10)} \begin{cases} \text{pentru orice } \varepsilon > 0, \text{există un interval } (p,q) \text{ din } (a,b) \text{ aşa ca} \\ |\int_u^v f(x)dx| < \varepsilon \text{ pentru orice interval } (u,v) \subseteq (a,b) \setminus (p,q). \end{cases}$$

(D) Să studiem în continuare alternativele secunde din (C7) +(C8) (respectiv, $a, b =$ finite). Condiţia (C9) din Teorema 4.5.7 va fi verificată şi deci f este sumabilă pe (a,b), dacă

$$\text{(C11)} \begin{cases} 0 \leq \lim\limits_{x\to a+} (x-a)^\lambda |f(x)| < \infty \text{ şi } 0 \leq \lim\limits_{x\to b-} (b-x)^\mu |f(x)| < \infty, \\ \text{pentru un } \lambda < 1 \text{ şi un } \mu < 1. \end{cases}$$

Dimpotrivă, dacă una din condiţiile limită

$$(C12)\ \begin{cases} 0 < \lim\limits_{x\to a+}(x-a)^{\lambda}|f(x)| \le \infty \text{ sau } 0 < \lim\limits_{x\to b-}(b-x)^{\mu}|f(x)| \le \infty, \\ \text{pentru un } \lambda \ge 1, \text{ respectiv, un } \mu \ge 1, \end{cases}$$

are loc, atunci condiţia în chestiune nu se verifică; şi,deci, f nu este sumabilă pe (a,b).
Să dovedim afirmaţia referitoare la (C11). În baza primei jumătăţi a acesteia, există $\gamma > 0$ şi $c \in (a,b)$ aşa ca

$$(x-a)^{\lambda}|f(x)| \le \gamma \ \ (\text{deci } \ |f(x)| \le \frac{\gamma}{(x-a)^{\lambda}}),\ x \in (a,c).$$

Integrând aceasta şi trecând la limită, avem

$$\lim_{u\to a+}\int_u^c |f(x)|dx \le \lim_{u\to a+}\int_u^c \frac{\gamma}{(x-a)^{\lambda}}dx =$$
$$= \frac{\gamma}{1-\lambda}(x-a)^{1-\lambda}|_a^c = \frac{\gamma}{1-\lambda}(c-a)^{1-\lambda}.$$

Analog,din a doua jumătate a lui (C11) există $\delta > 0$ şi $d \in (c,b)$ cu

$$\lim_{v\to b-}\int_v^b |f(x)|dx \le \lim_{v\to b-}\int_v^b \frac{\delta}{(b-x)^{\mu}}dx =$$
$$= \frac{-\delta}{1-\mu}(b-x)^{1-\mu}|_d^b = \frac{\delta}{1-\mu}(b-d)^{1-\mu}.$$

Iar de aici, combinând cu aditivitatea faţă de interval a integralei, afirmaţia reiese. La fel se tratează şi situaţia descrisă de (C12).

Exemplu. Să se studieze existenţa integralei

$$\int_0^1 x^{\lambda}(1-x)^{\mu}dx, \text{ unde } \lambda, \mu \text{ sunt parametri.}$$

Funcţia de integrat este pozitivă. Avem evident

$$\begin{array}{lclcl} \lim\limits_{x\to 0+} x^{-\lambda}(x^{\lambda}(1-x)^{\mu}) & = & \lim\limits_{x\to 0+}(1-x)^{\mu} & = & 1 \in (0,\infty) \\ \lim\limits_{x\to 1-}(1-x)^{-\mu}(x^{\lambda}(1-x)^{\mu}) & = & \lim\limits_{x\to 1-} x^{\lambda} & = & 1 \in (0,\infty). \end{array}$$

Aşadar, pentru $-\lambda < 1$, $-\mu < 1$ (deci $\lambda > -1$, $\mu > -1$) integrala există (este finită) iar pentru $-\lambda \ge 1$ sau $-\mu \ge 1$ (deci $\lambda \le -1$ sau $\mu \le -1$) integrala nu există (este infinită).

Referitor la existenţa integralei generalizate pe un asemenea interval, următorul rezultat (datorat lui Abel) este de semnalat aici. Fie f, g două funcţii continue pe (a,b).

Teorema 4.5.9 *Să presupunem că*
(C13) *fg este mărginită pe fiecare interval (c, b), $c \in (a, b)$,*
(C14) *mulţimea integralelor $\{\int_u^v f(x)dx;\ a < u < v < b\}$ este mărginită.*

$$\text{(C15)} \begin{cases} g \text{ este monotonă pe unul din intervalele } (a, c) \\ \text{unde } a < c < b \text{ şi } \lim\limits_{x \to a+} g(x) = 0. \end{cases}$$

Atunci, obligatoriu, integrala generalizată $\int_a^b f(x)g(x)dx$ există.

Demonstraţia se bazează pe teorema a doua de medie şi criteriul de existenţă (C10) stabilit anterior. Un rezultat asemănător are loc prin schimbarea rolului punctelor a şi b în condiţiile de mai sus.

Exemplu. Să se studieze integrala generalizată

$$\int_0^\pi x^\lambda \cos\frac{1}{x}dx, \quad \text{unde } \lambda > -2 \ \text{ este un parametru.}$$

Funcţia de integrat satisface evident (C13), deoarece este continuă în $x = \pi$. Condiţia (C14) rezultă din evaluarea

$$\left|\int_u^v \frac{1}{x^2}\cos\frac{1}{x}dx\right| = \left|\sin\frac{1}{u} - \sin\frac{1}{v}\right| \le 2,\ 0 < u < v.$$

În fine, funcţia $g(x) = x^{\lambda+2}$ satisface evident (C15) deoarece $\lambda > -2$. Deci, conform celor spuse, integrala anterioară există. Să observăm cu această ocazie că integrala (standard) $\int_0^\pi x^\lambda \left|\cos\frac{1}{x}\right| dx$ nu există: căci altfel (din $\cos^2\alpha = \frac{1}{2}(1 + \cos 2\alpha)$), ar exista şi

$$\int_0^\pi x^\lambda \cos^2\frac{1}{x} = \frac{1}{2}\left(\int_0^\pi x^\lambda dx + \int_0^\pi x^\lambda \cos\frac{2}{x}dx\right).$$

Cum însă (prin transformarea $x = 2t$)

$$\int_0^\pi x^\lambda \cos\frac{2}{x}dx = 2^{\lambda+1}\int_0^{\pi/2} t^\lambda \cos\frac{1}{t}dt$$

deducem (conform concluziei anterioare) că $\int_0^\pi x^\lambda dx$ ar exista pentru toţi $\lambda > -2$, ceea ce este absurd.

Proprietăţile de bază ale acestor integrale generalizate sunt, în fond, acelea ale integralelor standard (la care se aplică trecerea la limită). Astfel, sunt valabile şi aici proprietatea de aditivitate faţă de interval, de liniaritate şi monotonie. Cât despre formulele de calcul, menţionăm existenţa unei formule de tip Leibniz–Newton, în varianta

(R8) $\int_a^b f(x)dx = F(b-) - F(a+) \ \left(= F(x)|_{a+}^{b-}\right).$

(Aici, F este o primitivă a funcţiei f pe (a, b).) În fine, sunt şi aici valabile formulele de integrare prin părţi şi de schimbare a variabilei.

(E) Să studiem în continuare alternativa secundă din (C7) combinată cu alternativa primă din (C8) (respectiv, a =finit, $b = \infty$). (Cazul simetric acestuia se va discuta în acelaşi mod.) Condiţia (C9) din Teorema 4.5.7 va fi verificată şi, deci, f este sumabilă pe (a, b), dacă

$$\text{(C16)} \begin{cases} 0 \leq \lim\limits_{x \to a+} (x-a)^{\lambda}|f(x)| < \infty \text{ şi } 0 < \lim\limits_{x\to\infty} x^{\mu}|f(x)| < \infty \\ \text{pentru un } \lambda < 1, \text{ respectiv un } \mu > 1. \end{cases}$$

Dimpotrivă, dacă una din condiţiile limită

$$\text{(C17)} \begin{cases} 0 \leq \lim\limits_{x \to a+} (x-a)^{\lambda}|f(x)| \leq \infty \text{ sau } 0 < \lim\limits_{x\to\infty} x^{\mu}|f(x)| \leq \infty \\ \text{pentru un } \lambda \geq 1, \text{ respectiv un } \mu \leq 1 \end{cases}$$

are loc, atunci condiţia în chestiune nu se verifică; şi, deci, f nu este sumabilă pe (a, b).

Să dovedim afirmaţia referitoare la (C16). În baza jumătăţii secunde din (C16), va exista un $\delta > 0$ şi un $d > a$, suficient de mare (deci, în particular, $d > 0$), aşa încât

$$x^{\mu}|f(x)| \leq \delta \ \ (\text{sau } |f(x)| \leq \frac{\delta}{x^{\mu}}), \ x \geq d.$$

Integrând aceasta şi trecând la limită, avem

$$\lim_{v\to\infty} \int_d^v |f(x)|dx \leq \lim_{v\to\infty} \int_d^v \frac{\delta}{x^{\mu}}dx = \frac{\delta}{1-\mu}x^{1-\mu}|_d^{\infty} = \frac{\delta}{\mu - 1}d^{1-\mu}.$$

Pe de altă parte, în baza primei jumătăţi din (C16), există un $\gamma > 0$ şi un $c \in (a, d)$, aşa încât (conform unor calcule deja făcute la discuţia alternativei precedente)

$$\lim_{u\to a+} \int_u^c |f(x)|dx \leq \frac{\gamma}{1-\lambda}(c-a)^{1-\lambda}.$$

Iar de aici, combinând cu aditivitatea faţă de interval a integralei, afirmaţia reiese. Analog se tratează şi situaţia descrisă prin (C17).

Exemplu. Să se studieze existenţa integralei

$$\int_0^{\infty} x^{\lambda}e^{-x}dx, \ \text{unde } \lambda \text{ este un parametru.}$$

Funcţia de integrat este pozitivă. Avem, evident,

$$\lim_{x\to\infty} x^2(x^\lambda e^{-x}) = \lim_{x\to\infty} x^{\lambda+2}e^{-x} = 0 \in [0,\infty).$$

Pe de altă parte,

$$\lim_{x\to 0+} x^{-\lambda}(x^\lambda e^{-x}) = \lim_{x\to 0+} e^{-x} = 1 \in [0,\infty).$$

Şi deci, pentru $-\lambda < 1$ (adică $\lambda > -1$) integrala există; iar pentru $-\lambda \geq 1$ (adică $\lambda \leq -1$), integrala nu există (este infinită).

Ca şi la discuţia anterioară ne interesează în ce condiţii există integralele generalizate pe intervalul (a,∞). În acest sens, următorul rezultat de existenţă (de tip Abel) este de semnalat. (Demonstraţia acestuia se reduce şi aici la criteriul de existenţă de tip (C10) stabilit anterior). Fie f, g două funcţii continue pe (a,∞).

Teorema 4.5.10 *Să presupunem că*

(C18) fg *este mărginită pe intervalele* (a,c), $\forall c \in (a,\infty)$

(C19) *mulţimea integralelor* $\{\int_u^v f(x)dx; a<u<v<\infty\}$ *este mărginită*

(C20) $\begin{cases} g \text{ este monotonă pe unul din intervalele } (c,\infty) \\ \text{unde } a < c < \infty, \text{ şi } \lim\limits_{x\to\infty} g(x) = 0. \end{cases}$

Atunci, obligatoriu, integrala generalizată $\int_a^\infty f(x)g(x)dx$ *există.*

Exemplu. Să se studieze integrala generalizată

$$\int_0^\infty \frac{1}{x^\lambda}\sin x\,dx, \text{ unde } \lambda \in (0,1) \text{ este un parametru.}$$

Funcţia de integrat satisface (C18), deoarece

$$\lim_{x\to 0+} \frac{1}{x^\lambda}\sin x = \lim_{x\to\infty} x^{1-\lambda}\frac{\sin x}{x} = \begin{cases} 0, & \text{dacă } 0 < \lambda < 1 \\ 1, & \text{dacă } \lambda = 1. \end{cases}$$

Condiţia (C19) rezultă imediat din

$$\left|\int_u^v \sin x\,dx\right| = |\cos u - \cos v| \leq 2, \ 0 < u < v < \infty.$$

Iar, în fine, $g(x) = \dfrac{1}{x^\lambda}$ satisface evident (C20). Deci, conform rezultatului precedent, integrala generalizată definită anterior există. Se poate şi aici arăta că integrala standard din această funcţie nu există; nu mai dăm alte detalii.

Proprietăţile de bază ale acestor integrale sunt aceleaşi ca pentru integralele standard; deoarece ele rezultă din proprietăţile cunoscute, prin trecere la limită. Tot astfel se prezintă lucrurile cu formulele de calcul.

(F) A mai rămas de discutat alternativa primă din (C7)+(C8) $(a = -\infty,\ b = \infty)$. Cum, însă , toate raţionamentele apar drept combinaţie a celor anterioare, se poate spune că acest caz este clarificat. O concluzie analogă este valabilă şi pentru integrala generalizată.

(G) Conceptul anterior introdus de integrală se poate uşor extinde la funcţii cu valori vectoriale. Mai exact, fie (a, b) un interval al axei reale $(-\infty \leq a < b \leq \infty)$ şi p un număr natural. Considerăm dată o funcţie $f : (a, b) \longrightarrow R^p$. După cum ştim,

$$f(x) = (f_1(x), ..., f_p(x)),\ x \in (a, b),$$

unde, pentru fiecare $i \in \{1, ..., p\}$, f_i este o funcţie cu valori reale definită pe intervalul (a, b). În acest caz, vom conveni să notăm

(D11) $\int_a^b f(x)dx = \left(\int_a^b f_1(x)dx, ..., \int_a^b f_p(x)dx\right)$

dacă, desigur, fiecare din integralele componente există; iar funcţia însăşi, f, se va zice *sumabilă* pe (a, b). Proprietăţile de bază ale acestei integrale sunt, formal, aceleaşi ca la cazul $p = 1$. Menţionăm doar că, în formularea proprietăţii de monotonie , ordinea uzuală este cea vectorială. De asemenea, proprietatea de continuitate are aici forma

(P16) $\left\|\int_a^b f(x)dx\right\| \leq \int_a^b \|f(x)\|dx.$

Din formulele de calcul rămân valabile formula Leibniz–Newton şi cea de schimbare a variabilei.

În particular, dacă p, q sunt două numere naturale şi $f : (a, b) \rightarrow \mathcal{M}(p, q)$ este o funcţie cu valori matriciale, reprezentată în forma

$$f(x) = (f_{ij}(x)),\ x \in (a, b),$$

atunci, prin definiţie, vom pune

(D12) $\int_a^b f(x)dx = \left(\int_a^b f_{ij}(x)dx\right)$ (element din $\mathcal{M}(p, q)$).

Să adăugăm că, în unele cazuri (când produsul matricial este posibil), avem valabilă şi formula de integrare prin părţi. (Mai precis, dacă $f : (a, b) \longrightarrow \mathcal{M}(p, q)$, $g : (a, b) \longrightarrow \mathcal{M}(q, r)$ sunt de clasă C^1 pe $[a, b]$, atunci acest lucru are loc.) În fine, tot prin intermediul funcţiilor componente se pot introduce şi integrale generalizate pentru asemenea funcţii (vectoriale sau matriciale).

4.5.3 Integrala dublă (pentru funcţii de două variabile)

Ne vom ocupa în continuare de extinderea conceptului de integrală pentru funcţiile de două variabile. Vom face aceasta în mai multe etape.

Începem prin a preciza convenţiile şi noţiunile de bază. Să numim *dreptunghi* al planului R^2 orice produs cartezian de forma $K = I \times J$ unde I, J sunt intervale pe axa reală. Pentru orice asemenea dreptunghi definim *aria* sa prin

(D1) $\text{aria}\,(K) = \text{lung}\,(I) \cdot \text{lung}\,(J), \quad \text{dacă} \;\; K = I \times J.$

Vom zice că mulţimea A din R^2 este de *arie zero* dacă pentru orice $\varepsilon > 0$ există un şir (K_n) de dreptunghiuri care acoperă A cu suma ariilor inferioară lui ε :

(D2) $A \subseteq \bigcup\limits_{n=1}^{\infty} K_n, \quad \sum\limits_{n=1}^{\infty} \text{aria}\,(K_n) < \varepsilon.$

Vom scrie această constatare prin $\text{aria}\,(A) = 0$. Este uşor de văzut că

(P1) $\text{aria}\,(A) = 0, \; B \subseteq A \Longrightarrow \text{aria}\,(B) = 0.$

(P2) $\text{aria}\,(A_n) = 0, \; n = 1, 2, ..., \Longrightarrow \text{aria}\left(\bigcup\limits_{n=1}^{\infty} A_n\right) = 0.$

În particular, de aici rezultă constatarea

(P3) orice mulţime numărabilă are aria zero.

Există însă şi mulţimi ne–numărabile de arie zero; ca, de pildă,

$$A = \{(\varphi(t), \psi(t)); \; t \in (t_1, t_2)\},$$

unde funcţiile $\varphi : (t_1, t_2) \longrightarrow R, \; \psi : (t_1, t_2) \longrightarrow R$ sunt de clasă C^1 pe (t_1, t_2); nu mai dăm alte detalii.

Prin definiţie, vom zice că o proprietate $P(x, y)$ de variabile $(x, y) \in R^2$, are loc *aproape peste tot*, dacă

(D3) $\{(x, y) \in R^2; \; P(x, y) \text{ nu are loc}\}$ este de arie zero.

Să numim funcţia $g : R^2 \longrightarrow R$ *etajată* dacă există o diviziune a planului (prin dreptunghiuri)

$$\begin{array}{ll} -\infty < x_1 < x_2 < \cdots < x_n < \infty & (n \geq 2) \\ -\infty < y_1 < y_2 < \cdots < y_m < \infty & (m \geq 2) \end{array}$$

şi un sistem de constante $\{\lambda_{ij};\ 1 \leq i \leq n-1,\ 1 \leq j \leq m-1\}$ aşa încât

$$g(x,y) = \begin{cases} \lambda_{ij}, & \text{dacă } (x,y) \in (x_i, x_{i+1}) \times (y_j, y_{j+1}), \\ 0, & \text{în celelalte puncte.} \end{cases}$$

Nu este greu de văzut că

$$\mathcal{E}(R^2, R) = \text{mulţimea tuturor funcţiilor etajate } g : R^2 \longrightarrow R$$

are toate proprietăţile stabilite la cazul unei singure variabile.

Să introducem *aplicaţia de integrare* Λ pe clasa acestor funcţii prin

$$\text{(D4)} \quad \begin{cases} \Lambda(g) = \sum\limits_{i,j} \lambda_{ij} \text{aria}\,((x_i, x_{i+1}) \times (y_j, y_{j+1})), \text{ dacă } g \text{ este dată} \\ \text{de diviziunea } ((x_i, y_j)) \text{ şi constantele } (\lambda_{ij}). \end{cases}$$

Proprietăţile de bază ale acestei aplicaţii de integrare sunt în fond identice cu acelea din cazul funcţiilor de o singură variabilă.

(A) Vom introduce acum noţiunea de integrală (dublă) a unei funcţii pe un dreptunghi $D = (a,b) \times (c,d)$, unde $-\infty \leq a < b \leq \infty$, $-\infty \leq c < d \leq \infty$. Notăm deocamdată

$$\mathcal{F}(D, R) = \text{mulţimea tuturor funcţiilor } f : D \longrightarrow R.$$

O clasă importantă de asemenea funcţii este

$$\mathcal{E}(D, R) = \{g \in \mathcal{E}(R^2, R);\ g(x,y) = 0,\ \forall (x,y) \in R^2 \setminus D\}$$

(mulţimea tuturor funcţiilor etajate care se anulează înafara dreptunghiului D).Definim *integrala (Lebesgue)* a unei asemenea funcţii prin

(D5) $\iint_D g(x,y)dx\,dy = \Lambda(g), \quad \forall g \in \mathcal{E}(D, R)$.

Fie acum $f : D \longrightarrow [0, \infty)$ o funcţie cu valori pozitive. Vom zice că aceasta este *măsurabilă (Lebesgue)* dacă există un şir (g_n) de funcţii etajate cu valori pozitive care, aproape peste tot, este crescător şi convergent către f; adică

$$\text{(D6)} \begin{cases} (g_n(x,y)) \text{ este monoton crescător şi } g_n(x,y) \longrightarrow f(x,y), \\ \text{pentru toţi } (x,y) \in D \text{ cu excepţia unei mulţimi de arie zero.} \end{cases}$$

În acest caz, vom pune prin definiţie

$$\text{(D7)} \ \iint_D f(x,y)dx\,dy = \lim_{n\to\infty} \iint_D g_n(x,y)dx\,dy$$

şi vom numi aceasta *integrala (Lebesgue)* a lui f pe D. În cazul când integrala în discuţie este finită vom zice că f este *sumabilă (în sens Lebesgue)* pe dreptunghiul D.

În fine, fie $f : D \longrightarrow R$ o funcţie cu valori oarecare. Construim funcţiile cu valori pozitive

$$\text{(D8)} \begin{cases} f_+(x,y) = \max(0, f(x,y)), (x,y) \in D \text{ (partea pozitivă din } f) \\ f_-(x,y) = \max(0, -f(x,y)), (x,y) \in D \text{ (partea negativă din } f). \end{cases}$$

Vom zice că f este *măsurabilă (Lebesgue)* pe D dacă f_+ şi f_- sunt simultan măsurabile pe D; dacă acestea sunt şi sumabile, f se va zice *sumabilă* pe D. Notăm într-un asemenea caz

$$\text{(D9)} \ \iint_D f(x,y)dx\,dy = \iint_D f_+(x,y)dx\,dy - \iint_D f_-(x,y)dx\,dy$$

şi vom numi aceasta *integrala (Lebesgue)* a lui f pe dreptunghiul D. Vom mai pune

(D10) $\mathcal{L}(D,R) = \{f : D \longrightarrow R;\ f \text{ este sumabilă pe } D\}$.

Structura acestei clase şi proprietăţile de bază ale integralei introduse sunt aceleaşi ca la cazul funcţiilor de o singură variabilă. Menţionăm doar forma specifică a proprietăţii ereditar aditive

$$\text{(P4)} \begin{cases} f \in \mathcal{L}(D,R) \Longrightarrow f \in \mathcal{L}(D',R) \text{ pentru orice sub–dreptunghi} \\ D' = (a',b') \times (c',d') \text{ al lui } D; \text{ şi, în plus,} \\ \iint_D f(x,y)dx\,dy = \iint_{D_1} f(x,y)dx\,dy + \iint_{D_2} f(x,y)dx\,dy \text{ dacă} \\ D = D_1 \cup D_2 \ (D_1, D_2 = \text{dreptunghiuri cu o latură comună}). \end{cases}$$

În fine, Teorema 4.5.4 rămâne valabilă şi în acest context; nu mai dăm alte detalii.

(B) Fie în continuare dreptunghiul $D = (a,b) \times (c,d)$ mărginit (adică, $-\infty < a < b < \infty, -\infty < c < d < \infty$). Ne interesează şi aici principalele clase de funcţii sumabile pe D. Fie $f : D \longrightarrow R$ o funcţie deocamdată arbitrară. Pentru orice divizare a lui D

$$(\Delta) \qquad \begin{cases} a = x_0 < x_1 < \cdots < x_n < x_{n+1} = b \\ c = y_0 < y_1 < \cdots < y_m < y_{m+1} = d, \end{cases}$$

şi orice sistem de puncte intermediare

$$(\xi) \qquad \xi_{ij} \in (x_i, x_{i+1}) \times (y_j, y_{j+1}),\ 0 \leq i \leq n,\ 0 \leq j \leq m,$$

asociem funcţia etajată $g = g_{((\Delta),(\xi))}$ prin convenţia

$$g(x, y) = \begin{cases} f(\xi_{ij}), & \text{dacă } (x, y) \in (x_i, x_{i+1}) \times (y_j, y_{j+1}), \\ 0, & \text{în celelalte puncte.} \end{cases}$$

Integrala acestei funcţii etajate

$$\iint_D g(x, y) dx\, dy = \sum_{i,j} f(\xi_{ij})\, \text{aria}\, ((x_i, x_{i+1}) \times (y_j, y_{j+1}))$$

se va numi *suma Riemann* asociată funcţiei f,diviziunii (Δ) şi punctelor intermediare (ξ). Să mai definim

$$\text{norma}\,(\Delta) = \max\{\text{aria}\,((x_i, x_{i+1}) \times (y_j, y_{j+1}));\ 0 \leq i \leq n, 0 \leq j \leq m)\}$$

(*norma* diviziunii considerate). Considerăm ipoteza

(C1) f este continuă aproape peste tot şi mărginită pe D.

Teorema 4.5.11 *În condiţiile date, f este sumabilă pe D.*

Demonstraţia este asemănătoare cu aceea a Teoremei 4.5.5. Anume, se ia un şir $((\Delta_n), (\xi^{(n)}))$ de diviziuni şi puncte intermediare cu norma $(\Delta_n) \longrightarrow 0$ şi se verifică valabilitatea variantei bi–dimensionale a Teoremei 4.5.4 pentru şirul (g_n) asociat acestora. Ajungem la concluzia cerută, cu informaţia suplimentară

(R1) $\iint_D f(x, y) dx\, dy = \lim\limits_{n\to\infty} \iint_D g_n(x, y) dx\, dy.$

Prin definiţie, orice funcţie mărginită $f : D \longrightarrow R$ pentru care limita din dreapta relaţiei (R1) există, se va zice funcţie *integrabilă Riemann* pe D iar limita însăşi se va numi *integrala Riemann* pe D a funcţiei f. Cum avem şi aici proprietatea structurală

(P5) $\begin{cases} f = \text{mărginită şi integrabilă Riemann pe } D \implies \\ f \text{ este continuă aproape peste tot pe } D, \end{cases}$

va rezulta, din cele precedente, că noua integrală coincide cu integrala (Lebesgue) deja introdusă. În particular, dacă

(C2) f este continuă pe D şi mărginită pe D,

atunci f este sumabilă pe D. Următorul aspect este de semnalat aici. În ipoteza de mai sus, funcţia

$$F(x,y) = \iint_{D(x,y)} f(u,v)du\,dv, \ \ (x,y) \in D$$

unde, prin definiţie,

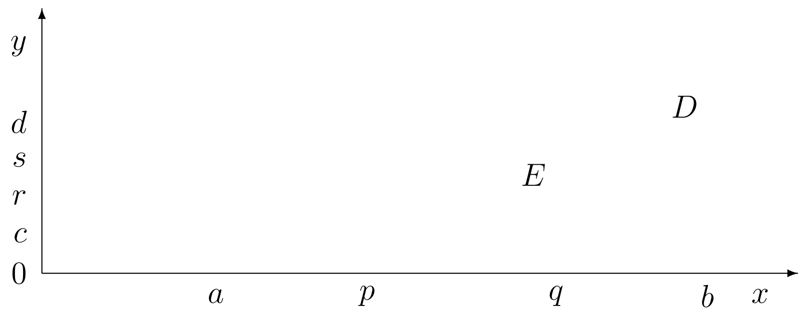
$$D(x,y) = [a,x] \times [c,y], \text{ oricare ar fi } (x,y) \in D,$$

poate fi construită. Numim aceasta *primitiva bi–dimensională* a funcţiei f. Justificarea acestei convenţii rezultă din

Teorema 4.5.12 *Oricare ar fi sub–dreptunghiul* $E = (p,q) \times (r,s)$ *din* D, *are loc formula Leibniz–Newton*

(R2) $\iint_E f(x,y)dx\,dy = F(p,r) - F(p,s) + F(q,s) - F(q,r)$.

Demonstraţie. Se ţine cont de aditivitatea integralei relativ la domeniu. (A se vedea şi figura alăturată.)



În particular, o condiţie suficientă pentru (C2) este

(C3) f este continuă pe $\operatorname{ad}(D) = [a,b] \times [c,d]$.

Menţionăm că, într-un asemenea context, avem formula de medie

(R3) $\iint_D f(x,y)dx\,dy = f(\xi,\eta)\,\mathrm{aria}\,(D)$, pentru un $(\xi,\eta) \in D$.

(C) Să ne întoarcem la cazul general relativ la D, în ipoteza

(C4) f este continuă pe D.

Dacă D este mărginit iar f este mărginită pe D atunci, conform celor spuse anterior, f este sumabilă pe D. Ne interesează ce se întâmplă când una din aceste condiţii nu mai are loc. Sunt deschise atunci alternativele:

(C5) dreptunghiul D nu este mărginit (în general)

(C6) f este nemărginită pe dreptunghiul D.

Un răspuns util în acest sens poate fi dat după cum urmează. Pentru fiecare submulţime E din D, să punem

(D11) $k_E(x,y) = \begin{cases} 1, & \text{dacă } (x,y) \in E \\ 0, & \text{dacă } (x,y) \in R^2 \setminus E. \end{cases}$

(Vom numi aceasta *funcţia caracteristică* a mulţimii E.) Să numim mulţimea E *măsurabilă (Lebesgue)* dacă funcţia sa caracteristică k_E este măsurabilă pe D. În acest caz, vom pune

(D12) $\mathrm{aria}\,(E) = \iint_D k_E(x,y)dx\,dy$ (*aria* lui E).

Dacă în plus k_E este chiar sumabilă, atunci integrala din dreapta este finită; vom spune atunci că mulţimea E este *sumabilă* (adică, de arie finită). Să facem notaţia

$$\mathcal{M}(D) = \{E \subseteq D;\ E \text{ este măsurabilă}\}.$$

Observăm că această clasă conţine toate sub–dreptunghiurile

$$E = (p,q) \times (r,s),\ a < p < q < b,\ c < r < s < d$$

sau, toate reuniunile finite de asemenea obiecte. Există însă şi alte mulţimi de acest fel; ca, de pildă,

$$E = \text{parte din } D \text{ cu aria}\,(\mathrm{fr}(E)) = 0.$$

Un caz particular de asemenea mulţime poate fi dat de aşa–numitele *domenii simple* faţă de $0y$:

$$E = \{(x,y);\ p < x < q;\ \varphi(x) \le y \le \psi(x)\}$$

unde (p,q) este un subinterval din (a,b), iar funcţiile $\varphi : (p,q) \to (c,d)$, $\psi(p,q) \to (c,d)$ sunt de clasă C^1 pe intervalul (p,q). (Analog, se pot introduce şi domeniile simple în raport cu $0x$.) Zicem că funcţia f este *sumabilă* pe $E \in \mathcal{M}(D)$ dacă funcţia $f \cdot k_E$ este sumabilă pe D. În acest caz vom defini

(D13) $\iint_E f(x,y)dx\,dy = \iint_D f(x,y)k_E(x,y)dx\,dy.$

Este util de precizat că, în acest context, proprietatea ereditar aditivă (P4) poate fi extinsă astfel

$$\text{(P6)}\quad \begin{cases} f \in \mathcal{L}(D,R) \Longrightarrow f \in \mathcal{L}(E,R), \text{ pentru orice parte măsurabilă} \\ E \text{ din } D\text{; mai mult, } \iint_E f(x,y)dx\,dy = \sum\limits_{n=1}^{\infty} \iint_{E_n} f(x,y)dx\,dy \\ \text{dacă } E = \bigcup\limits_{n=1}^{\infty} E_n \text{ este o descompunere a lui } E \text{ în părţi} \\ \text{măsurabile cu } E_i \cap E_j = \emptyset,\ i \neq j. \end{cases}$$

Putem acum formula următorul rezultat.

Teorema 4.5.13 *În ipoteza* (C4), *funcţia f este sumabilă pe D, dacă şi numai dacă*

(C7) $\lim\limits_{E\to D} \iint_E |f(x,y)|dx\,dy$ *este finită.*

(*Limita se ia în raport cu toate mulţimile măsurabile E din D, încadrabile în sub–dreptunghiuri*). *Iar atunci*

(R4) $\iint_D f(x,y))dx\,dy = \lim\limits_{E\to D} \iint_E f(x,y)dx\,dy.$

(*Adică, limita din dreapta există şi egalează integrala în chestiune.*)

Este util de menţionat că, prin modul de definiţie al acestei noţiuni, este exclusă posibilitatea ca limita din dreapta lui (R4) să existe fără ca funcţia să fie sumabilă pe D. (Aceasta, desigur, nu se mai întâmplă dacă , de pildă, în (C7) sau (R4) se lucrează numai cu dreptunghiuri; adică, într-un asemenea caz, am putea introduce un concept de integrală generalizată analog cu acela de la funcţiile de o variabilă. (Această posibilitate, însă , nu se ia în considerare aici.)

Înafară de calea indicată, mai există încă una de studiu a sumabilităţii unei funcţii pe dreptunghiul considerat, D. Începem cu elezentaiea unei formule de calcul pentru integrala dublă (datorate lui Fubini).

Teorema 4.5.14 *Dacă f este sumabilă pe D, atunci*

(R5) $\iint_D f(x,y)dx\,dy = \int_a^b \left(\int_c^d f(x,y)dy\right) dx.$

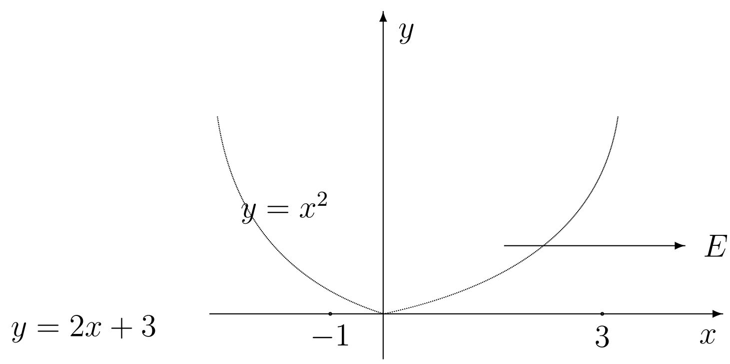
Mai precis:integrarea repetată din partea dreaptă a formulei (R5) *este posibilă, iar rezultatul său este tocmai integrala în discuţie.*
(Un rezultat analog are loc prin schimbarea rolului variabilelor x şi y.)

Să notăm aici că, din această formulă rezultă modalitatea de calcul a integralei pe orice parte măsurabilă E, din D. Astfel, de pildă, dacă E apare ca domeniu simplu faţă de $0y$, atunci (folosind scrierea sa indicată prin formula (D13)) avem reprezentarea

$$\iint_E f(x,y)dx\,dy = \int_p^q \left(\int_{\varphi(x)}^{\psi(x)} f(x,y)dy\right) dx.$$

(O concluzie analogă are loc pentru domeniile simple faţă de $0x$.)

Exemplu. Să se calculeze integrala $\iint_E xy\,dx\,dy$, unde E este domeniul plan limitat de parabola $y = x^2$ şi dreapta $y = 2x + 3$:



După cum rezultă şi din figură, domeniul E este simplu în raport cu $0y$, fiind reprezentat în forma

$$E = \{(x,y);\ -1 \le x \le 3, x^2 \le y \le 2x + 3\}.$$

Aplicând formula anterioară avem:

$$\iint_E xy\,dx\,dy = \int_{-1}^{3}\left(\int_{x^2}^{2x+3} xy\,dy\right)dx = \frac{1}{2}\int_{-1}^{3}\left(xy^2|_{y=x^2}^{y=2x+3}\right)dx =$$
$$= \frac{1}{2}\int_{-1}^{3} x[(2x+3)^2 - x^4]dx = \frac{1}{2}\int_{-1}^{3}(9x + 12x^2 + 4x^3 - x^4)dx = 53 + \frac{1}{3}\cdot$$

Suntem acum în măsură să dăm încă un criteriu de sumabilitate pe dreptunghiul considerat. (Ca şi mai înainte, prevalează ipoteza de continuitate (C4).) Avem astfel formula (de tip Tonelli):

Teorema 4.5.15 *Să admitem că*

(C8) *expresia* $\displaystyle\int_a^b \left(\int_c^d |f(x,y)|dy\right) dx$ *este finită.*

Atunci, f este sumabilă pe D, iar integrala sa poate fi calculată conform celor spuse anterior.
(Un rezultat analog are loc prin schimbarea rolului variabilelor x şi y.)

Exemplu. Să se studieze existenţa integralei duble

$$\iint_D e^{-(x^2+y^2)}dx\,dy, \text{ unde } D = (0,\infty)\times(0,\infty).$$

Avem evident evaluarea

$$\int_0^\infty \left(\int_0^\infty e^{-(x^2+y^2)}dy\right)dx = \int_0^\infty \left(e^{-x^2}\int_0^\infty e^{-y^2}dy\right)dx =$$
$$= \left(\int_0^\infty e^{-x^2}dx\right)^2 < \infty,$$

deoarece funcţia $h(x) = e^{-x^2}$ este sumabilă pe $(0,\infty)$. (Se aplică criteriile de existenţă cu limită). Şi atunci, cu teorema precedentă, afirmaţia rezultă de îndată.

Încheiem aceste consideraţii cu prezentarea unei formule de schimbare a variabilei, utilă în practică.

Teorema 4.5.16 *Fie dată transformarea*

$$x = \varphi(u,v), \quad y = \psi(u,v)$$

care aplică bijectiv domeniul M din planul (u,v) pe domeniul E din planul (x,y), cu funcţiile componente φ,ψ de clasă C^1 pe M şi

(C9) $\dfrac{D(\varphi,\psi)}{D(u,v)} = \begin{pmatrix} \dfrac{\partial\varphi}{\partial u} & \dfrac{\partial\varphi}{\partial v} \\ \dfrac{\partial\psi}{\partial u} & \dfrac{\partial\psi}{\partial v} \end{pmatrix}$ *este inversabilă,* $\forall (u,v) \in M.$

Are loc atunci egalitatea

(R6) $\iint_E f(x,y)dx\,dy = \iint_M f(\varphi(u,v),\psi(u,v)) \left| \det \dfrac{D(\varphi,\psi)}{D(u,v)} \right| du\,dv.$

Mai precis: dacă funcţia $(x,y) \vdash f(x,y)$ *este sumabilă pe* E*, atunci* $(u,v) \vdash f(\varphi(u,v),\psi(u,v)) \left| \det \dfrac{D(\varphi,\psi)}{D(u,v)} \right|$ *este sumabilă pe* M *şi are loc formula scrisă.*

Exemplu. Să se calculeze integrala dublă

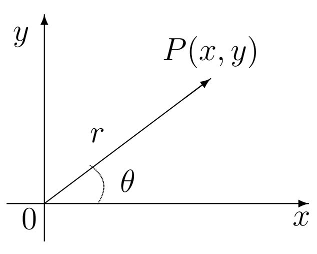
$$\iint_E e^{-(x^2+y^2)}dx\,dy, \text{ unde } E = (0,\infty)\times(0,\infty).$$

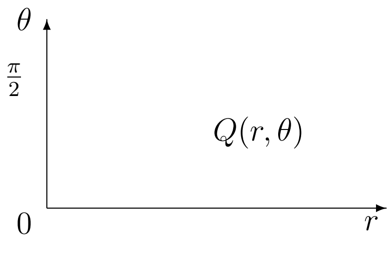
Vom trece la coordonatele polare (r,θ) definite de

$$x = r\cos\theta, \quad y = r\sin\theta.$$

Să observăm că, prin intermediul acestei transformări, domeniul E apare ca imagine a domeniului

$$M = (0,\infty)\times(0,\frac{\pi}{2}) \text{ (dreptunghi din planul } (r,\theta)).$$





Funcţiile componente ale transformării

$$\varphi(r,\theta) = r\cos\theta, \quad \psi(r,\theta) = r\sin\theta$$

sunt de clasă C^1 pe dreptunghiul M, cu

$$\begin{cases} \dfrac{D(\varphi,\psi)}{D(r,\theta)} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix},\ (r,\varphi)\in M \\ \left(\Longrightarrow \det \dfrac{D(\varphi,\psi)}{D(r,\theta)} = r(cos^2\theta + \sin^2\theta) = r \neq 0,\ (r,\varphi)\in M\right). \end{cases}$$

În fine, funcţia

$$f(x,y) = e^{-(x^2+y^2)},\quad x,y\in R$$

este sumabilă pe domeniul E, după cum s-a dovedit deja, într-un loc anterior. Aplicând acum formula de schimare a variabilei deducem

$$\begin{cases} \iint_E e^{-(x^2+y^2)}dx\,dy = \iint_M e^{-r^2} r\,dr\,d\theta = \int_0^\infty \left(\int_0^{\frac{\pi}{2}} e^{-r^2} r\,d\theta\right) dr = \\ = \dfrac{\pi}{2}\displaystyle\int_0^\infty e^{-r^2} r\,dr = -\dfrac{\pi}{4} e^{-r^2}|_0^\infty = \dfrac{\pi}{4}\cdot \end{cases}$$

Să observăm cu această ocazie că în baza unei evaluări anterioare a aceleiaşi integrale, obţinem

$$\left(\int_0^\infty e^{-x^2}dx\right)^2 = \frac{\pi}{4};\quad \text{deci}\quad \int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}\cdot$$

(D) Toate consideraţiile expuse pot fi extinse uşor la funcţii cu valori vectoriale. Anume, dacă p este un număr natural, fie $f : D \longrightarrow R^p$ o funcţie vectorială de două variabile. Folosind reprezentarea

$$f(x,y) = (f_1(x,y),...,f_p(x,y)),\quad (x,y)\in D$$

vom pune prin definiţie
(D14) $\int_D f(x,y)dx\,dy = (\iint_D f_1(x,y)dx\,dy, ..., \iint_D f_p(x,y)dx\,dy)$,
dacă fiecare integrală componentă există. Proprietăţile de bază ale integralei astfel definite rămân aceleaşi.

4.5.4 Integrala multiplă (pentru funcţii de mai multe variabile)

Conceptul de integrală dublă definit anterior poate fi extins, fără modificări esenţiale, la funcţiile de mai multe variabile. Anume, fie $m \geq 3$ un număr natural dat. Numim *paralelipiped* în spaţiul R^m, orice produs cartezian de forma $K = I_1 \times \cdots \times I_m$, unde $I_1, ..., I_m$ sunt intervale pe axa reală. Pentru orice asemenea paralelipiped, definim *volumul* său, prin

(D1) $\mathrm{vol}\,(K) = \mathrm{lung}\,(I_1) \cdots \mathrm{lung}\,(I_m)$, dacă $K = I_n \times \cdots \times I_m$.

Noţiunile de *mulţime de volum nul* şi *proprietate valabilă aproape peste tot* au acum un înţeles evident. Tot astfel apare şi noţiunea de *funcţie etajată* şi *aplicaţia de integrare* asociată.

Fie acum $D = (a_1, b_1) \times \cdots \times (a_m, b_m)$ un paralelipiped dat. Noţiunea de *integrală (Lebesgue)* a unei funcţii etajate care se anulează înafara lui D) se introduce prin

(D2) $\int \cdots \int_D g(x_1 ... x_m) dx_1, ..., dx_m = \Lambda(g)$.

Iar, pornind de aici, se introduc noţiunile de *măsurabilitate* şi *sumabilitate (Lebesgue)* ale unei funcţii $f : D \longrightarrow R$, prin definiţiile (formal aceleaşi) din Secţiunea 4.5.3. Proprietăţile de bază ale integralei introduse şi clasele de funcţii integrabile rămân şi aici valabile (în noul context). O menţiune specială trebuie făcută în legătură cu formula de tip Fubini care, aici, capătă forma

Teorema 4.5.17 *Dacă f este sumabilă pe D, atunci*

$$\text{(R)} \begin{cases} \displaystyle\int \cdots \int_D f(x_1, ..., x_m) dx_1 ... dx_m = \\ \displaystyle\int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\cdots \int_{a_m}^{b_m} f(x_1 ... x_{m-1}, x_m) dx_m \cdots \right) dx_2 \right) dx_1. \end{cases}$$

Adică: integrarea repetată din partea dreaptă e posibilă, iar rezultatul acesteia reprezintă tocmai valoarea integralei în discuţie.
(*Rezultate analoge pot fi scrise permutând variabilele* $x_1, ..., x_m$.)

Exemplu. Să se calculeze integrala triplă

$$\iiint_D xy^2 z^2 dx, dy\, dz, \text{ unde } D = (0, a) \times (0, b) \times (0, c).$$

Avem, conform formulei de tip Fubini,

$$\int\int\int_D xy^2z^2dx, dy\, dz = \int_0^a \left[\int_0^b \left(\int_0^c xy^2z^2dz\right) dy\right] dx =$$
$$= \int_0^a \left[\int_0^b \left(\frac{1}{3}xy^2z^3|_{z=0}^{z=c}\right) dy\right] dx = \int_0^a \left(\int_0^b \frac{c^3}{3}xy^2dy\right) dx =$$
$$= \int_0^a \left(\frac{c^3}{9}xy^3|_{y=0}^{y=6}\right) dx = \frac{c^3b^3}{9}\int_0^a x\, dx = \frac{c^3b^3}{18}x^2|_0^9 = \frac{c^3b^3a^2}{18}.$$

Să notăm cu această ocazie că formula anterioară permite reducerea integralei multiple la integrale simple pe orice parte măsurabilă E din D (introdusă în acelaşi mod); în particular,, de pildă, , la domeniile simple faţă de una din axe ($0x_m$ să zicem):

$$E = \{(x_1, ..., x_{m-1}, x_m); (x_1, ..., x_{m-1}) \in E_m,$$
$$\varphi(x_1, ..., x_{m-1}) \leq x_m \leq \psi(x_1, ..., x_{m-1})\}$$

unde E_m = proiecţia domeniului E pe spaţiul variabilelor $(x_1, ..., x_{m-1})$, iar $\varphi : E_m \longrightarrow R$, $\psi : E_m \longrightarrow R$ sunt funcţii de clasă C^1 pe E_m. De asemenea, criteriul de existenţă de tip Tonelli rămâne valabil. În plus, vom mai menţiona că formula de schimbare a variabilei admite o extensie directă la aceste integrale multiple.

Încheiem cu observaţia că după modelul deja cunoscut, se poate introduce integrala multiplă dintr-o funcţie cu valori vectoriale.

Probleme propuse la §4.5

1. Să se discute existenţa primitivelor pe R pentru funcţiile

a) $f(x) = \begin{cases} \cos\frac{1}{x}, \text{pentru } x \neq 0 \\ \lambda, \text{pentru } x = 0. \end{cases}$; b) $f(x) = \begin{cases} \frac{\sin x}{x}, \text{pentru } x \neq 0 \\ \mu, \text{pentru } x = 0. \end{cases}$.

2. Să se calculeze : a) $\int x^2e^x \sin x dx$; b) $\int \frac{1}{x^2} \ln^3 x dx$.

3. Să se determine o relaţie de recurenţă pentru calculul integralelor:
a) $I_n = \int x^ne^xdx$; b) $I_n = \int \mathrm{tg}^n x\, dx$; c) $I_n = \int e^{ax} \cos^n x\, dx$.

4. Să se calculeze : a) $\int \frac{2x+3}{x^3+2x^2-3x}dx$; b) $\int \frac{x^2dx}{(x+2)(x^2+1)^2}$.

5. Să se calculeze : a) $\int \frac{dx}{3+5\cos x}$; b) $\int \frac{\sin^2 x}{\cos^3 x}dx$.

6. Să se calculeze : a) $\int \frac{dx}{x\sqrt{-x^2+5x-6}}$, b) $\int \frac{x\,dx}{(x-1)\sqrt{1+x-x^2}}\cdot$
(Indicaţie. Se vor folosi substituţiile lui Euler.)

7. Să se calculeze integralele de tip binomial:
a) $\int \sqrt{\frac{1-\sqrt{x}}{x}}dx$; b) $\int \frac{\sqrt[3]{1+\sqrt[4]{x}}}{\sqrt{x}}dx$;
c) $\int x^3(1+x^2)^{-1/2}dx$; d) $\int \frac{x\,dx}{\sqrt{1+\sqrt[3]{x^2}}}\cdot$

8. Fără a calcula integrala, să se demonstreze inegalitatea:

$$0 \leq \int_{-\frac{1}{2}}^{0} x\ln(1-x^2)dx \leq \frac{1}{4}\ln\frac{4}{3}.$$

9. Să se calculeze integralele:
a) $\int_0^{\ln 2}\sqrt{e^x-1}dx$; b) $\int_1^3 \frac{dx}{x\sqrt{x^2+5x+1}}$; c) $\int_0^1 x^3e^{2x}dx$;
d) $\int_{-1}^0 \arccos\frac{1-x^2}{1+x^2}dx$; e) $\int_0^2 e^x\max(1,x^2)dx$; f) $\int_1^3 \frac{dx}{|x-2|+1}\cdot$

10. Să se studieze existenţa integralelor (şi să se calculeze):
a) $\displaystyle\int_0^\infty \frac{dx}{1+x^2}$; b) $\displaystyle\int_0^\infty e^{-ax}\sin x\,dx,\ a>0$; c) $\displaystyle\int_{-\infty}^\infty \frac{x\,dx}{1+c^4}\cdot$

11. Pornind de la definiţie, să se studieze convergenţa integralelor:
a) $\displaystyle\int_0^1 \frac{dx}{x\ln^2 x}$; b) $\displaystyle\int_0^\infty \frac{\operatorname{arctg} x}{x}dx$; c) $\displaystyle\int_0^\infty \frac{dx}{(x+1)\sqrt{|x^2-1|}}$;
d) $\displaystyle\int_0^\infty \sin x^2dx$; e) $\displaystyle\int_{-a}^a \frac{x^4}{\sqrt{a^2-x^2}}dx$; f) $\displaystyle\int_0^\infty e^{\sin x}\frac{\sin 2x}{x^\alpha},\ \alpha>0.$

12. Să se calculeze integralele duble:
a) $\iint_D \ln(x+y)dx\,dy,\quad D=(0,1)\times(1,2)$;
b) $\iint_D \sqrt{xy}dx\,dy$, D fiind delimitat de $x^3\leq y\leq x^2,\ x\geq 0$.

13. Să se calculeze , prin coordonate polare,integralele duble:
a) $\iint_D \sqrt{x^2+y^2}dx\,dy,\quad (D): x^2+y^2\leq 4$;
b) $\iint_D (a^2-x^2-y^2),\quad (D): x^2+y^2\leq a^2,\ y\leq 0.$

14. Să se studieze existenţa şi să se calculeze integrala
$\iint_D e^{-a\sqrt{x^2+y^2}}dx\,dy,\quad D=[0,\infty)\times[0,\infty),\ a>0.$

15. Să se arate că integrala dublă $\iint_D e^{-xy}\sin x\,dx\,dy$, unde $D=(0,\infty)\times(0,\infty)$ nu există, dar integralele iterate există şi sunt egale:

$$\int_0^\infty\left(\int_0^\infty e^{-xy}\sin x\,dx\right)dy = \int_0^\infty\left(\int_0^\infty e^{-xy}\sin x\,dy\right)dx = \frac{\pi}{2}\cdot$$

16. Să se calculeze integrala triplă $\int\int\int_D \frac{dx\,dy\,dz}{(1+x+y+z)^2}$, unde D este delimitat de planele $x=0$, $y=0$, $z=0$, $x+y+z=1$.

17. Să se studieze integrala triplă $\int\int\int_D e^{-(x^2+y^2+z^2)}dx\,dy\,dz$, unde $D = (0,\infty)\times(0,\infty)\times(0,\infty)$.

18. Să se calculeze integrala multiplă $\int\cdots\int_D x_1 x_2^2 \cdots x_m^m\, dx_1 \dots dx_m$, unde $D=[0,a_1]\times[0,a_2]\times\cdots\times[0,a_m]$.

4.6 Funcţii speciale

4.6.1 Integrale depinzând de parametru

Fie (a,b) un interval pe dreapta reală $(-\infty \le a < b \le \infty)$ şi Λ un interval deschis al axei reale. Considerăm funcţia de două variabile $f:(a,b)\times\Lambda \longrightarrow R$, care satisface

(C1) $x \vdash f(x;\xi)$ este sumabilă pe (a,b), pentru orice $\xi\in\Lambda$.

Putem atunci construi funcţia $h:\Lambda\longrightarrow R$ prin convenţia

(D1) $h(\xi)=\displaystyle\int_a^b f(x;\xi)dx,\ \forall\xi\in\Lambda.$

Numim aceasta, *integrală (simplă) depinzând de un parametru*. Problema de bază care apare în acest context este cum se transmit proprietăţile pe care funcţia f le are faţă de parametru (ξ), în proprietăţi de acelaşi tip ale funcţiei–integrală asociate, h. Desigur, proprietăţile vizate sunt continuitatea şi derivabilitatea.

(A) Să studiem problema continuităţii în raport cu parametrul a funcţiei $h:\Lambda\longrightarrow R$. Admitem ipoteza specifică

(C2) $\xi \vdash f(x,\xi)$ este continuă pe Λ, pentru orice $x\in(a,b)$.

Aceasta, însă , nu este suficient pentru concluzia pe care o avem în vedere. Mai este necesară o ipoteză specifică, legată de procesul trecerii la limită sub semnul integralei. Anume, aceasta poate fi de forma

$$(\mathrm{C3})\ \begin{cases} \text{pentru orice } \lambda \in \Lambda \text{ şi orice } \varepsilon > 0 \ \text{ cu } (\lambda - \varepsilon, \lambda + \varepsilon) \subseteq \Lambda, \\ \text{există o funcţie } g = g_{\lambda,\varepsilon}, \text{ sumabilă pe } (a, b) \\ \text{cu } |f(x, \xi)| \leq g(x), \ \forall x \in (a, b), \ \forall \xi \in (\lambda - \varepsilon, \lambda + \varepsilon). \end{cases}$$

Răspunsul la chestiunea formulată este următorul:

Teorema 4.6.1 *Dacă toate condiţiile* (C1), (C2), (C3) *sunt acceptate, atunci integrala parametrică* $\xi \vdash h(\xi)$ *este continuă pe intervalul* Λ.

Demonstraţie. Fie λ fixat în Λ şi (ξ_n) un şir din Λ cu $\xi_n \longrightarrow \lambda$. Se poate, desigur, presupune

$$\xi_n \in (\lambda - \varepsilon, \lambda + \varepsilon), \ \forall n \geq 1, \ \text{unde } \varepsilon > 0 \ \text{ este suficient de mic.}$$

Combinând cu (C3), avem evaluarea (cu notaţia folosită acolo)

$$|f(x, \xi_n)| \leq g(x), \ \forall x \in (a, b), \ n = 1, 2, ...$$

Iar, pe de altă parte, din condiţia (C2),

$$f(x, \xi_n) \longrightarrow f(x, \lambda) \text{ când } n \to \infty, \ \forall x \in (a, b).$$

Nu avem decât să ţinem acum cont de teorema Lebesgue (Secţiunea 4.5.2) pentru a încheia argumentul. ∎

După cum s-a spus deja, intervalul (a, b) este arbitrar (poate fi şi nemărginit). Când însă (a, b) este mărginit, există o cale de simplificare a condiţiilor puse. Anume, dacă se acceptă o condiţie de forma

(C4) $(x, \xi) \vdash f(x, \xi)$ este continuă pe $[a, b] \times \Lambda$,

atunci, toate condiţiile (C1), (C2), (C3) au loc. Faptul este evident pentru (C1) şi (C2); iar pentru (C3) se ţine cont de proprietatea de mărginire a unei funcţii continue (Secţiunea 4.2.4) pe compactele de forma $[a, b] \times [\lambda - \varepsilon, \lambda + \varepsilon]$.

Exemplu. Să se studieze continuitatea integralei parametrice

$$h(\xi) = \int_0^\infty \frac{x^{\xi-1}}{1+x} dx, \quad 0 < \xi < 1.$$

Funcţia de sub semnul integralei

$$f(x, \xi) = \frac{x^{\xi-1}}{1+x}, \ (x, \xi) \in (0, \infty) \times (0, 1)$$

satisface în primul rând condiţia de sumabilitate (C1). Aceasta rezultă imediat din existenţa limitelor

$$\lim_{x\to 0+} x^{1-\xi} f(x) = \lim_{x\to 0+} \frac{1}{1+x} = 1, \text{ când } \ 0 < 1-\xi < 1,$$
$$\lim_{x\to \infty} x^{2-\xi} f(x) = \lim_{x\to \infty} \frac{x}{1+x} = 1, \text{ când } \ 1 < 2-\xi < 2.$$

Pe de altă parte, funcţia considerată verifică evident condiţia (C2). Rămâne să mai verificăm (C3). Fie λ arbitrar în $(0,1)$ şi μ, ν un cuplu de numere cu $0 < \mu < \lambda < \nu < 1$. Funcţia $g : (0,\infty) \longrightarrow R$ dată de

$$g(x) = \begin{cases} \dfrac{x^{\mu-1}}{1+x}, & \text{dacă } \ 0 < x \le 1 \\ \dfrac{x^{\nu-1}}{1+x}, & \text{dacă } \ 1 < x < \infty \end{cases}$$

este, din cele spuse mai înainte, sumabilă pe $(0,\infty)$. Şi cum, evident,

$$f(x,\xi) \le g(x), \quad (x,\xi) \in (0,\infty) \times (\mu,\nu),$$

condiţia în chestiune este verificată. Deci, integrala parametrică $\xi \vdash h(\xi)$ este continuă pe $(0,\infty)$. Să mai notăm că , din formula (uşor de dovedit prin descompunere în fracţii simple)

$$\int_0^\infty \frac{z^{2m}}{1+z^{2n}} dz = \frac{\pi}{2n} \frac{1}{\sin \dfrac{2m+1}{2n}\pi}, \quad m,n = 1,2,...,m<n,$$

ajungem, cu substituţia $z = x^{1/2n}$, la concluzia

$$h\left(\frac{2m+1}{2n}\right) = \frac{\pi}{\sin\left(\dfrac{2m+1}{2n}\pi\right)}, \quad m,n = 1,2,...,m<n.$$

Combinând acest lucru cu continuitatea lui h ajungem (prin trecere la limită) la formula importantă

$$\int_0^\infty \frac{x^{\xi-1}}{1+x} dx = \frac{\pi}{\sin(\xi\pi)}, \ 0 < \xi < 1.$$

(B) Trecem acum la problema derivabilităţii în raport cu parametrul a funcţiei $h : \Lambda \longrightarrow R$. O condiţie specifică în acest sens ar fi

$$\text{(C5)}\quad \begin{cases} f \text{ este derivabilă parţial faţă de } \xi, \text{ iar aplicaţia} \\ x \vdash \dfrac{\partial f}{\partial \xi}(x,\xi) \text{ este sumabilă pe } (a,b),\ \forall \xi \in \Lambda. \end{cases}$$

Ca şi mai înainte, ea nu este însă suficientă pentru concluzia avută în vedere. Ar fi şi aici necesară o ipoteză specifică legată de procesul trecerii la limită sub semnul integralei. Anume, aceasta ar fi de forma

$$\text{(C6)}\quad \begin{cases} \text{pentru orice } \lambda \in \Lambda \text{ şi orice } \varepsilon > 0 \ \text{ cu } (\lambda-\varepsilon,\lambda+\varepsilon) \subseteq \Lambda \\ \text{există o funcţie } g = g_{\lambda,\varepsilon}, \text{ sumabilă pe } (a,b), \text{ cu} \\ \left|\dfrac{\partial f}{\partial \xi}(x,\xi)\right| \leq g(x),\ \forall x \in (a,b),\ \forall \xi \in (\lambda-\varepsilon,\lambda+\varepsilon). \end{cases}$$

Putem acum formula următorul răspuns la întrebarea pusă mai înainte.

Teorema 4.6.2 *Dacă toate condiţiile* (C1), (C5), (C6) *au loc, atunci integrala parametrică* $\xi \vdash h(\xi)$ *este derivabilă pe intervalul* Λ, *cu*

$$h'(\xi) = \int_a^b \frac{\partial f}{\partial \xi}(x,\xi)dx, \quad \forall \xi \in \Lambda.$$

Demonstraţie. Fie λ un element fixat în Λ şi (λ_n) un şir din Λ cu $\lambda_n \longrightarrow \lambda$. Putem şi aici presupune că

$$\lambda_n \in (\lambda-\varepsilon,\lambda+\varepsilon),\ \forall n \geq 1, \ \text{unde } \varepsilon > 0 \ \text{ este suficient de mic.}$$

Combinând cu Teorema de medie Lagrange şi cu (C6), avem

$$\left|\frac{1}{\lambda_n-\lambda}(f(x,\lambda_n)-f(x,\lambda))\right| \leq g(x),\ \forall x \in (a,b),\ n = 1,2,\dots\ .$$

Pe de altă parte, oricare ar fi $x \in (a,b)$

$$\frac{1}{\lambda_n-\lambda}(f(x,\lambda_n)-f(x,\lambda)) \longrightarrow \frac{\partial f}{\partial \xi}(x,\lambda) \text{ când } n \to \infty.$$

Din toate acestea găsim,ţinând cont din nou de Teorema Lebesgue (relativ la trecerea la limită sub semnul integralei)

$$\lim_{n\to\infty} \frac{1}{\lambda_n-\lambda}(h(\lambda_n)-h(\lambda)) = \lim_{n\to\infty} \int_a^b \frac{1}{\lambda_n-\lambda}(f(x,\lambda_n)-f(x,\lambda))dx$$
$$= \int_a^b \left[\lim_{n\to\infty} \frac{1}{\lambda_n-\lambda}(f(x,\lambda_n)-f(x,\lambda))\right] dx = \int_a^b \frac{\partial f}{\partial \xi}(x,\lambda)dx.$$

Iar,de aici, afirmaţia reiese. ■

Ca şi mai înainte,intervalul (a, b) este arbitrar (deci el poate fi şi nemărginit).Dar,dacă (a, b) este mărginit, există o cale de a simplifica ipotezele anterioare. Anume, dacă se acceptă o condiţie de forma

(C7) $\dfrac{\partial f}{\partial \xi}$ există şi este continuă pe $(a, b) \times \Lambda$,

atunci, obligatoriu, condiţiile (C5) + (C6) au loc; verificarea revine la o remarcă făcută anterior.

Exemplu. Să se studieze derivabilitatea integralei parametrice

$$h(\xi) = \int_0^1 \operatorname{arctg} \frac{x}{\xi} dx,\ \xi > 0.$$

Funcţia de două variabile

$$f(x, \xi) = \operatorname{arctg} \frac{x}{\xi},\ (x, \xi) \in [0, 1] \times (0, \infty)$$

este continuă pe domeniul scris.În plus, f este şi derivabilă parţial faţă de variabila ξ, cu

$$\frac{\partial f}{\partial \xi} = \frac{-x}{x^2 + \xi^2},\quad (x, \xi) \in [0, 1] \times (0, \infty)$$

continuă pe acelaşi domeniu. Deci, funcţia h este derivabilă, cu

$$h'(\xi) = -\int_0^1 \frac{x\, dx}{x^2 + \xi^2} = \frac{1}{2} \ln \frac{\xi^2}{1 + \xi^2},\quad \xi > 0.$$

(C) Rezultatele anterioare se pot uşor extinde la cazul integralelor multiple depinând de mai mulţi parametri. Anume, fie $D = (a, b) \times \cdots \times (a_m, b_m)$, un paralelipiped deschis în spaţiul R^m şi Λ, un domeniu din spaţiul R^p. (Aici, $m \geq 1$ şi $p \geq 1$ sunt două numere naturale fixate.) Fie, de asemenea, $f : D \times \Lambda \longrightarrow R$ o funcţie care satisface

(C8) $\begin{cases} (x_1, ..., x_m) \vdash f(x_1, ..., x_m; \xi_1, ..., \xi_p) \\ \text{este sumabilă pe } D, \text{ oricare ar fi } (\xi_1, ..., \xi_p) \in \Lambda. \end{cases}$

Putem atunci să considerăm funcţia $h : \Lambda \longrightarrow R$ dată de

(D2) $h(\xi_1, ..., \xi_p) = \int \cdots \int_D f(x_1, ..., x_m; \xi_1, ..., \xi_p) dx_1 \cdots dx_m, (\xi_1, ..., \xi_p) \in \Lambda.$

Ca şi mai înainte,vom numi aceasta,*integrală (multiplă) depinzând de parametri.*Problemele care se pun în legătură cu funcţiile în chestiune sunt aceleaşi ca la cazul $m = 1$, $p = 1$. Iar răspunsurile la ele pot fi formulate după o metodologie asemănătoare cu cea deja prezentată. Astfel, dintr-o ipoteză de forma

$$\text{(C9)} \begin{cases} (\xi_1, ..., \xi_p) \vdash f(x_1, ..., x_m; \xi_1, ..., \xi_p) \\ \text{este continuă pe } \Lambda, \text{ oricare ar fi } (x_1, ..., x_m) \in D \end{cases}$$

rezultă (în combinaţie cu o ipoteză tehnică de tip (C3)) continuitatea funcţiei $(\xi_1, ..., \xi_p) \vdash h(\xi_1, ..., \xi_p)$. În plus, dintr-o ipoteză de forma

$$\text{(C10)} \begin{cases} f \text{ este derivabilă parţial faţă de } \xi_1, ..., \xi_p, \text{ iar} \\ (x_1, ..., x_m) \vdash \dfrac{\partial f}{\partial \xi_i}(x_1, ..., x_m; \xi_1, ..., \xi_p) \text{ este sumabilă pe } D, \\ \text{oricare ar fi } i \in \{1, ..., p\}, \ (\xi_1, ..., \xi_p) \in \Lambda \end{cases}$$

rezultă (combinând şi aici cu o ipoteză tehnică de tip (C8)) derivabilitatea parţială a funcţiei h faţă de toate variabilele $\xi_1, ..., \xi_p$, cu

$$\frac{\partial h}{\partial \xi_i}(\xi_1, ..., \xi_p) = \int \cdots \int_D \frac{\partial f}{\partial \xi_i}(x_1, ..., x_m; \xi_1, ..., \xi_p) dx_1 \cdots dx_m, \ 1 \leq i \leq p.$$

Aceste formule se pot dovedi utile în diverse cazuri.

4.6.2 Funcţia Γ (Gamma) a lui Euler

Ne vom ocupa în continuare de următoarea integrală depinzând de parametru (introdusă de Euler)

(D1) $\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx, \ r > 0.$

O vom numi în continuare "funcţia Γ (Gamma)". Faptul că mulţimea de definiţie a acesteia este $(0, \infty)$ a fost deja stabilit într-un loc anterior (Secţiunea 4.5.2).

(A) Vom începe expunerea cu unele formule de bază relativ la această funcţie. În primul rând,avem formula de recurenţă

(P1) $\Gamma(r + 1) = r\Gamma(r), \ \forall r > 0.$

Într-adevăr,integrând prin părţi,deducem

$$\Gamma(r+1)=\int_0^\infty x^r e^{-x}dx=-x^r e^{-x}|_0^\infty+r\int_0^\infty x^{r-1}e^{-x}dx=r\Gamma(r)$$

(deoarece, evident, $\lim\limits_{x\to\infty} x^r e^{-x}=0$). Această formulă ne arată că este suficient să cunoaştem valorile funcţiei noastre pe $(0,1]$ pentru a le cunoaşte pe $(0,\infty)$. În particular, cum

$$\Gamma(1)=\int_0^\infty x^0 e^{-x}dx=\int_0^\infty e^{-x}dx=-e^{-x}|_0^\infty=1,$$

deducem succesiv, prin formula de recurenţă (cu convenţia $0!=1$)

$$\Gamma(n)=(n-1)!\ ,\quad \text{pentru } n=1,2,...\ .$$

Aceasta ne spune că funcţia Γ reprezintă extensia la semiaxa strict pozitivă a funcţiei factorial (definită, după cum ştim, numai pe mulţimea numerelor naturale). Pe de altă parte,

$$\Gamma(\frac{1}{2})=\int_0^\infty x^{-\frac{1}{2}}e^{-x}dx=2\int_0^\infty e^{-t^2}dt=2\frac{\sqrt{\pi}}{2}=\sqrt{\pi}.$$

(S-a făcut substituţia $x=t^2$ şi am ţinut cont de un rezultat stabilit în Secţiunea 4.5.3). Avem, deci, din aproape în aproape prin intermediul formulei de recurenţă (pentru $n=0,1,...$)

$$\Gamma\left(\frac{2n+1}{2}\right)=\frac{(2n-1)(2n-3)\cdots 3\cdot 1}{2^n}\sqrt{\pi}=\frac{(2n)!}{2^{2n}\cdot n!}\sqrt{\pi}.$$

O altă proprietate importantă a funcţiei Γ este formula complementelor (pe care o vom verifica ulterior):

(P2) $\Gamma(r)\cdot\Gamma(1-r)=\dfrac{\pi}{\sin(r\pi)},\ 0<r<1.$

Aceasta arată că este destul să cunoaştem valorile funcţiei Γ pe intervalul $(0,\frac{1}{2})$ pentru a le determina pe acelea din $(\frac{1}{2},1)$. Să notăm cu această ocazie că, pentru $r=\frac{1}{2}$, formula complementelor ne dă

$$\Gamma(\frac{1}{2})\cdot\Gamma(\frac{1}{2})=\frac{\pi}{\sin\dfrac{\pi}{2}}=\pi;\ \text{deci}\ \ \Gamma(\frac{1}{2})=\sqrt{\pi}.$$

Am regăsit astfel o valoare anterioară a acestei funcţii.

Menţionăm ,în acelaşi cadru, şi formula lui Legendre

(P3) $\Gamma(r) \cdot \Gamma(r + \frac{1}{2}) = \dfrac{\sqrt{\pi}}{2^{2r-1}}\Gamma(2r),\ \forall r > 0.$

(Aceasta va fi dedusă ulterior.)

De aici mai pot fi obţinute multe alte proprietăţi.Dar, acestea care au fost date, sunt caracteristice funcţiei Γ; adică, orice funcţie de clasă C^1 ce satisface (P1)–(P3) trebuie să coincidă cu funcţia Γ.

Să mai dăm, în fine,şi o altă reprezentare secvenţială pentru funcţia noastră (formula Euler–Gauss)

(D2) $\Gamma(r) = \lim\limits_{n\to\infty} \dfrac{n^r}{r} \dfrac{1 \cdot 2 \cdots (n-1)}{(r+1)(r+2)\cdots(r+n-1)},\ r > 0.$

Aceasta se dovedeşte utilă în multe cazuri concrete.

(B) Trecem acum la studiul proprietăţilor analitice ale funcţiei Γ. În primul rând, funcţia Γ este continuă şi are derivate de orice ordin pe $(0, \infty)$ dată de formula

$$\Gamma^{(n)}(r) = \int_0^\infty x^{r-1}e^{-x}(\ln x)^n dx,\ x \in (0, \infty), n = 1, 2, \dots .$$

Justificarea acestui fapt se obţine prin intermediul celor două rezultate anterioare. Comportarea funcţiei Γ înspre origine este dată de formula (deductibilă imediat din relaţia de recurenţă)

(R1) $\lim\limits_{r\to 0+} r\Gamma(r) = 1.$

De asemenea, comportarea înspre infinit a funcţiei Γ poate fi apreciată prin intermediul formulei lui Stirling

(R2) $\lim\limits_{r\to\infty} \sqrt{\dfrac{r}{2\pi}} \left(\dfrac{e}{r}\right)^r \Gamma(r) = 1.$

O consecinţă imediată a acestei formule este

(R3) $\lim\limits_{r\to\infty} \dfrac{\Gamma(r+a)}{r^a\Gamma(r)} = 1,\ \forall a > 0.$

(Se scrie relaţia (R2) pentru r şi $r + a$, iar apoi se împart egalităţile.) În particular, pentru $r = n + 1$ şi $a = m$ (numere naturale) avem

$$\lim_{n\to\infty} \frac{(n+m)!}{n^m \cdot n!} = 1,\ m = 1, 2, \dots$$

Iar pentru $r = n + 1,\ a = \frac{1}{2}$, se obţine

(R4) $\lim\limits_{n\to\infty} \dfrac{\Gamma(n+\frac{3}{2})}{\sqrt{n}\cdot n!} = 1$ (formula lui Wallis).

(C) Multe integrale uzuale din analiză se pot acum calcula utilizând proprietăţile funcţiei Γ.

Exemplu. Să se calculeze integralele

$$J_n = \int_0^\infty x^n e^{-x^2/2} dx,\ n = 0, 1, ...$$

Vom face substituţia $\frac{x^2}{2} = t$; adică

$$x = \sqrt{2}\ t^{1/2} \text{ (şi deci } dx = \tfrac{\sqrt{2}}{2}\ t^{-1/2}\ dt).$$

Înlocuind în integrală, ajungem la evaluarea

$$J_n = (\sqrt{2})^{n-1} \int_0^\infty t^{\frac{n-1}{2}} e^{-t} dt == (\sqrt{2})^{n-1}\Gamma(\frac{n+1}{2}).$$

Să reţinem faptul că funcţia Γ este considerată elementară; adică, este pusă pe acelaşi plan cu funcţiile elementare tradiţionale (puteri, radicali, exponenţiale logaritmice, trigonometrice şi inversele lor). Deci, şi derivatele acesteia vor avea aceeaşi calitate. Adică, dacă într-un calcul apar asemenea derivate se poate considera că acesta este încheiat.

Exemplu. Să se calculeze integrala $\int_0^\infty e^{-x} \ln x\, dx$. Din formula de derivare a funcţiei Γ avem

$$\Gamma'(r) = \int_0^\infty x^{r-1} e^{-x} \ln x\, dx,\ \forall r > 0.$$

Şi atunci, făcând $r = 1$, avem $\displaystyle\int_0^\infty e^{-x} \ln x\, dx = \Gamma'(1)$. Prin definiţie, mărimea $C = -\Gamma'(1)$ se numeşte *constanta lui Euler*. O posibilitate de calcul aproximativ al acesteia rezultă din formula

$$C = \lim_{n\to\infty} (1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln(n+1)).$$

4.6.3 Funcţia B (Beta) a lui Euler

Să considerăm în continuare următoarea integrală depinzând de doi parametri (introdusă tot de Euler)

$$\text{(D)}\;\; B(p,q)=\int_0^1 x^{p-1}(1-x)^{q-1}dx,\;\; (p,q)\in(0,\infty)\times(0,\infty).$$

O vom numi în continuare "funcţia B (Beta)". Faptul că mulţimea de definiţie a acesteia este $(0,\infty)\times(0,\infty)$ a fost deja stabilit într-un loc anterior (a se vedea Secţiunea 4.5.2).

(A) Se poate spune că studiul acestei funcţii se reduce la acela al funcţiei Γ. Justificarea acestui fapt este legată de rezultatul următor:

Teorema 4.6.3 *Are loc relaţia*

$$\text{(R1)}\;\; B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)},\;\; \textit{oricare ar fi}\;\; p,q>0.$$

Demonstraţie. În baza Formulei de tip Tonelli (Secţiunea 4.5.3) avem

$$\Gamma(p)\Gamma(q)=\int_0^\infty x^{p-1}e^{-x}dx\int_0^\infty y^{q-1}e^{-y}dy=\iint_E x^{p-1}y^{q-1}e^{-(x+y)}dx\,dy$$

unde, prin definiţie $E=(0,\infty)\times(0,\infty)$. Aplicăm integralei duble obţinute schimbarea de variabile

$$x=u(1-v),\quad y=uv.$$

Aceasta aplică bijectiv domeniul $M=(0,\infty)\times(0,1)$ din planul (u,v) pe domeniul E; în plus, funcţiile sale componente (notate cu φ,ψ) sunt de clasă C^1 pe domeniul M, cu

$$\frac{D(\varphi,\psi)}{D(u,v)}=\begin{pmatrix}\dfrac{\partial\varphi}{\partial u} & \dfrac{\partial\varphi}{\partial v}\\[2mm] \dfrac{\partial\psi}{\partial u} & \dfrac{\partial\psi}{\partial v}\end{pmatrix}=\begin{pmatrix}1-v & -u\\ v & u\end{pmatrix},\quad \forall(u,v)\in M$$

$$\left(\Longrightarrow \det\frac{D(\varphi,\psi)}{D(u,v)}=u(1-v)+uv=u\neq 0,\quad \forall(u,v)\in M\right).$$

Utilizând formula de schimbare a variabilei (Secţiunea 4.5.3) avem

$$\begin{aligned}\Gamma(p)\Gamma(q) &= \iint_M u^{p+q-1}(1-v)^{p-1}v^{q-1}e^{-u}du\,dv = \\ &= \int_0^\infty \left(\int_0^1 u^{p+q-1}(1-v)^{p-1}v^{q-1}e^{-u}dv\right) du = \\ &= \left(\int_0^\infty u^{p+q-1}e^{-u}du\right)\left(\int_0^1 (1-v)^{p-1}v^{q-1}dv\right) = \\ &= \Gamma(p+q)B(q,p).\end{aligned}$$

Aceasta încheie argumentul. ■

Ca o aplicaţie imediată a acestui fapt, să dovedim valabilitatea unor formule relative la funcţia Γ. Astfel, de pildă, formula complementelor se reduce aici la următoarea (din $\Gamma(1) = 1$)

$$\int_0^1 x^r(1-x)^{-r}dx = \frac{\pi}{\sin(r\pi)}, \quad 0 < r < 1,$$

sau echivalent (cu substituţia $x = \dfrac{y}{1+y}$)

$$\int_0^\infty \frac{y^{r-1}}{1+y}dy = \frac{\pi}{\sin(r\pi)}, \quad 0 < r < 1.$$

Şi cum această egalitate are loc (conform unui calcul anterior), concluzia este clară. Trecând la formula lui Legendre, să pornim de la evaluarea

$$B(r,r) = \int_0^1 x^{r-1}(1-x)^{r-1}dx = \int_0^1 \left[\frac{1}{4} - \left(\frac{1}{2} - x\right)^2\right]^{r-1} dx.$$

Făcând aici substituţia $\frac{1}{2} - x = \frac{1}{2}t^{1/2}$, ajungem la

$$B(r,r) = \frac{1}{2^{2r-1}}\int_0^1 t^{-\frac{1}{2}}(1-t)^{r-1}dt = \frac{1}{2^{2r-1}}B\left(\frac{1}{2}, a\right).$$

Şi atunci, folosind rezultatul anterior, afirmaţia este dovedită.

În fine, din formula (R3) (Secţiunea 4.6.2) rezultă aici

(R2) $\lim\limits_{r\to\infty} r^a B(r,a) = \Gamma(a),\ \forall a > 0.$

În particular, pentru $r = n$, $a = \frac{1}{2}$, deducem

(R3) $\lim\limits_{n\to\infty} \sqrt{n} B\left(n, \frac{1}{2}\right) = \sqrt{\pi}$,

care înseamnă de fapt tocmai formula lui Wallis.

(B) Vom pune în evidenţă acum utilitatea funcţiei B a lui Euler în calculul unor integrale.

Exemplu. Să se calculeze integralele parametrice

$$I(\lambda, \mu) = \int_0^{\frac{\pi}{2}} \sin^\lambda x \cos^\mu x \, dx, \quad \lambda > -1, \ \mu > -1.$$

Cu substituţia $\sin x = t^{1/2}$, sau, echivalent,

$$x = \arcsin\left(t^{1/2}\right) \ (\text{deci } dx = \tfrac{1}{2} t^{-1/2}(1-t)^{-1/2} dt)$$

integrala noastră devine

$$I(\lambda, \mu) = \frac{1}{2} \int_0^1 t^{\frac{\lambda-1}{2}} (1-t)^{\frac{\mu-1}{2}} dt = \frac{1}{2} B\left(\frac{\lambda+1}{2}, \frac{\mu+1}{2}\right).$$

În particular, de pildă, pentru $m = 0, 1, ...,$

$$I(m, 0) = \int_0^{\pi} \sin^m x \, dx = \frac{1}{2} B\left(\frac{m+1}{2}, \frac{1}{2}\right) = \frac{\sqrt{\pi}}{2} \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m+2}{2})}.$$

O clasă de integrale reductibilă la acestea este dată de următorul

Exemplu. Să se calculeze integralele parametrice

$$J(p, q) = \int_0^\infty x^p \left(1 + \frac{x^2}{a^2}\right)^{-q} dx, \quad p > -1, \ 1q > p + 1.$$

(Aici, $a > 0$ este arbitrar fixat.) Cu substituţia

$$x = a \ tg(y) \ (\ \text{deci } dx = \frac{a}{\cos^2 y} dy)$$

integrala în chestiune devine

$$J(p, q) = a^{p+1} \int_0^{\frac{\pi}{2}} \left(\frac{\sin y}{\cos y}\right)^p \left(\frac{1}{\cos^2 y}\right)^{-q} \frac{1}{\cos^2 y} dy =$$
$$= a^{p+1} \int^{\frac{\pi}{2}} (\sin y)^p (\cos y)^{2q-p-2} dy = \frac{a^{p+1}}{2} B\left(\frac{p+1}{2}, \frac{2q-p-1}{2}\right).$$

Aceste integrale se dovedesc utile în teoria probabilităţilor.

Probleme propuse la §4.6

1. Fie dată funcţia $f(x,\xi) = e^{-\frac{x^2}{\xi^2}}\dfrac{x}{\xi^2}$, $(x,\xi) \in [0,1] \times (0,\infty)$. Să se arate că

$$\lim_{\xi\to 0}\int_0^1 f(x,\xi)dx \neq \int_0^1 (\lim_{\xi\to 0} f(x,\xi))dx.$$

2. Să se calculeze derivatele de ordinul 1 şi 2 ale integralei parametrice $h(\xi) = \int_0^a f(x+\xi)dx$.

3. Dacă f este continuă pe $[0,a]$, atunci integrala parametrică

$$\varphi(u,v,w) = \int_0^a \frac{f(t)}{\sqrt{(u-t)^2 + v^2 + w^2}}$$

verifică ecuaţia lui Laplace: $\dfrac{\partial^2\varphi}{\partial u^2} + \dfrac{\partial^2\varphi}{\partial v^2} + \dfrac{\partial^2\varphi}{\partial w^2} = 0.$

4. Folosind derivarea în raport cu parametrul, să se calculeze integralele parametrice

a) $h(\xi) = \displaystyle\int^{\frac{\pi}{2}} \frac{1}{\cos x}\ln\frac{1+\xi\cos x}{1-\xi\cos x}, \quad -1 < \xi < 1$

b) $h(\xi) = \displaystyle\int^{\frac{\pi}{2}} \frac{\operatorname{arctg}(\xi \operatorname{tg} x)}{\operatorname{tg} x}dx, \quad \xi \in R.$

c) $h(\xi) = \displaystyle\int_0^{\infty} e^{-\xi x}\frac{\sin bx - \sin cx}{x}dx, \quad \xi \in R.$

5. Ştiind că $\displaystyle\int_0^b \frac{dx}{1+ax} = \frac{1}{a}\ln(1+ab)$, $a,b>0$, să se calculeze $\displaystyle\int_0^b \frac{dx}{(1+ax)^2}$, folosind posibilitatea unei derivări faţă de parametru.

6. Să se calculeze integrala parametrică de forma $h(\xi) = \displaystyle\int_0^{\infty} \frac{\sin x}{x}e^{-\xi x}dx, \xi > 0$, utilizând principiul derivării. Să se deducă apoi egalitatea $\displaystyle\int_0^{\infty} \frac{\sin x}{x}dx = \frac{\pi}{2}$, prin trecere la limită.

7. Să se calculeze următoarele integrale reductibile la funcţiile Γ şi B ale lui Euler. (În paranteză se indică substituţiile corespunzătoare.)

a) $\displaystyle\int_0^1 x^{p-1}(1-x^m)^{q-1}dx$ $\qquad (x^m = t)$

b) $\displaystyle\int_0^\pi \frac{dx}{\sqrt{3-\cos x}}$ $\qquad (\cos x = 1-2\sqrt{t})$

c) $\displaystyle\int_0^1 \frac{x^{a-1}(1-x)^{b-1}}{(x+p)^{a+b}}$ $\qquad (t = (1+p)\dfrac{x}{x+p})$

d) $\displaystyle\int_0^\infty x^m e^{-ax^n}dx$ $\qquad (t = ax^n)$.

Capitolul 5

Programare neliniară

5.1 Funcţii convexe şi concave

5.1.1 Funcţii convexe de o variabilă

Fie I un interval deschis din R. Funcţia $f : I \longrightarrow R$ se va zice *convexă pe intervalul* I, dacă oricare ar fi $x_1, x_2 \in I$, avem

(D1) $f(\lambda x_1 + \mu x_2) \leq \lambda f(x_1) + \mu f(x_2),\ \ \lambda, \mu \in [0,1],\ \ \lambda + \mu = 1.$

De asemenea, vom numi f *concavă* pe I, dacă $-f$ este convexă: adică,

(D2) $f(\lambda x_1 + \mu x_2) \geq \lambda f(x_1) + \mu f(x_2),\ \ \lambda, \mu \in [0,1],\ \ \lambda + \mu = 1.$

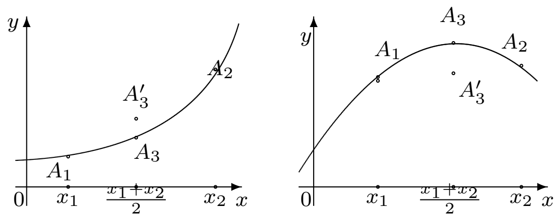
Dacă în (D1) inegalitatea e strictă:adică,pentru orice $x_1, x_2 \in I, x_1 \neq x_2$

(D3) $f(\lambda x_1 + \mu x_2) < \lambda f(x_1) + \mu f(x_2),\ \ \lambda, \mu \in (0,1),\ \ \lambda + \mu = 1,$

funcţia f se va zice *strict convexă* pe I, iar dacă $-f$ ar avea această proprietate, f se va numi *strict concavă* (pe I).

Interpretarea geometrică a noţiunilor introduse e simplă. De pildă, f este convexă pe I dacă şi numai dacă :

oricare ar fi punctele $A_1 = (x_1, f(x_1))$, $A_2 = (x_2, f(x_2))$, de pe graficul funcţiei f, segmentul de dreaptă A_1A_2 este situat deasupra graficului funcţiei cuprins între punctele A_1 şi A_2.
(În mod corespunzător se prezintă lucrurile la funcţia concavă).



Proprietatea de convexitate (sau concavitate) introdusă anterior se poate da şi sub forma următoare (aparent mai generală):

Teorema 5.1.1 *Funcţia f este convexă pe I dacă şi numai dacă oricare ar fi punctele x_1,...,$x_n \in I$,avem evaluarea*

$$(C1) \quad \begin{cases} f(\lambda_1 x_1 + \cdots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n), \\ \textit{pentru orice } \lambda_1, ..., \lambda_n \in [0,1] \textit{ cu } \lambda_1 + \cdots + \lambda_n = 1. \end{cases}$$

Demonstraţie. Se face prin inducţie relativ la numărul n de puncte utilizate. Anume, să zicem că se acceptă relaţia (C1) pentru un anume n şi fie $x_1, ..., x_n, x_{n+1}$ puncte date din I. Dacă $\lambda_1, ..., \lambda_n, \lambda_{n+1}$ sunt numere din $[0,1]$ cu $\lambda_1 + \cdots + \lambda_n + \lambda_{n+1} = 1$, să facem notaţia

$$\mu_1 = \frac{\lambda_1}{\lambda_1 + \cdots + \lambda_n}, ..., \mu_n = \frac{\lambda_n}{\lambda_1 + \cdots + \lambda_n} \text{ (deci } \mu_1 + \cdots + \mu_n = 1).$$

Desigur, aceasta cere $\lambda_1 + \cdots + \lambda_n \neq 0$ (ceea ce nu este restrictiv).Dacă mai punem $y = \mu_1 x_1 + \cdots + \mu_n x_n$, atunci, cu (C1),

$$f(y) = f(\mu_1 x_1 + \cdots + \mu_n x_n) \leq \mu_1 f(x_1) + \cdots + \mu_n f(x_n).$$

Ţinând acum cont de relaţia

$$\lambda_1 x_1 + \cdots + \lambda_n x_n + \lambda_{n+1} x_{n+1} = (1 - \lambda_{n+1}) y + \lambda_{n+1} x_{n+1}$$

obţinem, din convexitatea funcţiei f, evaluarea

$$\begin{aligned} & f(\lambda_1 x_1 + \cdots + \lambda_n x_n + \lambda_{n+1} x_{n+1}) \leq (1 - \lambda_{n+1}) f(y) + \lambda_{n+1} f(x_{n+1}) \leq \\ & \leq (1 - \lambda_{n+1})(\mu_1 f(x_1) + \cdots + \mu_n f(x_n)) + \lambda_{n+1} f(x_{n+1}) = \\ & = \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n) + \lambda_{n+1} f(x_{n+1}). \end{aligned}$$

Aceasta încheie argumentul. ■

Desigur, pentru funcţiile concave, caracterizarea anterioară este cu inegalităţi de sens opus:

$$\text{(C2)} \begin{cases} f(\lambda_1 x_1 + \cdots + \lambda_n x_n) \geq \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n), \\ \text{pentru orice } \lambda_1, ..., \lambda_n \in [0,1] \text{ cu } \lambda_1 + \cdots + \lambda_n = 1. \end{cases}$$

În plus, pentru convexitatea sau concavitatea strictă, relaţiile dinainte sunt stricte dacă elementele $\{x_1, ..., x_n\}$ nu se reduc la unul singur, iar numerele $\{\lambda_1, ..., \lambda_n\}$ sunt din $(0,1)$.

Să trecem acum la o caracterizare a funcţiilor convexe care va fi utilă mai departe.

Teorema 5.1.2 *Funcţia f este convexă pe I dacă şi numai dacă , oricare ar fi $a \in I$, funcţia raport*

$$R(x) = \frac{f(x) - f(a)}{x - a}, \ x \in I, x \neq a,$$

este monoton crescătoare pe $I \setminus \{a\}$.

Demonstraţie. Verificăm doar prima parte. Fie $a \in I$ arbitrar fixat şi x, y două puncte din $I \setminus \{a\}$ cu $x < y$. Avem posibilităţile:

$$a < x < y, \ \ x < a < y, \ \ x < y < a.$$

Vom discuta prima din acestea; celelelalte se tratează analog. Din

$$x = \frac{y - x}{y - a} a + \frac{x - a}{y - a} y, \ \ \text{cu} \ \ \frac{y - x}{y - a} + \frac{x - a}{y - a} = 1,$$

deducem, cu convexitatea funcţiei f, evaluarea

$$f(x) \leq \frac{y - x}{y - a} f(a) + \frac{x - a}{y - a} f(y).$$

Pe de altă parte, avem relaţia evidentă

$$f(a) = 1 \cdot f(a) = \frac{y - x}{y - a} f(a) + \frac{x - a}{y - a} f(a).$$

Scăzând cele două relaţii obţinute, găsim

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(y) - f(a)}{y - a} \quad (\text{deci } R(x) \leq R(y)).$$

Teorema este astfel dovedită. ■

Pentru cazul concav, funcţia raport anterioară este monoton descrescătoare; iar, pentru funcţiile strict convexe (sau concave) monotonia este şi ea strictă.

O consecinţă imediată a acestor fapte este

Teorema 5.1.3 *Dacă f este convexă pe I, atunci, în orice punct $a \in I$, există derivatele laterale finite $f'_s(a), f'_d(a)$ cu $f'_s(a) \leq f'_d(a)$. În particular, orice funcţie convexă pe un interval este continuă pe acel interval.*

Demonstraţia se reduce la a observa că, în orice $a \in I$,

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(y) - f(a)}{y - a}, \quad x < a < y.$$

Nu avem decât să trecem la limită pentru $x \to a-, y \to a+$ şi afirmaţia reiese.

Evident, concluzia teoremei este valabilă şi pentru funcţii concave (cu $f'_s(a) \geq f'_d(a)$). Să mai notăm că, oricare ar fi $a, b \in I$ cu $a < b$,

$$f'_d(a) \leq \frac{f(b) - f(a)}{b - a} \leq f'_s(b).$$

O consecinţă a acestei proprietăţi este următoarea. Să admitem ipoteza

(C3) f este derivabilă pe I.

Teorema 5.1.4 *Funcţia* (*derivabilă*) *f este convexă* (*respectiv, concavă*) *pe I dacă şi numai dacă derivata acesteia este crescătoare* (*respectiv, descrescătoare*) *pe I.*

Evident, pentru funcţii strict convexe (sau concave), monotonia derivatei este şi ea strictă.

În fine, să înlocuim (C3) cu o ipoteză mai tare

(C4) f este derivabilă de două ori pe I.

Teorema 5.1.5 *Funcţia f* (*cu proprietatea anunţată*) *este convexă pe I dacă* (*şi numai dacă*)

$$f''(x) \geq 0, \quad \textit{oricare ar fi } x \in I,$$

şi concavă pe I dacă (*şi numai dacă*)

$$f''(x) \leq 0, \quad \textit{oricare ar fi } x \in I.$$

Pentru funcţiile strict convexe (respectiv concave), inegalităţile respective sunt stricte. (În plus, ele au doar un rol de suficienţă.)

Prin definiţie, un punct $a \in I$ cu proprietatea că derivata a doua se anulează (cu schimbare de semn în jurul acestuia) se va numi *punct de inflexiune* al lui f.

Vom prezenta în continuare două aplicaţii utile ale acestor fapte.

(A) Fie dată funcţia $f : (0, \infty) \to R$ prin

$$f(x) = x^p, \ x > 0, \ \text{unde } p > 1 \text{ este fixat.}$$

Avem, evident,

$$f'(x) = px^{p-1}, \ f''(x) = p(p-1)x^{p-2}, \quad x \in (0, \infty).$$

Cum derivata a doua este strict pozitivă, funcţia f va fi obligatoriu strict convexă pe $(0, \infty)$. Urmează că, oricare ar fi elementele $x_1, ..., x_n$ din $(0, \infty)$, avem, cu Teorema 5.1.1

$$\begin{cases} (\lambda_1 x_1 + \cdots + \lambda_n x_n)^p \le \lambda_1 \cdot x_1^p + \cdots + \lambda_n \cdot x_n^p, \\ \text{pentru orice numere } \lambda_1, ..., \lambda_n \text{ din } [0,1] \text{ cu } \lambda_1 + \cdots + \lambda_n = 1. \end{cases}$$

Mai mult, dacă $\lambda_1, ..., \lambda_n$ sunt chiar din $(0, 1)$, atunci avem egalitate doar dacă $x_1 = \cdots = x_n$. În particular, să punem

$$q = \frac{p}{p-1} \quad (\text{deci } \frac{1}{p} + \frac{1}{q} = 1)$$

şi

$$\begin{cases} x_1 = u_1^{1-q} \cdot v_1, ..., x_n = u_n^{1-q} \cdot v_n \\ \lambda_1 = u_1^q/(u_1^q + \cdots + u_n^q), ..., \lambda_n = u_n^q/(u_1^q + \cdots + u_n^q), \end{cases}$$

unde $(u_1, ..., u_n), (v_1, ..., v_n)$ sunt două sisteme de numere din $(0, \infty)$. Obţinem, după calcule, relaţia

$$\text{(IH)} \ \ u_1 v_1 + \cdots + u_n v_n \le (u_1^q + \cdots + u_n^q)^{\frac{1}{q}} (v_1^p + \cdots + v_n^p)^{\frac{1}{p}},$$

numită inegalitatea lui Hölder. Pentru $p = 2$, relaţia precedentă devine cunoscuta inegalitate Cauchy–Schwarz

$$\text{(R1)} \ \ u_1 v_1 + \cdots + u_n v_n \le \sqrt{(u_1^2 + \cdots + u_n^2)(v_1^2 + \cdots + v_n^2)}.$$

Să notăm că relaţia poate fi scrisă pentru orice sistem de numere; dar atunci trebuie de folosit modulul în stânga.

(B) Fie dată acum funcţia $f : (0, \infty) \to R$ prin

$$f(x) = \ln x, \quad x \in (0, \infty).$$

Cum, evident,

$$f'(x) = \frac{1}{x}, \quad f''(x) = -\frac{1}{x^2}, \quad x \in (0, \infty),$$

funcţia noastră apare drept strict concavă pe intervalul $(0, \infty)$. Rezultă şi aici, conform Teoremei 5.1.1 că, oricare ar fi $x_1, ..., x_n$ din $(0, \infty)$

$$\ln(\lambda_1 x_1 + \cdots + \lambda_n x_n) \geq \lambda_1 \ln x_1 + \cdots + \lambda_n \ln x_n,$$

sau, echivalent (cu proprietăţile funcţiei logaritmice)

$$\begin{cases} x_1^{\lambda_1} \cdots x_n^{\lambda_n} \leq \lambda_1 x_1 + \cdots + \lambda_n x_n, \\ \text{pentru orice } \lambda_1, ..., \lambda_n \text{ din } [0, 1] \text{ cu } \lambda_1 + \cdots + \lambda_n = 1. \end{cases}$$

Mai mult, dacă numerele $\lambda_1, ..., \lambda_n$ sunt chiar din $(0, 1)$, avem egalitate doar dacă $x_1 = \cdots = x_n$. În particular, cu $\lambda_1 = \cdots = \lambda_n = \frac{1}{n}$,obţinem

(R2) $\sqrt[n]{x_1 \cdots x_n} \leq \dfrac{1}{n}(x_1 + \cdots + x_n)$

adică cunoscuta inegalitate a mediilor. Aceasta devine egalitate dacă toate numerele care apar aici sunt egale între ele.

5.1.2 Funcţii convexe de mai multe variabile

Fie $m \geq 1$ un număr natural dat. Să fixăm o mulţime convexă şi deschisă D din R^m. Vom zice că funcţia $f : D \longrightarrow R$ este *convexă pe mulţimea* D dacă, oricare ar fi $x, y \in D$,

(D1) $f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y), \quad \lambda, \mu \in [0, 1], \ \lambda + \mu = 1.$

De asemenea, vom numi f,*concavă pe mulţimea* D dacă $-f$ este convexă; adică

(D2) $f(\lambda x + \mu y) \geq \lambda f(x) + \mu f(y), \quad \lambda, \mu \in [0, 1], \ \lambda + \mu = 1.$

Pentru $m = 1$, noţiunile astfel definite se reduc la cele deja introduse (la funcţiile de o variabilă). Această implicaţie poate fi, într-un anume sens, inversată. Mai precis, să introducem notaţia

$$I_D(x,y) = \{\tau \in R;\ \tau x + (1-\tau)y \in D\}, \quad x, y \in D.$$

Nu este greu de văzut că mulţimea în cauză apare ca un interval deschis al axei reale, care include obligatoriu $[0,1]$. (Se ţine cont de alegerea lui D.) Avem atunci

Teorema 5.1.6 *Funcţia f este convexă dacă şi numai dacă oricare ar fi $x, y \in D$, funcţia de o variabilă $g_{x,y} : I_D(x,y) \longrightarrow R$ dată de*

$$g_{x,y}(\tau) = f(\tau x + (1-\tau)y),\ \ \tau \in I_D(x,y)$$

este convexă pe intervalul deschis $I_D(x,y)$.
(*O concluzie analogă este valabilă şi pentru concavitate.*)

Demonstraţia se sprijină direct pe definiţie; şi, de aceea, nu mai dăm alte detalii.

Să mai notăm şi aici că, proprietatea de convexitate sau concavitate introdusă anterior poate fi dată în forma (aparent mai generală) descrisă de Teorema 5.1.1. Cum, însă , aceasta nu va fi efectiv folosită mai departe, nu mai insistăm asupra ei.

Dintre proprietăţile ne–diferenţiale ale funcţiilor convexe, menţionăm (fără demonstraţie):

Teorema 5.1.7 *Dacă funcţia $f : D \longrightarrow R$ este convexă pe D atunci*
(a) *f este continuă pe D (şi, ca atare, mărginită pe orice parte compactă a lui D);*
(b) *f este mărginită inferior pe orice parte mărginită a lui D.*

Să trecem acum la proprietăţile diferenţiale ale funcţiilor convexe. Fie deci funcţia (de n variabile) $f : D \longrightarrow R$, unde D este o parte convexă şi deschisă din R^m. Acceptăm în continuare ipoteza

$$\text{(C1)} \quad \begin{cases} f \text{ este derivabilă parţial pe } D \text{ (adică, în orice } a \in D, \\ \text{există gradientul } \nabla f(a) = \left(\dfrac{\partial f}{\partial x_1}(a), ..., \dfrac{\partial f}{\partial x_m}(a)\right)). \end{cases}$$

Începem deocamdată cu următorul rezultat (pe care îl dăm fără demonstraţie).

Teorema 5.1.8 *Dacă funcţia f este convexă (sau concavă) pe D şi satisface* (C1)*, atunci ea este în mod necesar diferenţiabilă Fréchet în orice punct a din D.*

Cu alte cuvinte, convexitatea impusă lui f poate înlocui (în contextul lui (C1)) o ipoteză globală de forma

(C2) f este de clasă C^1 pe D,

pentru a deduce concluzia rezultatului anterior.

Următoarea caracterizare (diferenţială) a convexităţii (care poate fi considerată ca o extindere a Teoremei 5.1.2) se va dovedi utilă.

Teorema 5.1.9 *Fie $f : D \longrightarrow R$ o funcţie derivabilă parţial pe D. Atunci f este convexă pe D dacă şi numai dacă oricare ar fi $a \in D$*

(C3) $f(x) - f(a) \geq \nabla f(a)(x - a),\ x \in D.$

(*Un rezultat corespunzător are loc pentru concavitate, prin schimbarea sensului inegalităţii anterioare.*)

Demonstraţie. Necesitatea. Să admitem că f este convexă pe D şi fie $a \in D$ arbitrar fixat. Din proprietatea de convexitate, avem, în orice punct $x \in D$,evaluarea

$$f(a + \tau(x - a)) \leq (1 - \tau)f(a) + \tau f(x),\ 0 < \tau < 1.$$

Scăzând $f(a)$ şi împărţind cu $\tau > 0$, avem

$$\frac{1}{\tau}[f(a + \tau(x - a)) - f(a)] \leq f(x) - f(a),\ 0 < \tau < 1.$$

Conform ipotezei (C1) şi rezultatului precedent, limita părţii stângi pentru $\tau \to 0+$ este $\nabla f(a)(x - a)$. Iar de aici, concluzia este clară.

Suficienţa. Să admitem că (C3) are loc. Fie x, y două elemente din D şi λ un număr din $(0, 1)$. Conform proprietăţii anterioare, avem, cu notaţia $z = \lambda x + (1 - \lambda)y$,

$$\begin{array}{rcccl} f(x) - f(z) & \geq & \nabla f(z)(x - z) & = & (1 - \lambda)\nabla f(z)(x - y) \\ f(y) - f(z) & \geq & \nabla f(z)(y - z) & = & (-\lambda)\nabla f(z)(x - y). \end{array}$$

Înmulţim prima inegalitate cu λ, a doua cu $1-\lambda$ şi le adunăm. Obţinem

$$\lambda f(x)+(1-\lambda)f(y)-f(z)\geq 0 \text{ (adică } f(z)\leq \lambda f(x)+(1-\lambda)f(y)).$$

Şi atunci, cu notaţia făcută, afirmaţia reiese. ∎

Făcând din nou comparaţia cu cazul $m=1$, ar fi de aşteptat acum să existe o posibilitate de caracterizare a convexităţii sau concavităţii prin intermediul derivatelor parţiale de ordinul doi. Răspunsul este pozitiv şi poate fi dat în termenii următori. Fie $f : D \longrightarrow R$ o funcţie ce satisface

(C4) f este de clasă C^2 pe D.

Să reamintim o noţiune cu caracter algebric (Secţiunea 1.4.5). Anume, dată matricea $A=(a_{ij})$ din $\mathcal{M}(m)$, vom zice că aceasta este *de tip pozitiv*, dacă forma pătratică asociată

$$\psi_A(h)=h^\top Ah,\quad h=(h_1,...,h_m)^\top \text{ (element din } R^m)$$

este pozitiv definită, în sensul

$$\psi_A(h)>0, \text{ oricare ar fi } h\in R^m,\ h\neq 0.$$

Corespunzător, vom zice că A este *de tip negativ* dacă $-A$ este de tip pozitiv. Putem acum formula următorul rezultat, (comparabil cu Teorema 5.1.5 dinainte):

Teorema 5.1.10 *Dată funcţia f ce satisface* (C4), *să mai admitem că*

$$\text{(C5)}\ \begin{cases} \textit{matricea hessiană } Hf(a)=\left(\dfrac{\partial^2 f}{\partial x_i\partial x_j}(a)\right) \\ \textit{este de tip pozitiv, oricare ar fi } a\in D. \end{cases}$$

Atunci f este convexă pe D.
(*Un rezultat asemănător are loc şi pentru concavitate; dar atunci, "pozitiv" se va înlocui cu "negativ", în condiţia scrisă.*)

Demonstraţie. Vom arăta că, în fond, totul se reduce la rezultatul precedent. Fie $a\in D$, un element fixat. În baza Formulei de medie Lagrange (vezi Secţiunea 4.3.3), putem scrie pentru orice x arbitrar fixat din D, evaluara

$$f(x)-f(a)=\nabla f(a)(x-a)+\frac{1}{2}(x-a)^\top[Hf(a+\theta(x-a))](x-a),$$

unde $\theta = \theta(a, x)$ este un număr din intervalul $(0, 1)$. Ţinând acum cont de (C5), urmează că

$$f(x) - f(a) \geq \nabla f(a)(x - a).$$

Iar aceasta, combinată cu cele spuse în cadrul Teoremei 5.1.9, încheie argumentul. ∎

Exemplu. Să se studieze convexitatea funcţiei

$$f(x, y, z) = x + \frac{y^2}{4x} + \frac{z^2}{y} + \frac{2}{z}, \quad x, y, z > 0.$$

Calculăm derivatele parţiale de ordinul întâi şi doi

$$\frac{\partial f}{\partial x} = 1 - \frac{y^2}{4x^2}, \ \frac{\partial f}{\partial y} = \frac{y}{2x} - \frac{z^2}{y^2}, \ \frac{\partial f}{\partial z} = \frac{2z}{y} - \frac{2}{z^2},$$
$$\begin{cases} \dfrac{\partial^2 f}{\partial x^2} = \dfrac{y^2}{2x^3}, & \dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{-y}{2x^2}, \quad \dfrac{\partial^2 f}{\partial x \partial z} = 0, \\ \dfrac{\partial^2 f}{\partial y^2} = \dfrac{1}{2x} + \dfrac{2z^2}{y^3}, & \dfrac{\partial^2 f}{\partial y \partial z} = \dfrac{-2z}{y^2}, \quad \dfrac{\partial^2 f}{\partial z^2} = \dfrac{2}{y} + \dfrac{4}{z^3}. \end{cases}$$

Matricea hessiană a funcţiei noastre are deci forma

$$Hf(x, y, z) = \begin{pmatrix} \dfrac{y^2}{2x^3} & \dfrac{-y}{2x^2} & 0 \\ \dfrac{-y}{2x^2} & \dfrac{1}{2x} + \dfrac{2z^2}{y^3} & \dfrac{-2z}{y^2} \\ 0 & \dfrac{-2z}{y^2} & \dfrac{2}{y} + \dfrac{4}{z^3} \end{pmatrix}, \quad x, y, z > 0.$$

Calculăm acum determinanţii principali ai acesteia

$$\begin{aligned} \Delta_1 &= \frac{y^2}{2x^3} \\ \Delta_2 &= \det \begin{pmatrix} \dfrac{y^2}{2x^3} & \dfrac{-y}{2x^2} \\ \dfrac{-y}{2x^2} & \dfrac{1}{2x} + \dfrac{2z^2}{y^3} \end{pmatrix} = \frac{z^2}{y}, \\ \Delta_3 &= \det\left(Hf(x, y, z)\right) = \frac{4}{x^3 yz}. \end{aligned}$$

Cum toţi aceşti determinanţi sunt strict pozitivi pe domeniul (convex deschis) $D = \{(x, y, z); x, y, z > 0\}$, urmează că matricea hessiană este de tip pozitiv pe acest domeniu.(A se vedea Teorema 1.4.6). Deci, funcţia f este convexă pe D.

Probleme propuse la §5.1

1. Fie $f : R \longrightarrow R$ funcţie convexă şi $g : R \longrightarrow R$ funcţie convexă şi monoton crescătoare. Să se arate că funcţia compusă $g \circ f$ este convexă.

2. Dacă $f : R \longrightarrow R$ este continuă şi monoton crescătoare, atunci funcţia $F : I \longrightarrow R$ dată de $F(x) = \int_a^x f(t)dt$, $x \in I$ (unde a este arbitrar fixat în I) este convexă.

3. Să se studieze convexitatea şi concavitatea funcţiei $f(x) = x^2e^{-x}$, pe diferitele intervale ale mulţimii numerelor reale.

4. Să se arate că funcţia (de m variabile)

$$f(x_1 ... x_m) = |x_1| + \cdots + |x_m|, \ (x_1, ..., x_m) \in R^m$$

este convexă pe spaţiul R^m.

5. Să se studieze convexitatea funcţiei

$$\begin{aligned} f(x, y, z) \ &= \ x^2 + (x-1)^2 + (x+1)^2 + y^2 + (y-2)^2 + (y+2)^2 + \\ &+ z^2 + (z-3)^2 + (z+3)^2. \end{aligned}$$

6. Să se demonstreze inegalitatea

$$\sin a + \sin b + \sin c \leq 3\frac{\sqrt{3}}{2}, \quad \forall a, b, c \in [0, \pi].$$

(Indicaţie. Se va studia concavitatea funcţiei $f(x) = \sin x$ pe intervalul $[0, \pi]$.)

7. Să se studieze convexitatea funcţiei

$$f(x, y) = \lambda x^2 + y^2 + x + y, \quad x, y \in R$$

pentru diferitele valori ale parametrului λ şi pe diferite regiuni ale planului R^2.

8. Să se studieze convexitatea funcţiei

$$f(x_1, ..., x_m) = (c_1x_1 + \cdots + c_mx_m)^2,\ (x_1, ..., x_m) \in R^m,$$

unde $c_1, ..., c_m$ sunt nişte constante.

5.2 Extreme libere ale funcţiilor

5.2.1 Extreme ale funcţiilor de o variabilă

Fie I un interval (deschis) al axei reale şi $f : I \longrightarrow R$ o funcţie dată. Să ne dăm un punct $a \in I$ şi un sub–interval deschis J (al lui I), care conţine pe a. Zicem că a este *punct de minim* (respectiv,*maxim*) *global* al funcţiei f relativ la sub–intervalul (deschis) J, dacă

(D) $f(x) \geq f(a)$ (respectiv, $f(x) \leq f(a)$), pentru toţi $x \in J$.

Când una din aceste proprietăţi are loc, vom zice că $a \in I$ este *punct de extrem global* al funcţiei f *relativ* la subintervalul J. Când $J = I$, vom spune că avem de-a face cu un *punct de extrem global* al funcţiei f. Iar, în fine, când subintervalul J (ce conţine a) este subînţeles în definiţia (D), vom zice că $a \in I$ este un *punct de extrem local* al funcţiei f.

În legătură cu noţiunile introduse, putem formula următoarele probleme de bază:

(i) determinarea *condiţiilor necesare* de extrem (adică, stabilirea de proprietăţi – de natură diferenţială în special – pe care le au punctele de extrem);

(ii) determinarea *condiţiilor suficiente* de extrem (adică, găsirea de condiţii suplimentare – pe lângă cele stabilite deja – care să ne asigure o proprietate de extrem).

În ce priveşte prima problemă, să admitem ipoteza

(C1) f este derivabilă pe intervalul I.

Următorul rezultat (cunoscut sub numele de "Teorema lui Fermat") ne va da răspunsul cerut.

Teorema 5.2.1 *Fie a un punct de extrem (global sau local) al funcţiei f. Atunci, în mod necesar avem $f'(a) = 0$.*

Demonstraţie. Pentru precizarea cadrului, să considerăm , de pildă, cazul în care punctul a este de minim local. Avem, deci, asigurat un $\varepsilon > 0$ şi un subinterval deschis $J = (a - \varepsilon, a + \varepsilon)$ al lui I, cu

(R1) $f(x) - f(a) \geq 0, \quad x \in (a - \varepsilon, a + \varepsilon).$

În particular, de aici deducem

$$\frac{1}{x-a}(f(x) - f(a)) \geq 0, \quad \forall x \in (a, a + \varepsilon).$$

Şi atunci, trecând la limită pentru $x \to a+$, găsim $f'(a) \geq 0$. Pe de altă parte, aceeaşi relaţie (R1) ne dă

$$\frac{1}{x-a}(f(x) - f(a)) \leq 0, \quad \forall x \in (a - \varepsilon, a).$$

De unde, trecând la limită, $f'(a) \leq 0$. Combinând relaţiile găsite,avem $f'(a) = 0$ şi concluzia reiese. ∎

Prin definiţie, un punct $a \in I$ cu $f'(a) = 0$ se va numi *staţionar* pentru funcţia f. Avem deci implicaţia

(P) $a =$ punct de extrem (global sau local) $\Longrightarrow a =$ punct staţionar.

Se pune întrebarea în ce condiţii suplimentare implicaţia scrisă poate fi inversată. Acestea vor alcătui desigur răspunsul la cea de-a doua problemă formulată anterior; şi, în esenţă, ele se reduc la o analiză a derivatelor de ordin superior. Anume, să admitem că (C1) se înlocuieşte cu ipoteza mai tare

(C2) f este de clasă C^{n+1} pe I, pentru un anumit $n \geq 1$.

Avem, atunci, următorul rezultat (deseori aplicabil în practică).

Teorema 5.2.2 *Fie $a \in I$ un punct staţionar al funcţiei f, cu*

(C2) $f'(a) = \cdots = f^{(n)}(a) = 0, \quad f^{(n+1)}(a) \neq 0.$

Au loc atunci concluziile:

(a) *Dacă n este impar, atunci a este punct de minim local când $f^{(n+1)}(a) > 0$; şi, respectiv, punct de maxim local dacă $f^{(n+1)}(a) < 0$.*

(b) *Dacă n este par, atunci a nu este punct de extremum pentru funcţia f.*

Demonstraţie. Din ipoteza de continuitate impusă derivatelor, va exista un sub–interval deschis $J = (a - \varepsilon, a + \varepsilon)$ al lui I cu

(R2) $f^{(n+1)}(x)$ are acelaşi semn cu $f^{(n+1)}(a)$, $\forall x \in J$.

Să fixăm un asemenea x din J şi să scriem formula de medie Taylor de ordinul n (vezi Secţiunea 4.3.3). Avem, atunci, combinând şi cu ipoteza (C2) (acceptată aici)

$$\begin{cases} f(x) - f(a) = \dfrac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(a + \theta(x-a)), \\ \text{pentru un anume } \theta = \theta(a,x) \text{ din } (0,1). \end{cases}$$

Să notăm cu această ocazie că, obligatoriu, $a + \theta(x-a) \in J$ (deoarece $a, x \in J$). În acest caz, dacă, de pildă, $f^{(n+1)}(a) > 0$, trebuie să avem $f^{(n+1)}(a + \theta(x-a)) > 0$. Iar, dacă, în plus, n este impar, atunci este clar că $f(x) - f(a) \geq 0$, deoarece $(x-a)^{n+1} \geq 0$. Celelalte cazuri se discută analog. ■

Să reţinem din acest raţionament că, în fapt, la alternativa (a), punctul în cauză este totuşi un punct de extrem global al funcţiei considerate în raport cu acele sub–intervale deschise J (din I) care conţin acel punct şi satisfac proprietatea (R2) (care, în baza unui rezultat anterior, devine o condiţie de convexitate (sau de concavitate) a funcţiei pe un asemenea interval).

Exemplu. Să se determine extremele funcţiei

$$f(x) = x^3 - 3x + 1, \quad x \in R.$$

Calculăm derivatele de ordinul unu şi doi ale funcţiei noastre

$$f'(x) = 3x^2 - 3, \quad f''(x) = 6x, \qquad x \in R.$$

Determinăm acum punctele staţionare (pe baza definiţiei)

$$f'(x) = 0 \Longrightarrow x^2 - 1 = 0 \Longrightarrow x_1 = -1, \; x_2 = 1.$$

Natura acestora se precizezază conform tabelului de mai jos:

x	$f''(x)$	Natura punctului x
$x_1 = -1$	-6	punct de maxim (local)
$x_2 = 1$	$+6$	punct de minim (local)

Din expresia derivatei a doua rezultă acum clar că punctul $x_1 = -1$ este de maxim global al funcţiei f în raport cu sub–intervalul $J_1 = (-\infty, 0)$; iar punctul $x_2 = 1$ este de minim global în raport cu sub–intervalul $J_2 = (0, \infty)$. Desigur, aceste intervale mai pot fi extinse prin studiul tabloului de variaţie al funcţiei f

x	$-\infty$		-1		1		∞
$f'(x)$		$+$	0	$-$	0	$+$	
$f(x)$	$-\infty$	$\nearrow$	3 (max)	$\searrow$	-1 (min)	$\nearrow$	$+\infty$

De exemplu, din acest tablou se vede uşor că, în realitate, $x_1 = -1$ este de maxim global în raport cu sub–intervalul $J_1' = (-\infty, 1)$. Dar, şi informaţia anterioară este utilă în practică, mai ales atunci când derivata întâia are o expresie relativ complicată.

5.2.2 Extreme ale funcţiilor de mai multe variabile

Fie $m \geq 1$ un număr natural dat. Să fixăm o mulţime convexă şi deschisă D din R^m şi fie $f : D \longrightarrow R$ o funcţie dată. Mai luăm apoi un punct a din D şi o parte convexă deschisă E din D care conţine a. Zicem că a este *punct de minim* (respectiv, *maxim*) *global* al funcţiei f relativ la submulţimea (convexă deschisă) E, dacă

(D1) $f(x) \geq f(a)$ (respectiv, $f(x) \leq f(a)$), pentru toţi $x \in E$.

Când una din aceste proprietăţi are loc, vom zice că punctul $a \in D$ este *punct de extrem global* al funcţiei f *relativ* la submulţimea E. În particular, pentru $E = D$, vom spune că avem de-a face cu un *punct de extrem global* al funcţiei f. Iar când partea convexă deschisă E (ce conţine a) este subînţeleasă în (D1), vom zice că $a \in D$ este un *punct de extrem local* al funcţiei f.

Ca şi la cazul $m = 1$, suntem interesaţi în stabilirea condiţiilor necesare de extremum şi apoi (când acestea au fost găsite), în stabilirea condiţiilor suficiente de extremum. În ce priveşte prima problemă, să facem ipoteza

$$\text{(C1)} \quad \left\{ \begin{array}{l} f \text{ este derivabilă parţial pe } D \text{ (adică, există gradientul} \\ \nabla f(x) \text{ în orice punct } x \text{ al domeniului } D). \end{array} \right.$$

Următorul rezultat (cunoscut şi acesta sub denumirea ”Teorema lui Fermat”) este răspunsul cerut.

Teorema 5.2.3 *Fie a un punct de extrem (global sau local) al funcţiei f. Atunci, în mod necesar,*

(R1) $\nabla f(a) = 0$ *(adică* $\dfrac{\partial f}{\partial x_1}(a) = 0, ..., \dfrac{\partial f}{\partial x_m}(a) = 0$*).*

Demonstraţie. Să considerăm, de pildă, cazul în care punctul a este de minim local. (Celelalte situaţii se vor discuta în acelaşi mod.) Avem, deci, asigurat un $\varepsilon > 0$ şi o sferă (deschisă) $E = S(a, \varepsilon)$ din D, cu

(R2) $f(x) - f(a) \geq 0, \quad x \in E.$

Notând ca de obicei prin $\{e_1, ..., e_m\}$ baza canonică din R^m, să fixăm un $i \in \{1, ..., m\}$. În baza relaţiei anterioare, putem scrie

(R3) $f(a + \tau e_i) - f(a) \geq 0, \ \forall \tau \in (-\varepsilon, \varepsilon).$

În particular, aceasta ne dă

$$\frac{1}{\tau}(f(a + \tau e_i) - f(a)) \geq 0, \ \forall \tau \in (0, \varepsilon).$$

Şi, deci, trecând la limită pentru $\tau \to 0+$, găsim $\dfrac{\partial f}{\partial x_i}(a) \geq 0$. Pe de altă parte, aceeaşi relaţie (R3) ne dă

$$\frac{1}{\tau}(f(a + \tau e_i) - f(a)) \leq 0, \ \forall \tau \in (-\varepsilon, 0).$$

Trecând şi aici la limită, găsim $\dfrac{\partial f}{\partial x_i}(a) \leq 0$; şi, deci, $\dfrac{\partial f}{\partial x_i}(a) = 0$. Cum $i \in \{1, ..., m\}$ a fost arbitrar ales, concluzia reiese. ■

Prin definiţie, un punct $a \in D$ cu $\nabla f(a) = 0$ se va zice *staţionar* pentru funcţia f. Am demonstrat deci că implicaţia (P) (Secţiunea 5.2.1) este valabilă şi în acest context.

Ne interesează acum în ce condiţii suplimentare putem inversa această implicaţie; acestea vor fi desigur tocmai condiţiile suficiente de extremum pe care le căutăm. În esenţă, totul revine la studiul derivatelor parţiale de ordin doi. Anume, să admitem că (C1) se înlocuieşte cu ipoteza mai tare

(C2) $\left\{ \begin{array}{l} f \text{ este de clasă } C^2 \text{ pe } D \text{ (şi, deci, în particular} \\ Hf(x) \text{ există, în orice punct } x \text{ al domeniului } D). \end{array} \right.$

Să notăm, pentru simplitate,în fiecare $p \in \{1, ..., m\}$

$$\text{(D2)} \left\{ \begin{array}{lcl} \Delta_p(x) & = & \text{determinantul format cu primele } p \text{ linii şi} \\ & & p \text{ coloane din matricea } Hf(x),\ x \in D. \end{array} \right.$$

Avem atunci valabilă

Teorema 5.2.4 *Fie $a \in D$ un punct staţionar al funcţiei f ,cu*

(C3) $\Delta_1(a) \neq 0, ..., \Delta_m(a) \neq 0.$

Au loc atunci concluziile

(a) *Dacă* $\Delta_i(a) > 0, 1 \leq i \leq m$ *(sau,echivalent, $Hf(a)$ este de tip pozitiv) atunci a este un punct de minim local pentru funcţia f.*

(b) *Dacă* $(-1)^i \Delta_i(a) > 0, 1 \leq i \leq m$ *(sau,echivalent, $Hf(a)$ este de tip negativ) atunci a este un punct de maxim local pentru funcţia f.*

(c) *În toate celelalte cazuri rămase relative la semnele lui $\Delta_1(a), ...,$ $\Delta_m(a)$,punctul a nu este de extrem (local sau global) pentru funcţia f.*

Demonstraţie. Din ipoteza de continuitate impusă derivatelor parţiale de ordin doi, va exista o sferă deschisă $E = S(a, \varepsilon)$ din D cu

(R4) $\Delta_p(x)$ are acelaşi semn cu $\Delta_p(a),\ p \in \{1, ..., m\},\ x \in E.$

Să fixăm un asemenea x în E şi să scriem formula de medie Taylor de ordin unu; avem, deci (combinând şi cu ipoteza de staţionaritate)

$$\left\{ \begin{array}{l} f(x) - f(a) = \dfrac{1}{2}(x-a)^\top [Hf(a+\theta+\theta(x-a))](x-a), \\ \text{pentru un anume } \ \theta = \theta(a, x) \text{ din } (0, 1). \end{array} \right.$$

Să notăm că, obligatoriu, $a+\theta(x-a) \in E$, deoarece mulţimea E este convexă şi $a, x \in E$. Acum, pentru a face o alegere, admitem că alternativa (a) are loc. În baza lui (R4), ajungem deci la

$$\Delta_i(a + \theta(x-a)) > 0, \quad 1 \leq i \leq m$$

sau, cu alte cuvinte,

$$Hf(a + \theta(x-a)) \quad \text{este de tip pozitiv.}$$

Dar atunci, din formula Taylor anterioară este clar că $f(x) - f(a) \geq 0$; adică, a este punct de minim local. Teorema este dovedită. ■

Ca şi la cazul $m = 1$, să reţinem faptul că în cadrul alternativelor (a) sau (b), punctul în cauză este totuşi de extrem global al funcţiei considerate pe acele submulţimi convexe deschise E (din D) care conţin acel punct şi satisfac proprietatea (R4). În particular, pentru $E = D$, va rezulta că punctul a este de extrem global pentru funcţia f. Iar, în fine, punctele $a \in D$ ce satisfac alternativa (c) se vor numi *puncte şa* ale funcţiei considerate. Aceasta provine din faptul că, în raport cu anumite direcţii din R^m, ele apar ca puncte de minim (local) ale funcţiei f, iar, în raport cu alte direcţii, ele au proprietatea de a fi puncte de maxim (local) pentru funcţia în cauză.

Exemplu. Să se studieze extremele funcţiei

$$f(x, y) = x^3 + 3xy^2 - 15x - 12y, \quad x, y \in R.$$

Calculăm derivatele parţiale de ordinul întâi

$$\frac{\partial f}{\partial x} = 3x^2 + 3y^2 - 15, \ \frac{\partial f}{\partial y} = 6xy - 12.$$

Punctele staţionare ale funcţiei se obţin prin anularea simultană a derivatelor parţiale; deci, din sistemul

$$x^2 + y^2 - 5 = 0, \ xy - 2 = 0.$$

Soluţiile acestui sistem sunt punctele (din plan)

$$P_1 = (1, 2), \ P_2 = (2, 1), \ P_3 = (-1, -2), \ P_4 = (-2, -1).$$

Pentru a stabili natura acestora, vom construi derivatele parţiale de ordin doi ale funcţiei

$$\frac{\partial^2 f}{\partial x^2} = 6x, \ \frac{\partial^2 f}{\partial x \partial y} = 6y, \ \frac{\partial^2 f}{\partial y^2} = 6x.$$

Deci, matricea hessiană are forma

$$Hf(x, y) = \begin{pmatrix} 6x & 6y \\ 6y & 6x \end{pmatrix}, \ (x, y \in R^2.$$

Determinanţii principali ai acesteia sunt

$$\Delta_1(x, y) = 6x, \quad \Delta_2(x, y) = 36(x^2 - y^2), \quad (x, y) \in R^2.$$

Sintetizăm acum concluziile în tabelul de mai jos:

Punctul (P)	$\Delta_1(P)$	$\Delta_2(P)$	Natura punctului (P)
$P_1 = (1, 2)$	$+6$	-108	punct şa
$P_2 = (2, 1)$	$+12$	$+108$	punct de minim (local)
$P_3 = (-1, -2)$	-6	-108	punct şa
$P_4 = (-2, -1)$	-12	$+108$	punct de maxim (local)

Din expresia acestor determinanţi, rezultă acum că, de fapt, punctul $P_2 = (2, 1)$ este de minim global în raport cu regiunea $E_1 = \{(x, y);\ x > 0, x^2-y^2 > 0\}$; şi, analog, punctul $P_4 = (-2, -1)$ este de maxim local în raport cu regiunea $E = \{(x, y);\ x < 0, x^2 - y^2 > 0\}$. Aceasta, desigur, ar putea fi util în diverse cazuri concrete când se cere o comparaţie a valorilor funcţiei în punctele de extrem cu valorile în puncte mai mult sau mai puţin apropiate de acestea.

Să ne întoarcem la alternativele (a), (b) deja discutate. După cum se ştie, condiţiile impuse acolo matricei hessian antrenează convexitatea (respectiv, concavitatea) funcţiei f pe domeniul considerat. Ar fi de aşteptat deci ca, pentru funcţii convexe (sau concave) să regăsim concluzia teoremei anterioare, chiar în absenţa unei ipoteze de tip (C2) privind derivatele parţiale de ordin doi. Că aceasta este adevărat, rezultă din următoarea teoremă de tip global. Fie $f : D \longrightarrow R$ o funcţie care satisface doar condiţia (C1).

Teorema 5.2.5 *Să presupunem că $a \in D$ este un punct staţionar al funcţiei f. Atunci, acesta este un*

(a) *punct de minim global, dacă f este convexă pe D.*

(b) *punct de maxim global, dacă f este concavă pe D.*

Demonstraţie. Vom discuta doar prima alternativă; cealaltă se tratează complet analog. Conform unei caracterizări a convexităţii dată într-un loc anterior, avem

$$f(x) - f(a) \geq \nabla f(a)(x - a),\ \forall x \in D.$$

Cum însă $\nabla f(a) = 0$, ajungem la concluzia

$$f(x) - f(a) \geq 0, \quad \text{pentru orice} \quad x \in D.$$

Şi cu aceasta, teorema este dovedită. ∎

Să remarcăm că, dacă în rezultatul anterior, ipoteza de convexitate (sau concavitate) are caracter local (adică, se impune doar pe o parte convexă deschisă E din D care conţine a), atunci punctul în chestiune va apare ca extrem local pentru funcţia f. Deci, această variantă a Teoremei 5.2.5 ar constitui o generalizare efectivă a Teoremei 5.2.4. În realitate, însă , convexitatea (sau concavitatea) funcţiei este greu de verificat prin alte metode decât cele care se referă la matricea hessian. Şi, deci, Teorema 5.2.4 îşi păstrează valabilitatea în multe cazuri concrete.

Exemplu. Să se studieze extremele funcţiei

$$f(x,y,z) = x + \frac{y^2}{4x} + \frac{z^2}{y} + \frac{2}{z},\ x,y,z > 0.$$

Am văzut deja (Secţiunea 5.1.2) că această funcţie este convexă pe $D = \{(x,y,z);\ x,y,z > 0\}$. Să calculăm derivatele parţiale de ordinul întâi ale funcţiei noastre:

$$\frac{\partial f}{\partial x} = 1 - \frac{y^2}{4x^2},\ \frac{\partial f}{\partial y} = \frac{y}{2x} - \frac{z^2}{y^2},\ \frac{\partial f}{\partial z} = \frac{2z}{y} - \frac{2}{z^2}.$$

Punctele staţionare ale funcţiei f se vor obţine prin anularea simultană a derivatelor parţiale; deci, din sistemul

$$4x^2 - y^2 = 0,\ y^3 - 2xz^2 = 0,\ z^3 - y = 0.$$

Singura soluţie din domeniul D a acestui sistem este punctul $P = \left(\frac{1}{2}, 1, 1\right)$. În baza celor spuse anterior, urmează că acesta este punct de minim global al funcţiei f pe mulţimea D.

5.2.3 Metoda celor mai mici pătrate

Fie x şi y două variabile care definesc un anumit fenomen. Presupunem că, în urma unor investigaţii de natură teoretică, am reuşit să specificăm legitatea fenomenului în cauză, prin intermediul unei relaţii de forma

$$\text{(E)} \qquad \psi(y) = \lambda_1\varphi_1(x) + \cdots + \lambda_m\varphi_m(x).$$

Aici, funcţiile $\varphi_1, ..., \varphi_m, \psi$ sunt bine determinate, iar $\lambda_1, ..., \lambda_m$ sunt nişte parametri care urmează a fi determinaţi prin intermediul unui anumit grup de observaţii asupra fenomenului, definite de tabelul alăturat

Variabila x	x_1	x_2	$\cdots$	x_n
Variabila y	y_1	y_2	$\cdots$	y_n

(Se subînţelege aici că toate valorile $x_1, ..., x_n$ ale variabilei independente x sunt distincte între ele.) În acest sens, există suficiente motive să afirmăm că o metodă utilă de determinare a parametrilor în cauză este de a le atribui valorile care se deduc din determinarea punctului de minim al sumei de pătrate (în variabilele $\lambda_1, ..., \lambda_m$)

$$F(\lambda_1, ..., \lambda_m) = \sum_{i=1}^{n}(\lambda_1\varphi_1(x_i) + \cdots + \lambda_m\varphi_m(x_i) - \psi(y_i))^2.$$

Pentru a nu complica prea mult calculele, vom lua în discuţie cazul mai simplu (dar important) când ecuaţia fenomenului este de forma liniară

$$\text{(EL)} \qquad y = \lambda x + \mu.$$

(Dar,trebuie de subliniat că raţionamentele pe care le dezvoltăm au un caracter general). Într-un asemenea caz,funcţia de minimizat are forma

$$F(\lambda, \mu) = \sum_{i=1}^{n}(\lambda x_i + \mu - y_i)^2, \ \lambda, \mu \in R.$$

Forma pătratică (în variabilele λ şi μ) extrasă din această funcţie

$$G(\lambda, \mu) = \sum_{i=1}^{n}(\lambda x_i + \mu)^2, \quad \lambda, \mu \in R$$

este obligatoriu pozitiv definită; deoarece

$$G(\lambda, \mu) = 0 \Longrightarrow \lambda x_i + \mu = 0, \quad 1 \leq i \leq n;$$

iar aceasta implică neapărat $\lambda = \mu = 0$. (Altfel, polinomul $P(x) = \lambda x + \mu$ ar avea n rădăcini distincte $x_1, ..., x_n$, contradicţie). Ca urmare, matricea formei pătratice (sau, echivalent, hessianul acesteia

$$HG(\lambda, \mu) = 2\begin{pmatrix} \sum\limits_i x_i^2 & \sum\limits_i x_i \\ \sum\limits_i x_i & n \end{pmatrix}$$

este de tip pozitiv; adică

$$\Delta_1 = 2(\sum_i x_i^2) > 0, \quad \Delta_2 = 4\left(n(\sum_i x_i^2) - (\sum_i x_i)^2\right) > 0.$$

Cum însă , evident, F şi G au acelaşi hessian

$$HF(\lambda,\mu) = HG(\lambda,\mu),\ \forall(\lambda,\mu) \in R^2,$$

urmează că funcţia F este neapărat convexă pe R^2.

Să analizăm punctele staţionare ale acesteia. Avem deocamdată

$$\begin{cases} \dfrac{\partial F}{\partial \lambda} & = 2\displaystyle\sum_{i=1}^{n}(\lambda x_i + \mu - y_i)x_i = \\ & = 2\left[\lambda\left(\displaystyle\sum_{i=1}^{n} x_i^2\right) + \mu\left(\displaystyle\sum_{i=1}^{n} x_i^2\right) - \displaystyle\sum_{i=1}^{n} y_i x_i\right] \\ \dfrac{\partial F}{\partial \mu} & = 2\displaystyle\sum_{i=1}^{n}(\lambda x_i + \mu - y_i) = 2\left[\lambda\left(\displaystyle\sum_{i=1}^{n} x_i\right) + \mu\cdot n - \displaystyle\sum_{i=1}^{n} y_i\right] \end{cases}$$

Sistemul care dă punctele staţionare are, deci, forma

$$\text{(S)} \qquad \begin{cases} \lambda\left(\displaystyle\sum_i x_i^2\right) + \mu\left(\displaystyle\sum_i x_i\right) = \displaystyle\sum_i y_i x_i, \\ \lambda\left(\displaystyle\sum_i x_i\right) + \mu\cdot n = \displaystyle\sum_i y_i. \end{cases}$$

(Vom numi acesta *sistem normal* ataşat datelor noastre.)
Determinantul acestui sistem este tocmai $\frac{1}{4}\Delta_2 > 0$. Deci (S) are soluţie unică; fie ea (λ_0, μ_0). Acestea reprezintă tocmai valorile căutate ale parametrilor.

Exemplu. Se cunoaşte că un fenomen, de variabile componente (x, y) are o specificare de tip liniar (EL). Se cere o evaluare a parametrilor λ, μ care apar în ecuaţia modelului, prin intermediul următoarelor 5 observaţii asupra acestuia

Variabila x	-1	0	2	3	5
Variabila y	3	1	4	7	9

Avem, evident,

$$\sum_i x_i^2 = 39;\ \sum_i x_i = 9;\ \ \sum_i y_i x_i = 71;\ \sum_i y_i = 24.$$

Sistemul normal ataşat datelor problemei este

$$39\lambda + 9\mu = 71;\ 9\lambda + 5\mu = 24.$$

Soluţia sistemului în cauză este

$$\lambda = \frac{139}{114}; \quad \mu = \frac{297}{114}.$$

Deci, ecuaţia fenomenului considerat este

$$y = \frac{139}{114}x + \frac{297}{114}.$$

După cum am spus deja, metoda celor mai mici pătrate se aplică la fenomene specificate în forma (E). Din clasa acestora fac de exemplu parte, toate fenomenele descrise de o lege exponenţială; deoarece, evident,

$$y = ae^{\lambda x} \Longleftrightarrow \ln y = \lambda x + \ln a.$$

Mai menţionăm aici şi clasa tuturor fenomenelor de tip polinomial

$$y = \lambda_1 x^m + \lambda_2 x^{m-1} + \cdots + \lambda_m x + \lambda_{m+1}.$$

Cu alte cuvinte, clasa de fenomene la care se poate aplica metoda în cauză este suficient de largă, pentru a include toate cazurile particulare cu relevanţă practică.

Probleme propuse la §5.2

1. Să se studieze , pornind de la definiţie, extremele funcţiei $f(x) = |x-1|^m,\ x \in R$, unde m este un număr natural.

2. Să se studieze extremele funcţiei $f(x) = e^x - (1+x)$. Ca aplicaţie, să se demonstreze inegalitatea $e^x \geq 1 + x$, pentru orice $x \in R$.

3. Să se studieze extremele funcţiilor

$$(a) \quad f(x) = x(x-1)^2(x-2)^3; \quad (b) \quad f(x) = \frac{1}{x(4-x)^2};$$

$$(c) \quad f(x) = \sqrt[3]{(x^2-1)^2}; \quad (d) \quad f(x) = x \ln x.$$

4. Să se găsească punctele de extrem ale funcţiilor de două variabile

$$(a)\ f(x,y)=x^3+y^3-3xy; \quad (b)\ f(x,y)=3xy^2-x^3-15x-36y;$$
$$(c)\ f(x,y)=xy+\frac{50}{x}+\frac{20}{y}; \quad (d)\ f(x,y)=xy\ln(x^2+y^2).$$

5. Să se determine extremele funcţiilor de trei sau mai multe variabile

(a) $f(x,y,z)=x^2+y^2+z^2-xy+x-2z$;

(b) $f(x,y,z)=\frac{1}{x}+\frac{x}{y}+\frac{y}{z}+\frac{z}{16},\ x,y,z>0$;

(c) $f(x,y,z,t)=xyzt(1-x-y-z-t)$;

(d) $f(x_1,...,x_m)=x_1^2+x_2^2+\cdots+x_m^2+x_1+x_2+\cdots+x_m$.

6. Să se determine, prin metoda celor mai mici pătrate, dreapta $y=ax+b$ care este cel mai aproape situată de punctele $M_1=(-1,3)$, $M_2=(1,4), M_3=(2,2), M_4=(5,0)$.

7. Ecuaţia unui fenomen este de tip exponenţial $y=a\cdot e^{bx}$, unde a şi b sunt parametri. Să se determine, prin metoda celor mai mici pătrate, valorile acestor parametri, utilizând cuplurile de valori observate $M_1=(2,5), M_2=(3,1), M_3=(4,7)$.

5.3 Funcţii definite implicit

5.3.1 Funcţii implicite de o variabilă

Fie D un domeniu (mulţime conexă şi deschisă) din plan şi $F:D\longrightarrow R$ o funcţie continuă de două variabile. Să considerăm ecuaţia funcţională

$$\text{(E)} \qquad F(x,y)=0.$$

Zicem că funcţia continuă (de o variabilă) $\varphi:I\longrightarrow R$, unde I este un interval al axei reale, apare ca *soluţie de tip implicit, a ecuaţiei* (E), dacă

(D1) $(x, \varphi(x)) \in D$ şi $F(x, \varphi(x)) = 0$, pentru toţi $x \in I$.

În general, pentru un acelaşi interval I, pot exista mai multe asemenea funcţii (sau nici una). (De pildă, pentru

$$F(x, y) = x^2 + y^2 - 1, \quad (x, y) \in R^2$$

şi intervalul $I = [-1, 1]$, funcţiile

$$\varphi_1(x) = -\sqrt{1 - x^2}, \ \varphi_2(x) = \sqrt{1 - x^2}, \quad -1 \leq x \leq 1,$$

sunt două soluţii de tip implicit ale ecuaţiei (E).) Din acest motiv, pentru a le individualiza, este necesar să mai asociem ecuaţiei noastre o condiţie de forma următoare (numită *condiţie iniţială*)

(I) $\qquad \varphi(a) = b, \ \ \text{unde } (a, b) \in D \text{ satisface } F(a, b) = 0.$

În acest caz, funcţia în chestiune se va zice *soluţie de tip implicit a problemei* (E) + (I). (În principiu, aceasta ne fixează în mod unic soluţia; de pildă, singura soluţie (din cele două) care mai verifică condiţia iniţială $\varphi(\frac{1}{2}) = \frac{\sqrt{3}}{2}$ este $\varphi_2 : I \longrightarrow R$.)

Cele spuse mai înainte ne sugerează cum ar trebui să fie formulată problema pentru a obţine soluţii viabile. Mai precis, fie din nou D un domeniu plan şi $F : D \longrightarrow R$ o funcţie de două variabile. Admitem ca ipoteză generală

$$\text{(C1)} \begin{cases} F \text{ este de clasă } C^1 \text{ pe } D \text{ (adică este continuă pe } D, \\ \text{împreună cu derivatele sale parţiale } \dfrac{\partial F}{\partial x}, \dfrac{\partial F}{\partial y}). \end{cases}$$

Avem în acest caz

Teorema 5.3.1 *Fie $(a, b) \in D$ un punct fixat cu*

(C2) $F(a, b) = 0, \quad \dfrac{\partial F}{\partial y}(a, b) \neq 0.$

Există atunci un dreptunghi $D(a, b) = [a - \varepsilon, a + \varepsilon] \times [b - \eta, b + \eta]$ din D şi o funcţie $\varphi : [a - \varepsilon, a + \varepsilon] \longrightarrow [b - \eta, b + \eta]$, cu proprietăţile
(P1) *φ este continuă pe intervalul $[a - \varepsilon, a + \varepsilon]$*
(P2) *φ este o soluţie implicită a problemei* (E)+(I)

(P3) φ *este de clasă* C^1 *pe* $(a-\varepsilon, a+\varepsilon)$, *cu derivata dată de*

$$\varphi'(x) = -\frac{\partial F}{\partial x}(x, \varphi(x)) / \frac{\partial F}{\partial y}(x, \varphi(x)), \quad x \in (a-\varepsilon, a+\varepsilon)$$

(P4) φ *este unica funcţie de la* $[a-\varepsilon, a+\varepsilon]$ *la* R *care să verifice simultan condiţiile date.*

Demonstraţie. Fie $\alpha > 0$ un număr cu proprietatea

$$[a-\alpha, a+\alpha] \times [b-\alpha, b+\alpha] \subset D.$$

(Asemenea numere există, deoarece D este o mulţime deschisă.) Fie $G : D \longrightarrow R$ definită astfel

$$G(x,y) = y - \frac{1}{\theta} F(x,y), \ (x,y) \in D, \ \text{unde } \theta = \frac{\partial F}{\partial y}(a,b).$$

Avem, evident (cu notaţia făcută)

$$\frac{\partial G}{\partial y} = 1 - \frac{1}{\theta}\frac{\partial F}{\partial y} \quad (\text{deci } \frac{\partial G}{\partial y}(a,b) = 1 - \frac{1}{\theta} \cdot \theta = 0).$$

Iar aceasta, combinată cu (C1) ne arată că , pentru λ fixat din $(0,1)$ există un η în $(0, \alpha)$ aşa încât

(R1) $\left| \frac{\partial G}{\partial y}(x,y) \right| \leq \lambda, \ x \in [a-\eta, a+\eta], \ y \in [b-\eta, b+\eta].$

Să notăm că, din această evaluare se ajunge uşor (prin Teorema de medie Lagrange) la condiţia Lipschitz

(R2) $|G(x,y_1) - G(x,y_2)| \leq \lambda |y_1 - y_2|$, oricare ar fi
$x \in [a-\eta, a+\eta], y_1, y_2 \in [b-\eta, b+\eta]$.

Pe de altă parte,

$$G(a,b) = b - \frac{1}{\theta} F(a,b) = b;$$

şi, deci, se poate indica un număr ε în $(0, \eta)$ aşa încât , cu notaţia $K = [a-\varepsilon, a+\varepsilon]$, să avem

(R3) $|G(x,b) - b| \leq (1-\lambda)\eta, \ \forall x \in K.$

Treecem acum la partea efectivă a demonstraţiei noastre. Să punem $M = [b-\eta, b+\eta]$ şi să definim şirul de funcţii

$$\varphi_0 : K \longrightarrow M, \varphi_1 : K \longrightarrow M, ..., \varphi_n : K \longrightarrow M, ... ,$$

prin formulele de recurenţă (în orice $x \in K$)

(D2) $\varphi_0(x) = b,\ \varphi_{n+1}(x) = G(x, \varphi_n(x)),\quad n = 0, 1, ...$

Vom arăta, prin inducţie după n, că

(P5) funcţia φ_n este bine definită, pentru toţi n.

(P5) $|\varphi_{n+1}(x) - \varphi_n(x)| \leq \lambda^n \eta,\ x \in K, n = 0, 1, ...$

Vom folosi în acest sens relaţiile (R2) şi (R3). De pildă, (P5) are loc (cu $n = 1$) din (R3). Deci, φ_2 se poate construi ca funcţie de la K la M; şi (P6) (cu $n = 2$) are loc, deoarece

$$|\varphi_2(x) - \varphi_1(x)| = |G(x, \varphi_1(x)) - G(x, \varphi_0(x))| \leq \lambda\eta,\ \forall x \in K.$$

Iar de aici, combinând cu (R3),

$$|\varphi_2(x) - b| \leq |\varphi_2(x) - \varphi_1(x)| + |\varphi_1(x) - b| \leq \lambda\eta + (1-\lambda)\eta = \eta,\ x \in K;$$

adică, (P5) (cu $n = 2$) are loc, etc. În baza criteriului de comparaţie Weierstrass (Teorema 4.4.7) rezultă , din (P6), că seria de funcţii $\sum\limits_n (\varphi_{n+1} - \varphi_n)$ este uniform convergentă pe K. Echivalent, asta spune că şirul de funcţii (φ_n) converge uniform (pe intervalul K) înspre o funcţie $\varphi : K \longrightarrow M$ care satisface

φ este continuă pe K (deci,(P1) are loc).

Mai mult,este clar că $\varphi(a) = b$. Pe de altă parte, în baza condiţiei Lipschitz (R2),

$$|G(x, \varphi_n(x)) - G(x, \varphi(x))| \leq \lambda|\varphi_n(x) - \varphi(x)|,\ x \in K,\ n = 1, 2, ...$$

Aceasta arată că

$$\lim_{n\to\infty} G(x, \varphi_n(x)) = G(x, \varphi(x)),\quad \text{uniform faţă de } x \in K.$$

Trecând deci la limită în relaţia de recurenţă, găsim imediat

$$\varphi(x) = \lim_{n\to\infty} \varphi_{n+1}(x) = \lim_{n\to\infty} G(x, \varphi_n(x)) = G(x, \varphi(x)),\ \forall x \in K,$$

sau, echivalent,

$$F(x, \varphi(x)) = 0, \ \text{ oricare ar fi } \ x \in K.$$

Cu alte cuvinte, φ satisface concluzia (P2) din enunţ.
Să fixăm acum un $u \in K$ şi să scriem formula lui Lagrange pentru funcţia F, în raport cu punctele $(u, \varphi(u)), (x, \varphi(x))$, unde $x \in K$ este arbitrar fixat. Avem

$$0 = F(x, \varphi(x)) - F(u, \varphi(u)) = (x-u)\frac{\partial F}{\partial x}(\tau, \sigma) + (\varphi(x) - \varphi(u))\frac{\partial F}{\partial y}(\tau, \sigma)$$

sau, echivalent,

$$\frac{\varphi(x) - \varphi(u)}{x-u} = -\frac{\partial F}{\partial x}(\tau, \sigma)/\frac{\partial F}{\partial y}(\tau, \sigma),$$

unde (τ, σ) este un punct intermediar pe segmentul din plan care uneşte punctele $(u, \varphi(u))$ şi $(x, \varphi(x))$. Trecând la limită pentru $x \to u$ ajungem la formula (de derivare din (P3) (căci, evident, $(\tau, \sigma) \to (u, \varphi(u))$); deci (P3) însăşi are loc.
Pentru ultima parte, să presupunem că ar mai exista o funcţie $\psi : [a-\varepsilon, a+\varepsilon] \longrightarrow R$ cu proprietăţile indicate. Avem cu necesitate

(R4) $\varphi(t) = \psi(t)$, în orice $t \in [a-\varepsilon, a+\varepsilon]$ cu $\psi(t) \in [b-\eta, b+\eta]$;

căci, din cele dinainte

$$\begin{cases} |\varphi(x) - \psi(x)| = |G(x, \varphi(x)) - G(x, \psi(x))| \leq \lambda|\varphi(x) - \psi(x)| \\ \Longrightarrow \quad \varphi(x) = \psi(x). \end{cases}$$

Mai departe,să punem , de pildă,

$$a^* = \sup\{t \in [a, a+\varepsilon]; \varphi(s) = \psi(s), \forall s \in [a, t]\},\ b^* = \varphi(a^*)(= \psi(a^*)).$$

Evident,

$$\frac{\partial F}{\partial y}(a^*, b^*) \neq 0, \text{ (în baza lui (C2).}$$

Deci, enunţul anterior se poate aplica datelor noastre. Ajungem la concluzia că pe un anume dreptunghi $D(a^*, b^*) = [a^* - \varepsilon^*, a^* + \varepsilon^*] \times [b^* - \eta^*, b^* + \eta^*]$ există o funcţie $\varphi^* : [a^* - \varepsilon^*, a^* + \varepsilon^*] \longrightarrow [b^* - \eta^*, b^* + \eta^*]$ care să verifice proprietăţile (P1)–(P3) (cu (a^*, b^*) în loc de (a, b)). În plus,aceasta verifică şi (R4) (în noile condiţii);adică, oricare ar fi funcţia $\psi^* : [a^* - \varepsilon^*, a^* + \varepsilon^*] \longrightarrow R$ cu aceleaşi proprietăţi,

(R5) $\varphi^*(t) = \psi^*(t)$, în orice $t \in [a^* - \varepsilon^*, a^* + \varepsilon^*]$ cu $\psi(t) \in [b^* - \eta^*, b^* + \eta^*]$.

Dar atunci (cu alegerea funcţiilor φ, ψ),există $\omega > 0$ cu

$$\varphi^*(t) = \varphi(t) = \psi(t), \text{ pentru orice } t \in [a^*, a^* + \omega],$$

în contradicţie cu alegerea punctului a^*.Deci,(P4) are loc. Teorema este astfel dovedită. ■

Rezultatul obţinut are caracter local, deoarece toate raţionamentele se derulează într-o vecinătate a punctului de plecare (a, b). Se poate ajunge însă şi la o soluţie globală a problemei, prin prelungiri succesive. De exemplu, să punem $a_1 = a + \varepsilon, b_1 = \varphi(a + \varepsilon)$. Cum $\frac{\partial F}{\partial y}(a_1, b_1) \neq 0$, se poate aplica teorema dinainte pentru acest punct. Ajungem astfel la un dreptunghi

$$D(a_1, b_1) = [a_1 - \varepsilon_1, a_1 + \varepsilon_1] \times [b_1 - \eta_1, b_1 + \eta_1] \subset D$$

şi la o funcţie $\varphi_1 : [a_1 - \varepsilon_1, a_1 + \varepsilon_1] \longrightarrow [b_1 - \eta_1, b_1 + \eta_1]$ care să satisfacă condiţiile (P1)–(P4) date acolo. Funcţia $\psi_1 : [a - \varepsilon, a + \varepsilon + \varepsilon_1] \longrightarrow R$ dată de

$$\psi_1(x) = \begin{cases} \varphi(x), & a - \varepsilon \leq x \leq a + \varepsilon \\ \varphi_1(x), & a + \varepsilon \leq x \leq a + \varepsilon + \varepsilon_1, \end{cases}$$

apare ca prelungire a funcţiei φ care îndeplineşte, de asemenea, condiţiile teoremei anterioare. Aceasta se mai poate în continuare prelungi în $a_2 = a + \varepsilon + \varepsilon_1, b_2 = \varphi_1(a + \varepsilon + \varepsilon_1)$ (dacă $\frac{\partial F}{\partial y}(a_2, b_2) \neq 0$), şi aşa mai departe. (Analog se petrec lucrurile cu prelungirile înspre stânga). Ajungem astfel în final la un interval deschis $I = I(a, b)$ al axei reale şi o funcţie $\psi : I \longrightarrow R$ cu proprietăţile indicate în fapt de concluziile (P1)–(P4) din rezultatul precedent.

Exemplu. Să se studieze funcţia $y = \varphi(x)$ definită de ecuaţia

$$x^3 + y^3 - 3x^2y - 3 = 0, \quad y(1) = 2.$$

Avem, evident,

$$\frac{\partial F}{\partial x} = 3x^2 - 6xy,\ \frac{\partial F}{\partial y} = 3y^2 - 3x^2\ \left(\Longrightarrow \frac{\partial F}{\partial y}(1,2) = 9 > 0\right).$$

Rezultatul anterior poate fi deci aplicat. Se obţine o funcţie $\varphi : [1-\varepsilon, 1+\varepsilon] \longrightarrow R$, definită implicit de ecuaţia anterioară şi care, în plus, verifică $\varphi(1) = 2$. Derivata acestei funcţiei pe interval va fi dată de

$$\varphi'(x) = -\frac{\partial F}{\partial x}/\frac{\partial F}{\partial y} = -\frac{x^2 - 2xy}{y^2 - x^2}.$$

Ţinând cont de regiunea în care se găseşte punctul iniţial $(1,2)$, se poate spune că $\varphi'(1) > 0$. Deci, funcţia φ creşte în vecinătatea punctului 1. Se poate prelungi domeniul de definiţie al acestei funcţii după modelul alăturat, la un interval deschis al axei reale. Pe acest nou interval, derivata funcţiei implicite este dată de formula dinainte. Să notăm că, de aici, putem deduce şi formula pentru derivata a doua din funcţia implicită.

5.3.2 Funcţii implicite de mai multe variabile

Fie m, p două numere naturale $(m > p)$ şi D un domeniu al spaţiului R^{m+p}. Să ne dăm un număr de funcţii continue $(F_1 : D \longrightarrow R, ..., F_p : D \longrightarrow R)$ şi să considerăm sistemul de ecuaţii

$$\text{(E)} \qquad F_1(x_1, ..., x_m; y_1, ..., y_p) = 0, ..., F_p(x_1, ..., x_m; y_1, ..., y_p) = 0.$$

Zicem că sistemul de funcţii continue $(\varphi_1 : A \longrightarrow R, ..., \varphi_p : A \longrightarrow R)$, unde A este un continuum în R^m, apare ca o *soluţie de tip implicit* a sistemului de ecuaţii (E),dacă

$$\text{(D1)} \quad \begin{cases} (x_1, ..., x_m; \varphi_1(x_1, ..., x_m), ..., \varphi_p(x_1, ..., x_m)) \in D \ \text{ şi} \\ F_i(x_1, ..., x_m; \varphi_1(x_1, ..., x_m), ..., \varphi_p(x_1, ..., x_m)) = D, \\ \text{pentru orice } (x_1, ..., x_m) \in A,\ 1 \le i \le p. \end{cases}$$

În general, pe un acelaşi continuum A din R^m, pot exista mai multe soluţii pentru asemenea sisteme de ecuaţii (sau nici una). De aceea, pentru a le individualiza, este necesar să mai asociem sistemului nostru de ecuaţii încă un grup de condiţii (numite *iniţiale*)

$$\text{(I)} \begin{cases} \varphi_1(a_1,...,a_m) = b_1, ..., \varphi_p(a_1,...,a_m) = b_p, & \text{unde} \\ F_i(a_1,...,a_m; b_1,...,b_p) = 0, \ 1 \le i \le p. \end{cases}$$

Să notăm că, pentru $m = 1, p = 1$, dezvoltările acestea coincid cu cele precedente. Aceasta, de altfel, ne sugerează modul de rezolvare a problemei propuse. Mai precis, fie din nou domeniul D din R^{m+p} şi funcţiile continue $F_1 : D \longrightarrow R, ..., F_p : D \longrightarrow R$. Acceptăm ipoteza

(C1) F_i este de clasă C^1 pe D, pentru $i = 1, ..., p$.

Avem atunci

Teorema 5.3.2 *Fie $a = (a_1,...,a_m)$,$b = (b_1,...,b_p)$ un sistem de puncte fixate din R^m, respectiv R^p,cu*

$$\text{(C2)} \ (a,b) \in D, \ F_i(a;b) = 0, \ 1 \le i \le p, \quad \det\left(\frac{\partial F_i}{\partial y_j}(a;b)\right) \ne 0.$$

Există atunci un continuum $D(a;b) = S[a,\varepsilon] \times S[b,\eta]$ din D şi un sistem de funcţii $\varphi = (\varphi_1,...,\varphi_p)$; de la $S[a,\varepsilon]$ la $S[b,\eta]$ cu
(P1) *$\varphi_1, ..., \varphi_p$ sunt continue pe $S[a,\varepsilon]$*
(P2) *$\varphi = (\varphi_1,...,\varphi_p)$ este o soluţie implicită a problemei* (E)+(I)
(P3) *$\varphi_1,...,\varphi_p$ sunt de clasă C^1 pe $S(a,\varepsilon)$, cu derivatele parţiale în variabilele $(x_1,...,x_m)$ date de sistemul*

$$\text{(SD)} \qquad \sum_{j=1}^{p} \frac{\partial F_i}{\partial y_j}(x,\varphi(x))\frac{\partial \varphi_j}{\partial x_k}(x) = -\frac{\partial F_i}{\partial x_k}(x,\varphi(x)), \ 1 \le i \le p.$$

(P4) *$\varphi = (\varphi_1,...,\varphi_p)$ este singura funcţie de la $S[a,\varepsilon]$ la R^p care să îndeplinească simultan condiţiile date.*

Demonstraţie. (Schiţă) Pentru a nu complica prea mult notaţiile, vom lua în discuţie cazul $m = 1, p = 2$. Fie $\alpha > 0$, un număr cu

$$[a-\alpha, a+\alpha] \times [b_1-\alpha, b_1+\alpha] \times [b_2-\alpha, b_2+\alpha] \subseteq D.$$

(Asemenea numere există, deoarece D este o mulţime deschisă.) Să notăm apoi, pentru simplitate,

$$\left[\frac{D(F_1,F_2)}{D(y_1,y_2)}(a;b)\right]^{-1} = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}.$$

Definim perechea de funcţii $G_1 : D \longrightarrow R$, $G_2 : D \longrightarrow R$ prin

$$\begin{cases} G_1(x; y_1, y_2) = y_1 - \gamma_{11} F_1(x; y_1, y_2) - \gamma_{12} F_2(x; y_1, y_2) \\ G_2(x; y_1, y_2) = y_2 - \gamma_{21} F_1(x; y_1, y_2) - \gamma_{22} F_2(x; y_1, y_2) \end{cases}$$

Avem, evident, cu alegerea matricei noastre

$$\frac{\partial G_1}{\partial y_1}(a; b) = 0, \frac{\partial G_1}{\partial y_2}(a; b) = 0, \frac{\partial G_2}{\partial y_1}(a; b) = 0, \frac{\partial G_2}{\partial y_2}(a; b) = 0.$$

Iar aceasta, combinată cu ipoteza (C1), ne asigură că, pentru λ fixat în $(0, 1)$, există un η în $(0, \alpha)$ aşa încât

$$\text{(R1)} \quad \begin{cases} \left|\dfrac{\partial G_i}{\partial y_j}(x; y_1, y_2)\right| \leq \dfrac{\lambda}{2}, \ i, j \in \{1, 2\}, \ \text{oricare ar fi} \\ (x, y_1, y_2) \in [a - \eta, a + \eta] \times [b_1 - \eta, b_1 + \eta] \times [b_2 - \eta, b_2 + \eta]. \end{cases}$$

Să notăm că, din această evaluare se ajunge la condiţia Lipschitz

$$\text{(R2)} \quad \begin{cases} |G_i(x, y_1, y_2) - G_i(x, y_1', y_2')| \leq \frac{\lambda}{2}(|y_1 - y_1'| + |y_2 - y_2'|), \\ i \in \{1, 2\}, \ \text{pentru orice } x \in [a - \eta, a + \eta] \ \text{şi orice} \\ y_1, y_1' \in [b_1 - \eta, b_1 + \eta], y_2, y_2' \in [b_2 - \eta, b_2 + \eta], \end{cases}$$

Pe de altă parte,

$$G_1(a; b_1, b_2) = b_1, \qquad G_2(a, b_1, b_2) = b_2.$$

Se poate atunci indica un număr ε în (D, η) aşa încât cu notaţia $K = [a - \varepsilon, a + \varepsilon]$, să avem

$$\text{(R3)} \ |G_1(x, b_1, b_2) - b_1| \leq (1 - \lambda)\eta, \ |G_2(x, b_1, b_2) - b_2| \leq (1 - \lambda)\eta, \ x \in K.$$

Trecem la partea efectivă a demonstraţiei. Să punem
$M = [b_1 - \eta, b_1 + \eta] \times [b_2 - \eta, b_2 + \eta]$ şi să definim şirul de funcţii

$$(\varphi_0, \psi_0) : K \longrightarrow M, \quad (\varphi_1, \psi_1) : K \longrightarrow M, ...,$$

prin formulele de recurenţă (pentru $x \in K$)

$$\text{(D2)} \quad \begin{cases} \varphi_0(x) = b_1, \psi_0(x) = b_2, \\ \varphi_{n+1}(x) = G_1(x, \varphi_n(x), \psi_n(x)), \psi_{n+1}(x) = G_2(x, \varphi_n(x), \psi_n(x)), \ n \geq 0. \end{cases}$$

Se arată şi aici (prin inducţie faţă de n) că

(P4) şirul (φ_n, ψ_n) este bine definit, pentru toţi n.

(P5) $|\varphi_{n+1}(x) - \varphi_n(x)| \leq \lambda^n \eta, |\psi_{n+1}(x) - \psi_n(x)| \leq \lambda^n \eta,\ x \in K, n \geq 0$.

Acestea ne spun că şirul (φ_n, ψ_n) converge uniform pe mulţimea K, înspre o funcţie $(\varphi, \psi) : K \longrightarrow M$. Avem (din proprietăţile convergenţei uniforme)

(φ, ψ) este continuă pe K (deci,verifică (P1)).

Apoi,este clar că $\varphi(a) = b_1, \psi(a) = b_2$. Pe de altă parte, trecând la limită în relaţia de recurenţă,

$$\varphi(x) = G_1(x, \varphi(x), \psi(x)), \psi(x) = G_2(x, \varphi(x), \psi(x)),\ x \in K.$$

Dar atunci, prin definiţia celor două funcţii,

$$\begin{cases} \gamma_{11} F_1(x, \varphi(x), \psi(x)) + \gamma_{12} F_2(x, \varphi(x), \psi(x)) = 0, & x \in K \\ \gamma_{21} F_1(x, \varphi(x), \psi(x)) + \gamma_{22} F_2(x, \varphi(x), \psi(x)) = 0, & x \in K. \end{cases}$$

Şi, cum matricea (γ_{ij}) este nesingulară, deducem

$$F_1(x, \varphi(x), \psi(x)) = 0,\ F_2(x, \varphi(x), \psi(x)) = 0,\ x \in K.$$

Aceasta ne spune că (φ, ψ) verifică (P2). Celelalte concluzii se deduc acum imediat după modelul anterior. ∎

Exemplu. Funcţiile $u = \varphi(x_1, x_2, x_3),\ v = \psi(x_1, x_2, x_3)$ sunt definite implicit de sistemul funcţional

$$x_1 + x_2 + x_3 - u - v = 0,\ x_1 u^2 + x_2 uv + x_3 v^2 - 1 = 0.$$

Să se calculeze $\dfrac{\partial u}{\partial x_1}, \dfrac{\partial v}{\partial x_1}$ în vecinătatea punctului $(0, 0, 1; 0, 1)$ faţă de care se explicitează sistemul în chestiune.
Pornim de la matricea jacobiană

$$\frac{D(F_1, F_2)}{D(u, v)} = \begin{pmatrix} \dfrac{\partial F_1}{\partial u} & \dfrac{\partial F_1}{\partial v} \\ \dfrac{\partial F_2}{\partial u} & \dfrac{\partial F_2}{\partial v} \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 2x_1 u + x_2 v & x_2 u + 2x_3 v \end{pmatrix}.$$

Aceasta este nesingulară în punctul nostru. Putem deci explicita sistemul anterior pentru a obţine funcţiile indicate. Derivatele lor se obţin prin derivarea faţă de x_1 a ecuaţiilor sistemului

$$1 - \frac{\partial u}{\partial x_1} - \frac{\partial v}{\partial x_1} = 0,\ u^2 + 2x_1 u\frac{\partial u}{\partial x_1} + x_2\left(v\frac{\partial u}{\partial x_1} + u\frac{\partial v}{\partial x_1}\right) + 2x_3 v\frac{\partial v}{\partial x_1} = 0.$$

Matricea acestui sistem în necunoscutele $\left(\frac{\partial u}{\partial x_1}, \frac{\partial v}{\partial x_1}\right)$ este nesingulară, cum s-a văzut mai sus. Prin rezolvare găsim

$$\begin{cases} \dfrac{\partial u}{\partial x_1} = -(u^2 + x_2 u + 2x_3 v)/(2x_1 u - 2x_3 v + x_2 v - x_2 u) \\ \dfrac{\partial v}{\partial x_1} = (u^2 + 2x_1 u + x_2 v)/(2x_1 u - 2x_3 v + x_2 v - x_2 u). \end{cases}$$

5.3.3 Curbe şi suprafeţe în spaţii finit dimensionale

Fie m număr natural dat şi q un alt număr natural cu $q < m$. Dat un domeniu E din spaţiul R^q şi funcţiile (de clasă C^1), $\varphi_1 : E \longrightarrow R, ...,$ $\varphi_m : E \longrightarrow R$ cu

(C1) rang $\left(\dfrac{D(\varphi_1, ..., \varphi_m)}{D(t_1, ..., t_q)}\right) = q$, în orice $(t_1, ..., t_q) \in E$,

să numim *suprafaţă q–dimensională* în R^m generată de acestea, mulţimea tuturor punctelor din R^m de forma

(D1) $x_1 = \varphi_1(t_1, ..., t_q), ..., x_m = \varphi_m(t_1, ..., t_q), ..., (t_1, ..., t_q) \in E.$

Relaţiile care o definesc se vor chema *ecuaţiile parametrice ale suprafeţei* considerate. În particular, pentru $q = 1$, suprafaţa în chestiune se va numi *curbă* în spaţiul R^m.

În baza rezultatelor anterioare, mai există o modalitate "implicită" de introducere a unei suprafeţe. Mai exact, dat numărul natural m şi numărul natural $p < m$, fie D un domeniu în spaţiul R^m. Considerăm date funcţiile (de clasă C^1) $g_1 : D \longrightarrow R, ..., g_p : D \longrightarrow R$, cu

(C2) rang $\left(\dfrac{D(g_1, ..., g_p)}{D(x_1, ..., x_m)}\right) = p$, în orice $(x_1, ..., x_m) \in D$.

Vom numi *suprafaţă* $(m-p)$–*dimensională* în R^m generată de acestea, mulţimea tuturor punctelor $x = (x_1, ..., x_m)$ din R^m, de forma

(D2) $g_1(x_1, ..., x_m) = 0, ..., g_p(x_1, ..., x_m) = 0.$

Relaţiile care o definesc le vom numi *ecuaţiile implicite ale suprafeţei* considerate. În particular, pentru $p = m - 1$, suprafaţa în chestiune se va numi o *curbă* în spaţiul considerat R^m.

Se poate trece de la scrierea implicită a unei suprafeţe la scrierea parametrică a acesteia prin intermediul Teoremei Funcţiilor Implicite. Astfel, dacă $c = (c_1, ..., c_m)$ este un punct situat pe suprafaţa definită implicit de (D2), cu, de pildă,

$$\det\left(\frac{D(g_1, ..., g_p)}{D(x_1, ..., x_p)}(c_1, ..., c_m)\right) \neq 0$$

va exista, cu notaţia $b = (c_{p+1}, ..., c_m)$, un domeniu $E = S(b, \varepsilon)$ în spaţiul R^{m-p} şi un sistem de funcţii

$$\varphi_1 : E \longrightarrow R, ..., \varphi_p : E \longrightarrow R$$

care sunt definite implicit de sistemul (D2); adică

$$x_1 = \varphi_1(t_1, ..., t_{m-p}), ..., x_p = \varphi_p(t_1, ..., t_{m-p}), x_{p+1} = t_1, ..., x_m = t_{m-p}.$$

Am ajuns astfel la o reprezentare parametrică a suprafeţei în cauză, pe o vecinătate a punctului $c = (c_1, ..., c_m)$ dat iniţial. Reciproca este, de asemenea, valabilă; adică, orice suprafaţă dată parametric poate fi, de asemenea, reprezentată şi implicit.

Exemplu. Fie dată suprafaţa sferică (bi–dimensională) din R^3 :

$$x^2 + y^2 + z^2 - a^2 = 0, \quad a > 0 \ \text{ număr fixat.}$$

Avem în orice punct al acestei suprafeţe

$$\text{rang}\begin{pmatrix} \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} & \dfrac{\partial g}{\partial z} \end{pmatrix} = \text{rang}\,(2x\ \ 2y\ \ 2z) = 1.$$

Deci, aceasta se poate reprezenta parametric, în baza celor spuse anterior. Există mai multe moduri de a face acest lucru. Unul din ele constă în folosirea coordonatelor sferice. Se ajunge la reprezentarea

$$x = a\cos t\cos s,\ y = a\cos t\sin s,\ z = a\sin t,\ -\pi < t < \pi,\ 0 < s < 2\pi.$$

Procedeul se poate extinde la un număr arbitrar de dimensiuni.

În legătură cu noţiunile introduse, următoarea problemă se poate dovedi utilă în multe situaţii. Fie D un domeniu din spaţiul R^m şi $g_1 : D \longrightarrow R, ..., g_p : D \longrightarrow R$ un grup de funcţii (de clasă C^1) ce satisfac (C2). Vom asocia acestora suprafaţa (dată implicit)

$$(\Sigma) \qquad g_1(x_1, ..., x_m) = 0, ..., g_p(x_1, ..., x_m) = 0.$$

Fie $a = (a_1, ..., a_m)$ un punct al acestei suprafeţe. Zicem că vectorul $h = (h_1, ..., h_m)$ din R^m este *tangent* suprafeţei (Σ) în punctul a dacă

$$\text{(D3)} \quad \frac{\partial g_i}{\partial x_1}(a)h_1 + \cdots + \frac{\partial g_i}{\partial x_m}(a)h_m = 0, \ 1 \le i \le p;$$

sau, echivalent (cu definiţia gradientului)

$$\text{(D4)} \quad (\nabla g_i(a)) \cdot h^\top = 0, \ 1 \le i \le p.$$

Suntem în măsură acum să formulăm

Teorema 5.3.3 *Fie $a = (a_1, ..., a_m)$ un punct al suprafeţei (Σ) şi $h = (h_1, ..., h_m)$ un vector din R^m tangent în a la suprafaţa (Σ). Există $\delta > 0$ suficient de mic şi o curbă de clasă C^1*

$$(\Gamma) \qquad x_i = \varphi_i(t), \ -\delta < t < \delta, \ 1 \le i \le m,$$

cu proprietăţile
(P1) $\varphi_i(0) = a_i, \ \varphi_i'(0) = h_i, \ 1 \le i \le m$; *adică, curba trece prin punctul a şi are direcţia h.*
(P2) $g_i(\varphi_1(t), ..., \varphi_m(t)) = 0, \ -\delta < t < \delta, \ i = 1, ..., p$; *adică, curba în chestiune este situată pe suprafaţa (Σ).*

Demonstraţie. Să definim domeniul din R^{p+1}

$$E = \{(t; y_1, ..., y_p); \ a + th + y_1\nabla g_1(a) + \cdots + y_p\nabla g_p(a) \in D\}.$$

Desigur, originea lui R^{p+1} este un punct al lui E. Introducem acum funcţiile (de la E la R)

$$G_i(t; y_1, ..., y_p) = g_i(a + th + y_1\nabla g_1(a) + \cdots + y_p\nabla g_p(a)),$$
$$(t; y_1, ..., y_p) \in E, \quad 1 \le i \le p.$$

Avem, din observaţia precedentă

$$G_i(0;0,...,0) = g_i(a) = 0,\ 1 \le i \le p.$$

Pe de altă parte,în baza formulei de derivare a funcţiilor compuse (vezi Secţiunea 4.3.2) avem, pentru $1 \le i, j \le p$,

$$\frac{\partial G_i}{\partial y_j}(t; y_1, ..., y_p) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial x_k}(a + th + y_1 \nabla g_1(a) + \cdots + y_p \nabla g_p(a)) \frac{\partial g_j}{\partial x_k}(a).$$

Şi, deci, cu $t = 0,\ y_1 = \cdots = y_p = 0$,

$$\frac{\partial G_i}{\partial y_j}(0;0,0) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial x_k}(a) \frac{\partial g_j}{\partial x_k}(a),\ 1 \le i, j \le p,$$

sau, matricial,

$$\frac{D(G_1, ..., G_p)}{D(y_1, ..., y_p)}(0;0,0) = \left[\frac{D(g_1, ..., g_p)}{D(x_1, ..., x_m)}(a)\right] \left[\frac{D(g_1, ..., g_p)}{D(x_1, ..., x_m)}(a)\right]^{\top}.$$

Pe baza unor observaţii asupra matricilor de tip pozitiv (Secţiunea 1.4.5), deducem de aici că matricea din stânga acestei relaţii este în mod necesar nesingulară. Avem îndeplinite astfel toate condiţiile pentru aplicarea Teoremei Funcţiilor Implicite (cazul general). În baza acestuia va exista un $\delta > 0$ şi un sistem de funcţiii (toate de clasă C^1 pe $(-\delta, \delta)$)

$$y_1 = \psi_1(t), .., y_p = \psi_p(t),\ -\delta < t < \delta,$$

care să verifice
(P3) $\psi_1(0) = 0, ..., \psi_p(0) = 0$
(P4) $(\psi_1, ..., \psi_p)$ este definită implicit de $(G_1, ..., G_p)$ în sensul

$$G_i(t, \psi_1(t), ..., \psi_p(t)) = 0,\ -\delta < t < \delta,\ \ 1 \le i \le p.$$

Să notăm că, în acest caz, derivatele în origine ale funcţiilor $\psi_1, ..., \psi_p$ sunt date de sistemul

$$\frac{\partial G_i}{\partial t}(0) + \sum_{j=1}^{p} \frac{\partial G_i}{\partial y_j}(a)\psi_j'(0) = 0,\ 1 \le i \le p.$$

Aceasta, combinată cu

$$\frac{\partial G_i}{\partial t}(0) = \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j}(a) h_j = 0,\ 1 \le i \le p,$$

ne conduce imediat la relaţiile

$$\psi_1'(0) = \cdots = \psi_p'(0) = 0.$$

Să definim acum funcţiile de la $(-\delta, \delta)$ la R

$$\varphi_i(t) = a_i + th_i + \psi_1(t)\frac{\partial g_1}{\partial x_i}(a) + \cdots + \psi_p(t)\frac{\partial g_p}{\partial x_i}(a), -\delta < t < \delta, 1 \le i \le m.$$

Nu este greu de văzut, utilizând şi relaţiile găsite, că sistemul de funcţii $(\varphi_1, ..., \varphi_m)$ este cel căutat. Teorema este dovedită. ■

Exemplu. Fie dată din nou suprafaţa sferică

$$x^2 + y^2 + z^2 - a^2 = 0, \quad a > 0 \ \text{ fixat.}$$

Să considerăm un punct $M = \left(\frac{a\sqrt{2}}{2}, 0, \frac{a\sqrt{2}}{2}\right)$ pe această sferă şi un vector $h = \frac{a\sqrt{2}}{2}(0, 1, 0)$, tangent la suprafaţa sferică în punctul M (după cum se vede uşor). Conform teoremei precedente, există o curbă pe această suprafaţă

$$x = \varphi_1(t),\ y = \varphi_2(t),\ z = \varphi_3(t)$$

care să treacă prin punctul M şi care să aibă drept tangentă în acest punct vectorul h. De pildă, o asemenea curbă este o porţiune dintr-un cerc situat pe suprafaţa sferică, de ecuaţii parametrice

$$x = a\frac{\sqrt{2}}{2}\cos t,\ y = a\frac{\sqrt{2}}{2}\sin t,\ z = a\frac{\sqrt{2}}{2},\ \frac{-\pi}{4} < t < \frac{\pi}{4}.$$

Probleme propuse la §5.3

1. Fie funcţia $y = f(x)$, definită implicit prin relaţia $x^2 + y^2 + 2axy = 0$, unde $a > 1$ este un parametru. Să se calculeze primele două derivate pentru

această funcţie şi să se arate că $f''(x) = 0$. Să se încerce şi o rezolvare directă a problemei, prin explicitare.

2. Funcţiile $y = f(x)$, $z = g(x)$ sunt definite implicit prin intermediul sistemului $xyz = a$, $x + y + z = b$. Să se calculeze derivatele $f'(x)$ şi $g'(x)$. (Se va încerca şi o rezolvare directă, prin explicitare.)

3. Funcţia $z = f(x, y)$ este definită implicit prin relaţia $x \cos y + y \cos z + z \cos x - 1 = 0$, $f(1, 0) = 0$. Să se calculeze mărimile $\frac{\partial f}{\partial x}(1, 0)$, $\frac{\partial f}{\partial y}(1, 0)$.

4. Fie funcţia $z = f(x, y)$, unde $y = \varphi(x)$ este definită implicit de ecuaţia $x^2 + y^2 - 1 = 0$. Să se calculeze derivata funcţiei compuse $z = f(x, \varphi(x))$.

5. Să se arate că funcţia $z = f(x, y)$ dată de ecuaţia $F(x - ay, y - bz) = 0$, satisface o condiţie de forma $a\frac{\partial f}{\partial x} + b\frac{\partial f}{\partial y} = 0$.

6. Fie suprafaţa sferică (din R^4), $x^2 + y^2 + z^2 + t^2 - 1 = 0$. Să se arate că se poate indica o reprezentare parametrică a ei, de forma

$$x = \cos\alpha\cos\beta\cos\gamma, y = \cos\alpha\cos\beta\sin\gamma, z = \cos\alpha\sin\beta, t = \sin\alpha$$

unde α, β, γ sunt parametri reali.

7. Fie suprafaţa definită implicit de ecuaţia $x^3 + y^3 + z^3 - 2xyz = 0$ (din spaţiul R^3).Să se găsească o curbă $x = \alpha(t), y = \beta(t), z = \gamma(t)$ pe această suprafaţă, care să îndeplinească condiţiile

$$\alpha(0) = \beta(0) = \gamma(0) = 0,\ \alpha'(0) = \beta'(0) = 1, \gamma'(0) = \sqrt[3]{2}.$$

(Se vor lua funcţii liniare în variabila t).

5.4 Extreme condiţionate ale funcţiilor

5.4.1 Cazul restricţiilor de tip egalitate

Fie m număr natural dat şi D un domeniu convex din R^m. Să dăm sistemul de funcţii

$$g_1 : D \longrightarrow R, ..., g_p : D \longrightarrow R \ \ (\text{unde } p < m).$$

Desigur, aceasta echivalează cu a da funcţia $g : D \longrightarrow R^p$, prin

$$g(x_1, ..., x_m) = (g_1(x_1, ..., x_m), ..., g_p(x_1, ..., x_m)),\ (x_1, ..., x_m) \in D.$$

Conform unei convenţii anterioare, sistemul de ecuaţii

$$\text{(RE)} \qquad g_1(x_1, ..., x_m) = 0, ..., g_p(x_1, ..., x_m) = 0$$

defineşte o suprafaţă în spaţiul R^m, notată prin $D(g)$; vom numi aceasta *suprafaţa generată* de ecuaţiile (sau restricţiile) în cauză (RE).

Fie în continuare $f : D \longrightarrow R$ o funcţie dată. Mai luăm un punct $a = (a_1, ..., a_m)$ din D care satisface (RE) şi un domeniu convex E al lui D care conţine a. Zicem că punctul în cauză este *punct de minim* (respectiv, *maxim*) *global* al funcţiei f relativ la E, *condiţionat* de restricţiile (RE) dacă

$$\text{(D1)} \quad \begin{cases} f(x) \geq f(a) \text{ (respectiv } f(x) \leq f(a)\text{) pentru toţi } x \in E \\ \text{care satisfac restricţiile (RE).} \end{cases}$$

Când una din aceste proprietăţi are loc, vom zice că punctul a din D, care satisface (RE), este un *punct de extrem global* al funcţiei f relativ la E, *condiţionat* de restricţiile (RE). În particular, pentru $E = D$, vom spune că avem de-a face cu un *punct de extrem global* al funcţiei f, *condiţionat* de restricţiile (RE); sau, pe scurt, un *punct de extrem global condiţionat* al funcţiei considerate. Iar, când domeniul E (din D) care conţine a, este subînţeles în definiţia (D1), vom zice că punctul $a \in D$ ce satisface (RE) apare ca *punct de extrem local* al funcţiei f, *condiţionat* de restricţiile (RE); sau, pe scurt, un *punct de extrem local condiţionat* al funcţiei în cauză.

În legătură cu noţiunile introduse, putem formula următoarele probleme de bază:

(i) determinarea condiţiilor necesare de extrem condiţionat (adică, stabilirea de proprietăţi diferenţiale pe care le au punctele de extrem condiţionat);

(ii) determinarea condiţiilor suficiente de extrem condiţionat (adică, găsirea de condiţii suplimentare care să ne asigure o proprietate de extrem condiţionat).

După cum se vede, aceste probleme se formulează oarecum la fel cu analogele lor din teoria extremelor libere ale funcţiilor. Vom vedea însă că, prezenţa restricţiilor de egalitate (RE) măreşte în mod semnificativ gradul de dificultate al acestora.

În legătură cu prima problemă, următorul răspuns este de menţionat. Admitem ca ipoteză de bază

(C1) $f, g_1, ..., g_p$ sunt de clasă C^1 pe domeniul D.

De asemenea, relativ la sistemul de funcţii $g = (g_1, ..., g_p)$, vom accepta o ipoteză de regularitate de forma

$$\text{(C2)} \begin{cases} \text{vectorii gradient } \{\nabla g_1, ..., \nabla g_m\} \text{ sunt liniar independenţi} \\ \text{în orice punct } x = (x_1, ..., x_m) \text{ din } D. \end{cases}$$

Pentru fiecare sistem de numere $\lambda_1, ..., \lambda_p$, să definim o funcţie $L : D \longrightarrow R$ prin formula

$$\text{(D2)} \quad L(x) = f(x) + \sum_{i=1}^{p} \lambda_i g_i(x), \quad x = (x_1, ..., x_m) \in D.$$

Vom numi aceasta *funcţia Lagrange* asociată sistemului $\{f, g_1, ..., g_p\}$; iar $\lambda_1, ..., \lambda_p$, *multiplicatorii Lagrange* corespunzători. Fie acum $a = (a_1, ..., a_m)$ un punct din domeniul D care mai satisface restricţiile de egalitate (RE); sau (cu notaţiile noastre) un punct al suprafeţei $D(g)$.

Teorema 5.4.1 *Să presupunem că punctul $a \in D(g)$ este de extrem condiţionat (local sau global) pentru funcţia f. Există un sistem unic de multiplicatori Lagrange $\lambda_1, ..., \lambda_m$ astfel încât , notând prin $L : D \longrightarrow R$ funcţia Lagrange asociată, să avem*

$$\text{(R1)} \quad \nabla L(a) = 0 \quad \left(\textit{adică } \frac{\partial L}{\partial x_1}(a) = 0, ..., \frac{\partial L}{\partial x_m}(a) = 0\ \right).$$

Demonstraţie. Fie $h = (h_1, ..., h_m)$ un vector din R^m tangent la suprafaţa $D(g)$ în punctul a, în sensul

$$(\nabla g_i(a)))h^\top = 0, \quad 1 \leq i \leq p.$$

Există (vezi Teorema 5.3.3), un număr $\delta > 0$ (suficient de mic) şi o curbă de clasă C^1 (în R^m)

$$\varphi(t) = (\varphi_1(t), ..., \varphi_m(t)), \quad -\delta < t < \delta$$

cu proprietăţile

(P1) $\varphi(0) = a$, $\varphi'(0) = h$; adică, curba în chestiune trece prin punctul a şi are direcţia h, (P2) $\varphi(t) \in D(g)$, $-\delta < t < \delta$ (curba astfel construită este situată pe suprafaţa $D(g)$).

Presupunem acum, pentru fixarea cadrului, că punctul $a \in D(g)$ ar fi de minim global condiţionat. (Celelalte cazuri se vor trata absolut la fel.) Avem, deci, valabilă relaţia (prin construcţia anterioară)

$$f(\varphi(t)) - f(\varphi(0)) \geq 0, \ -\delta < t < \delta.$$

Aceasta ne spune că $0 \in (-\delta, \delta)$ apare ca punct de minim global al funcţiei compuse (de clasă C^1)

$$b(t) = f(\varphi(t)), \ -\delta < t < \delta.$$

Dar atunci, conform teoremei lui Fermat, $b'(0) = 0$. Să calculăm acum derivata funcţiei în cauză. Avem ,cu formula de derivare compusă (Secţiunea 4.3.2)

$$b'(t) = [\nabla f(\varphi(t))](\varphi'(t))^{\top}, \ -\delta < t < \delta.$$

Şi, deci (cu proprietăţile anterioare)

$$b'(0) = [\nabla f(\varphi(0))](\varphi'(0))^{\top} = [\nabla f(a)] \cdot h^{\top}.$$

Rezumând, avem implicaţia

$$[\nabla g_i(a)]h^{\top} = 0, \ 1 \leq i \leq p \Longrightarrow [\nabla f(a)]h^{\top} = 0.$$

Dar, în acest caz, conform unui rezultat anterior (vezi Secţiunea 1.5.2), există constantele $\mu_1, ..., \mu_p$ cu

$$\nabla f(a) = \mu_1 \nabla g_1(a) + \cdots + \mu_p \nabla g_p(a).$$

Nu avem decât să punem acum $\lambda_1 = -\mu_1, ..., \lambda_p = -\mu_p$, ca să încheiem argumentul. ■

Cu alte cuvinte, într-un punct de extrem condiţionat pentru funcţia f avem o proprietate de tip Fermat (R1); aceasta însă – spre deosebire de cazul extremelor libere – nu face să intervină doar funcţia în cauză (f), ci şi legăturile $(g_1, ..., g_p)$ prin intermediul funcţiei (de sinteză) Lagrange, L.

Prin definiţie, punctul $a = (a_1, ..., a_m)$ din D se va zice *punct staţionar* al funcţiei f, *condiţionat* de restricţiile (RE) – sau, pe scurt, *punct staţionar condiţionat* – dacă există un sistem de multiplicatori Lagrange $\lambda_1, ..., \lambda_p$ aşa încât , pentru funcţia Lagrange asociată $L : D \longrightarrow R$, să aibă loc relaţiile (R1). Am demonstrat deci implicaţia

(P3) a=punct de extrem condiţionat $\Longrightarrow$ a=punct staţionar condiţionat.

Ne interesează acum, în ce condiţii suplimentare putem inversa această implicaţie; acestea vor fi, desigur, tocmai condiţiile suficiente de extrem condiţionat pe care le căutăm. În esenţă, totul revine la studiul derivatelor parţiale de ordin doi. Anume, să admitem că (C1) se înlocuieşte cu ipoteza mai tare

(C3) $f, g_1, ..., g_p$ sunt de clasă C^2 pe D.

Să ne aducem aminte de o convenţie anterioară (Secţiunea 1.4.5). Fie $A = (\alpha_{ij})$ matrice (simetrică) din $\mathcal{M}(m)$ şi

$$\psi_A(h_1, ..., h_m) = \sum_{i,j} \alpha_{ij} h_i h_j,\ (h_1, ..., h_m) \in R^m$$

forma pătratică în m variabile asociată acesteia. Dată matricea $B = (\beta_{ij})$ din $\mathcal{M}(p, m)$ cu $\operatorname{rang}(B) = p$, vom spune că forma pătratică ψ_A este *de tip pozitiv (negativ) faţă de* B dacă ea are acest caracter pe spaţiul liniar $(m-p)$–dimensional al soluţiilor sistemului liniar omogen

(S$_0$) $$\beta_{i1} h_1 + \cdots + \beta_{im} h_m = 0,\ 1 \leq i \leq p.$$

Concret, aceasta revine, prin intermediul rezolvării sistemului omogen în chestiune

$$h_i = \nu_{i1} k_1 + \cdots + \nu_{iq} k_q, \quad 1 \leq i \leq m,$$

la studiul formei pătratice (în $q = m - p$ variabile)

$$\psi_{AB}(k_1, ..., k_q) = \sum_{r=1}^{q} \sum_{s=1}^{q} \gamma_{rs} k_r k_s,\ (k_1, ... k_q) \in R^q,$$

obţinută din precedenta prin înlocuirile de mai sus; deci, o formă pătratică de matrice $C = (\gamma_{rs})$, unde

(D3) $$\gamma_{rs} = \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_{ij} \nu_{ir} \nu_{js},\ 1 \leq r, s \leq q.$$

Avem, în aceste condiţii,

Teorema 5.4.2 *Fie $a = (a_1, ..., a_m)$ un punct staţionar condiţionat al funcţiei f, relativ la un anumit sistem de multiplicatori Lagrange $\lambda_1, ..., \lambda_p$ şi o funcţie Lagrange asociată L. Să mai notăm*

$$A = \left(\frac{\partial^2 L}{\partial x_i \partial x_j}(a) \right) \qquad \text{(\textit{hessianul funcţiei Lagrange în a})}$$

$$B = \frac{D(g_1, ..., g_p)}{D(x_1, ..., x_m)}(a) \quad \text{(\textit{matricea jacobiană a restricţiilor în a})}$$

În acest caz, punctul considerat este un

(a) *punct de minim condiţionat pentru f dacă forma pătratică de matrice A este de tip pozitiv faţă de B;*

(b) *punct de maxim condiţionat pentru f dacă forma pătratică de matrice A este de tip negativ faţă de B.*

Demonstraţie. Fie $E = S(a, \rho)$ o sferă de rază suficient de mică în jurul punctului a, inclusă în domeniul D. Vom lua punctul $x = (x_1, ..., x_m)$ arbitrar fixat în sfera E, de aşa manieră încât restricţiile de egalitate (RE) au loc. Notăm, pentru simplitate,

$$h_1 = x_1 - a_1, ..., h_m = x_m - a_m.$$

Să scriem formula de medie Taylor (Secţiunea 4.3.3) în punctele considerate, pentru funcţia Lagrange L. Există punctul intermediar $b = a + \theta(x - a)$ (unde $0 < \theta < 1$) aşa încât , cu notaţia

$$\alpha^*_{ij} = \frac{\partial^2 L}{\partial x_i \partial x_j}(b),\ 1 \le i, j \le m,\ A^* = (\alpha^*_{ij})$$

următoarea evaluare să fie valabilă

$$f(x) - f(a) = \frac{1}{2} \sum_{i,j} \alpha^*_{ij} h_i h_j.$$

Pe de altă parte, să transformăm egalităţile

$$g_i(x) - g_i(a) = 0,\ 1 \le i \le p,$$

în sensul formulei de medie Lagrange. Avem asigurate punctele intermediare

$$b_1 = a + \theta_1(x - a), ..., b_p = a + \theta_p(x - a)$$

aşa încât , cu notaţia

$$\beta_{ij}^* = \frac{\partial g_i}{\partial x_j}(b_i),\ 1 \le i \le p, 1 \le j \le m,\ B^* = (\beta_{ij}^*)$$

să fie valabile relaţiile

$$(\mathrm{S}_0^*) \qquad \beta_{i1}^* h_1 + \cdots + \beta_{im}^* h_m = 0,\ 1 \le i \le q.$$

Să rezolvăm acest sistem omogen (lucru posibil, în baza ipotezei de independenţă (C2)). Ajungem să exprimăm soluţia acestuia prin $q = m - p$ parametri $(k_1, ...k_q)$:

$$h_i = \nu_{i1}^* k_1 + \cdots + \nu_{iq}^* k_q,\ 1 \le i \le m.$$

Şi atunci, înlocuind în expresia anterioară, ajungem la exprimarea diferenţei $f(x) - f(a)$ ca formă pătratică în variabilele $k_1, ..., k_q$:

$$f(x) - f(a) = \frac{1}{2} \sum_{r=1}^{q} \sum_{s=1}^{q} \gamma_{rs}^* k_r k_s$$

unde, prin definiţie,

$$\gamma_{rs}^* = \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_{ij}^* \nu_{ir}^* \nu_{js}^*,\ 1 \le r, s \le q.$$

În baza ipotezei de regularitate (C3), determinanţii principali $(\Delta_1^*, ..., \Delta_q^*)$ ai acestei matrici au acelaşi semn cu determinanţii principali ai matricei $C = (\gamma_{rs})$, definită prin formula anterioară (D3) (unde matricile $A = (\alpha_{ij})$ şi $B = (\beta_{ij})$ sunt luate ca în enunţul teoremei). Aceasta încheie argumentul. ■

Rezultatele expuse constituie baza metodologică a următorului algoritm de rezolvare a problemei de extremum cu restricţii egalitate

$$(\mathrm{P}) \qquad \left| \begin{array}{l} \text{(extr) } f(x_1, ..., x_m) \\ \hline g_i(x_1, ..., x_m) = 0,\ 1 \le i \le p \\ (x_1, ..., x_m) \in D \end{array} \right.$$

care, din punct de vedere formal, poate fi asimilată unei probleme de programare (neliniară):

Pasul 1. Se formează funcţia Lagrange

$$L(x_1, ..., x_m) = f(x_1, ..., x_m) + \sum_{i=1}^{p} \lambda_i g_i(x_1, ..., x_m)$$

şi se rezolvă sistemul cu $m+p$ necunoscute $(x_1,...,x_m;\lambda_1,...\lambda_p)$

$$\frac{\partial L}{\partial x_i}=0,\ 1\le i\le m,\quad g_j(x_1,...,x_m)=0,\ 1\le j\le p.$$

Presupunem că am efectuat rezolvarea şi fie

$$x_1=a_1,...,x_m=a_m,\ \lambda_1,...,\lambda_p$$

una din soluţiile sistemului în cauză.

Pasul 2. Se construiesc matricile

$$A=\left(\frac{\partial^2 L}{\partial x_i\partial x_j}(a)\right),\quad B=\left(\frac{\partial g_i}{\partial x_j}(a)\right)\cdot$$

Dacă matricea A este de tip pozitiv în raport cu B, atunci punctul $a=(a_1,...,a_m)$ este o soluţie (locală) a problemei (P) în sensul minimului; iar dacă A este de tip negativ în raport cu B, atunci punctul în chestiune este o soluţie (locală) a problemei (P), în sensul maximului.

Să notăm că, dacă matricea A este de tip pozitiv (negativ) atunci ea este de tip pozitiv (negativ) faţă de B; aceasta se poate dovedi util în unele situaţii practice.

Exemplu. Să se rezolve problema

$$(\mathrm{P})\quad \left|\begin{array}{l}(\mathrm{extr})\ f(x,y,z)=xy+xz+yz\\ \hline xyz-1=0\\ x>0,y>0,z>0.\end{array}\right.$$

Formăm funcţia Lagrange

$$L(x,y,z)=xy+xz+yz+\lambda(xyz-1),\ x,y,z>0.$$

Rezolvăm apoi sistemul

$$\frac{\partial L}{\partial x}=0,\ \frac{\partial L}{\partial y}=0,\ \frac{\partial L}{\partial z}=0,\ xyz-1=0,$$

adică

$$y+z+\lambda yz=0,\ z+x+\lambda zx=0,\ x+y+\lambda xy=0,\ xyz-1=0.$$

Aceasta se face prin câteva artificii simple (ca, de pildă, scăderi şi substituţii). Obţinem soluţia (care corespunde şp domeniului luat)

$$x = 1,\ y = 1,\ z = 1,\ \lambda = -2.$$

Funcţia Lagrange asociată acestui λ este deci

$$L(x, y, z) = xy + xz + yz - 2(xyz - 1),\ x > 0, y > 0, z > 0.$$

Avem imediat

$$\begin{array}{lll}
\dfrac{\partial L}{\partial x} = y + z - 2yz, & \dfrac{\partial L}{\partial y} = x + z + 2xz, & \dfrac{\partial L}{\partial x} = x + y - 2xy, \\
\dfrac{\partial^2 L}{\partial x^2} = 0, & \dfrac{\partial^2 L}{\partial y^2} = 0, & \dfrac{\partial^2 L}{\partial z^2} = 0, \\
\dfrac{\partial^2 L}{\partial x \partial y} = 1 - 2z, & \dfrac{\partial^2 L}{\partial x \partial z} = 1 - 2y, & \dfrac{\partial^2 L}{\partial y \partial z} = 1 - 2x.
\end{array}$$

De aici rezultă expresia matricei hessian în $(1, 1, 1)$

$$A = \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}.$$

Forma pătratică ataşată acestei matrici este

$$\psi_A(h_1, h_2, h_3) = -2(h_1h_2 + h_1h_3 + h_2h_3).$$

Avem, pe de altă parte,

$$\frac{\partial g}{\partial x} = yz,\ \frac{\partial g}{\partial y} = xz,\ \frac{\partial g}{\partial z} = xy.$$

Rezultă de aici expresia matricii jacobian

$$B = (1\ 1\ 1)\ \ \text{(matrice linie)}$$

şi, deci, sistemul omogen verificat de variabilele (h_1, h_2, h_3) este

$$h_1 + h_2 + h_3 = 0\ \ (\Longrightarrow h_3 = -h_1 - h_2).$$

Înlocuind în forma pătratică anterioară găsim o altă formă pătratică, în două variabile
$$\psi_{AB}(h_1, h_2) = 2[h_1^2 + h_1h_2 + h_2^2].$$
Această formă pătratică, având matricea
$$C = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \text{ (element din } \mathcal{M}(2,2))$$
este pozitiv definită, deoarece
$$\Delta_1 = 2 > 0,\ \Delta_2 = 3 > 0.$$
Ca atare, punctul $(1, 1, 1)$ este de minim condiţionat pentru funcţia noastră.

5.4.2 Cazul restricţiilor de tip inegalitate

Fie m număr natural dat şi D un domeniu convex din R^m. Să dăm sistemul de funcţii
$$g_1 : D \longrightarrow R, ..., g_p : D \longrightarrow R \quad (p = \text{număr natural fixat}).$$
Aceasta, desigur, echivalează cu a da funcţia $g : D \longrightarrow R^p$, prin
$$g(x_1, ..., x_m) = (g_1(x_1, ..., x_m), ..., g_p(x_1, ..., x_m)), \quad (x_1, ..., x_m) \in D.$$
Prin definiţie, să numim sistemul de inecuaţii
$$\text{(RI)} \qquad g_1(x_1, ..., x_m) \leq 0, ..., g_p(x_1, ..., x_m) \leq 0$$
domeniul negativ din R^m generat de funcţia g; îl vom nota prin $D(g; -)$. Evident, orice asemenea domeniu poate fi privit şi ca unul pozitiv, dacă schimbăm semnul funcţiilor componente $g_1, ..., g_p$; aceasta însă nu se dovedeşte esenţial în continuare.

Fie acum $f : D \longrightarrow R$ o funcţie dată. Mai luăm un punct $a = (a_1, ..., a_m)$ din D care satisface (RI) şi un subdomeniu convex E al lui D, care conţine a. Zicem că punctul în cauză este *punct de minim* (respectiv, *maxim*) global al funcţiei f relativ la E, *condiţionat* de restricţiile (RI), dacă
$$\text{(D1)} \quad \begin{cases} f(x) \geq f(a) \text{ (respectiv, } f(x) \leq f(a)\text{), pentru toţi } x \in E \\ \text{care satisfac restricţiile (RI).} \end{cases}$$

Când una din aceste proprietăţi are loc, vom zice că punctul a din D, care satisface (RI), este un *punct de extrem global* al funcţiei f, relativ la E, *condiţionat* de restricţiile (RI). În particular, pentru $E = D$, vom spune că avem de-a face cu un *punct de extrem global* al funcţiei f, *condiţionat* de restricţiile (RI); sau, pe scurt, un *punct de extrem global condiţionat* al funcţiei considerate. Iar, când domeniul E (din D) care conţine a este subînţeles în definiţia (D1), vom zice că punctul $a \in D$ ce satisface (RI) apare ca *punct de extrem local* al funcţiei f, *condiţionat* de restricţiile (RI); sau, pe scurt, ca *punct de extrem local condiţionat* al funcţiei în cauză.

Acum, la fel ca la cazul restricţiilor de tip egalitate, problemele importante ce apar în acest context sunt:

(i) determinarea condiţiilor necesare de extrem condiţionat;

(ii) determinarea condiţiilor suficiente de extrem condiţionat.

În legătură cu prima problemă, următorul răspuns poate fi dat. Admitem şi aici ipoteza (C1) (scrisă pentru cazul restricţiilor egalitate). Pentru fiecare punct x din $D(g;-)$, să notăm

$$\text{(D2)} \quad \begin{cases} I(x) = \{i \in \{1,...,p\}; g_i(x) = 0\} & \text{(indicii activi în } x\text{)} \\ J(x) = \{i \in \{1,...,p\}; g_i(x) < 0\} & \text{(indicii inactivi în } x\text{).} \end{cases}$$

De asemenea, pentru fiecare sistem de numere $\lambda_1, ...\lambda_p$, să definim funcţia $L : D \longrightarrow R$, prin convenţia (D2) (scrisă la cazul restricţiilor egalitate). O vom numi şi aici *funcţia Lagrange* asociată sistemului $\{f, g_1, ..., g_p\}$; iar $\lambda_1, ...\lambda_p$, *multiplicatorii Lagrange* corespunzători. Fie acum $a = (a_1, ..., a_m)$ un punct din domeniul D care mai satisface restricţiile de inegalitate (RI); sau (cu notaţiile noastre), un punct al domeniului $D(g;-)$.

Teorema 5.4.3 *Să presupunem că punctul $a \in D(g;-)$ este de minim (maxim) condiţionat –local sau global– pentru funcţia f şi că,în plus,*

(C1) *vectorii $\{\nabla g_i(a);\ i \in I(a)\}$ sunt liniar independenţi.*

Există atunci un sistem unic de multiplicatori Lagrange $\lambda_1, ...\lambda_p$ cu

(P1) $\lambda_i \geq 0$ $(\lambda_i \leq 0)$, $\lambda_i g_i(a) = 0$, $1 \leq i \leq p$,

astfel încât , notând cu L funcţia Lagrange asociată să avem asigurată concluzia (R1) *din Teorema* 5.4.1 .

Demonstraţie. (Schiţă) Vom studia doar cazul de minim (căci perechea lui se reduce la acesta prin simpla trecere la funcţia opusă, $-f$). Fie $h = (h_1, ..., h_m)$ un vector din R^m ce satisface

(C2) $(\nabla g_i(a))h^\top < 0,\ i \in I(a)$.

Deoarece D este deschisă, va exista un $\delta > 0$, suficient de mic, aşa încât

$$a + th \in D, \qquad \text{pentru toţi} \quad t \in (-\delta, \delta).$$

Pentru fiecare $i \in \{1, ..., p\}$, să definim

$$\varphi_i(t) = g_i(a + th), \quad -\delta < t < \delta.$$

Dacă $i \in J(a)$, atunci $\varphi_i(a) = g_i(a) < 0$. Şi deci, din continuitatea funcţiei g_i, va exista un $\delta_i > 0$ cu $\delta_i \leq \delta$ şi

(R1) $\varphi_i(t) = g_i(a + th) < 0, \quad 0 < t < \delta_i$.

Dacă $i \in I(a)$, $\varphi_i(a) = g_i(a) = 0$. În baza formulelor de derivare compusă (Secţiunea 4.3.2) şi a alegerii vectorului h,

$$\varphi_i'(0) = (\nabla g_i(a))h^\top < 0.$$

Şi atunci, cu definiţia derivatei există un $\delta_i > 0$ cu $\delta_i \leq \delta$ aşa încât (R1) să fie din nou valabilă. Punând $\varepsilon = \min\{\delta_1, ..., \delta_p\}$, ajungem la concluzia

$$a + th \in D(g; -),\ 0 \leq t < \varepsilon.$$

În acest caz, o condiţie de forma

(C3) $(\nabla f(a))h^\top < 0$

nu poate fi acceptată. Căci, dacă punem

$$\psi(t) = f(a + th), \quad -\varepsilon < t < \varepsilon$$

am avea, din aceleaşi formule de derivare compusă , $\psi'(0) < 0$. Şi deci, din definiţia derivatei, ar exista un $\varepsilon^* > 0$ cu $\varepsilon^* \leq \varepsilon$ şi

$$\psi(t) < \psi(0),\ 0 < t < \varepsilon^*;$$

adică, de fapt,

$$f(a + th) < f(a),\ 0 < t < \varepsilon^*,$$

în contradicţie cu alegerea punctului $a \in D(g; -)$.

Am arătat, deci, că (C2)+(C3) nu pot fi simultan adevărate. Dar atunci, conform Teoremei lui Jordan, (Secţiunea 1.5.2), există numerele reale μ_0 şi $\{\mu_i;\ i \in I(a)\}$ cu

(P2) $\mu_0 \geq 0,\ \mu_i \geq 0,\ i \in I(a),\ \mu_0 + \sum\limits_{i \in I(a)} \mu_i > 0$

(P3) $\mu_0 \nabla f(a) + \sum\limits_{i \in I(a)} \mu_i \nabla g_i(a) = 0.$

Acum, din ipoteza (C1), avem obligatoriu $\mu_0 \neq 0$. Căci altfel, cu (P3) s-ar obţine o contradicţie. Nu avem acum decât să împărţim cu μ_0 şi să notăm

$$\lambda_i = \frac{\mu_i}{\mu_0},\ i \in I(a);\quad \lambda_i = 0,\ i \in J(a),$$

pentru a ajunge la concluzia cerută. ■

Aşadar, într-un punct de extrem condiţionat pentru funcţia f avem proprietatea de staţionaritate de tip (R1) (Secţiunea 5.4.1), în care intervin atât legăturile $(g_1, ..., g_p)$, cât şi proprietatea (P1) a aceluiaşi (intermediate de mulţimea indicilor activi). Prin definiţie, relaţiile în chestiune se vor numi *condiţii necesare Kuhn–Tucker* de extrem condiţionat; iar punctul respectiv $a = (a_1, ..., a_m)$, din $D(g; -)$, un *punct critic în sens Kuhn–Tucker.* Am demonstrat deci implicaţia

(P4) a=punct de extrem condiţionat $\Longrightarrow$ a=punct critic Kuhn–Tucker.

Este natural acum să ne întrebăm dacă şi reciproca este valabilă; adică, dacă – şi în ce măsură – condiţiile necesare de extrem Kuhn–Tucker sunt şi suficiente. Răspunsul poate fi dat după cum urmează. Să admitem ipoteza

(C4) $f, g_1, ..., g_p$ sunt simultan convexe (concave) pe D.

Teorema 5.4.4 *Să presupunem că $a = (a_1, ..., a_m)$ din $D(g; -)$ este un punct critic în sens Kuhn–Tucker al funcţiei f.Atunci,acesta este punct de minim (maxim) global pentru funcţia f, condiţionat de restricţiile* (RI) *(iniţial date).*

Demonstraţie. Va fi suficient să studiem cazul convex. Fie $x = (x_1, ..., x_m)$ un element arbitrar fixat din $D(g; -)$. Din proprietăţile de convexitate ale funcţiilor $g_1, ..., g_p$, avem (cu Teorema 5.1.9)

$$(\nabla g_i(a))(x - a)^\top \leq g_i(x) - g_i(a) = g_i(x) \leq 0,\ x \in D(g; -),\ i \in I(a).$$

Fie $\lambda_1, ..., \lambda_p$ multiplicatorii Lagrange care intervin în alegerea punctului critic $a \in D(g; -)$. Putem deci scrie

$$\left(\sum_{i \in I(a)} \lambda_i \nabla g_i(a)\right)(x-a)^\top \leq 0, \quad x \in D(g;-).$$

De aici, în baza condiţiilor Kuhn–Tucker (presupuse a avea loc cu aceşti multiplicatori) se obţine

$$(-\nabla f(a))(x-a)^\top \leq 0 \ \ (\text{sau } (\nabla f(a))(x-a)^\top \geq 0).$$

Aceasta, combinată cu convexitatea lui f, ne va da (prin rezultatul deja citat)

$$f(x) \geq f(a), \text{ pentru orice } x \in D \text{ care satisface (RI)}$$

de unde concluzia. Celălalt caz se studiază analog. ■

Rezultatele expuse ne permit să abordăm următoarea problemă de programare neliniară cu restricţii inegalităţi:

$$(\mathrm{P}) \qquad \left|\begin{array}{l} (\text{extr})\ f(x_1,...,x_m) \\ \hline g_1(x_1,...,x_m) \leq 0, \quad 1 \leq i \leq p \\ (x_1,...,x_m) \in D. \end{array}\right.$$

Aici funcţiile $f, g_1, ..., g_p$ se iau în conformitate cu ipoteza (C1) (din Secţiunea 5.4.1).

Pasul 1. Se formează funcţia Lagrange

$$L(x_1,...,x_m) = f(x_1,...,x_m) + \sum_{i=1}^{p} \lambda_i g_i(x_1,...,x_m)$$

şi se rezolvă sistemul de $m+p$ ecuaţii cu $m+p$ necunoscute

$$(\mathrm{SL}) \qquad \begin{cases} \dfrac{\partial L}{\partial x_1} = 0, ..., \dfrac{\partial L}{\partial x_m} = 0, \\ \lambda_1 g_1(x_1,...,x_m) = 0, ..., \lambda_p g_p(x_1,...,x_m) = 0. \end{cases}$$

Se obţine un grup de soluţii ale acestuia, căruia îi mai impunem condiţiile suplimentare

$$(\mathrm{CL}) \qquad g_i(x_1,...,x_m) \leq 0, \ i \leq i \leq p; \quad \lambda_i \geq 0, \ 1 \leq i \leq p.$$

Presupunem efectuată rezolvarea şi fie

$$x_1 = a_1, ..., x_m = a_m; \quad \lambda_1, ..., \lambda_p$$

una din soluţiile obţinute.

Pasul 2. Se studiază convexitatea funcţiilor $f, g_1, ..., g_p$ în punctul $a = (a_1, ..., a_m)$; de pildă, prin

$$\text{(C5)} \quad \begin{cases} \text{matricile hessiene } Hf(a), Hg_1(a), ..., Hg_p(a) \\ \text{sunt (simultan) de tip pozitiv (negativ) pe } D, \end{cases}$$

dacă,în plus,acceptăm o ipoteză de forma (C3)(Secţiunea 5.4.1). În cazul când această proprietate (simultană) are loc, punctul în cauză este de minim (maxim) global pentru funcţia f, condiţionat de restricţiile (RI)(iniţial date).

Să mai notăm în final că, în conformitate cu principiile generale dezbătute în cazul liniar (Secţiunea 2.1.2), orice problemă de programare neliniară de forma

$$\text{(P)} \quad \left| \begin{array}{l} \text{(extr) } f(x_1, ..., x_m) \\ \hline g_i(x_1, ..., x_m) \leq 0, \quad 1 \leq i \leq r, \ g_i(x_1, ..., x_m) \geq 0, \ r+1 \leq i \leq p \\ (x_1, ..., x_m) \in D \end{array} \right.$$

poate fi adusă la forma precedentă. Deci, algoritmul anterior este aplicabil unei întregi clase de asemenea probleme.

Exemplu. Să se rezolve problema de programare neliniară

$$\text{(P)} \quad \left| \begin{array}{l} \text{(extr) } f(x, y) = x^3 + y^3 - 12x - 3y \\ \hline x - y + 1 \leq 0, \ -x \leq 0 \\ x > 0, \ y > 0. \end{array} \right.$$

Formăm funcţia Lagrange

$$L(x, y) = x^3 + y^3 - 12x - 3y + \lambda(x - y + 1) + \mu(-x)$$

şi rezolvăm sistemul de condiţii necesare Kuhn–Tucker

$$\text{(KT1)} \qquad \frac{\partial L}{\partial x} = 0, \ \frac{\partial L}{\partial y} = 0, \ \lambda \cdot g_1(x, y) = 0, \ \mu \cdot g_2(x, y) = 0$$

$$\text{(KT2)} \qquad \lambda \geq 0, \ \mu \geq 0, \ x > 0, \ y > 0, \ g_1(x, y) \leq 0, \ g_2(x, y) \leq 0.$$

Se obţine din prima parte sistemul

$$3x^2 - 12 + \lambda - \mu = 0,\ 3y^2 - 3 - \lambda = 0,\ \lambda(x - y + 1) = 0,\ \mu(-x) = 0.$$

Prin rezolvarea acestui sistem se obţin 7 soluţii

$$\begin{array}{lllll} M_1 & : x_1 = 2, & y_1 = 1, & \lambda = 0, & \mu = 0 \\ M_2 & : x_2 = 2, & y_2 = -1, & \lambda = 0, & \mu = 0 \\ M_3 & : x_3 = -2, & y_3 = 1, & \lambda = 0, & \mu = 0 \\ M_4 & : x_4 = -2, & y_4 = -1, & \lambda = 0, & \mu = 0 \\ M_5 & : x_5 = 0, & y_5 = 1, & \lambda = 0, & \mu = -12 \\ M_6 & : x_6 = 0, & y_6 = -1, & \lambda = 0, & \mu = -12 \\ M_7 & : x_7 = 1, & y_7 = 2, & \lambda = 9, & \mu = 0 \end{array}$$

Din toate aceste 7 soluţii ale sistemului (KT1), numai ultima verifică şi inegalităţile (KT2). Deci $M_7 = (1, 2)$ apare ca punct staţionar Kuhn–Tucker al problemei noastre. Remarcăm acum că funcţiile

$$g_1(x, y) = x - y + 1,\ g_2(x, y) = -x$$

sunt evident convexe, datorită liniarităţii. În ce priveşte funcţia obiectiv, avem

$$\frac{\partial f}{\partial x} = 3x^2 - 12, \quad \frac{\partial f}{\partial y} = 3y^2 - 3$$
$$\frac{\partial^2 f}{\partial x^2} = 6x,\ \frac{\partial^2 f}{\partial x \partial y} = 0,\ \frac{\partial^2 f}{\partial y^2} = 6y.$$

Deci, hessianul funcţiei noastre este

$$Hf(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}, \quad (x, y) \in R^2.$$

Determinanţii principali ai acestuia sunt

$$\Delta_1(x, y) = 6x,\ \Delta_2(x, y) = 36xy.$$

Constatăm că

$$\Delta_1(x, y) > 0, \quad \Delta_2(x, y) > 0, \text{ pentru toţi } x, y > 0.$$

Deci, f este convexă pe domeniul

$$D = \{(x, y) \in R^2;\ x > 0, y > 0\}.$$

Şi, ca atare, punctul obţinut apare ca soluţie a problemei noastre.

Probleme propuse la §5.4

1. Să se studieze extremele funcţiei

$$f(x,y) = x + y \quad \text{dacă} \;\; x^2 + y^2 - 1 = 0, \;\; x, y \in R.$$

(Indicaţie. Se poate folosi şi reprezentarea parametrică $x = \cos t$, $y = \sin t$ pentru curba în cauză.)

2. Să se studieze extremele funcţiei $f(x,y) = x^2+y^2-x-y$, cu variabilele supuse la legătura $x + y = 1$.
(Indicaţie. Se va folosi funcţia Lagrange; dar se poate aborda şi direct problema, prin explicitarea legăturii.)

3. Să se studieze extremele funcţiei $f(x,y,z) = xyz$, cu legăturile $x^2 + y^2 + z^2 = 1$, $x + y + z = 0$.

4. Să se determine extremele funcţiei $f(x_1, ..., x_m) = x_1 + \cdots + x_m$, cu legătura $x_1 \cdots x_m = 1$, $x_1 > 0, ..., x_m > 0$.
(Indicaţie. Se va folosi metoda multiplictorilor Lagrange. Putem însă da o rezolvare mai directă şi prin intermediul inegalităţii mediilor.)

5. Să se studieze extremele funcţiei $f(x,y) = x^2+y^2-3x-2y$, cu legătura $x^2 + y^2 \leq 1$, $x, y \in R$.

6. Să se studieze extremele funcţiei $f(x,y,z) = x^2 + y^2 + z^2$ cu legăturile $y^2 \leq 4$, $x^2 + z^2 \leq 1$.

Capitolul 6

Procese dinamice

6.1 Procese de tip secvenţial (Şiruri recurente)

6.1.1 Generalităţi

După cum am convenit deja (Secţiunea 1.1.10), un *şir* de numere reale înseamnă o funcţie $x : N \longrightarrow R$. Valoarea acestei funcţii în punctul generic n se va indica prin x_n sau $x(n)$; vom folosi în continuare cea de-a doua notaţie.

Fie dată funcţia $f : N \times R \longrightarrow R$ şi numerele $n_0 \in N$, $x_0 \in R$. Să considerăm relaţiile

$$(E) \qquad x(n+1) = f(n, x(n))$$

$$(I). \qquad x(n_0) = x_0$$

Prima dintre ele se zice *ecuaţie cu diferenţe de ordinul întâi.* Sensul lui $(E) + (I)$ este următorul: Să se găsească acele şiruri $n \vdash x(n)$ care satisfac condiţia iniţială (I) şi

$$x(n+1) = f(n, x(n)), \text{ pentru toţi } n \in N, \ n \geq n_0.$$

Vom numi o asemenea soluţie *şir recurent de ordinul întâi.* Analog, fie $f : N \times R^k \longrightarrow R$, funcţia dată. Considerăm relaţiile

$$(E_k) \qquad x(n+k) = f(n, x(n), ..., x(n+k-1))$$

$$(I_k) \qquad x(n_0) = \alpha_0, ..., x(n_0 + k - 1) = \alpha_{k-1}$$

Vom numi (E_k) *ecuaţie cu diferenţe de ordinul* k. Interpretarea ei este analogă cu cea precedentă. Orice soluţie a problemei $(E_k) + (I_k)$ se va zice *şir recurent de ordin* k. În fine, fie date funcţiile $f_i : N \times R \longrightarrow R$, $1 \leq i \leq k$, numărul natural n_0 şi numerele reale x_i^0, $1 \leq i \leq k$. Considerăm relaţiile

$$(S_k) \qquad x_i(n+1) = f_i(n, x_1(n), ..., x_k(n)), \ 1 \leq i \leq k$$

$$(J_k) \qquad x_i(n_0) = x_i^0, \ 1 \leq i \leq k.$$

Primul ansamblu de egalităţi se numeşte *sistem cu diferenţe de ordinul întâi.* Interpretarea lui $(S_k) + (J_k)$ este următoarea: Să se găsească şirurile $n \vdash x_i(n)$, $1 \leq i \leq k$, ce satisfac condiţiile iniţiale (J_k) şi (relativ la orice $i \in \{1, ..., k\}$)

$$x_i(n+1) = f_i(n, x_1(n), ..., x_k(n)), \text{ pentru toţi } n \in N, \ n \geq n_0.$$

Grupul soluţiilor acestei probleme îl vom numi *sistem de şiruri recurente de ordinul întâi.*

Este evident că problemea $(E_k) + (I_k)$ apare ca o extensie a problemei $(E)+(I)$ (la care se reduce pentru $k = 1$). La rândul ei, problema $(E_k)+(I_k)$ poate fi (tehnic) redusă la problema $(S_k) + (J_k)$. În adevăr, dat sistemul cu diferenţe de ordin întâi

$$(S_k'), \qquad \begin{cases} x_1(n+1) = x_2(n), ..., x_{k-1}(n+1) = x_k(n), \\ x_k(n+1) = f(n, x_1(n), ..., x_k(n)) \end{cases}$$

componenta $n \vdash x_1(n)$ a soluţiei acestuia verifică (E_k), după cum se poate uşor vedea. Reciproc, dată ecuaţia cu diferenţe (E_k) este clar că şirurile de forma

$$n \vdash x(n), \ n \vdash x(n+1), ..., n \vdash x(n+k-1)$$

sunt componentele soluţiei sistemului cu diferenţe (S_k').

Din punct de vedere tehnic, problema existenţei şi unicităţii soluţiei unei ecuaţii cu diferenţe sau sistem cu diferenţe este în general pozitiv rezolvabilă (se determină din aproape în aproape componentele soluţiei). Ceea ce contează însă în teoria acestor ecuaţii (sau sisteme) este expresia soluţiei problemei în cauză.

Este util de subliniat faptul că problemele de tip recurent au aplicaţii diverse în diferite domenii ale matematicilor aplicate.

Noţiunea de ecuaţie cu diferenţe este legată de aceea de operator diferenţă. Aceasta necesită unele convenţii de notaţie. Să punem

$$F(N, R) = \text{spaţiul tuturor şirurilor de numere reale.}$$

Vom introduce un operator pe acesta, $\Delta : F(N, R) \longrightarrow F(N, R)$, prin

(D1) $\Delta x(n) = x(n+1) - x(n),\ n \in N,\ x \in F(N, R);$

îl vom numi *operatorul diferenţă* (pe spaţiul respectiv). Evident, un asemenea operator este liniar, în sensul

(P1) $\Delta(x + y) = \Delta(x) + \Delta(y), \Delta(\lambda x) = \lambda\Delta(x), \quad x, y \in F(N, R), \lambda \in R.$

Ne interesează acum iteratele acestui operator. Avem

$$\begin{aligned} \Delta^2 x(n) = \Delta(\Delta x)(n) = \Delta x(n+1) - \Delta x(n) = \\ = x(n+2) - 2x(n+1) + x(n),\ x \in N. \end{aligned}$$

În general, are loc formula (pentru orice $k = 1, 2, ...$)

(P2) $\Delta^k x(n) = \sum_{h=0}^{k} (-1)^{k-h} C_k^h x(n+h),\ n \in N.$

Aceasta poate fi scrisă şi sub formă operatorială. Mai întâi să introducem *operatorul identic* pe $F(N, R)$

(D2) $Ix(n) = x(n),\ n \in N,\ x \in F(N, R),$

iar apoi *operatorul de translaţie* (pe acelaşi spaţiu)

(D3) $Qx(n) = x(n+1),\ n \in N,\ x \in F(N, R).$

Iteratele acestui operator au forma (pentru $k \geq 1$)

$$Q^k x(n) = x(n+k),\ n \in N,\ x \in F(N, R).$$

În această perspectivă, formula (P1) va primi forma operatorială

(P3) $\Delta^k = \sum_{h=0}^{k} (-1)^{k-h} C_k^h Q^h = (Q - I)^k,\ k = 1, 2...,$

unde, prin definiţie, am pus $Q^0 = I$. În particular $\Delta = Q - I$, de unde $Q = \Delta + I$; şi atunci (cu $\Delta^0 = I$),

$$Q^k = (\Delta + I)^k = \sum_{h=0}^{k} C_k^h \Delta^{k-h}, \ k = 1, 2, ... \ .$$

Adică, trecând la şiruri, avem (pentru orice k=1,2,...)

$$\text{(P4)} \ \ x(n+k) = \sum_{h=0}^{k} C_k^h \Delta^{k-h} x(n), \ n \in N.$$

Cu alte cuvinte, pentru orice $k \geq 1$, operatorul diferenţă Δ^k se exprimă liniar prin translatele $\{Q^0, ..., Q^k\}$; şi, reciproc, operatorul de translaţie Q^k se exprimă liniar prin diferenţele $\{\Delta^0, ..., \Delta^k\}$. Aceasta ne permite ca orice ecuaţie cu diferenţe dată prin operatorii de translaţie să o scriem prin operatorii diferenţă asociaţi. De pildă, (E) se poate scrie în mod echivalent

$$(E) \qquad \Delta x(n) = g(n, x(n))$$

ecuaţia (E_k), va avea forma

$$(E_k) \qquad \Delta^k x(n) = g(n, \Delta x(n), ..., \Delta^{k-1} x(n)).$$

iar, în fine, sistemul (S_k) devine aici

$$(S_k) \qquad \Delta x_i(n) = g_i(n, x_1(n), ..., x_k(n)), \ 1 \leq i \leq k.$$

În felul acesta, scrierea ar fi mai aproape de ceea ce în mod obişnuit ar sugera noţiunea de diferenţă. Dar, din punct de vedere tehnic, aceasta ar genera unele complicaţii. De aceea nu vom folosi această terminologie, ci pe aceea deja propusă.

6.1.2 Ecuaţii liniare cu diferenţe finite

Numim *ecuaţie liniară cu diferenţe finite de ordin* k orice ecuaţie cu diferenţe de forma

$$(E) \qquad a_0(n)x(n+k) + a_1(n)x(n+k-1) + \cdots + a_k(n)x(n) = b(n).$$

Aici, $n \vdash a_0(n), ..., n \vdash a_k(n), n \vdash b(n)$ sunt nişte şiruri date. Asociem acesteia condiţiile iniţiale

$$(I) \qquad x(n_0) = \alpha_0, x(n_0+1) = \alpha_1, ..., x(n_0+k-1) = \alpha_{k-1}.$$

Aici, n_0 este un număr natural dat, iar $p = (\alpha_0, ..., \alpha_{k-1})^\top$ un vector din R^k. În ce priveşte existenţa şi unicitatea soluţiei este clar că în condiţiile (acceptate în continuare)

(C1) $a_0(n) \neq 0, a_k(n) \neq 0$, pentru orice $n \in N$,

problema (E)+(I) are o singură soluţie $n \vdash x(n)$; respectiv, un singur şir care să verifice $(E) + (I)$. Mai exact, prima parte din (C1) ne ajută să construim această soluţie pentru $n \geq n_0 + k$, iar a doua parte să o construim pentru $n < n_0$. În ce priveşte structura acestei soluţii, să pornim de la ecuaţia omogenă asociată lui (E); adică

$$(E_0) \qquad a_0(n)x(n+k) + a_1(n)x(n+k-1) + \cdots + a_k(n)x(n) = 0.$$

Notăm cu X mulţimea tuturor şirurilor $n \vdash x(n)$ din $F(N, R)$ care verifică (E_0). Pentru a vedea cum este structurată mulţimea X, introducem operatorul cu diferenţe $L : F(N, R) \longrightarrow F(N, R)$ prin

(D1) $Lz(n) = a_0(n)z(n+k) + a_1(n)z(n+k-1) + \cdots + a_k(n)z(n)$,
$n \in N,\ z \in F(N, R)$.

Nu este greu de văzut că acesta este un operator liniar, în sensul

(P1) $L(x + y) = L(x) + L(y), L(\lambda x) = \lambda L(x), \quad x, y \in F(N, R), \lambda \in R.$

De aici şi din

(P2) $X = \{z \in F(N, R);\ L(z) = 0\}$,

deducem că X este un subspaţiu liniar al lui $F(N, R)$. Ne interesează acum dimensiunea acestui subspaţiu. În acest sens, mai definim un operator $W : F(N, R) \longrightarrow F(N, R^k)$ prin

(D2) $Wz(n) = (z(n), z(n+1), ..., z(n+k-1))^\top,\ n \in N, z \in F(N, R)$.

Următorul rezultat (dat fără demonstraţie) se dovedeşte util aici.

Teorema 6.1.1 *Fie $y_1, ..., y_k$ un număr de k soluţii din $F(N, R)$ ale ecuaţiei (E_0). Atunci, afirmaţiile următoare sunt echivalente:*

(P3) *$\{y_1, ..., y_k\}$ este liniar independent în $F(N, R)$*

(P4) *matricea $[Wy_1(n), ..., Wy_k(n)]$ este nesingulară,pentru toţi $n \in N$*

(P5) *matricea* $[Wy_1(n_0), ..., Wy_k(n_0)]$ *este nesingulară,pentru* $n_0 \in N$.

Ca o imediată aplicaţie a acestui rezultat, avem:

Teorema 6.1.2 *Mulţimea X a soluţiilor ecuaţiei omogene (E_0) formează un subspaţiu k–dimensional al lui $F(N, R)$.*

Demonstraţie. Fie $\{e_1, ..., e_k\}$ baza canonică din R^k. Pentru fiecare $i \in \{1, ..., k\}$, notăm prin y_i soluţia din $F(N, R)$ a problemei

$$L(y_i) = 0,\ W(y_i)(0) = e_i.$$

Arătăm că $\{y_1, ..., y_k\}$ formează o bază în spaţiul X; o vom numi aceasta *sistem fundamental de soluţii* ale ecuaţiei omogene (E_0). Din

$$[W(y_1)(0), ..., W(y_k)(0)] = I \text{ (matricea unitate din } \mathcal{M}(k))$$

şi teorema precedentă, $\{y_1, ..., y_k\}$ este liniar independent. Fie apoi $y \in X$ o soluţie arbitrară a lui (E_0). Avem

$$Wy(0) = y(0)W(y_1)(0) + \cdots + y(k-1)W(y_k)(0).$$

Iar, de aici, împreună cu observaţiile privind unicitatea soluţiei problemei $(E) + (I)$,

$$y = y(0) \cdot y_1 + y(1) \cdot y_2 + \cdots + y(k-1) \cdot y_k.$$

Adică y este o combinaţie liniară a sistemului de funcţii $\{y_1, ..., y_k\}$. Teorema este dovedită. ■

Putem acum răspunde la problema privind structura mulţimii soluţiilor ecuaţiei omogene (E).

Teorema 6.1.3 *Formula*

(R1) $x(n) = C_1y_1(n) + \cdots + C_ky_k(n) + z(n),\ n \in N$

unde $C_1, ..., C_k$ sunt constante arbitrare, $\{y_1, ..., y_k\}$ este un sistem fundamental de soluţii ale lui (E_0), iar z o soluţie particulară a lui (E), reprezintă expresia soluţiei generale a ecuaţiei (E).

Demonstraţie. N-avem decât să remarcăm că diferenţa a două soluţii ale lui (E) reprezintă o soluţie a ecuaţiei omogene (E_0). ■

Este natural să ne întrebăm cum se poate determina o asemenea soluţie particulară a lui (E). Aceasta este în general posibil dacă se presupune cunoscut sistemul fundamental $\{y_1, ..., y_k\}$ de soluţii ale ecuaţiei omogene (E_0). Anume, se poate căuta soluţia în chestiune, z, prin metoda variaţiei constantelor, în forma

$$z(n) = C_1(n)y_1(n) + \cdots + C_k(n)y_k(n), \; n \in N,$$

unde şirurile $n \vdash C_1(n), ..., n \vdash C_k(n)$ urmează a fi determinate din sistemul (în variaţii)

$$(SV) \qquad \begin{cases} y_1(n)D_1(n) + \cdots + y_k(n)D_k(n) = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_1(n+k-2)D_1(n) + \cdots + y_k(n+k-2)D_k(n) = 0 \\ y_1(n+k-1)D_1(n) + \cdots + y_k(n+k-1)D_k(n) = \dfrac{b(n)}{a_0(n)}. \end{cases}$$

Aici, prin definiţie,

(D3) $D_h(n) = \Delta C_h(n), \; n \in N, \; 1 \leq h \leq k.$

Să notăm că matricea acestui sistem este nesingulară în baza unui rezultat anterior. Nu mai dăm detalii.

Un caz particular important este acela în care şirurile coeficient $n \vdash a_0(n), ..., n \vdash a_k(n)$ sunt constante. Este vorba deci de ecuaţia (cu coeficienţi constanţi)

$$(E) \qquad a_0x(n+k) + a_1x(n+k-1) + \cdots + a_kx(n) = b(n), \; n \in N$$

unde $\{a_0, ..., a_k\}$ sunt numere reale date, cu

(C2) $a_0 \neq 0, \; a_k \neq 0.$

În acest caz, chestiunea determinării unui sistem fundamental de soluţii ale ecuaţiei omogene asociate

$$(E_0) \qquad a_0x(n+k) + a_1(n+k-1) + \cdots + a_kx(n) = 0$$

cât şi aceea a determinării unei soluţii particulare ale ecuaţiei (E) capătă unele aspecte specifice.

(A) Relativ la ecuaţia cu diferenţe omogenă (E_0), să asociem acesteia următoarea ecuaţie de tip algebric (pe care o vom numi *caracteristică*) peste corpul complex (C)

$$(EC) \qquad a_0\lambda^k + a_1\lambda^{k-1} + \cdots + a_k = 0.$$

Notăm cu P polinomul din partea stângă

(D4) $P(\lambda) = a_0\lambda^k + a_1\lambda^{k-1} + \cdots + a_k, \ \lambda \in K.$

Dacă γ este o rădăcină reală de multiplicitate q, a ecuaţiei caracteristice (EC), adică

$$P(\gamma) = 0, P'(\gamma) = 0, ..., P^{(q-1)}(\gamma) = 0, P^{(q)}(\gamma) \neq 0,$$

atunci îi asociem şirurile (de numere reale)

$$x_1(n) = \gamma^n, x_2(n) = n\gamma^n, ..., x_q(n) = n^{q-1}\gamma^n.$$

Iar dacă $\zeta = \rho(\cos\theta + i\sin\theta)$ (deci şi conjugata acesteia, $\bar{\zeta} = \rho(\cos\theta - i\sin\theta)$) este o rădăcină complexă de multiplicitate q, a ecuaţiei caracteristice (EC), adică

$$\begin{cases} P(\zeta) = 0, P'(\zeta) = 0, \cdots, P^{q-1}(\zeta) = 0, P^{(q)}(\zeta) \neq 0 \\ P(\bar{\zeta}) = 0, P'(\bar{\zeta}) = 0, \cdots, P^{q-1}(\bar{\zeta}) = 0, P^{(q)}(\bar{\zeta}) \neq 0, \end{cases}$$

atunci îi asociem şirurile (reale)

$$\begin{cases} x_1(n) = \rho^n\cos n\theta, x_2(n) = n\rho^n\cos n\theta, ..., x_q(n) = n^{q-1}\rho^n\cos n\theta \\ x_{q+1}(n) = \rho^n\sin n\theta, x_{q+2}(n) = n\rho^n\sin n\theta, ..., x_{2q}(n) = n^{q-1}\rho^n\sin n\theta. \end{cases}$$

Deci,la cele k rădăcini ale ecuaţiei caracteristice (EC) (fiecare luată de atâtea ori cât arată multiplicitatea sa), ajungem să asociem k şiruri $x_1(n), ..., x_k(n)$ (printr-o re–numerotare potrivită a celor anterioare).

Teorema 6.1.4 *Sistemul de şiruri astfel determinat este un sistem fundamental de soluţii pentru ecuaţia cu diferenţe omogenă* (E_0).

Demonstraţie. (Schiţă). Pentru prima parte, ne vom limita doar la rădăcini reale. Fie deci γ rădăcină reală de multiplicitate q a ecuaţiei caracteristice (EC); condiţiile care o definesc se pot scrie şi sub forma

$$\sum_{h=0}^{k} a_{k-h}h^s\gamma^h = 0, \ 0 \leq s \leq q-1; \ \sum_{h=0}^{k} a_{k-h}h^q\gamma^h \neq 0.$$

Fie r număr natural arbitrar, Definim şirul

$$z(n) = n^r \gamma^n, \ n \in N.$$

Avem deocamdată, pentru $h \in \{0, ..., k\}$,

$$z(n+h) = (n+h)^r \gamma^{n+h} = ((n+h)^r \gamma^h)\gamma^n = \left(\sum_{s=0}^{r} C_r^s n^{r-s} h^s \gamma^h \right) \gamma^n.$$

Iar de aici

$$\begin{cases} L(z)(n) = \sum\limits_{h=0}^{k} a_{k-h} z(n+h) = \left(\sum\limits_{h=0}^{k} a_{k-h} \sum\limits_{s=0}^{r} C_r^s n^{r-s} h^s \gamma^h \right) \gamma^n = \\ = \sum\limits_{s=0}^{r} C_r^s \left(\sum\limits_{h=0}^{k} a_{k-h} h^s \gamma^h \right) n^{r-s} \gamma^n, \quad \text{pentru toţi} \ \ n \in N. \end{cases}$$

Ţinând cont de cele dinainte spuse, rezultă că funcţiile

$$z_r(n) = n^r \gamma^n, \ n \in N, \ 0 \le r \le q-1$$

sunt soluţii ale ecuaţiei omogene (E_0); iar aceasta probează prima parte. Pentru a doua parte vom admite şi aici că avem de-a face doar cu rădăcini reale, pentru simplitate. Să presupunem că soluţiile scrise $x_1(n), ..., x_k(n)$ nu ar fi independente. Ar exista deci constantele $\delta_1, ..., \delta_k$, nu toate nule, aşa ca

(R2) $\delta_1 x_1(n) + \cdots + \delta_k x_k(n) = 0, \quad \forall n \in N.$

Putem transforma această relaţie în

(R3) $Q_1(n)\gamma_1^n + \cdots + Q_h(n)\gamma_h^n = 0, \quad \forall n \in N,$

unde polinoamele $Q_1(n), ..., Q_h(n)$ nu sunt identic nule, iar $\gamma_1, ..., \gamma_h$ sunt rădăcinile reale (distincte două câte două) ale ecuaţiei caracteristice, de multiplicităţi $m_1, ..., m_h$ respectiv (cu $m_1 + \cdots + m_h = k$). Să reţinem că, din $a_k \neq 0$, nici una din aceste rădăcini nu este zero. Împărţim relaţia (R3) cu γ_1^n şi notăm

$$\gamma_2/\gamma_1 = \eta_2, ..., \gamma_h/\gamma_1 = \eta_h.$$

Avem, deci,

$$Q_1(n) + Q_2(n)\eta_2^n + \cdots + Q_h(n)\eta_h^n = 0, \quad \forall n \in N.$$

Aplicând operatorul diferenţă Δ de un număr suficient de ori (respectiv $\mathrm{grad}\,(Q_1) + 1$ ori), obţinem

(R4) $R_2(n)\eta_2^n + \cdots + R_h(n)\eta_h^n = 0,\ n \in N,$

unde $R_2(n), ..., R_h(n)$ sunt polinoame de grad respectiv egal cu al polinoamelor $Q_2(n), ..., Q_h(n)$. (Am folosit aici faptul că $\eta_2, ..., \eta_h$ sunt diferite de 1.) Procedeul se poate relua, ajungând la

(R5) $S_h(n)\nu_h^n = 0,\ n \in N,$

unde $S_h(n)$ este un polinom de grad egal cu acela al lui $Q_h(n)$, iar ν_h este un număr distinct de 1. Aceasta este însă imposibil. Contradicţia la care am ajuns probează afirmaţia. ■

Exemplu. Să rezolvăm ecuaţia cu diferenţe omogenă

$$x(n+2) - 3x(n+1) + 2x(n) = 0.$$

Ecuaţia caracteristică asociată este $\lambda^2 - 3\lambda + 2 = 0$, cu rădăcinile $\lambda_1 = 1, \lambda_2 = 2$. Soluţia generală a ecuaţiei în cauză este deci

$$x(n) = C_1\lambda_1^n + C_2\lambda_2^n = C_1 + C_2 \cdot 2^n,\ n \in N,$$

unde C_1, C_2 sunt constante arbitrare. Acestea pot fi unic determinate printr-o condiţie iniţială. Astfel, de pildă, dacă se mai cere în plus

$$x(0) = 1,\ x(1) = -1$$

atunci sistemul ce dă constantele în cauză este

$$C_1 + C_2 = 1,\ C_1 + 2C_2 = -2,$$

cu soluţia $C_1 = 4,\ C_2 = -3$. Rezultă soluţia

$$x(n) = 4 - 3 \cdot 2^n,\ n \in N$$

care satisface ecuaţia şi condiţia iniţială (impusă mai sus).

(B) În ce priveşte determinarea unei soluţii particulare a ecuaţiei cu diferenţe neomogene (E), metoda generală de lucru este după cum am arătat, aceea a variaţiei constantelor; aceasta este posibil aici, deoarece dispunem de un sistem fundamental de soluţii ale ecuaţiei omogene. Să ilustrăm acest fapt prin următorul

Exemplu. Să se rezolve ecuaţia cu diferenţe

$$x(n+1) - 3x(n) = \ln(1+n).$$

Trecem deocamdată la rezolvarea ecuaţiei omogene asociate

$$x(n+1) - 3x(n) = 0.$$

Ecuaţia caracteristică este

$$\lambda - 3 = 0, \quad \text{cu rădăcina } \lambda = 3.$$

Deci, soluţia generală a ecuaţiei omogene este

$$y(n) = C \cdot 3^n, \quad \text{unde } C = \text{constantă}.$$

Căutăm acum o soluţie particulară pentru ecuaţia nesingulară, de forma

$$z(n) = C(n)3^n, \ C(n) = \text{şir nedeterminat}.$$

Scriind că aceasta verifică ecuaţia în cauză, obţinem

$$C(n+1)3^{n+1} - 3C(n) \cdot 3^n = \ln(1+n), \ n \in N.$$

Iar, de aici, obţinem

$$C(n+1) - C(n) = 3^{-n-1} \ln(1+n), \ n \in N.$$

Aceasta ne permite să găsim uşor expresia şirului $C(n)$. De exemplu, dacă punem $C(0) = 0$, atunci

$$C(1) - C(0) = 3^{-1} \ln 1, ..., C(n) - C(n-1) = 3^{-n} \ln n;$$

şi, deci, adunând relaţiile (cu reduceri de termeni asemenea)

$$C(n) = 3^{-1} \ln 1 + 3^{-2} \ln 2 + \cdots + 3^{-n} \ln n = \sum_{k=1}^{n} 3^{-k} \ln k, \ n \geq 1.$$

Obţinem astfel soluţia particulară $z(n)$ a ecuaţiei neomogene. Şi atunci, soluţia generală a ecuaţiei neomogene iniţiale este

$$x(n) = y(n) + z(n) = \left(C + \sum_{k=1}^{n} 3^{-k} \ln k \right) 3^n, \ n \geq 1,$$

unde C este o constantă reală.

În legătură cu aceste fapte, este de subliniat că, în anumite situaţii, o soluţie particulară a ecuaţiei cu diferenţe neomogene poate fi găsită într-un mod mai direct. Acestea se referă la termeni liberi de forma

$$b(n) = (R(n)\cos n\theta + S(n)\sin n\theta)\rho^n,\ n \in N,$$

unde $R(n), S(n)$ sunt polinoame cu coeficienţi reali iar $\rho > 0$, $\theta \in R$ sunt constante date. În adevăr, să notăm

$$m = \max\{\operatorname{grad}(R), \operatorname{grad}(S)\}$$

şi să admitem că $\lambda = \rho(\cos\theta + i\sin\theta)$ este o rădăcină de multiplicitate q, a ecuaţiei caracteristice (EC). (Dacă $q = 0$, asta desigur înseamnă că λ dinainte nu este rădăcină a ecuaţiei în cauză.) Există atunci o soluţie particulară a ecuaţiei neomogene (E), de forma

$$z(n) = n^q[U(n)\cos n\theta + V(n)\sin n\theta]\rho^n,$$

unde $U(n)$ şi $V(n)$ sunt polinoame de grad cel mult m, care urmează a fi determinate (din condiţia ca $z(n)$ să verifice ecuaţia în cauză).

Exemplu. Să se rezolve ecuaţia cu diferenţe

$$x(n+3) - x(n) = (n^2 + n + 1)2^n.$$

Ecuaţia cu diferenţe omogenă asociată

$$x(n+3) - x(n) = 0$$

are ecuaţia caracteristică $\lambda^3 - 1 = 0$, cu rădăcinile

$$\lambda_1 = 1, \lambda_2 = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}, \lambda_3 = \cos\frac{2\pi}{3} - i\sin\frac{2\pi}{3}.$$

Ca atare, soluţia generală a ecuaţiei omogene este

$$y(n) = C_1 + C_2\cos\frac{2n\pi}{3} + C_3\sin\frac{2n\pi}{3},\ n \in N.$$

Căutăm acum o soluţie particulară a ecuaţiei neomogene, de forma

$$z(n) = (\alpha n^2 + \beta n + \gamma)2^n,\ n \in N,$$

unde α, β, γ sunt constante ce urmează a fi găsite. Anume, înlocuind în ecuaţia iniţială, găsim

$$(\alpha(n+3)^2 + \beta(n+3) + \gamma)2^{n+3} - (\alpha n^2 + \beta n + \gamma)2^n = (n^2 + n - 1)2^n.$$

Făcând calculele în partea stângă ajungem la identitatea

$$(7\alpha n^2 + (48\alpha + 7\beta)n + 72\alpha + 24\beta + 7\gamma)2^n = (n^2 + n + 1)2^n.$$

Egalând coeficienţii termenilor asemenea se obţine sistemul (liniar)

$$7\alpha = 1, 48\alpha + 7\beta = 1, 72\alpha + 24\beta + 7\gamma = 1$$

cu soluţia

$$\alpha = \frac{1}{7}, \beta = -\frac{41}{49}, \gamma = \frac{529}{343}.$$

Am ajuns astfel la soluţia particulară

$$z(n) = \left(\frac{1}{7}n^2 - \frac{41}{49}n + \frac{529}{343}\right)2^n.$$

Adumând acum soluţiile găsite, obţinem tocmai soluţia generală a ecuaţiei iniţiale

$$x(n) = C_1 + C_2 \cos\frac{2n\pi}{3} + C_3 \sin\frac{2n\pi}{3} + \left(\frac{1}{7}n^2 - \frac{41}{49}n + \frac{529}{343}\right)2^n.$$

6.1.3 Sisteme liniare cu diferenţe finite

Numim *sistem liniar de k ecuaţii cu diferenţe în k şiruri necunoscute* orice sistem cu diferenţe de forma

$$(S) \qquad \begin{cases} x_1(n+1) = a_{11}(n)x_1(n) + \cdots + a_{1k}(n)x_k(n) + b_1(n) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_k(n+1) = a_{k1}(n)x_1(n) + \cdots + a_{kk}(n)x_k(n) + b_k(n) \end{cases}$$

unde şirurile $n \vdash a_{ij}(n),\ 1 \leq i, j \leq k$, $n \vdash b_i(n),\ 1 \leq i \leq k$ sunt date. Ca de obicei, un asemenea sistem se asociază cu condiţia iniţială

$$(J) \qquad x_1(n_0) = x_1^0, ..., x_k(n_0) = x_k^0,$$

unde n_0 este un număr natural fixat, iar $x^0 = (x_1^0, ..., x_k^0)^\top$, un vector fixat din spaţiul R^k. Ne interesează în primul rând existenţa şi unicitatea soluţiei problemei $(S) + (J)$. În acest sens, să introducem notaţiile

(D1) $A(n) = (a_{ij}(n)), b(n) = (b_1(n), ..., b_k(n))^\top,$
$x(n) = (x_1(n), ..., x_k(n))^\top, \quad n \in N.$

Problema în chestiune devine atunci

$$(E) \qquad x(n+1) = A(n)x(n)$$

$$(I) \qquad x(n_0) = x^0.$$

Se deduce de aici că, dacă

(C1) $A(n)$ este nesingulară, pentru toţi $n \in N$,

problema considerată admite soluţie unică; adică, un singur sistem de şiruri $x(n) = (x_1(n), ..., x_k(n))$ care să o verifice. Trecem acum la structura mulţimii soluţiilor lui $(S) + (J)$. Pentru aceasta, vom asocia sistemul omogen

$$(S_0) \qquad \begin{cases} x_1(n+1) = a_{11}(n)x_1(n) + \cdots + a_{1k}(n)x_k(n) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_k(n+1) = a_{k1}(n)x_1(n) + \cdots + a_{kk}(n)x_k(n) \end{cases}$$

sau, vectorial (cu notaţiile anterioare)

$$(E_0) \qquad x(n+1) = A(n)x(n).$$

Vom nota prin X mulţimea soluţiilor din $(F(N,R))^k$ ale sistemului omogen (S_0) (respectiv, ale ecuaţiei vectoriale omogene (E_0)). Aceasta formează un subspaţiu liniar. Ne interesează şi aici dimensiunea sa. În acest sens, următorul rezultat (dat fără demonstraţie) va fi util:

Teorema 6.1.5 *Fie date k soluţii vectoriale*

$$y_1 = (y_{11}, ..., y_{1k})^\top, ..., y_k = (y_{k1}, ..., y_{kk})^\top$$

ale ecuaţiei (E_0). *Afirmaţiile următoare sunt echivalente:*

(P1) $\{y_1, ..., y_k\}$ *sunt liniar independente în* $(F(N,R))^k$

(P2) *matricea* $[y_1(n), ..., y_k(n)]$ *este nesingulară, în orice* $n \in N$

(P3) *matricea* $[y_1(n_0), ..., y_k(n_0)]$ *este nesingulară, pentru un* $n_0 \in N$.

Pe baza acestui rezultat avem acum

Teorema 6.1.6 *Mulţimea X a soluţiilor din $(F(N,R))^k$ ale sistemului omogen (S_0) (respectiv, ale ecuaţiei vectoriale omogene (E_0)) formează un subspaţiu k–dimensional al lui $(F(N,R))^k$.*

Demonstraţie. (Schiţă) Fie $n_0 \in N$ fixat. Pentru fiecare $i \in \{1, ..., k\}$, fie y^i soluţia din $(F(N,R))^k$ a problemei (vectoriale)

$$(E^i) \qquad x(n+1) = A(n)x(n), \; x(n_0) = e_i$$

(Aici, ca de obicei, $\{e_1, ..., e_k\}$ este baza canonică din R^k.) Din

$$[y^1(n_0), ..., y^k(n_0)] = [e_1, ..., e_k] = I \text{ (matricea unitate din } \mathcal{M}(k))$$

şi rezultatul anterior, $\{y^1, ..., y^k\}$ apare ca un sistem liniar independent în X. (Vom numi acesta *sistem fundamental de soluţii* ale sistemului omogen (S_0)). Fie $y = (y_1, ..., y_k)^\top$ soluţie arbitrară a aceluiaşi sistem. Avem, evident (din formula dezvoltării după o bază)

$$y(n_0) = y_1(n_0)y^1(n_0) + \cdots + y_k(n_0)y^k(n_0).$$

De aici, şi din observaţia privind unicitatea,

$$y = y_1(n_0)y^1 + \cdots + y_k(n_0)y^k.$$

Teorema este dovedită. ∎

Putem acum răspunde la întrebarea privind structura mulţimii soluţiilor sistemului cu diferenţe (S).

Teorema 6.1.7 *Formula*

$$\text{(R1)} \quad x(n) = C_1y^1(n) + \cdots + C_ky^k(n) + z(n), \; n \in N$$

unde $\{y^1, ..., y^k\}$ este un sistem fundamental de soluţii ale sistemului omogen (S_0), $(C_1, ..., C_k)^\top$ un vector arbitrar din R^k iar z o soluţie particulară a sistemului (neomogen) (S), reprezintă expresia soluţiei generale a sistemului (S).

O problemă naturală este, desigur, aceea a găsirii unei soluţii particulare a lui (S). Fie $\{y^1, ..., y^k\}$ un sistem fundamental de soluţii pentru (S_0). Căutăm respectiva soluţie particulară sub forma

$$z(n) = C_1(n)y^1(n) + \cdots + C_k(n)y^k(n),\ n \in N,$$

unde şirul vectorial $(C_1(n), ..., C_k(n))^\top$ urmează a fi determinat. Mai exact, să impunem condiţia

(C2) $\Delta C_1(n)y^1(n) + \cdots + \Delta C_k(n)y^k(n) = b(n),\ n \in N.$

Matricea sistemului este nesingulară; şi, deci, acesta are soluţia unică

$$\Delta C_1(n) = C_1(n+1) - C_1(n), ..., \Delta C_k(n) = C_k(n+1) - C_k(n),$$

ceea ce permite determinarea şirului vectorial $(C_1(n), ..., C_k(n))^\top$. Se vede uşor că $z(n)$ astfel definită verifică sistemul considerat; nu mai dăm alte detalii.

Un caz particular important este acela în care şirurile coeficient $a_{ij}(n)$, $1 \le i, j \le k$, sunt constante. Este vorba deci de sistemele cu diferenţe

$$(S) \qquad \begin{cases} x_1(n+1) = a_{11}(n)x_1(n) + \cdots + a_{1k}(n)x_k(n) + b_1(n) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_k(n+1) = a_{k1}(n)x_1(n) + \cdots + a_{kk}(n)x_k(n) + b_k(n) \end{cases}$$

sau vectorial (cu notaţia $A = (a_{ij})$)

$$(E) \qquad x(n+1) = Ax(n) + b(n).$$

Aspectul care trebuie subliniat aici este acela al determinării directe a unui sistem fundamental de soluţii pentru sistemul omogen

$$(S_0) \qquad \begin{cases} x_1(n+1) = a_{11}x_1(n) + \cdots + a_{1k}x_k(n) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_k(n+1) = a_{k1}x_1(n) + \cdots + a_{kk}x_k(n) \end{cases}$$

respectiv, pentru ecuaţia vectorială omogenă

$$(E_0) \qquad x(n+1) = Ax(n)$$

Mai exact, din ecuaţia însăşi avem

$$x(n+k) = A^k x(n),\ n \in N,\ k \ge 1.$$

Introducem polinomul caracteristic al lui A:

$$P(\lambda) = \det(A - \lambda I) = (-1)^k \lambda^k + r_1 \lambda^{k-1} + \cdots + r_{k-1}\lambda + r_k.$$

(Este important de observat că $r_k \neq 0$; deoarece matricea A este nesingulară, conform cu ipoteza (C1).) Deducem de aici, conform cu Teorema Cayley–Hamilton (Secţiunea 1.1.3):

$$(-1)^k x(n+k) + r_1 x(n+k-1) + \cdots + r_k x(n) = 0, \; n \in N.$$

Adică, toate şirurile componente $(x_1(n), ..., x_k(n))$ ale soluţiei lui (S_0) satisfac una şi aceeaşi ecuaţie cu diferenţe de ordin k, cu coeficienţi constanţi. Aceasta permite determinarea unui sistem fundamental de soluţii pentru (S_0) şi deci a soluţiei acestuia.

Exemplu. Să se rezolve sistemul cu diferenţe

$$x(n+1) = x(n) + 4y(n), \quad y(n+1) = x(n) + y(n).$$

Matricea sistemului şi polinomul ei caracteristic sunt

$$A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}; \quad P(\lambda) = \det \begin{pmatrix} 1-\lambda & 4 \\ 1 & 1-\lambda \end{pmatrix} = \lambda^2 - 2\lambda - 3.$$

Rezultă că ambele componente satisfac

$$x(n+2) - 2x(n+1) - 3x(n) = 0; y(n+2) - 2y(n+1) - 3y(n) = 0.$$

Ecuaţia caracteristică (dată de polinomul caracteristic) este

$$\lambda^2 - 2\lambda - 3 = 0; \quad \text{cu rădăcinile } \lambda_1 = -1, \lambda_2 = 3.$$

Rezultă forma acestor componente ale sistemului

$$x(n) = C_1(-1)^n + C_2 \cdot 3^n; \quad y(n) = C_3(-1)^n + C_4 \cdot 3^n$$

unde C_1, C_2, C_3, C_4 sunt constante. Pentru a reduce (la două) numărul acestora, verificăm sistemul iniţial cu aceste şiruri şi avem

$$\begin{cases} C_1(-1)^{n+1} + C_2 \cdot 3^{n+1} = C_1(-1)^n + C_2 \cdot 3^n + 4[C_3(-1)^n + C_4 \cdot 3^n] \\ C_3(-1)^{n+1} + C_4 \cdot 3^{n+1} = C_1(-1)^n + C_2 \cdot 3^n + C_3(-1)^n + C_4 \cdot 3^n. \end{cases}$$

Identificând coeficienţii pe lângă $(-1)^n$ şi 3^n avem

$$-C_1 = C_1 + 4C_3, 3C_2 = C_2 + 4C_4, -C_3 = C_1 + C_3, 3C_4 = C_2 + C_4.$$

Se obţine astfel soluţia

$$C_1 = -2\alpha,\ C_2 = 2\beta,\ C_3 = \alpha,\ C_4 = \beta$$

cu α, β constante arbitrare. Şi de aici soluţia sistemului în cauză

$$x(n) = -2\alpha(-1)^n + 2\beta \cdot 3^n, \quad y(n) = \alpha(-1)^n + \beta \cdot 3^n$$

sau, matricial,

$$\begin{pmatrix} x(n) \\ y(n) \end{pmatrix} = \begin{pmatrix} -2(-1)^n & 2 \cdot 3^n \\ (-1)^n & 3^n \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Am obţinut astfel şi un sistem fundamental de soluţii ale sistemului

$$v(n) = \begin{pmatrix} -2 \\ 1 \end{pmatrix} (-1)^n, \quad w(n) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} 3^n.$$

În ce priveşte determinarea unei soluţii particulare a sistemului neomogen (S), metoda generală a fost deja expusă în cazul neconstant. Există şi aici posibilitatea de găsire mai directă a acesteia, dacă termenul liber este de forma

$$b(n) = (R(n)\cos n\theta + S(n)\sin n\theta)\rho^n$$

unde $R(n), S(n)$ sunt polinoame cu coeficienţi vectoriali. Anume, se va căuta acea soluţie de o formă apropiată celei deja scrise; nu mai dăm alte detalii.

6.1.4 Aplicaţii la calculul puterilor unei matrici

Fie k număr natural dat şi

$$A = (a_{ij}), \quad \text{matrice (pătrată) din } \mathcal{M}(k).$$

Să notăm puterile acesteia în modul următor

$$A^n = (a_{ij}(n)), \quad n = 1, 2, \ldots\ .$$

Ne propunem în continuare să găsim expresia şirurilor $(a_{ij}(n)), 1 \leq i, j \leq k$, apărute în această reprezentare. În acest sens, plecăm de la polinomul caracteristic al matricei noastre

$$P(\lambda) = \det(A - \lambda I) = (-1)^k \lambda^k + r_1 \lambda^{k-1} + \cdots + r_{k-1}\lambda + r_k.$$

Conform Teoremei Cayley–Hamilton, matricea în chestiune satisface propria sa ecuaţie caracteristică

(P1) $(-1)^k A^k + r_1 A^{k-1} + \cdots + r_{k-1}A + r_k I = 0;$

sau, înmulţind cu A^n, (pentru n arbitrar)

(P2) $(-1)^k A^{n+k} + r_1 A^{n+k-1} + \cdots + r_{k-1}A^{n+1} + r_k A^n = 0,\ n \in N.$

Trecând la elemente, avem deci (pentru $i \leq i, j \leq k$)

$$(-1)^k a_{ij}(n+k) + r_1 a_{ij}(n+k-1) + \cdots + r_{k-1} a_{ij}(n+1) + r_k a_{ij}(n) = 0, n \in N.$$

Cu alte cuvinte,şirurile componente ale puterilor matricii A satisfac una şi aceeaşi ecuaţie liniară cu diferenţe de ordin k, omogenă. Fie

$$z_1(n), ..., z_k(n),\ n \in N$$

un sistem fundamental de soluţii ale acestei ecuaţii.Rezultă de aici reprezentarea şirurilor componente

(R1) $a_{ij}(n) = C_{ij}^{(1)} z_1(n) + C_{ij}^{(2)} z_2(n) + \cdots + C_{ij}^{(k)} z_k(n), n \in N,\ 1 \leq i, j \leq k,$

unde

$$C^{(1)} = (C_{ij}^{(1)}), ..., C^{(k)} = (C_{ij}^{(k)})$$

sunt matrici cu elemente constante. Avem, deci,

(R2) $A^n = z_1(n)C^{(1)} + z_2(n)C^{(2)} + \cdots + z_k(n)C^{(k)},\ n \in N.$

Matricile constante $C^{(1)}, ..., C^{(k)}$ se pot determina prin identificarea expresiilor anterioare cu $A^0 = I, A^1, ..., A^{k-1}$:

$$\begin{cases} z_1(0)C^{(1)} + z_2(0)C^{(2)} + \cdots + z_k(0)C^{(k)} = I \\ z_1(1)C^{(1)} + z_2(1)C^{(2)} + \cdots + z_k(1)C^{(k)} = A \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ z_1(k-1)C^{(1)} + z_2(k-1)C^{(2)} + \cdots + z_k(k-1)C^{(k)} = A^{k-1}. \end{cases}$$

În felul acesta, ajungem la reprezentarea puterilor matricii noastre.

Exemplu. Să se găsească o reprezentare a puterilor matricii $A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$. Polinomul caracteristic al matricii noastre este

$$P(\lambda) = \det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 1 \\ -1 & 3-\lambda \end{pmatrix} = \lambda^2 - 4\lambda + 4.$$

Rădăcinile acestui polinom sunt $\lambda_1 = \lambda_2 = 2$. Un sistem fundamental de soluţii asociate acestora este

$$z_1(n) = 2^n,\ z_2(n) = n2^n, \quad n \in N.$$

Urmează de aici reprezentarea

$$A^n = 2^n P + n2^n Q, \quad n \in N,$$

unde

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}, \quad Q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$$

sunt matrici care trebuie determinate. Avem

$$2^0 \cdot P + 0 \cdot 2^0 \cdot Q = I,\ 2^1 \cdot P + 1 \cdot 2^1 \cdot Q = A.$$

De aici rezultă imediat expresia acestor matrici

$$P = I,\ Q = \frac{1}{2}(A - 2I).$$

În final, ajungem la reprezentarea

$$A^n = 2^n \cdot I + n2^{n-1}(A - 2I),\ n \in N.$$

Probleme propuse la §6.1

1. Dacă $P(n) = a_0 n^k + a_1 n^{k-1} + \cdots + a_{k-1} n + a_k$ este un polinom de grad k să se arate că $\Delta^{k+1} P(n) = 0, \quad \forall n \in N$.

2. Fie S_n suma primilor n termeni ai şirului $0, 1, 1, 2, 2, 3, 3, 4, 4, ..$ Să se găsească o formulă pentru (S_n) şi să se arate apoi că

$$S_{p+q} - S_{p-q} = pq, \text{ oricare ar fi } p, q \in N \text{ cu } p > q.$$

(Indicaţie.Se va lua cazul $n =$ par, $n =$ impar.)

3. Să se rezolve ecuaţia liniară $x(n+1)-2x(n) = b(n)$ pentru următoarele forme ale termenului liber:

$$b(n) = 5^n,\ b(n) = 2^n,\ b(n) = \operatorname{arctg}(n),\ b(n) = (2n+3)2^n.$$

4. Să se determine acea soluţie $x(n)$ a ecuaţiei liniare

$$x(n+2) + \frac{5}{2}x(n+1) - x(n) = 0, \quad , n \in N,$$

cu proprietatea $\lim\limits_{n\to\infty} x(n) = 1$.

5. Să se rezolve ecuaţia liniară $x(n+3)-3x(n+1)+2x(n)=b(n)$ pentru următoarele forme ale termenului liber

$$b(n) = n+1, \quad b(n) = (n^2-1)2^n, \quad b(n) = 3^n.$$

6. Să se rezolve sistemele cu diferenţe finite
(a) $x(n+1) = x(n) - y(n),\ y(n+1) = x(n)$
(b) $x(n+1) = x(n) + y(n) + 2n,\ y(n+1) = x(n) - y(n) + 1$
(c) $x(n+1) = y(n) + z(n), y(n+1) = x(n) + z(n), z(n+1) = x(n) + y(n)$
(d) $x(n+1) = y(n) + 1,\ y(n+1) = z(n) + n,\ z(n+1) = x(n) + n^2$.
(Indicaţie.La sistemele neomogene se va folosi metoda eliminării.)

7. Să se calculeze puterile A^n $(n \geq 1)$, dacă

$$\text{(a)} \quad A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \qquad \text{(b)} \quad A = \begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix}$$

$$\text{(c)} \quad A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \text{(d)} \quad A = \begin{pmatrix} 0 & \alpha & \alpha \\ 1 & 0 & \alpha \\ 1 & 1 & 0 \end{pmatrix}$$

6.2 Elemente de programare dinamică

6.2.1 Noţiuni de bază

Programarea dinamică este o tehnică de rezolvare a unei clase speciale de probleme de optimizare, numite *probleme de decizie.* Acestea apar în studiul unor *sisteme controlabile* (adică, sisteme la care putem interveni în evoluţia lor prin decizii sau comenzi).

Să admitem că am definit un anume sistem (notat (SD)) prin intermediul unui model matematic. În cadrul acestuia, oricare stare (poziţie) a sistemului în discuţie este caracterizată de un anumit număr de parametri, numiţi şi *coordonate de fază* ale sistemului. Dacă notăm aceştia prin $\lambda_1, ..., \lambda_m$, atunci

$$x = (\lambda_1, ..., \lambda_m)^\top \text{ (element din } R^m)$$

va defini *starea* sistemului. Mulţimea tuturor acestor stări va fi o parte X din spaţiul R^m (numit *spaţiul fazelor* sistemului).

O ipoteză esenţială care o vom face aici este că momentele de timp în care se observă (se dirijează) sistemul sunt momente discrete, notate convenţional $0, 1, ..., k, ...$; mulţimea acestora este deci N. Ele se vor numi *etapele de evoluţie* ale sistemului. Să mai punem, pentru $k \in N$,

(D1) X_k = mulţimea stărilor sistemului la momentul k.

În ce priveşte evoluţia sistemului (SD), vom admite că aceasta nu poate avea loc fără o intervenţie exterioară. Din punct de vedere matematic, ea este descrisă de un anumit număr de parametri (numiţi *variabile de decizie*). Notând aceştia prin $\mu_1, ..., \mu_p$, va rezulta că vectorul

$$a = (\mu_1, ..., \mu_p)^\top \text{ (element din } R^p)$$

caracterizează decizia în cauză. Mulţimea tuturor acestor decizii va defini o parte A din R^p (numit *spaţiul deciziilor*). Să introducem convenţia, pentru $k \in N,\ x \in X_k$

$$\text{(D2)} \quad \begin{cases} A(k; x) = \text{mulţimea deciziilor ce pot fi luate asupra sistemului} \\ \text{în momentul } k, \text{ dacă starea lui (în acel moment) este } x. \end{cases}$$

Fiecare decizie $a \in A(k; x)$ are ca efect aducerea sistemului la momentul $k + 1$ într-o altă stare, notată $E(a; k, x)$. Cele două expresii introduse ne permit acum să determinăm

(D3) $\begin{cases} X_{k+1}(k;x) = \text{mulţimea stărilor sistemului la momentul } k+1 \\ \text{în ipoteza că, la momentul } k\text{, acesta se afla în starea } x. \end{cases}$

Anume, se poate scrie

(R) $X_{k+1}(k;x) = \{E(a;k,x);\ a \in A(k;x)\}$.

Acum, date momentul k şi starea $x \in X_k$, vom conveni ca aplicaţia
$a \vdash E(a;k,x)$ (de la $A(k;x)$ la $X_{k+1}(k;x)$)
să identifice element cu element aceste mulţimi. Adică, vom considera că între decizia $a \in A(k;x)$ şi efectul acesteia, definit de starea $E(a;k,x) \in X_{k+1}(k;x)$ nu putem face vreo distincţie. În felul acesta, vom putea interpreta elementul y din $X_{k+1}(k;x)$ atât ca stare a sistemului (obţinută prin aplicarea unei decizii de transfer) cât şi ca decizie de transfer a sistemului (în această stare). În fine,să mai notăm

(D4) $X_k(k+1;z) = \{y \in X_k;\ z \in X_{k+1}(k;y)\},\ z \in X_{k+1}, k \in N$.

Adică, $X_k(k+1;z)$ reprezintă mulţimea tuturor stărilor din X_k de la care se poate ajunge printr-o decizie de transfer, în starea $z \in X_{k+1}$.

6.2.2 Principiul de optimalitate Bellman

După cum s-a mai spus deja, evoluţia unui sistem de tip secvenţial este condusă prin decizii luate de către un observator exterior sistemului (numit *decident*). Pentru a motiva deciziile luate, este necesar să introducem un criteriu care să reflecte interesul decidentului pentru tranziţiile de stare ale sistemului considerat. În acest scop, vom ataşa fiecărui moment $k \in N$, o *funcţie de utilitate parţială* v_k, definită prin convenţia

(D1) $\begin{cases} v_k(x,y) = \text{utilitatea aducerii sistemului din starea } x \in X_k \\ \text{(în momentul } k\text{) la starea } y \in X_{k+1}(k;x) \text{ (în momentul } k+1\text{).} \end{cases}$

Următoarea problemă apare acum ca importantă. Fie r, un număr natural dat şi $x_0 \in X_0, x_r \in X_r$ două stări date.Numim *politică* de la x_0 lax_r, orice succesiune de stări

$$\Delta(x_0, x_r) = \{x_0, x_1, ..., x_{r-1}, x_r\}$$

cu proprietăţile

$$x_{k+1} \in X_{k+1}(k;x_k),\ 0 \le k \le r-1.$$

Oricărei asemenea politici îi putem asocia o *utilitate totală*

(D2) $v(\Delta(x_0, x_r)) = v_0(x_0, x_1) + v_1(x_1, x_2) + \cdots + v_{r-1}(x_{r-1}, x_r)$.

Problema enunţată este acum următoarea

$$\text{(P)} \quad \begin{cases} \text{În mulţimea politicilor de la } x_0 \text{ la } x_r\text{, să se găsească} \\ \text{aceea de utilitate totală optimă (maximă sau minimă).} \end{cases}$$

Înainte de a indica modul de rezolvare a problemei este necesar în prealabil să punem în evidenţă o proprietate importantă a politicilor extremale. Dată politica $\Delta(x_0, x_r)$ ca mai sus, fie i, j cu $0 \leq i < j \leq r$ două momente intermediare şi $x_i \in X_i$, $x_j \in X_j$ stările corespunzătoare din $\Delta(x_0, x_r)$. Vom numi *sub–politică* (a politicii iniţiale) relativ la aceste stări (şi momente) succesiunea de stări (din $\Delta(x_0, x_r)$)

$$\Delta(x_i, x_j) = (x_i, x_{i+1}, ..., x_{j-1}, x_j).$$

(Cu alte cuvinte, o sub–politică nu este altceva decât un ”segment” al unei politici).

Următorul rezultat cunoscut sub numele de ”Principiul de Optimalitate Bellman” este fundamental în teoria pe care o prezentăm aici:

Teorema 6.2.1 *Dacă politica $\Delta(x_0, x_r)$ este optimală (între x_0 şi x_r), atunci orice sub–politică a acesteia, $\Delta(x_i, x_j)$ este, de asemenea, optimală (între x_i şi x_j).*

Demonstraţie. Să presupunem că sub–politica

$$\Delta(x_i, x_j) = (x_i, x_{i+1}, ..., x_{j-1}, x_j)$$

n-ar fi optimală, între x_i şi x_j. Ar exista atunci o altă sub–politică (între x_i şi x_j)

$$\Delta^*(x_i, x_j) = (x_i, x^*_{i+1}, ..., x^*_{j-1}, x_j)$$

pentru care utilitatea total 2 $v(\Delta^*(x_i, x_j))$ ar fi mai bună (respectiv: mai mică în cazul minimului şi mai mare în cazul maximului). Dar atunci, politica (între x_0 şi x_r)

$$\Delta^*(x_0, x_r) = (x_0, ..., x_i, x^*_{i+1}, ..., x^*_{j-1}, x_j, x_{j+1}, ..., x_r)$$

ar fi, la rândul ei, mai bună decât politica $\Delta(x_0, x_r)$, ceea ce este imposibil prin ipoteză. Deci, $\Delta(x_i, x_j)$ este optimală. ∎

Cu alte cuvinte, o politică optimală este compusă numai din sub–politici (segmente) optimale. Deci, dacă măcar una din sub–politicile unei politici date nu este optimală, atunci nici politica în ansamblu nu este optimală.

Prin aceleaşi tehnici ca mai sus se poate demonstra următoarea extensie a principiului anterior:

Teorema 6.2.2 *Fie $x_0 \in X_0,\ x_i \in X_i,\ x_j \in X_j,\ x_r \in X_r$, patru stări date, cu $0 \le i \le j \le r$ şi $\Gamma(x_i, x_j)$, o politică dată (optimală sau nu) de la x_i la x_j. În mulţimea politicilor de la x_0 la x_r care conţin $\Gamma(x_i, x_j)$ ca sub–politică, cea mai bună este aceea care se obţine completând $\Gamma(x_i, x_j)$ cu o sub–politică optimală de la x_0 la x_i şi o sub–politică optimală de la x_j la x_r.*

În particular, dacă $i = 0, j = r$, rezultatul enunţat se poate considera identic cu precedentul.

Să introducem în final încă o convenţie. Fie $x_0 \in X_0$ stare dată. Numim *politică cu punct de plecare* x_0, orice succesiune

$$\Delta(x_0) = (x_0, x_1, ..., x_n, ...)$$

în care

$$x_{k+1} \in X_{k+1}(k; x_k),\ \ k = 0, 1, 2, ...$$

Să notăm că, datorită infinităţii de stări cu care se operează, noţiunea de optimalitate definită anterior este aplicabilă aici doar în mod secvenţial. Anume, vom zice că o asemenea politică este *optimală* dacă, pentru orice n, sub–politica sa $\Delta(x_0, ..., x_n)$ este optimală (în sensul deja cunoscut). Alte aspecte vor fi discutate ulterior.

6.2.3 Formule de recurenţă. Cazul orizontului finit

Principiul de optimalitate formulat mai înainte ne permite să construim în mod recurent politicile optimale. Aceasta se poate face în două moduri posibile:

(i) în sens crescător al momentelor (analiză prospectivă)

(ii) în sens descrescător al momentelor (analiză retrospectivă).

(A) Cu privire la prima variantă, situaţia se prezintă astfel. Fie $r \ge 0$ un moment dat şi $S \subseteq X_r$, o mulţime de stări fixată. Pentru fiecare alt moment $k > r$ şi fiecare $x \in X_k$, să notăm

(D1) $\begin{cases} W_{r,k}(S;x) = \text{utilitatea optimă a politicilor de la una din} \\ \text{stările lui } S \text{ la starea } x. \end{cases}$

Ne interesează în continuare construcţia mărimilor

$$W_{r,k}(S;x),\ k > r,\ x \in X_k.$$

Desigur, pentru $k = r + 1$, avem

(R1) $W_{r,r+1}(S;x) = \text{opt}\{v_r(z,x); z \in S \cap X_r(r+1;x)\}$.

Se pune problema ce se întîmplă pentru $k \geq r + 2$. În acest sens, următorul rezultat se dovedeşte esenţial.

Teorema 6.2.3 *Avem relaţiile de recurenţă* (*pentru* $k \geq r + 2$)

(R2) $W_{r,k}(S;x) = \text{opt}\{W_{r,k-1}(S,y) + v_{k-1}(y,x); y \in X_{k-1}(k,x)\}$.

Demonstraţie. Fie (cu notaţiile anterioare)

$$\Delta(x_r, x) = (x_r, x_{r+1}, ..., x_{k-1}, x) \ \text{ cu } \ x_r \in S,$$

acea politică având ca punct de sosire starea x, pentru care se realizează optimul $W_{r,k}(S;x)$.Conform Principiului de Optimalitate Bellman, sub–politica

$$\Delta(x_r, x_{k-1}) = (x_r, x_{r+1}, ..., x_{k-1})$$

trebuie să fie, de asemenea, optimală printre toate sub–politicile care au $x_r \in S$ drept stare iniţială şi x_{k-1}, drept stare finală. Ca atare, valoarea ei este obligatoriu $W_{r,k-1}(S;x_{k-1})$ şi deci

$$v(\Delta(x_r, x)) = W_{r,k-1}(S;x_{k-1}) + v_{k-1}(x_{k-1}, x).$$

Pe de altă parte, oricare ar fi starea $x'_{k-1} \in X_{k-1}(k;x)$ şi sub–politica

$$\Delta'(x'_r, x'_{k-1}) = (x'_r, x'_{r+1}, ..., x'_{k-1}),\ x'_r \in S,$$

cu $v(\Delta'(x'_r, x'_{k-1})) = W_{r,k-1}(S;x'_{k-1})$, este clar că politica

$$\Delta'(x'_r, x) = (x'_r, x'_{r+1}, ..., x'_{k-1}, x)$$

are, în cel mai bun caz, utilitatea totală

$$v(\Delta'(x'_r, x)) = W_{r,k-1}(S;x'_{k-1}) + v_{k-1}(x'_{k-1}, x)$$

identică cu aceea a politicii $\Delta(x_r, x)$. Aceasta încheie argumentul. ∎

Rezultatul obţinut permite obţinerea, din aproape în aproape a mărimilor care ne interesează. Astfel, de pildă, pentru $k = r + 2$

(R3) $W_{r,r+2}(S;x) = \text{opt}\{W_{r,r+1}(S;y) + v_{r+1}(y,x); y \in X_{r+1}(r+2;x)\}$.

Să notăm cu această ocazie că, dacă $T \subseteq X_k$ este dată şi punem

(D2) $W_{r,k}(S;T)$ = valoarea otimă a politicilor de la una din stările lui S la una din stările lui T,

atunci, cu cele spuse dinainte, este clar că

(R4) $W_{r,k}(S;T) = \text{opt}\{W_{r,k}(S,x);\ x \in T\}$.

Avem astfel posibilitatea să calculăm şi astfel de valori optime între mulţimi de stări.

(B) Analog se prezintă lucrurile şi în sens retrospectiv.Anume, fie $k > 0$ un moment dat şi $T \subseteq X_k$ o mulţime de stări fixată. Pentru fiecare alt moment $r < k$ şi fiecare $x \in X_r$, să notăm

(D3) $\begin{cases} \overline{W}_{r,k}(x,T) = \text{utilitatea optimă a politicilor} \\ \text{de la starea } x \in X_r \text{ la o stare din } T. \end{cases}$

Ne interesează în continuare mărimile

$$\overline{W}_{r,k}(x,T),\ 0 \le r < k,\ x \in X_r.$$

Pentru $r = k-1$ avem, evident,

(R5) $\overline{W}_{k-1,k}(x,T) = \text{opt}\{v_{k-1}(x,z);\ z \in T \cap X_k(k-1;x)\}$.

Pentru $r \le k-2$ calculul mărimilor în cauză se face în conformitate cu următorul rezultat (simetric faţă de precedentul) pe care îl dăm fără demonstraţie:

Teorema 6.2.4 *Avem relaţiile de recurenţă*

(R6) $\overline{W}_{r,k}(x,T) = \text{opt}\{v_r(x,y) + \overline{W}_{r+1,k}(y,T);\ y \in X_{r+1}(r;x)\}$.

Deci, din aproape în aproape, toate mărimile în cauză pot fi calculate. În particular, dacă S este o mulţime de stări din X_r, atunci (cu notaţia deja introdusă) putem, de asemenea, scrie

(R7) $\overline{W}_{r,k}(S,T) = \text{opt}\{\overline{W}_{r,k}(x,T);\ x \in S\}$.

De aici tragem concluzia cu caracter general că, în fapt, analiza prospectivă are aceeaşi valoare ca şi analiza retrospectivă în studiul politicilor optimale.

Exemplu. Să considerăm sistemul secvenţial definit de mulţimea de stări (în fiecare moment)

$$X_k = [0, \infty), \text{ pentru } k = 0, 1, 2, ...$$

Mulţimea deciziilor asupra fiecărei stări (sau, echivalent, a rezultatelor acestora), va fi definită prin convenţia

$$X_{k+1}(k; x) = \{y \in X_{k+1}; 2^k x \leq y\}, \ k \in N, \ x \in X_k.$$

Să mai definim şi utilităţile tranziţiilor de stare. Anume, dat un şir (p_n) strict descrescător de numere din $(0, 1)$ să punem, pentru $k \in N$

$$v_k(x, y) = p_k|x - y|, \ x \in X_k, \ y \in X_{k+1}(k; x).$$

Fie acum $S = [0, 1]$ o mulţime de stări din X_0. Ne interesează, pentru fiecare $k \geq 1$, expresiile funcţiilor

$$W_k(x) \quad = \quad \begin{array}{l} \text{valoarea minimă a unei politici cu originea} \\ \text{într-o stare din } S \text{ şi cu extremitatea în starea } x \in X_k. \end{array}$$

În primul rând avem, cu convenţiile anterioare

$$W_1(x) = \min\{p_0(x - y); 0 \leq y \leq \min(1, x)\}, \quad x \in [0, \infty),$$

ceea ce ne dă

$$W_1(x) = \begin{cases} 0, & \text{dacă} \quad 0 \leq x \leq 1 \\ p_0(x - 1), & \text{dacă} \quad x > 1. \end{cases}$$

Apoi, cu formula de recurenţă din cazul prospectiv

$$W_2(x) = \min\{W_1(y) + p_1(x - y); \ 0 \leq y \leq x/2\}, \ x \in [0, \infty).$$

Avem, din cele spuse anterior,

$$W_1(y) - p_1 y = \begin{cases} -p_1 y, & 0 \leq y \leq 1 \\ (p_0 - p_1)y - p_0, & y > 1. \end{cases}$$

Şi atunci, obţinem formula de reprezentare

$$W_2(x) = \begin{cases} p_1 x/2, & 0 \leq x \leq 2 \\ p_1(x - 1), & x > 2. \end{cases}$$

Procedeul poate astfel continua în mod indefinit. Se găsesc astfel toate aceste funcţii (W_k).

6.2.4 Ecuaţia funcţională a programării dinamice

Ne vom ocupa în continuare de construirea unei politici optimale în cazul orizontului infinit. Fie din nou (SD) un sistem de tip secvenţial definit de elementele

(a) mulţimea stărilor: $X_0, X_1, ...,$

(b) mulţimea deciziilor ce se pot lua asupra unei stări la un moment dat: $X_{k+1}(k; x)$, $k = 0, 1, ..., x \in X_k$,

(c) utilitatea unei tranziţii de stare: $v_k(x, y)$, pentru $x \in X_k$, $y \in X_{k+1}(k; x)$, $k = 0, 1, ...,$.

Fie $x \in X_0$ stare fixată. Am numit politică având ca punct de plecare starea x, orice succesiune

$$\Delta(x) = (x, x_1, ..., x_n, ...)$$

în care

$$x_1 \in X_1(0; x), x_2 \in X_2(1; x_1), ..., x_n \in X_n(n-1; x_{n-1}), ...$$

De asemenea, am definit o asemenea politică drept optimală dacă

$$\Delta(x, x_n) = (x, x_1, ..., x_n) \text{ este optimală pentru orice } n.$$

În continuare, din motive de comoditate vom lucra în cazul staţionar, caracterizat de

(C1) $X_k = X_h$, pentru toţi k, h. (Mulţimile de stări ale sistemului sunt aceleaşi.) Fie X valoarea comună a lor.

(C2) $X_{k+1}(k; x) = X_{h+1}(h; x)$, pentru toţi k, h. (Efectul unei schimbări de stare nu depinde de momentul ales.) Fie $X(x; +)$ valoarea comună a acestora.

(C3) $v_k(x, y) = v_h(x, y)$, pentru toţi k, h. (Utilitatea unei tranziţii de stare nu depinde de momentul ales.) Fie $v(x, y)$ valoarea comună a lor.

Să introducem acum mărimile, pentru $k < n$,

$$\text{(D1)} \quad \begin{cases} W_{k,n}(x)\text{=valoarea optimă a politicilor de la starea} \\ \text{iniţială } x \in X_k \text{ la o stare din } X_n. \end{cases}$$

În urma acestor ipoteze, mărimile în cauză nu depind decât de diferenţa $n-k$ şi de starea x. Putem deci nota

(D2) $F_n(x) = W_{k,k+n}(x)$, pentru un $k \in N$.

Definiţia e consistentă (nu depinde de k). Avem deci pentru orice x

$$\text{(D3)} \quad \begin{cases} F_n(x) & = & \text{utilitatea optimă a unui număr de } n \text{ tranziţii} \\ & & \text{de stare, începând cu starea } x. \end{cases}$$

Fie acum dată o politică optimală $\Delta(x)$ (corespunzând unei mulţimi numărabile de tranziţii de stare). Conform definiţiei, avem pentru orice rang n

(R1) $F_n(x) =$ valoarea (optimală) a sub–politicii $\Delta(x, x_n)$.

Intuitiv, este normal să acceptăm că, dacă

(C4) $F(x) = \lim\limits_{n\to\infty} F_n(x)$ există, $\forall x \in X$

atunci, (prin convenţie)

(R2) $F(x) =$ valoarea optimală a politicii $\Delta(x)$, $\quad \forall x \in X$.

Să notăm că, datorită relaţiilor de recurenţă (din cazul orizontului finit), avem (cu notaţia $F_0(x) = 0,\ x \in X$)

(R3) $F_{n+1}(x) = \text{opt}\{v(x, y) + F_n(y);\ y \in X(x; +)\},\ x \in X,\ n \in N.$

Deci, trecând formal la limită (pentru $n \longrightarrow \infty$) ajungem la o caracterizare a valorii optimale în chestiune:

(RB) $F(x) = \text{opt}\{v(x, y) + F(y);\ y \in X(x; +)\},\ x \in X.$

Vom numi aceasta *ecuaţia funcţională a programării dinamice* datorată lui Bellman.

Să precizăm în continuare unele condiţii care dau legitimitate operaţiilor de mai sus. În acest sens, admitem deocamdată ipoteza

(C5) $0 \in X\ , X(0; +) = \{0\}\ , v(0, 0) = 0.$

Facem notaţia

(D4) $X[\lambda] = \{x \in X;\ \|x\| \leq \lambda\},\ \lambda \geq 0.$

Să mai admitem încă o ipoteză de forma

(C6) $\{v(x, y);\ x \in X[\lambda], y \in X(x; +)\}$ este mărginită, $\forall \lambda \geq 0$.

În acest caz, funcţia $\varphi : [0, \infty) \longrightarrow [0, \infty)$ dată de

(D5) $\varphi(\lambda) = \sup_{x \in X[\lambda]} \sup\{v(x, y); y \in X(x; +)\}, \ \lambda \geq 0$

este bine definită şi crescătoare pe $[0, \infty)$. În plus,avem şi $\varphi(0) = 0$ (datorită lui (C3)). Mai notăm,în fine,şi evaluarea utilă

(R4) $v(x, y) \leq \varphi(\|x\|), \quad x \in X, \ y \in X(x; +)$.

Putem acum da răspunsul la întrebarea pusă anterior.

Teorema 6.2.5 *Să presupunem că există un $\alpha \in (0, 1)$ cu proprietăţile*

(C7) $\sup_{y \in X(x;+)} \|y\| \leq \alpha \|x\|$, *pentru orice* $x \in X$

(C8) $\sum_n \varphi(\alpha^n \lambda)$ *este convergentă, pentru orice* $\lambda \geq 0$.

Au loc concluziile:
(P1) *Funcţiile* ($F_n : X \longrightarrow R; \ n = 1, 2, ...$) *introduse prin formula* (D2) *sunt continue în origine, şi nule în acest punct. Şirul* (F_n) *converge uniform pe* $X[\lambda]$ (*oricare ar fi* $\lambda \geq 0$) *către o funcţie* $F : X \longrightarrow R$, *continuă în origine şi nulă în acest punct, care satisface ecuaţia funcţională* (RB).
(P2) *Funcţia* F *astfel definită este unică în raport cu proprietăţile menţionate; adică ecuaţia* (RB) *are o singură soluţie în clasa funcţiilor de la* X *la* R *care sunt continue în origine şi nule în acest punct.*

Demonstraţie. (Schiţă) Se arată mai întâi că, dacă $G : X \longrightarrow R$, $H : X \longrightarrow R$ sunt două funcţii date, şi

$$G^*(x) = \text{opt}\{v(x, y) + G(y); \ y \in X(x; +)\}, \ x \in X,$$
$$H^*(x) = \text{opt}\{v(x, y) + H(y); \ y \in X(x; +)\}, \ x \in X,$$

atunci în mod necesar

(R5) $|G^*(x) - H^*(x)| \leq \sup_{y \in X(x;+)} |G(y) - H(y)|, \ x \in X$.

Acest fapt,combinat cu (R3)+(R4),ne dă (din (C7))

(R6) $|F_{n+1}(x) - F_n(x)| \leq \varphi(\alpha^n \|x\|), \quad x \in X, \ n \geq 0$.

Deci,cu necesitate,

(R7) $\sup_{x \in X[\lambda]} |F_{n+1}(x) - F_n(x)| \leq \varphi(\alpha^n \lambda), \quad \lambda \geq 0.$

Aceasta, cu Criteriul de comparaţie Weierstrass (Secţiunea 4.4.3) arată că şirul ($F_n : X \longrightarrow R;\ n = 1, 2, ...$) converge uniform pe $X[\lambda]$, oricare ar fi $\lambda \geq 0$. Pe de altă parte, pentru orice $n \geq 1$,

(R8) F_n este continuă în $x = 0$ şi $F_n(0) = 0$,

după cum rezultă imediat din (C8) (termenul general al unei serii convergente tinde la zero). Ca atare, funcţia limită

$$F(x) = \lim_{n \to \infty} F_n(x),\ x \in X$$

este, de asemenea, continuă în $x = 0$ şi nulă în acest punct. În plus, din proprietăţile operaţiei ”opt” este clar că F satisface (RB). În fine, dacă F şi G ar fi două soluţii ale lui (EB), este evident, în conformitate cu evaluarea (R5), că $F = G$. Teorema este astfel demonstrată. ■

Probleme propuse la §6.2

1. Fie un sistem dinamic de tip secvenţial definit prin

a) mulţimile de stare $X_i = [0, 1],\ i = 0, 1, 2, ...;$

b) mulţimile de tranziţie $X_{i+1}(i; x) = \{y \in X_{i+1};\ x^2 \leq y\}$, $x \in X_i$, unde $i = 0, 1, 2, ...;$

c) utilităţile tranziţiilor de stare $v_i(x, y) = \dfrac{1}{i+1}|2x - y|$, $x \in X_i$, $y \in X_{i+1}(i; x)$, $i = 0, 1, 2, ...$.

Se cere să se calculeze valorile optime (maximă sau minimă ale politicilor $\Delta = (x_0, x_1, x_2, x_3)$ în condiţiile

i) $x_0 \in [0, \frac{1}{2}]$ iar x_3 este nedeterminată;

ii) x_0 este nedeterminată iar $x_3 \in [0, 1]$;

iii) x_1, x_2 sunt stări fixate din X_1, respectiv X_2.

2. Fie un sistem dinamic de tip secvenţial definit prin elementele

a) mulţimile de stare $X_i = [0, \infty],\ i = 0, 1, 2, ...;$

b) mulţimile de tranziţie $X_{i+1}(i; x) = \{y \in X_{i+1};\ 2^i y \leq x\}$, $x \in X_i$, unde $i = 0, 1, 2, ...;$

c) utilităţile tranziţiilor de stare $v_i(x, y) = x + y$, $x \in X_i$, $y \in X_{i+1}(i; x)$, $i = 0, 1, 2, ...$.

Să se arate că, pentru orice $x \in X_0$, politica $\Delta = (x, 2^{-1}x, 2^{-2}x, ...)$ este maximală printre toate politicile de forma $\Gamma = (x, x_1, x_2, ...)$, $x_1 \in X_1(0; x), x_2 \in X_2(1; x_1), ...$ Să se calculeze efectiv valoarea politicii Δ.

6.3 Procese de tip continuu (Ecuaţii diferenţiale)

6.3.1 Punerea problemei

Să considerăm dată o funcţie $f : \Omega \longrightarrow R$, unde Ω este un anumit domeniu din R^2. Fie, de asemenea, $P(t_0, x_0)$ un punct din Ω. Considerăm relaţiile:

(E) $\quad x' = f(t, x)$

(I) $\quad x(t_0) = x_0.$

Prima dintre ele se numeşte *ecuaţie diferenţială de ordinul întâi*. Sensul acestora este următorul: Să se găsească acele funcţii derivabile $t \vdash x(t)$ care satisfac condiţia iniţială (I) şi

$$(t, x(t)) \in \Omega,\ x'(t) = f(t, x(t)), \quad \forall t \in \Gamma$$

unde $\Gamma = \Gamma(P)$ este un interval deschis al axei reale, ce depinde de punctul consderat. Din analiza unor exemple simple (ce le vom trata ulterior) se poate observa că mulţimea soluţiilor lui (E) depinde de o constantă. A fost, deci, natural să asociem lui (E) condiţia iniţială (I), pentru a individualiza acea funcţie. Ansamblul (E) + (I) se va numi *problemă Cauchy de ordinul întâi.*

Analog, fie dată funcţia $f : \Omega \longrightarrow R$, unde Ω este un domeniu din R^{k+1}. Considerăm relaţiile

$(\mathrm{E_k})$ $x^{(k)} = f(t, x, x', ..., x^{(k-1)})$

$(\mathrm{I_k})$ $x(t_0) = \alpha_0, x'(t_0) = \alpha_1, ..., x^{(k-1)}(t_0) = \alpha_{k-1}$.

Vom numi $(\mathrm{E_k})$ *ecuaţie diferenţială de ordinul* k. Interpretarea acesteia este analogă cu cea precedentă. Se poate constata şi aici prin exemple că soluţiile acestei ecuaţii depind de k constante arbitrare. De aceea, apare necesitatea impunerii unor condiţii iniţiale de tipul $(\mathrm{I_k})$, pentru a individualiza soluţiile. Ansamblul $(\mathrm{E_k})$ + $(\mathrm{I_k})$ se va numi *problemă Cauchy de ordinul* k.

În fine, ca o extensie a lui (E) + (I), fie date funcţiile $f_i : \Omega \longrightarrow R$, $1 \le i \le k$, unde Ω este un domeniu din R^{k+1} şi un punct $P(t_0, x_1^0, ..., x_k^0)$ din Ω. Considerăm relaţiile

$(\mathrm{S_k})$ $x_i' = f_i(t, x_1, ..., x_k),\ 1 \le i \le k$

$(\mathrm{J_k})$ $x_i(t_0) = x_i^0,\ 1 \le i \le k$.

Primul ansamblu de egalităţi se numeşte *sistem diferenţial de ordinul întâi.* Interpretarea lui $(\mathrm{S_k})$ + $(\mathrm{J_k})$ este următoarea: Să se găsească funcţiile derivabile $t \vdash x_i(t),\ 1 \le i \le k$, care satisfac $(\mathrm{J_k})$ şi

$$(t, x_1(t), ..., x_k(t)) \in \Omega,\ x_i'(t) = f_i(t, x_1(t), ..., x_k(t)),\ t \in \Gamma, 1 \le i \le k,$$

unde $\Gamma = \Gamma(P)$ este un interval deschis al axei reale ce depinde de punctul P. Soluţiile $(x_1, ..., x_k)$ ale acestui sistem depind de k constante arbitrare. De aceea, pentru găsirea unui anume ansamblu de soluţii a fost necesar să cuplăm $(\mathrm{S_k})$ cu $(\mathrm{J_k})$; vom numi şi aici acest cuplu, *problemă Cauchy.* Este important de subliniat că $(\mathrm{S_k})$ reprezintă de asemenea şi o extensie a lui $(\mathrm{E_k})$. În adevăr, dat sistemul diferenţial

$(\mathrm{S_k^*})$ $x_1' = x_2, x_2' = x_3, ..., x_{k-1}' = x_k, x_k' = f(t, x_1, ..., x_k)$

componenta x_1 din soluţia acestuia verifică $(\mathrm{E_k})$. Reciproc, dată ecuaţia diferenţială $(\mathrm{E_k})$, este clar că funcţiile $(x, x', ..., x^{(k-1)})$ sunt componentele soluţiei lui $(\mathrm{S_k^*})$.

Ecuaţiile diferenţiale reprezintă un instrument de bază al modelării diferitelor fenomene cu caracter dinamic. O parte a acestora va fi abordată şi în cele ce urmează.

6.3.2 Câteva tipuri elementare de ecuaţii diferenţiale

Să trecem acum în revistă unele tipuri ”elementare” de ecuaţii diferenţiale; adică, ecuaţii rezolvabile prin intermediul unor metode de calcul integral. Desigur, nu este în intenţia noastră să prezentăm în detaliu toate condiţiile tehnice care permit aceste rezolvări; dorim numai să punem în evidenţă metoda ca atare.

(A). Începem cu ecuaţiile diferenţiale de ordinul întâi de forma

$$(ES) \qquad x' = a(t)b(x)$$

numite *ecuaţii cu variabile separabile.* Aici, a şi b sunt funcţii continue; iar a doua dintre ele este nenulă. Pentru rezolvare, scriem

$$\frac{dx}{dt} = a(t)b(x) \quad \text{sau} \quad \frac{dx}{b(x)} = a(t)dt.$$

Notând acum

$$B(x) = \int \frac{dx}{b(x)}, \quad A(t) = \int a(t)dt,$$

ecuaţia devine, prin integrare,

$$B(x) = A(t) + C; \quad \text{deci} \quad x = B^{-1}(A(t) + C).$$

(Aici, B^{-1} desemnează inversa funcţiei B.) Constanta C se determină dacă se mai adaugă la ecuaţia în cauză şi o condiţie iniţială.

Exemplu. Să se rezolve problema Cauchy

$$x' = -\frac{1}{t}\,\frac{1+x^2}{x}, \ (t > 0, x > 0), \quad x(1) = 1.$$

Vom scrie ecuaţia sub forma

$$\frac{x\,dx}{1+x^2} = -\frac{dt}{t} \Longrightarrow \int \frac{x\,dx}{1+x^2} = -\int \frac{dt}{t}.$$

Efectuând integralele, obţinem

$$\frac{1}{2}\ln(1+x^2) = -\ln t + C \Longrightarrow \ln(1+x^2) = \ln\frac{1}{t^2} + 2C.$$

Se deduce $1+x^2 = K\dfrac{1}{t^2}$, unde $K = e^{2C}$. Folosind şi condiţia iniţială avem $K = 2$. Deci soluţia ecuaţiei este dată de formula

$$1+x^2 = \frac{2}{t^2} \text{ sau } x(t) = \sqrt{\frac{2}{t^2} - 1}, 0 < t \le \sqrt{2}.$$

(B). Să trecem acum la ecuaţiile de forma

$$(EO) \qquad x' = f\left(\frac{x}{t}\right),$$

numite *ecuaţii omogene*. Aici, f este o funcţie continuă. Pentru rezolvarea lor se face schimbarea de funcţie

$$x = ty \text{ (deci } x' = y + ty').$$

Prin înlocuire se obţine o ecuaţie (în funcţia y)

$$y + ty' = f(y) \text{ sau } y' = \frac{1}{t}(f(y) - y),$$

care, după cum se vede, este cu variabile separabile. Rezolvând aceasta după metodologia deja indicată şi revenind la substituţie, ajungem să găsim soluţia ecuaţiei.

Exemplu. Să se rezolve ecuaţia

$$x' = \frac{x^2 + t^2}{xt} = \frac{x}{t} + \frac{t}{x}, \quad (t > 0, x > 0).$$

Cu substituţia $x = ty$ (deci $x' = y + ty'$), ecuaţia devine

$$y + ty' = y + \frac{1}{y}, \quad \text{sau} \quad y' = \frac{1}{t} \cdot \frac{1}{y}\cdot$$

Scriem acum ecuaţia sub forma

$$\frac{dy}{dt} = \frac{1}{t} \cdot \frac{1}{y} \Longrightarrow y\,dy = \frac{dt}{t} \Longrightarrow \int y\,dy = \int \frac{dt}{t}\cdot$$

Prin integrare găsim $\frac{y^2}{2} = \ln t + C$ adică, $y^2 = 2\ln t + K$, unde $K = 2C$. De aici (cu condiţia pusă), $y = \sqrt{2\ln t + K}$; şi, revenind la susbstituţie,

$$x(t) = t\sqrt{2\ln t + K}, \ t > 0, \ K = \text{constantă arbitrară}.$$

(C). Un exemplu important de ecuaţie diferenţială de ordinul întâi este

$$(EL) \qquad x' = a(t)x + b(t)$$

numită *ecuaţie diferenţială liniară.* Aici, a şi b sunt două funcţii continue pe un acelaşi interval. Să notăm pentru simplitate

$$A(t) = \int a(t)dt \quad \text{(primitiva funcţiei } a\text{)}.$$

Înmulţind ecuaţia cu $e^{-A(t)}$ obţinem

$$x'(t)e^{-A(t)} - x(t)a(t)e^{-A(t)} = e^{-A(t)}b(t)$$

sau, observând că în stânga avem o derivată,

$$\frac{d}{dt}(x(t)e^{-A(t)}) = e^{-A(t)}b(t).$$

Integrând termen cu termen egalitatea, avem

$$x(t)e^{-A(t)} = \int e^{-A(t)}b(t)dt + C.$$

(Am pus separat constanta de integrare în dreapta.) Deci

$$x(t) = e^{A(t)}\left(C + \int e^{-A(t)}b(t)dt\right).$$

Exemplu. Să se rezolve problema Cauchy

$$x' = \frac{1}{t}x + t^3, \ (t > 0), \quad x(1) = 0.$$

Aplicăm formula anterioară de rezolvare. Avem

$$A(t) = \int \frac{1}{t}dt = \ln t \Longrightarrow e^{A(t)} = e^{\ln t} = t, \ e^{-A(t)} = \frac{1}{t}\cdot$$

Deci, soluţia generală este

$$x = t\left(C + \int \frac{1}{t}t^3 dt\right) = t\left(C + \int t^2 dt\right) = t\left(C + \frac{t^3}{3}\right)\cdot$$

Din $x(1) = 0$ avem $C + \frac{1}{3} = 0$ deci $C = -\frac{1}{3}\cdot$ Şi atunci, soluţia problemei este funcţia $x(t) = \frac{t}{3}(t^3 - 1), \quad t > 0.$

(D). Un exemplu important de ecuaţie reductibilă la cea liniară este

$$(EB) \qquad x' = a(t)x + b(t)x^\alpha$$

numită *ecuaţie Bernoulli.* Aici, a , b au aceeaşi semnificaţie ca mai sus, iar α este un număr real. Împărţind cu x^α, obţinem

$$\frac{x'}{x^\alpha} = a(t)x^{1-\alpha} + b(t).$$

Aceasta sugerează să facem schimbarea de funcţie

$$x^{1-\alpha} = y; \text{ deci } (1-\alpha)\frac{x'}{x^\alpha} = y'.$$

Înlocuind în ecuaţia iniţială (după ce în prealabil înmulţim cu factorul $1-\alpha$), obţinem

$$y' = (1-\alpha)a(t)y + (1-\alpha)b(t),$$

care nu este altceva decât o ecuaţie liniară în noua funcţie necunoscută. Aplicând metodologia anterioară, găsim soluţia ecuaţiei noastre.

Exemplu. Să se rezolve ecuaţia diferenţială

$$x' = -\frac{1}{t}x + \frac{1}{t^2x^2} \quad (t > 0, \ x > 0).$$

Înmulţim cu x^2 şi găsim

$$x^2x' = -\frac{1}{t}x^3 + \frac{1}{t^2}\cdot$$

Facem schimbarea de funcţie

$$y = x^3; \quad \text{deci} \quad y' = 3x^2x'.$$

Se ajunge la ecuaţia diferenţială liniară în y

$$y' = -\frac{3}{t}y + \frac{3}{t^2}\cdot$$

Aplicăm formula de rezolvare a acesteia

$$A(t) = -\int \frac{3}{t}dt = -3\ln t \Longrightarrow e^{A(t)} = \frac{1}{t^3}\cdot$$

Obţinem soluţia ecuaţiei în y (cu formula indicată)

$$y = \frac{1}{t^3}\left(C + \int t^3 \cdot \frac{3}{t^2}dt\right) = \frac{1}{t^3}\left(C + \frac{3}{2}t^2\right).$$

Revenind la substituţie, găsim astfel soluţia

$$x(t) = \sqrt[3]{\frac{1}{t^3}\left(C + \frac{3}{2}t^2\right)},\ t > 0,\quad C = \text{constantă arbitrară}.$$

6.3.3 Rezultate generale de existenţă şi unicitate

Vom trece în continuare la studiul unor condiţii standard în care problemele formulate la început admit soluţie (unică). Acestea vor avea rolul de a da o imagine corectă a dificultăţilor tehnice care apar în acest cadru cât şi a metodelor specifice utilizate aici.

(A). Fie Γ un interval deschis al axei reale şi $\Omega = \Gamma \times R$. Considerăm dată o funcţie $f : \Omega \longrightarrow R$ cu

(C1) f este continuă pe Ω.

Fie, de asemenea, $P(t_0, x_0)$ un punct din Ω. Se poate lua în discuţie problema $(E) + (I)$ formulată anterior. Pentru o rezolvare a acesteia, mai avem nevoie de convenţia următoare. Zicem că f satisface o *condiţie Lipschitz locală* în raport cu intervalul Γ dacă pentru orice interval compact (închis şi mărginit) Θ din Γ există o constantă strict pozitivă $L(\Theta)$ aşa ca

(D1) $|f(t,x) - f(t,y)| \leq L(\Theta)|x - y|,\quad t \in \Theta,\ x, y \in R.$

Să acceptăm în continuare ipoteza

(C2) f este local lipschitziană în raport cu Γ.

Rezultatul fundamental de existenţă şi unicitate relativ la problema Cauchy (E)+(I) este

Teorema 6.3.1 *În ipotezele admise, pentru fiecare punct $P(t_0, x_0)$ din Ω există o funcţie derivabilă unică $t \vdash x(t)$ de la Γ la R care să verifice ecuaţia* (E) *şi condiţia iniţială* (I).

Demonstraţie. (Schiţă) Începem prin a nota că problema (E)+(I) este echivalentă cu ecuaţia integrală (numită *ecuaţie de tip Volterra*)

$$(EV) \qquad x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds, \ t \in \Gamma.$$

Demonstraţia se bazează pe proprietatea de derivare a integralei definite; nu mai dăm detalii. Să trecem acum la etapele efective ale construcţiei noastre.

Etapa 1. Fie Θ un interval compact al lui Γ care conţine t_0 în interior. Dorim să arătăm că ecuaţia integrală (EV) admite o (singură) soluţie pe intervalul Θ. În acest scop, construim şirul de aproximante succesive (ale problemei noastre)

$$\text{(D2)} \quad \begin{cases} x_0(t) = x_0, & t \in \Theta \\ x_{n+1}(t) = x_0 + \int_{t_0}^{t} f(s, x_n(s))ds, & t \in \Theta, \ n = 0, 1, ... \end{cases}$$

Să evaluăm diferenţele dintre două aproximante succesive. Fie $L = L(\Theta)$ constanta Lipschitz introdusă de ipoteza (C2). Să mai notăm

$$\mu = \sup\{|x_1(t) - x_0(t)|; \ t \in \Theta\}, \ \rho = \sup\{|t - t_0|; \ t \in \Theta\}.$$

Avem, cu condiţia Lipschitz, pentru toţi $t \in \Theta$

$$\begin{cases} |x_2(t) - x_1(t)| = \left|\int_{t_0}^{t}(f(s, x_1(s)) - f(s, x_0(s)))ds\right| \le \\ \le \left|\int_{t_0}^{t} |f(s, x_1(s)) - f(s, x_0(s))|ds\right| \le L\left|\int_{t_0}^{t} |x_1(s) - x_0(s)|ds\right| \le \\ \le L\left|\int_{t_0}^{t} \mu ds\right| = \mu L|t - t_0|. \end{cases}$$

În general, se obţine evaluarea

$$|x_{n+1}(t) - x_n(t)| \le \mu L^n \frac{|t - t_0|^n}{n!}, \quad t \in \Theta, \ n \in N.$$

Din inegalitatea

$$\mu L^n \frac{|t - t_0|^n}{n!} \le \mu \frac{(L\rho)^n}{n!}, \quad t \in \Theta, \ n \in N$$

şi convergenţa seriei numerice din dreapta rezultă, cu Criteriul de comparaţie Weierstrass, că seria de funcţii $\sum_n \mu L^n \frac{|t-t_0|^n}{n!}$ converge uniform pe Θ.

Combinând acest fapt cu majorările obţinute anterior, deducem că seria de funcţii $\sum_n (x_{n+1}-x_n)$ este uniform convergentă pe intervalul respectiv; sau, echivalent, că şirul de funcţii (x_n) converge uniform pe intervalul Θ. Să notăm

$$z(t) = \lim_{n\to\infty} x_n(t),\ t \in \Theta \ \ \text{(limita uniformă a şirului).}$$

Avem deci

$$\sigma_n \longrightarrow 0 \ \ \text{pentru} \ \ n \to \infty$$

unde, prin definiţie,

$$\sigma_n = \sup\{|x_n(t) - z(t)|; \ \ t \in \Theta\}, \ \ \ n \in N.$$

Să mai facem notaţia (pentru simplitate)

$$w(t) = x_0 + \int_{t_0}^{t} f(s, z(s))ds, \ t \in \Theta.$$

Avem, din condiţia Lipschitz şi convenţiile făcute,

$$\begin{cases} |x_{n+1}(t) - w(t)| \leq \left| \int_{t_0}^{t} |f(s, x_n(s)) - f(s, z(s))| ds \right| \leq \\ \leq L \left| \int_{t_0}^{t} |x_n(s) - z(s)| ds \right| \leq L\rho\sigma_n, \qquad t \in \Theta, \ n \in N. \end{cases}$$

Aceasta arată că şirul de funcţii (x_n) converge uniform şi către funcţia continuă w. Cum însă limita uniformă este unică, urmează $z(t) = w(t), \ t \in \Theta$; adică, z este soluţie, pe intervalul Θ, a ecuaţiei Volterra (EV) (echivalentă după cum am spus, cu $(E) + (I)$).
Fie acum y, o altă soluţie, pe intervalul Θ, a lui (EV). Să punem

$$\sigma = \sup\{|z(t) - y(t)|; \ t \in \Theta\}.$$

Avem atunci (cu scrierea lui (EV) şi notaţia făcută)

$$\begin{cases} |z(t) - y(t)| \leq \left| \int_{t_0}^{t} |f(s, z(s)) - f(s, y(s))| ds \right| \leq \\ \leq L \left| \int_{t_0}^{t} |z(s) - y(s)| ds \right| \leq L \left| \int_{t_0}^{t} \sigma ds \right| = \sigma L |t - t_0|, \ t \in \Theta. \end{cases}$$

Ţinând cont de scrierea lui (EV) şi relaţia obţinută anterior,găsim

$$\begin{cases} |z(t)-y(t)| \le \left|\int_{t_0}^t |f(s,z(s))-f(s,y(s))|ds\right| \le \\ L\left|\int_{t_0}^t |z(s)-y(s)|ds\right| \le L\left|\int_{t_0}^t \sigma L|s-t_0|ds\right| = \sigma L^2\dfrac{|t-t_0|^2}{2},\ t\in\Theta. \end{cases}$$

În general, avem evaluările

$$|z(t)-y(t)| \le \sigma\frac{L^n|t-t_0|^n}{n!} \le \sigma\frac{(L\rho)^n}{n!},\ t\in\Theta,\ n\in N.$$

Cum însă $\lim\limits_{n\to\infty}\dfrac{(L\rho)^n}{n!}=0$, urmează că, necesar,

$$z(t)=y(t),\ \text{pentru toţi}\ t\in\Theta,$$

şi, deci, soluţia găsită este unică pe intervalul Θ.

Etapa 2. Vom prelungi acum succesiv soluţiile determinate până la o soluţie pe intervalul Γ. Fie (Θ_n) un şir de intervale compacte ce conţin t_0 în interior, cu

$$\Theta_0\subset\Theta_1\subset\Theta_2\subset\cdots;\ \bigcup_n\Theta_n=\Gamma.$$

Conform etapei precedente, problema $(E)+(I)$ are câte o soluţie unică, z_n, pe fiecare interval Θ_n. Definim atunci $z:\Gamma\longrightarrow R$ prin

$$z(t)=z_n(t),\quad \text{dacă}\ t\in\Theta_n.$$

Definiţia este consistentă, în baza proprietăţii de unicitate. În plus, ecuaţia (E) este verificată pe intervalul Γ, după cum se poate uşor constata. Aceasta încheie argumentul. ∎

Un caz particular remarcabil este acela al problemei Cauchy ataşată ecuaţiei diferenţiale liniare

$$(EL)\qquad x'=a(t)x+b(t).$$

Aici, $a,\ b$ sunt două funcţii continue definite pe un interval deschis Γ al axei reale. În adevăr, dacă Θ este un interval compact al lui Γ şi

$$L(\Theta)=\sup\{|a(t)|;\ t\in\Theta\},$$

atunci, evident

$$|f(t,x) - f(t,y)| = |a(t)|\,|x-y| \le L(\Theta|x-y|, \quad x,y \in R.$$

Rezultă, conform teoremei demonstrate că această problemă Cauchy are o soluţie unică pe intervalul Γ. Expresia ei a fost deja găsită într-un loc anterior.

Revenind la cazul general, o problemă importantă de studiu ar fi următoarea: ce se întâmplă când condiţia Lipschitz $(C2)$ nu mai are loc. Un răspuns posibil ar fi următorul. Zicem că funcţia $f : \Omega \longrightarrow R$ verifică o *condiţie Lipschitz locală* pe domeniul Ω dacă, pentru orice $P(t_0, x_0)$ din Ω există intervalele compacte $\Gamma(P)$ şi $D(P)$ conţinând t_0, respectiv x_0 în interior, şi numărul strict pozitiv $L(P)$, aşa încât

(D3) $|f(t,x) - f(t,y)| \le L(P)|x-y|,\ t \in \Gamma(P),\ x,y \in \Delta(P).$

Să admitem acum ipoteza (C1) şi

(C3) f este local lipschitziană pe domeniul Ω.

Dat punctul $P(t_0, x_0) \in \Omega$, există atunci un interval deschis maximal $\Gamma(P) \subseteq \Gamma$ şi o singură funcţie derivabilă $t \vdash x(t)$ de la $\Gamma(P)$ la R care să constituie o soluţie a problemei $(E)+(I)$ pe intervalul $\Gamma(P)$. Maximalitatea intervalului $\Gamma(P)$ înseamnă aici următorul lucru: nu există un alt interval deschis $\Gamma^*(P)$ din Γ care să includă strict $\Gamma(P)$ şi pe care să avem definită o altă soluţie a problemei $(E)+(I)$. Asta, în particular, ne spune că, dacă

$$\Gamma = (\alpha, \beta),\ \Gamma(P) = (\alpha(P), \beta(P)),\ \alpha \le \alpha(P) < \beta(P) \le \beta$$

atunci, avem fie $\beta(P) = \beta$, fie $\beta(P) < \beta$, caz în care cu necesitate

$$|x(t)| \longrightarrow \infty \text{ pentru } t \longrightarrow \beta(P).$$

(O concluzie analogă are loc şi în punctul $\alpha(P)$.)

Ca o ilustrare a celor spuse, să reluăm problema Cauchy asociată unei ecuaţii cu variabile separabile

$$(ES) \qquad x' = a(t)b(x), \quad x(t_0) = x_0.$$

Aici, a este o funcţie continuă pe un interval $\Gamma = (\alpha, \beta)$, iar b, o funcţie continuă şi strict pozitivă pe un interval deschis $\Delta = (\rho, \sigma)$. Presupunem că

b este local lipschitziană pe Δ; în acest caz, ipoteza (C3) se verifică, după cum se vede, uşor. Cu notaţiile făcute deja

$$A(t) = \int_{t_0}^{t} a(s)ds,\ t \in \Gamma;\ B(x) = \int_{x_0}^{x} \frac{d\xi}{b(\xi)},\ x \in \Delta$$

am văzut că soluţia problemei este

$$x(t) = B^{-1}(A(t)), \quad t \in \Gamma(P).$$

Intervalul pe care există soluţia este definit de

(C4) $A(t)$ aparţine domeniului de definiţie al lui B^{-1}

sau, echivalent,

$$\lim_{x \to \rho+} B(x) < A(t) < \lim_{x \to \sigma-} B(x).$$

Deci s-ar putea ca domeniul de definiţie al soluţiei să nu fie identic cu intervalul Γ de la început. Aceasta este de altfel vizibil şi pe cazul concret tratat anterior. (De fapt, condiţia Lipschitz pentru funcţia b nu era absolut necesară aici; s-a pus totuşi pentru a face legătura cu consideraţiile expuse anterior). În fine, este util de menţionat că rezultatul de mai sus poate fi extins chiar la domenii Ω care nu sunt neapărat de forma $\Gamma \times R$ dinainte; nu mai dăm aici alte detalii.

(B). Fie din nou Γ un interval deschis al axei reale şi $\Omega = \Gamma \times R^k$. Considerăm dat un sistem de k funcţii $f_i : \Omega \longrightarrow R$, $1 \le i \le k$ şi un punct $P(t_0, x_1^0, ..., x_k^0)$ din Ω. Ne interesează precizarea unor condiţii uzuale în care problema Cauchy $(S_k) + (J_k)$ are soluţie (unică). Pentru aceasta avem nevoie de următoarea convenţie. Zicem că funcţiile $f_1, ..., f_k$ satisfac o *condiţie Lipschitz locală* în raport cu intervalul Γ dacă, pentru orice interval compact Θ al lui Γ există o constantă strict pozitivă $L(\Theta)$ aşa încât să avem, pentru $i \in \{i, ..., k\}$,

$$\text{(D4)} \quad \begin{cases} |f_i(t, x_1, ..., x_k) - f_i(t, y_1, ..., y_k)| \le L(\Theta) \left(\sum_{i=1}^{n} |x_i - y_i| \right), \\ t \in \Theta, \quad (x_1, ..., x_k), (y_1, ..., y_k) \in R^k. \end{cases}$$

Ca o extensie directă a rezultatului precedent, dăm

Teorema 6.3.2 *Să admitem că funcţiile $f_1, ..., f_k$ sunt continue pe Ω şi local lipschitziene în raport cu intervalul Γ. Pentru fiecare punct $P(t_0, x_1^0, ..., x_k^0)$ din Ω există atunci un sistem unic de funcţii derivabile $x_1(t), ..., x_k(t)$, de la Γ la R, care să satisfacă ecuaţiile sistemului (S_k) şi condiţiile iniţiale (J_k).*

Demonstraţie. (Schiţă) Se urmează în principiu raţionamentul precedent. Observăm mai întâi că problema $(S_k) + (J_k)$ este echivalentă cu sistemul de ecuaţii integrale (de tip Volterra)

$$(SV) \qquad x_i(t) = x_i^0 + \int_{t_0}^{t} f_i(s, x_1(s), ..., x_k(s))ds, \; t \in \Gamma, \; 1 \leq i \leq k.$$

Fie Θ un interval compact al lui Γ care conţine t_0 în interior. Construim aproximantele succesive (ale problemei) prin

$$(\mathrm{D5}) \begin{cases} x_i^{[0]}(t) = x_i^0, \quad t \in \Theta, 1 \leq i \leq k; \;\; \text{iar,pentru} \;\; n \geq 0, \\ x_i^{[n+1]}(t) = x_i^0 + \int_{t_0}^{t} f_i(s, x_1^{[n]}(s), ..., x_k^{[n]}(s))ds, \; t \in \Theta, 1 \leq i \leq k. \end{cases}$$

Se arată, ca la Teorema 6.3.1, că şirurile de funcţii

$$(x_1^{[n]})_{n \in N}, ..., (x_k^{[n]})_{n \in N}$$

converg uniform pe intervalul compact Θ către un sistem de funcţii $(z_1, ..., z_k)$ despre care se arată că verifică sistemul de ecuaţii integrale Volterra (SV) (deci, şi problema $(S_k) + (J_k)$) pe Θ. În plus, funcţiile găsite sunt unice cu această proprietate. Prelungirea soluţiilor găsite la întreg intervalul Γ se face ca mai înainte. ■

(C). Fie Γ un interval deschis al axei reale şi $\Omega = \Gamma \times R^k$. Considerăm dată o funcţie $f : \Omega \longrightarrow R$ şi un punct $P(t_0, \alpha_0, \alpha_1, ..., \alpha_{k-1})$ din Ω. Ne interesează şi aici condiţiile uzuale în care problema Cauchy $(E_k) + (I_k)$ are soluţie (unică). Acestea pot fi direct deduse din rezultatul precedent. Mai precis, avem

Teorema 6.3.3 *Să presupunem că funcţia f este continuă pe Ω şi local lipschitziană în raport cu intervalul Γ. Atunci, dat punctul $P(t_0, \alpha_0, ..., \alpha_{k-1})$ din Ω, există o singură funcţie $t \vdash x(t)$ de la Γ la R, cu derivate până la ordinul k inclusiv, care să soluţioneze problema $(E_k) + (I_k)$.*

Demonstraţie. După cum s-a observat deja într-un loc anterior, dacă $t \vdash x(t)$ este o soluţie a problemei $(E_k)+(I_k)$, atunci sistemul de funcţii $(x_1, ..., x_k)$ definit de convenţiile

$$x_1 = x, x_2 = x', ..., x_k = x^{(k-1)}$$

apare ca soluţie a sistemului diferenţial

$$(S_k^*) \qquad x_1' = x_2, x_2' = x_3, ..., x_{k-1}' = x_k, x_k' = f(t, x_1, ..., x_k)$$

cu condiţiile iniţiale

$$(J_k^*) \qquad x_1(t_0) = \alpha_0, x_2(t_0) = \alpha_1, ..., x_k(t_0) = \alpha_{k-1}.$$

Dar, din ipoteză, funcţiile care apar în membrii secunzi ai lui (S_k^*) sunt continue pe Ω şi local lipschitziene în raport cu intervalul Γ. Există deci soluţie unică pe (Γ) a problemei Cauchy $(S_k^*) + (J_k^*)$ (se aplică aici Teorema 6.3.2). Ţinând cont acum că prima componentă a acestei soluţii verifică în mod necesar $(E_k) + (I_k)$, concluzia reiese. ■

(D). Ca şi mai înainte, se pune problema ce se întâmplă când condiţia Lipschitz anterioară nu are loc. Un răspuns posibil este următorul. Să numim funcţiile $f_1, ..., f_k$, *local lipschitziene* în raport cu domeniul Ω dacă, pentru orice $P(t_0, x_1^0, ..., x_k^0)$ din Ω există intervalele compacte $\Gamma(P)$ şi, respectiv, $\Delta_1(P), ..., \Delta_k(P)$, conţinând t_0 şi respectiv $x_1^0, ..., x_k^0$ în interior, precum şi numărul $L(P) > 0$ aşa încât să avem pentru toţi $i \in \{1, ..., k\}$

$$\text{(D6)} \quad \begin{cases} |f_i(t, x_1, ..., x_k) - f_i(t, y_1, ..., y_k)| \le L(P) \left(\sum_{i=1}^{n} |x_i - y_i| \right), \\ t \in \Gamma(P), \quad (x_1, ..., x_k), (y_1, ..., y_k) \in \Delta_1(P) \times \cdots \times \Delta_k(P). \end{cases}$$

Să admitem acum că funcţiile $f_1, ..., f_k$ sunt continue pe Ω şi local lipschitziene în raport cu acest domeniu. Dat punctul $P(t_0, x_1^0, ..., x_k^0)$ în Ω, există atunci un interval deschis maximal $\Gamma(P) \subseteq \Gamma$ şi un sistem unic de funcţii derivabile $(x_1(t), ..., x_k(t))$ de la $\Gamma(P)$ la R^k care să constituie o soluţie a problemei $(S_k) + (J_k)$ pe intervalul $\Gamma(P)$. Este important şi aici de subliniat că intervalul $\Gamma(P)$ nu coincide, în general, cu intervalul (deschis) Γ dat iniţial.

În particular, se poate formula un rezultat de acest fel şi pentru problema $(E_k) + (I_k)$. Menţionăm şi aici că toate aceste consideraţii se pot extinde la domenii Ω arbitrare.

6.3.4 Ecuaţii diferenţiale liniare de ordin superior

Numim *ecuaţie diferenţială liniară de ordinul* k, o relaţie de forma

$$(E) \qquad a_0(t)x^{(k)} + a_1(t)x^{(k-1)} + \cdots + a_{k-1}(t)x' + a_k(t)x = b(t)$$

Aici, funcţiile $a_0(t), a_1(t), ..., a_k(t), b(t)$ sunt continue pe un interval deschis Γ; iar, în plus, (C1) $a_0(t) \neq 0$, pentru toţi $t \in \Gamma$.

De obicei, ecuaţia scrisă se mai asociază cu condiţia iniţială

$$(I) \qquad x(t_0) = \alpha_0, x'(t_0) = \alpha_1, ..., x^{(k-1)}(t_0) = \alpha_{k-1},$$

unde t_0 este un element al lui Γ, iar $p = (\alpha_0, ..., \alpha_{k-1})^\top$ este un vector din R^k. Cu privire la problema Cauchy $(E) + (I)$, ne interesează două aspecte:

(i) existenţa şi unicitatea soluţiei (pentru un vector fixat p din R^k);

(ii) structura mulţimii soluţiilor (obţinute când $p \in R^k$ este luat arbitrar).

În ce priveşte primul aspect, să notăm că funcţia $f : \Gamma \times R^k \longrightarrow R$ definită prin

$$\begin{cases} f(t, x_1, ..., x_k) = -\frac{1}{a_0(t)}(b(t) - a_k(t)x_1 - a_{k-1}(t)x_2 - \cdots - a_1(t)x_k) \\ t \in \Gamma, \quad (x_1, ..., x_k) \in R^k, \end{cases}$$

este continuă şi local lipschitziană în raport cu intervalul Γ. Deci (ţinând cont de Teorema 6.3.3), problema Cauchy $(E) + (I)$ admite o singură soluţie x, definită pe intervalul Γ. Mai exact, notând

$$C^{(k)}(\Gamma) = \{z : \Gamma \longrightarrow R;\ z \text{ este de clasă } C^k \text{ pe } \Gamma\},$$

rezultatul în chestiune ne spune că există un singur element x în $C^{(k)}(\Gamma)$ care să verifice $(E) + (I)$.

Trecem acua la structura mulţimii soluţiilor lui (E); sau, ceea ce este acelaşi lucru, a soluţiilor lui $(E) + (I)$, pentru $p \in R^k$ arbitrar. În acest scop, vom porni de la ecuaţia omogenă ataşată lui (E); adică

$$(E_0) \qquad a_0(t)x^{(k)} + a_1(t)x^{(k-1)} + ccdots + a_{k-1}(t)x' + a_k(t)x = 0.$$

Să notăm cu $X(\Gamma)$ mulţimea soluţiilor lui (E_0); adică, mulţimea tuturor funcţiilor z din $C^{(k)}(\Gamma)$ care transformă (E_0) într-o egalitate de funcţii. Pentru a vedea cum este structurată această mulţime, să definim un operator diferenţial $L : C^{(k)}(\Gamma) \longrightarrow C(\Gamma)$ prin

(D1) $L(z)(t) = a_0(t)z^{(k)}(t) + \cdots + a_{k-1}(t)z'(t) + a_k(t)z(t)$,
$t \in \Gamma,\ z \in C^{(k)}(\Gamma)$.

Nu este greu de văzut că acesta este un operator liniar:

(P1) $L(x + y) = L(x) + L(yr, L(\lambda x) = \lambda L(x),\ \lambda \in R,\ x, y \in C^{(k)}(\Gamma)$.

Combinând aceasta cu faptul că

(P2) $X(\Gamma) = \{z \in C^{(k)}(\Gamma);\ L(z) = 0\}$

este clar că $X(\Gamma)$ apare ca subspaţiu liniar al lui $C^{(k)}(\Gamma)$. Chestiunea care se pune acum este aceea a dimensiunii acestui subspaţiu. Pentru aceasta, să mai definim un operator diferenţial $W : C^{(k)}(\Gamma) \longrightarrow (C(\Gamma))^k$, prin formula

(D2) $W(z)(t) = (z(t), z'(t), ..., z^{(k-1)}(t))^\top, t \in \Gamma,\ z \in C^{(k)}(\Gamma)$.

Următorul rezultat (dat fără demonstraţie) se va dovedi util pentru scopurile noastre:

Teorema 6.3.4 *Fie $y_1, .., y_k$ un număr de k soluţii din $C^{(k)}(\Gamma)$, ale ecuaţiei omogene (E_0). Afirmaţiile următoare sunt echivalente:*

(P3) $\{y_1, .., y_k\}$ *este liniar independent în* $C^{(k)}(\Gamma)$.

(P4) *matricea* $[Wy_1(t), .., Wy_k(t)]$ *este nesingulară pentru orice* $t \in \Gamma$.

(P5) *matricea* $[Wy_1(t_0), .., Wy_k(t_0)]$ *este nesingulară (pentru un* $t_0 \in \Gamma$).

Putem da un răspuns complet la întrebarea formulată anterior.

Teorema 6.3.5 *Mulţimea $X(\Gamma)$ a soluţiilor ecuaţiei omogene (E_0) formează un subspaţiu k–dimensional al spaţiului $C^{(k)}(\Gamma)$.*

Demonstraţie. Am arătat deja că $X(\Gamma)$ este un subspaţiu liniar. Fie $\{e_1, ..., e_k\}$ baza canonică din R^k. De asemenea, fie t_0 un element fixat din Γ. Pentru fiecare $i \in \{1, ..., k\}$ notăm cu y_i soluţia, din $C^{(k)}(\Gamma)$, a probdemei Cauchy

$$L(y_i) = 0,\ W(y_i)(t_0) = e_i.$$

Arătăm că $\{y_1, ..., y_k\}$ formează o bază în spaţiul $X(\Gamma)$; vom numi aceasta *sistem fundamental de soluţii* ale ecuaţiei omogene (E_0). În primul rând, din

$$[Wy_1(t_0), ..., Wy_k(t_0)] = I,\ \text{(matricea unitate din } \mathcal{M}(k))$$

şi teorema precedentă, $\{y_1, ..., y_k\}$ este un sistem liniar independent în $X(\Gamma)$. Fie acum $y \in X(\Gamma)$ o soluţie arbitrară a lui (E_0). Avem

$$Wy(t_0) = y(t_0)W(y_1)(t_0) + \cdots + y^{(k-1)}(t_0)W(y_k)(t_0)$$

(din descompunerea după o bază). De aici,(combinând cu partea de unicitate din Teorema 6.3.3),

$$y = y(t_0)y_1 + y'(t_0)y_2 + \cdots + y^{(k-1)}(t_0)y_k.$$

Aceasta încheie argumentul. ∎

Putem acum să ne întoarcem la problema structurii mulţimii soluţiilor ecuaţiei neomogene (E). Avem în acest sens

Teorema 6.3.6 *Formula*

(R1) $x(t) = C_1y_1(t) + \cdots + C_ky_k(t) + z(t),\ t \in \Gamma,$

unde $C_1, ..., C_k$ *sunt constante arbitrare,* $\{y_1(t), ..., y_k(t)\}$ *este un sistem fundamental de soluţii ale ecuaţiei omogene* (E_0)*, iar* $z(t)$ *o soluţie particulară a ecuaţiei neomogene* (E)*, reprezintă expresia soluţiei generale a ecuaţiei neomogene* (E)*.*

Demonstraţie. Nu avem decât să observăm că diferenţa dintre două soluţii ale lui (E) este o soluţie a ecuaţiei omogene (E_0). ∎

O problemă naturală care se pune în acest context este aceea a determinării unei soluţii particulare ale ecuaţiei neomogene (E). Se poate obţine o asemenea soluţie – prin aşa–numita metodă a variaţiei constantelor – dacă se cunoaşte un sistem fundamental de soluţii $\{y_1, ..., y_k\}$, ale ecuaţiei omogene (E_0). Anume, să căutăm soluţia în chestiune z, sub forma

(R2) $z(t) = C_1(t)y_1(t) + \cdots + C_k(t)y_k(t),$

unde funcţiile derivabile $C_1(t), ..., C_k(t)$ urmează a fi determinate. Mai exact, să impunem acestor funcţii condiţiile (pentru toţi $t \in \Gamma$)

$$\text{(C2)}\quad \begin{cases} y_1(t)C_1'(t) + \cdots + y_k(t)C_k'(t) = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_1^{(k-2)}(t)C_1'(t) + \cdots + y_k^{(k-2)}(t)C_k'(t) = 0 \\ y_1^{(k-1)}(t)C_1'(t) + \cdots + y_k^{(k-1)}(t)C_k'(t) = b(t)/a_0(t). \end{cases}$$

Matricea acestui sistem este nesingulară, în baza Teoremei 6.3.4. Deci sistemul are soluţia unică $\{C_1'(t), ..., C_k'(t)\}$ şi ca atare se poate găsi (prin integrare) sistemul de funcţii $\{C_1(t), ..., C_k(t)\}$ care să satisfacă aceste condiţii.

Se arată acum uşor că funcţia $z(t)$, dată de (R2), este o soluţie a ecuaţiei neomogene (E); nu mai dăm alte detalii.

Un caz particular important pentru aplicaţii este acela în care funcţiile coeficient $a_0(t), ..., a_k(t)$ sunt constante. Este vorba deci de ecuaţia

$$(E) \qquad a_0x^{(k)} + a_1x^{(k-1)} + \cdots + a_{k-1}x' + a_kx = b(t),\ \ t \in \Gamma,$$

unde $a_0, ..., a_k$ sunt numere reale date, cu $a_0 \neq 0$. În acest caz, chestiunea determinării unui sistem fundamental de soluţii ale ecuaţiei omogene asociate

$$(E_0) \qquad a_0x^{(k)} + a_1x^{(k-1)} + \cdots + a_{k-1}x' + a_kx = 0$$

cât şi aceea a determinării unei soluţii particulare ale ecuaţiei (neomogene) (E) capătă unele aspecte specifice, descrise mai jos.

(A). Relativ la ecuaţia omogenă (E_0), să asociem acesteia următoarea ecuaţie algebrică (numită *caracteristică*) peste corpul complex C

$$(EC) \qquad a_0\lambda^k + a_1\lambda^{k-1} + \cdots + a_{k-1}\lambda + a_k = 0.$$

Notăm cu $P(\lambda)$ polinomul din partea stângă

$$P(\lambda) = a_0\lambda^k + a_1\lambda^{k-1} + \cdots + a_{k-1}\lambda + a_k,\ \ \lambda \in C.$$

Dacă γ este o rădăcină reală de multiplicitate r a ecuaţiei caracteristice:

$$P(\gamma) = 0, ..., P^{(r-1)}(\gamma) = 0, P^{(r)}(\gamma) \neq 0$$

atunci, asociem acesteia funcţiile

$$x_1(t) = e^{\gamma t}, ..., x_r(t) = t^{r-1}e^{\gamma t},\ \ t \in R.$$

Iar dacă $\zeta = \alpha + i\beta$ (deci şi conjugata ei, $\bar{\zeta} = \alpha - i\beta$) este o rădăcină complexă de multiplicitate r, a ecuaţiei caracteristice:

$$\begin{array}{c} P(\zeta) = 0, P'(\zeta) = 0, ..., P^{(r-1)}(\zeta) = 0, P^{(r)}(\zeta) \neq 0 \\ P(\bar{\zeta}) = 0, P'(\bar{\zeta}) = 0, ..., P^{(r-1)}(\bar{\zeta}) = 0, P^{(r)}(\bar{\zeta}) \neq 0, \end{array}$$

atunci, asociem acesteia funcţiile

$$\begin{cases} x_1(t) = e^{\alpha t}\cos\beta t, ..., x_r(t) = t^{r-1}e^{\alpha t}\cos\beta t \\ x_{r+1}(t) = e^{\alpha t}\sin\beta t, ..., x_{2r}(t) = t^{r-1}e^{\alpha t}\sin\beta t, \quad t \in R. \end{cases}$$

Am ajuns astfel ca, la cele k rădăcini ale ecuaţiei caracteristice (EC) (fiecare numărată de atâtea ori cât arată multiplicitatea ei) să asociem un număr de k funcţii $\{x_1, ..., x_k\}$ (printr-o re–numerotare a celor scrise anterior).

Teorema 6.3.7 *Sistemul de funcţii astfel determinat este un sistem fundamental de soluţii ale ecuaţiei omogene* (E_0).

Demonstraţie. (Schiţă) Pentru prima parte ne vom limita doar la rădăcini reale. Reamintim formula lui Leibniz de derivare (succesivă) a unui produs de funcţii

$$(z \cdot w)^{(r)}(t) = \sum_{h=0}^{r} C_r^h z^{(r-h)}(t) w^{(h)}(t),\ t \in \Gamma,\ r = 0, 1, ...$$

Fie acum λ număr real arbitrar iar m, număr natural. Avem atunci, pentru orice $r \in \{0, ..., k\}$,

$$\frac{d^r}{dt^r}(t^m e^{\lambda t}) = \left(\sum_{h=0}^{r} C_r^h \frac{d^h}{dt^h}(t^m)\lambda^{r-h}\right) e^{\lambda t} = \left(\sum_{h=0}^{m} C_m^h t^{m-h} \frac{d^h}{d\lambda^h}(\lambda^r)\right) e^{\lambda t}.$$

(Se vor lua în discuţie cazurile $r \geq m$ şi $r < m$.) Înmulţind în această formulă cu a_{k-r} şi adunând rezultatele pentru toate valorile lui r, avem

$$L(t^m e^{\lambda t}) = \left(\sum_{h=0}^{m} C_m^h t^{m-h} P^{(h)}(\lambda)\right) e^{\lambda t}.$$

Deci, dacă $\lambda = \gamma$ este rădăcină reală de multiplicitate r, a ecuaţiei caracteristice (EC), atunci funcţiile asociate sunt soluţii ale ecuaţiei omogene (E_0). (Analog se va proceda şi la cazul complex.) Pentru a doua parte, vom admite şi aici – pentru simplitate – că avem de a face doar cu rădăcini reale. Să presupunem, prin reducere la absurd, că funcţiile $x_1(t), ..., x_k(t)$ nu ar fi liniar independente în $C^{(k)}(\Gamma)$. Ar exista deci constantele $\delta_1, ..., \delta_k$, nu toate nule, aşa ca

$$\delta_1 x_1(t) + \cdots + \delta_k x_k(t) = 0,\ t \in \Gamma.$$

Putem transforma această relaţie în

(R3) $Q_1(t)e^{\gamma_1 t} + \cdots + Q_r(t)e^{\gamma_r t} = 0,\ t \in \Gamma.$

unde polinoamele $Q_1(t), ..., Q_r(t)$ nu sunt identic nule, iar $\gamma_1, ..., \gamma_r$ sunt rădăcinile reale, distincte două câte două, ale ecuaţiei caracteristice, de multiplicităţi $m_1, ..., m_r$, respectiv (unde $m_1 + \cdots + m_r = k$). Înmulţim relaţia (R3) cu $e^{-\gamma_1 t}$ şi notăm $\gamma_2 - \gamma_1 = \mu_2$,..., $\gamma_r - \gamma_1 = \mu_r$. Avem, deci,

$$Q_1(t) + Q_2(t)e^{\mu_2 t} + \cdots + Q_r(t)e^{\mu_r t} = 0,\ t \in \Gamma.$$

Derivând această identitate de un număr suficient de ori (mai exact,de $\mathrm{gr}(Q_1)+1$ ori), obţinem

$$U_2(t)e^{\mu_2 t}+\cdots+U_r(t)e^{\mu_r t}=0,\ t\in\Gamma,$$

unde $U_2(t),...,U_r(t)$ sunt polinoame de grad respectiv egal cu al polinoamelor $Q_2(t),...,Q_r(t)$. (Am folosit aici faptul că $\mu_2,...,\mu_r$ sunt nenule.) Acum, numerele $\mu_2,...,\mu_r$ sunt din nou distincte două câte două. Se poate relua deci raţionamentul precedent etc. Ajungem în final la o relaţie de forma

(R4) $V_r(t)e^{\nu_r t}=0,\ t\in\Gamma$

unde $V_r(t)$ este un polinom de grad egal cu acela al lui $Q_r(t)$, iar $\nu_r\neq 0$. Acest lucru este însă imposibil. Contradicţia la care am ajuns demonstrează afirmaţia şi, odată cu ea, teorema. ∎

Exemplu. Să se rezolve ecuaţia diferenţială omogenă

$$x'''+x''-5x'+3x=0.$$

Ecuaţia caracteristică asociată acesteia este

$$\lambda^3+\lambda^2-5\lambda+3=0,\ \text{cu rădăcinile } \lambda_1=\lambda_2=1,\lambda_3=-3.$$

Soluţia generală a ecuaţiei noastre este deci

$$x(t)=C_1e^t+C_2te^t+C_3e^{-3t},\ t\in R,$$

unde C_1,C_2,C_3 sunt constante arbitrare. Acestea pot fi unic determinate printr-o condiţie iniţială impusă soluţiei Astfel, de pildă, dacă se mai cere în plus

$$x(2)=0,x'(2)=-1,x''(2)=1$$

atunci sistemul care dă constantele în cauză este

$$\begin{aligned}&C_1e^2+2C_2e^2+C_3e^{-6}=0,\\&C_1e^2+3C_2e^2-3C_3e^{-6}=-1,\\&C_1e^2+4C_2e^2+9C_3e^{-6}=1\end{aligned}$$

cu soluţiile

$$C_1=\frac{1}{20}e^{-2},C_2=-\frac{1}{10}e^{-2},C_3=\frac{3}{20}e^{6}.$$

Rezultă soluţia

$$x(t) = \frac{1}{20}e^{t-2} - \frac{1}{10}te^{t-2} + \frac{3}{20}e^{-3(t-2)},\ t \in R$$

care satisface ecuaţia şi condiţiile iniţiale impuse.

(B). În ce priveşte determinarea unei soluţii particulare a ecuaţiei diferenţiale neomogene (E), metoda generală de lucru este, după cum am arătat, aceea a variaţiei constantelor; aceasta este posibil aici, deoarece dispunem de un sistem fundamental de soluţii ale ecuaţiei omogene. Există însă anumite situaţii când soluţia particulară în chestiune poate fi determinată într-un mod mai simplu. Acestea se referă la termenii liberi de forma

$$b(t) = (M(t)\cos\beta t + N(t)\sin\beta t)e^{\alpha t},\ t \in R$$

unde $M(t), N(t)$ sunt polinoame cu coeficienţi reali iar α, β constante reale (cazurile $\alpha = 0$ sau $\beta = 0$ fiind permise). În adevăr, să notăm $m = \max\{\mathrm{gr}\,(M), \mathrm{gr}\,(N)\}$ şi să presupunem că numărul $\lambda = \alpha + i\beta$ este rădăcină, de multiplicitate r, pentru ecuaţia caracteristică asociată. (Desigur, când $r = 0$, aceasta ar însemna că numărul în cauză nu este de fapt rădăcină a ecuaţiei caracteristice). Există atunci posibilitatea de a determina o soluţie particulară a ecuaţiei neomogene (E), de forma

$$z(t) = t^r(U(t)\cos\beta t + V(t)\sin\beta t)e^{\alpha t},\ t \in R$$

unde $U(t)$ şi $V(t)$ sunt polinoame de grad cel mult m, care urmează a fi determinate. Anume, dacă ţinem cont de o formulă anterioară, avem

$$L(z)(t) = (U^*(t)\cos\beta t + V^*(t)\sin\beta t)e^{\alpha t},\ t \in R$$

unde $U^*(t), V^*(t)$ sunt polinoame de grad cel mult m, ai căror coeficienţi depind liniar de aceia ai polinoamelor iniţiale $U(t), V(t)$. Din identităţile

$$U^*(t) = M(t),\ V^*(t) = N(t),\ t \in R,$$

rezultă atunci coeficienţii polinoamelor $M(t), N(t)$, prin rezolvări de sisteme algebrice liniare. Şi astfel, soluţia particulară z, a ecuaţiei noastre a fost determinată.

Exemplu. Să se rezolve ecuaţia neomogenă

$$x''' - 2x'' + x' - 2x = (t+1)e^{3t}.$$

Ecuaţia neomogenă asociată

$$x''' - 2x'' + x' - 2x = 0$$

are ecuaţia caracteristică

$$\lambda^3 - 2\lambda^2 + \lambda - 2 = 0, \text{ cu rădăcinile } \lambda_1 = 2, \lambda_2 = i, \lambda_3 = -i.$$

Un sistem fundamental de soluţii pentru aceasta este

$$x_1(t) = e^{2t}, x_2(t) = \cos t, x_3(t) = \sin t.$$

Deci, soluţia generală a ecuaţiei omogene este

$$y(t) = C_1 e^{2t} + C_2 \cos t + C_3 \sin t$$

unde C_1, C_2, C_3 sunt constante arbitrare. Căutăm acum o soluţie particulară a ecuaţiei iniţiale de forma

$$z(t) = (\alpha t + \beta)e^{3t}, \ t \in R.$$

(Aceasta deoarece $\lambda = 3$ nu este rădăcină a ecuaţiei caracteristice.) Avem

$$\begin{aligned} z'(t) &= (3\alpha t + \alpha + 3\beta)e^{3t}, \\ z''(t) &= (9\alpha t + 6\alpha + 9\beta)e^{3t}, \\ z'''(t) &= (27\alpha t + 27\alpha + 27\beta)e^{3t}. \end{aligned}$$

Înlocuind în ecuaţia neomogenă iniţială, ajungem la identitatea

$$(10\alpha t + 16\alpha + 10\beta)e^{3t} = (t+1)e^{3t}.$$

Egalând coeficienţii termenilor asemenea, găsim

$$10\alpha = 1, 16\alpha + 10\beta = 1 \quad \left(\Longrightarrow \alpha = \frac{1}{10}, \beta = -\frac{3}{50}\right) \cdot$$

Deci soluţia particulară este de forma

$$z(t) = \left(\frac{1}{10}t - \frac{3}{50}\right)e^{3t}, \ t \in \Gamma.$$

În acest caz, soluţia generală a ecuaţiei noastre este următoarea

$$x(t) = C_1 e^{2t} + C_2 \cos t + C_3 \sin t + \left(\frac{1}{10}t - \frac{3}{50}\right)e^{3t}, \ t \in R,$$

unde C_1, C_2, C_3 sunt constante arbitrare.
Să observăm în final că, dacă ecuaţia iniţială s-ar fi modificat în

$$x''' - 2x'' + x' - 2x = (t+1)e^{2t}$$

atunci, o soluţie particulară a acesteia s-ar fi căutat de forma

$$z(t) = t(\alpha t + \beta)e^{2t},\ t \in R.$$

Aceasta deoarece $\lambda = 2$ apare ca rădăcină (simplă) a ecuaţiei caracteristice. Alte situaţii de acest fel pot fi tratate în mod asemănător.

6.3.5 Sisteme diferenţiale liniare

Numim *sistem diferenţial liniar* în k funcţii necunoscute, orice sistem diferenţial de forma

$$(S) \qquad \begin{cases} x_1' = a_{11}(t)x_1 + \cdots + a_{1k}(t)x_k + b_1(t) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_k' = a_{k1}(t)x_1 + \cdots + a_{kk}(t)x_k + b_k(t) \end{cases}$$

unde funcţiiile $a_{ij}(t), 1 \le i, j \le k$ şi $b_i(t), 1 \le i \le k$ sunt continue pe un interval deschis Γ. Ca de obicei, un asemenea sistem se asociază cu condiţia iniţială

$$(J) \qquad x_1(t_0) = x_1^0, ..., x_k(t_0) = x_k^0$$

unde t_0 este un element din Γ, iar $x^0 = (x_1^0, ..., x_k^0)^\top$, un vector din spaţiul R^k. Cu privire la problema Cauchy $(S)+(J)$ ne interesează aici două aspecte

(i) existenţa şi unicitatea soluţiei (pentru x^0 fixat în R^k)

(ii) structura mulţimii soluţiilor (pentru x^0 arbitrar în R^k).

În ce priveşte primul aspect, să notăm că funcţiile

$$f_i(t, x_1, ..., x_k) = a_{i1}(t)x_1 + \cdots + a_{ik}(t)x_k + b_i(t),$$
$$t \in \Gamma,\ (x_1, ..., x_k) \in R^k, 1 \le i \le k$$

sunt continue pe domeniul lor de definiţie $\Omega = \Gamma \times R^k$ şi local lipschitziene în raport cu intervalul Γ. Deci, conform rezultatelor anterioare, problema Cauchy $(S) + (J)$ admite un sistem unic de soluţii $t \vdash x_i(t)$, $1 \le i \le k$, definite pe întreg intervalul Γ. Mai exact, să introducem notaţiile

$$A(t) = (a_{ij}(t)), b(t) = (b_1(t), ..., b_k(t))^\top, x(t) = (x_1(t), ..., x_k(t))^\top, t \in \Gamma.$$

Problema Cauchy $(S) + (J)$ devine astfel o problemă vectorială

$$(E) \qquad x'(t) = A(t)x(t) + b(t)$$

$$(I) \qquad x(t_0 = x^0.$$

În acest caz, rezultatele citate ne arată că există o singură funcţie vectorială x în $(C^{(1)}(\Gamma))^k$, ce verifică $(E) + (I)$.

Trecem acum la structura mulţimii soluţiilor lui $(S)+(J)$ (sau, echivalent, $(E) + (I)$). Pentru aceasta, este necesar să începem cu sistemul omogen asociat lui (S)

$$(S_0) \qquad \begin{cases} x_1' = a_{11}(t)x_1 + \cdots + a_{1k}(t)x_k \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_k' = a_{k1}(t)x_1 + \cdots + a_{kk}(t)x_k. \end{cases}$$

sau, într-o scriere vectorială (cu notaţiile anterioare)

$$(E_0) \qquad x' = A(t)x.$$

Să notăm şi aici prin $X(\Gamma)$ mulţimea soluţiilor din $(C^{(1)}(\Gamma))^k$ ale sistemului (S_0). Nu este greu de văzut că aceasta apare ca spaţiu liniar (mai exact, un subspaţiu liniar al lui $(C^{(1)}(\Gamma))^k$). Chestiunea care se pune este precizarea dimensiunii acestui spaţiu. Următorul rezultat (dat fără demonstraţie) se va dovedi util în acest sens.

Teorema 6.3.8 *Fie*

$$y_1 = (y_1, ..., y_k)^\top, ..., y_k = (y_{k1}, ..., y_{kk})^\top$$

soluţii din $(C^{(1)}(\Gamma))^k$ *ale sistemului diferenţial omogen* (S_0).
Afirmaţiile următoare sunt echivalente

(P1) $\{y_1, ..., y_k\}$ *sunt liniar independente în* $(C^{(1)}(\Gamma))^k$.

(P2) *matricea* $[y_1(t), ..., y_k(t)]$ *este nesingulară pentru orice* $t \in \Gamma$.

(P3) *matricea* $[y_1(t_0), ..., y_k(t_0)]$ *este nesingulară pentru un* $t_0 \in \Gamma$.

Suntem acum în măsură să dăm răspunsul la întrebarea pusă anterior.

Teorema 6.3.9 *Mulţimea $X(\Gamma)$ a soluţiilor sistemului omogen (S_0) formează un subspaţiu k–dimensional al lui $(C^{(1)}(\Gamma))^k$.*

Demonstraţie. (Schiţă) Fie $t_0 \in \Gamma$ element fixat. Pentru fiecare $i \in \{1, ..., k\}$, să notăm cu y^i soluţia din $(C^{(1)}(\Gamma))^k$ a problemei Cauchy

$$x' = A(t)x, \ x(t_0) = e_i.$$

(Aici, $\{e_1, ..., e_k\}$ este baza canonică din R^k.) Din

$$[y^1(t_0), ..., y^k(t_0)] = I \ \ \text{(matricea unitate din } \mathcal{M}(k))$$

şi rezultatul precedent, $\{y^1, ..., y^k\}$ apare ca un sistem liniar independent în $X(\Gamma)$. Vom numi acesta *sistem fundamental de soluţii* ale lui (S_0). Fie acum $y = (y_1, ..., y_k)^\top$, o soluţie arbitrară a lui (S_0). Avem

$$y(t_0) = y_1(t_0)y^1(t_0) + \cdots + y_k(t_0)y^k(t_0).$$

De aici şi din partea de unicitate a Teoremei 6.3.3,

$$y = y_1(t_0)y^1 + \cdots + y_k(t_0)y^k.$$

Aceasta încheie demonstraţia. ■

Putem acum să ne întoarcem la problema structurii mulţimii soluţiilor sistemului neomogen (S). În acest sens avem imediat

Teorema 6.3.10 *Formula*

(R1) $x(t) = C_1 y^1(t) + \cdots C_k y^k(t) + z(t), \ t \in \Gamma$

unde $C_1, ..., C_k$ sunt nişte constante arbitrare, $\{y^1(t), ..., y^k(t)\}$ este un sistem fundamental de soluţii ale sistemului omogen (S_0), iar $z(t)$ o soluţie particulară a sistemului neomogen (S), reprezintă expresia soluţiei generale a sistemului neomogen (S).

O problemă naturală care se pune în acest cadru este aceea a determinării unei asemenea soluţii particulare. Se poate face aceasta, dacă dispunem de un sistem fundamental de soluţii ale sistemului omogen (S_0). Anume,

$$y^1(t) = \begin{pmatrix} y_{11}(t) \\ \cdots\cdots \\ y_{1k}(t) \end{pmatrix}, ..., y^k(t) = \begin{pmatrix} y_{k1}(t) \\ \cdots\cdots \\ y_{kk}(t) \end{pmatrix}$$

fiind un asemenea sistem, să căutăm respectiva soluţie sub forma

(R2) $z(t) = C_1(t)y^1(t) + \cdots C_k(t)y^k(t),\ t \in \Gamma$

unde funcţiile derivabile $C_1(t), ..., C_k(t)$ urmează a fi determinate din sistemul liniar derivat

$$(SD) \qquad C'_1(t)y^1(t) + \cdots + C'_k(t)y^k(t) = b(t),\ t \in \Gamma.$$

Matricea asociată sistemului acesta este nesingulară,conform alegerii funcţiilor $y^1(t), ..., y^k(t)$. Deci sistemul (SD) are soluţie unică $C'_1(t)$,...,$C'_k(t)$; iar aceasta, prin integrare, permite determinarea funcţiilor $C_1(t)$, ...,$C_k(t)$.Se poate acum arăta uşor că funcţia z dată de (R2) verifică într-adevăr sistemul neomogen (S); şi cu aceasta, afirmaţia reiese. ■

Un caz particular, important pentru aplicaţii este acela în care funcţiile coeficient $a_{ij}(t), 1 \le i, j \le k$ sunt constante. Este vorba deci de sistemele diferenţiale liniare

$$(S) \qquad \begin{cases} x'_1 = a_{11}x_1 + \cdots + a_{1k}x_k + b_1(t), \quad t \in \Gamma \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x'_k = a_{k1}x_1 + \cdots + a_{kk}x_k + b_k(t), \quad t \in \Gamma \end{cases}$$

sau vectorial (cu notaţia $A = (a_{ij})$)

$$(E) \qquad x' = Ax + b(t), \quad t \in \Gamma.$$

Aspectul care trebuie subliniat aici este acela al determinării directe a unui sistem fundamental de soluţii pentru sistemul omogen

$$(S_0) \qquad \begin{cases} x'_1 = a_{11}x_1 + \cdots + a_{1k}x_m \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x'_k = a_{k1}x_1 + \cdots + a_{kk}x_k \end{cases}$$

sau, echivalent, pentru ecuaţia vectorială omogenă

$$(E_0) \qquad x' = Ax.$$

Mai exact, din ultima ecuaţie avem, prin derivări succesive,

$$x^{(n)} = A^n x,\ n = 1, 2, ...$$

Să introducem polinomul caracteristic al lui A:

$$P(\lambda) = \det(A - \lambda I) = (-1)^k\lambda^k + r_1\lambda^{k-1} + \cdots + r_{k-1}\lambda + r_k,\ \lambda \in C.$$

Conform teoremei Cayley–Hamilton (Secţiunea 1.3.3), $P(A) = 0$; adică

$$(-1)^k A^k + r_1 A^{k-1} + \cdots + r_{k-1} A + r_k I = 0.$$

De aici, cu formulele deja scrise, deducem

$$(-1)^k x^{(k)} + r_1 x^{(k-1)} + \cdots + r_{k-1} x' + r_k x = 0.$$

Adică, toate funcţiile componente $(x_1(t), ..., x_k(t))$ ale funcţiei vectoriale $x(t)$ satisfac una şi aceeaşi ecuaţie diferenţială liniară de ordinul k cu coeficienţi constanţi. Aceasta permite determinarea unui sistem fundamental de soluţii pentru sistemul (S_0) şi deci scrierea soluţiei acestuia.

Exemplu. Să se rezolve sistemul diferenţial

$$x' = -x + 3y, \quad y' = -3x + 5y.$$

Matricea sistemului şi polinomul ei caracteristic sunt

$$A = \begin{pmatrix} -1 & 3 \\ -3 & 5 \end{pmatrix}; \; P(\lambda) = \det \begin{pmatrix} -1-\lambda & 3 \\ -3 & 5-\lambda \end{pmatrix} = \lambda^2 - 4\lambda + 4.$$

Rezultă că ambele componente satisfac ecuaţia

$$x'' - 4x' + 4x = 0, \; y'' - 4y' + 4y = 0.$$

Ecuaţia caracteristică (dată de polinomul caracteristic) este

$$\lambda^2 - 4\lambda + 4 = 0, \; \text{cu rădăcinile } \lambda_1 = \lambda_2 = 2.$$

Rezultă de aici forma componentelor sistemului

$$x(t) = C_1 e^{2t} + C_2 t e^{2t}, \; y(t) = C_3 e^{2t} + C_4 t e^{2t}$$

unde C_1, C_2, C_3, C_4 sunt constante. Pentru a reduce (la două) numărul acestora, verificăm sistemul iniţial cu aceste funcţii şi avem identităţile

$$\begin{cases} 2C_1 e^{2t} + C_2(1+2t)e^{2t} = -(C_1 e^{2t} + C_2 t e^{2t}) + 3(C_3 e^{2t} + C_4 t e^{2t}) \\ 2C_3 e^{2t} + C_4(1+2t)e^{2t} = -3(C_1 e^{2t} + C_2 t e^{2t}) + 5(C_3 e^{2t} + C_4 t e^{2t}) \end{cases}$$

Identificând coeficienţii pe lângă e^{2t} şi te^{2t}, rezultă sistemul

$$\begin{aligned} 2C_1 + C_2 &= -C_1 + 3C_3, \quad & 2C_2 &= -C_2 + 3C_4, \\ 2C_3 + C_4 &= -3C_1 + 5C_3, \quad & 2C_4 &= -3C_2 + 5C_4. \end{aligned}$$

Se obţine soluţia (depinzând de parametrii α, β)

$$C_1 = \alpha,\ C_2 = 3\beta - 3\alpha,\ C_3 = \beta,\ C_4 = 3\beta - 3\alpha.$$

Şi de aici găsim expresia soluţiei sistemului omogen

$$x(t) = \alpha e^{2t} + (3\beta - 3\alpha)te^{2t},\ y(t) = \beta e^{2t} + (3\beta - 3\alpha)te^{2t}$$

sau, sub o formă matricială,

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} (1-3t)e^{2t} & 3te^{2t} \\ -3te^{2t} & (1+3t)e^{2t} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Cu această ocazie am obţinut şi un sistem fundamental de soluţii ale sistemului omogen

$$v(t) = \begin{pmatrix} 1-3t \\ -3t \end{pmatrix} e^{2t},\quad w(t) = \begin{pmatrix} 3t \\ 1+3t \end{pmatrix} e^{2t}.$$

Să notăm cu această ocazie că, deoarece

$$v(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\quad w(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

matricea funcţională obţinută punând una lângă alta coloanele celor două soluţii este tocmai matricea exponenţială (vezi Secţiunea 4.4.4):

$$e^{At} = [v(t), w(t)] = \begin{pmatrix} (1-3t)e^{2t} & 3te^{2t} \\ -3te^{2t} & (1+3t)e^{2t} \end{pmatrix}.$$

În fine, relativ la determinarea unei soluţii particulare a sistemului nneomogen, metoda generală a fost deja expusă la cazul ne–constant. Există şi aici posibilitatea de găsire mai directă a acesteia, pentru termeni liberi de forma

$$b(t) = (P(t)\cos\beta t + Q(t)\sin\beta t)e^{\alpha t}$$

unde $P(t)$ şi $Q(t)$ sunt polinoame cu coeficienţi vectoriali. Anume, se va căuta respectiva soluţie de o formă apropiată celei deja scrise; nu mai dăm alte detalii.

Probleme propuse la §6.3

1. Să se integreze ecuaţiile diferenţiale

(a) $x' = \dfrac{-2t}{t^2-1}x^2$; (b) $x' = \dfrac{x}{t} - e^{\frac{x}{t}}$

(c) $x' = \dfrac{x+t-2}{x-t-4}$; (d) $x' + 2tx - 2te^{-t^2} = 0$

(e) $x' = -\dfrac{x}{t} - t''x^3e^t$; (f) $x' = x^2 - 2xe^t + e^{2t} + e^t$.

(Indicaţie.(c): Se face schimbarea de funcţie şi de variabilă $s = t+1, y = x-3$; (f):Se face schimbarea de funcţie $x = y + e^t$).

2. Să se calculeze primele trei aproximante succesive pentru soluţia problemei Cauchy

$$x' = 2t + 3x^4,\ x(0) = 0;\ t \in [-1, 1],\ x \in R.$$

3. Să se rezolve ecuaţia diferenţială

$$x'' - 5x' + 4x = b(t),\ t > 0$$

pentru următoarele alegeri de termen liber

$$b(t) = t^2 + 1;\ b(t) = te^t;\ b(t) = \ln(t + 1);\ b(t) = t + 4^t.$$

4. Să se rezolve problema

$$x'' - 9x = t,\ x(0) = 0,\ x'(1) = 0.$$

5. Să se arate că orice soluţie, x, a ecuaţiei

$$x'' + 2x' + 2x = (t^2 + t + 1)e^{-kt}, \quad \text{unde } k > 0,$$

are proprietatea $\lim\limits_{t\to\infty} x(t) = 0$.

6. Să se rezolve sistemele (omogene)

(a) $x' = 3x - 2y,\ y' = 4x - y$;

(b) $x' = 2x - y - z,\ y' = x - z,\ z' = 3x - y - 2z$.

7. Să se rezolve sistemele (neomogene)

(a) $x' = 2x + y + 2e^t,\ y' = x + 2y - 3e^{4t};$

(b) $x' = y - 5\cos t,\ y' = 2x + y.$

(Indicaţie. Se va folosi metoda eliminării.)

8. Să se rezolve sistemul diferenţial cu derivate de ordin superior

$$x'' + 5x' + 2y' + y = 0,\ 3x'' + 5x + y' + 3y = 0.$$

(Indicaţie. Se va folosi şi aici metoda eliminării.)

9. Să se calculeze matricea exponenţială e^{At} pentru matricile

$$\text{(a)}\ \ A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ \ \text{(b)}\ \ A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

(Indicaţie. Se va rezolva sistemul diferenţial care o defineşte.)

6.4 Gestiunea stocurilor

6.4.1 Terminologie şi notaţii

Orice activitate de producţie (respectiv, de desfacere) presupune asigurarea de resurse (respectiv, produse). Consumul acestora se face în mod continuu; dar, re–înnoirea lor, în mod discret (la anumite momente). Înţelegem prin *stoc* cantitatea de resurse (respectiv, produse) care există între momentele de re–înnoire a lor. Următoarele elemente principale intervin în analiza unei asemenea probleme

(i) Cererea de produse (din partea beneficiarilor)

(ii) Costuri asociate diferitelor operaţii.

Acestea din urmă sunt de mai multe feluri:

(A) *Costul de lansare a comenzii* (independent de cantitatea de produse comandată); îl vom nota prin K.

(B) *Costul de achiziţie* (valoarea la care este evaluată o unitate din produsul respectiv); se va nota prin c.

(C) *Costul de stocare* (valoarea stocării unei unităţi din acel produs în unitatea de timp); vom nota prin h această mărime.

(D) *Costul de penalizare* (valoarea care se pierde prin lipsa din stoc a unei unităţi de produs în unitatea de timp); îl vom nota prin b.

În analiza unui sistem de stoc se folosesc anumite reguli de decizie care stabilesc când şi cât de mari trebuie să fie comenzile de re–aprovizionare; vom numi acestea *politici de stoc*. Una din cele mai întâlnite politici de stoc este aceea de tip (u, U). Semnificaţia ei este următoarea: când stocul atinge nivelul u, se face o re–aprovizionare până la atingerea nivelului U. Variabilele u şi U se numesc *variabile de decizie*.

Obiectivele teoriei stocurilor sunt esenţial legate de timpul de funcţionare a sistemului, numit *perioadă orizont*. Anume, pentru cazul unei perioade orizont infinită, obiectivul de urmărit este minimizarea costului mediu sau maximizarea beneficiului mediu pe unitatea de timp. Iar, în cazul unei perioade orizont finită, se pune problema minimizării costului total sau maximizării beneficiului total pe acea perioadă.

Din punct de vedere dinamic, un sistem de stoc poate fi imaginat prin mişcarea unui produs către interiorul – şi, ulterior, către exteriorul – unui punct de stocare. Să facem notaţia

(D1) $S(t)$ = nivelul stocului la momentul $t \geq 0$.

O valoare pozitivă pentru $S(t)$ indică existenţa unui stoc (real); iar o valoare negativă pentru $S(t)$ arată lipsă de stoc. Dacă mai notăm

$$S^{(+)}(t) = \max(S(t), 0), \quad S^{(-)}(t) = \max(-S(t), 0)$$

avem identitatea

$$S(t) = S^{(+)}(t) - S^{(-)}(t), \ t \geq 0.$$

Aici, $S^{(+)}(t)$ este tocmai stocul disponibil la momentul t, iar $S^{(-)}(t)$, lipsa de stoc la momentul t. Momentul $t = 0$ îl considerăm, de regulă, momentul intrării în funcţie a sistemului.

O problemă importantă care se pune aici este aceea a evoluţiei în timp a nivelului de stoc. De obicei, aceasta se rezolvă prin

(a) modele discrete (ecuaţii cu diferenţe)

(b) modele continue (ecuaţii diferenţiale).

Să ne referim la cel de-al doilea caz. Cel mai simplu model de stoc este acela definit de ecuaţia (de balanţă)

$$S'(t) = p(t) - q(t),\ t \geq 0$$

unde, prin definiţie, am pus (pentru toţi $t \geq 0$)

(D2) $\begin{cases} p(t) = \text{rata intrării în stoc la momentul } t \\ q(t) = \text{rata ieşirii din stoc la momentul } t \end{cases}$

Se obţine formula stocului (raportată la timp)

(R1) $S(t) = S(0) + \int_0^t (p(\tau) - q(\tau))d\tau,\ t \geq 0.$

Un caz mai complicat este acela în care o anumită parte din produsul stocat se pierde, din cauze obiective. (Un exemplu ar fi acela al produselor perisabile.) În acest caz, modelul de stoc este definit de ecuaţia de balanţă

(R2) $S'(t) + \mu S(t) = p(t) - q(t),\ t \geq 0,$

unde $p(t), q(t)$ au semnificaţiile anterioare, iar $\mu > 0$ este *factorul de depreciere al produsului.* Prin rezolvarea ecuaţiei diferenţiale apărute, obţinem formula stocului

$$S(t) = S(0)e^{-\mu t} + \int_0^t e^{\mu(t-\tau)}(p(\tau) - q(\tau))d\tau,\ t \geq 0.$$

Desigur, alte situaţii pot implica şi modele de balanţă mai complicate (şi deci, formule de stoc mai complicate). Ne vom mărgini însă doar la cele deja scrise.

6.4.2 Model de stoc cu perioadă fixă şi cerere constantă

Considerăm un sistem de stoc în ipotezele următoare:

(C1) sistemul are perioadă orizont infinită, compusă din cicluri aprovizionare–consum identice, de lungime θ;

(C2) politica de stoc este de tip (u, U) cu $0 \leq u < U$ (deci, nu se admit lipsuri din stoc);

(C3) comanda de aprovizionare de mărime $z = U - u$ este satisfăcută imediat ($p(t) = 0,\ 0 \leq t \leq \theta$);

(C4) rata de consum din stoc este constantă ($q(t) = q,\ 0 \leq t \leq \theta$).

Se cunosc:

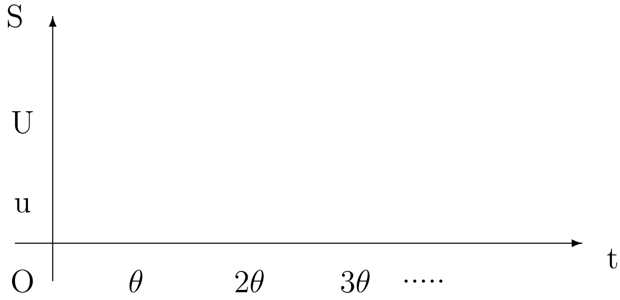
– costul de lansare a comenzii (K);

– costul de achiziţie (c);

– costul de stocare (h).

În aceste condiţii se cere să se determine variabilele de decizie (u, U) astfel încât costul total mediu (de aprovizionare şi stocare) pe unitatea de timp să fie minim.

Pentru rezolvare vom porni de la graficul nivelului de stoc



Din condiţiile problemei rezultă formula stocului

(R1) $S(t) = u + z - qt,\ 0 \leq t \leq \theta$.

Întrucât $S(\theta) = u$, trebuie să avem $z = q\theta$ sau $\theta = \dfrac{z}{q}$; iar aceasta ne dă lungimea ciclului. Costul de aprovizionare pe ciclu este $K + cz$; deci, pe unitate de timp, acesta este

(R2) $C_1 = \dfrac{1}{\theta}(K + cz) = \dfrac{K}{\theta} + c\dfrac{z}{\theta} = K\dfrac{q}{z} + cq.$

Costul de stocare pe ciclu este

$$h\int_0^\theta S(\tau)d\tau = h\int_0^\theta (u + z - q\tau)d\tau = h\left(u + \frac{z}{2}\right)\theta;$$

şi, deci, pe unitate de timp, acesta este

(R3) $C_2 = \frac{1}{\theta} h \left(u + \frac{z}{2} \right) \theta = h \left(u + \frac{z}{2} \right)$

Urmează de aici că, costul total mediu (de aprovizionare şi stocare) pe unitatea de timp este

$$F(u, z) = C_1 + C_2 = K\frac{q}{z} + cq + h \left(u + \frac{z}{2} \right),\ \ u \geq 0, z > 0.$$

Problema formulată iniţial se transcrie acum astfel

$$(P) \qquad \min\{F(u, z);\ \ u \geq 0, z > 0\}.$$

Aceasta este o problemă de programare neliniară cu restricţii inegalităţi (Secţiunea 5.4.2). Concret, restricţiile sunt date de funcţia

$$g(u, z) = -u,\ \ u, z \in R$$

iar domeniul comun al funcţiilor F şi g este

$$D = \{u, z \in R;\ \ z > 0\} \text{ (parte deschisă din } R^2\text{)}.$$

Suntem deci în condiţiile aplicării condiţiilor necesare Kuhn–Tucker. Putem însă proceda şi mai simplu. Anume, cu notaţia $F(0, z) = G(z)$ şi evaluarea

$$F(u, z) \geq G(z),\ \ u \geq 0, z > 0$$

se poate spune că (P) este reductibilă la problema

$$(P_0) \qquad \min\{G(z) = K\frac{q}{z} + cq + h\frac{z}{2};\ \ z > 0\}$$

care apare ca problemă de extrem liber. Avem

$$G'(z) = 0 \Longrightarrow -\frac{Kq}{z^2} + \frac{h}{2} = 0 \Longrightarrow z^* = \sqrt{\frac{2Kq}{h}}\cdot$$

În concluzie, soluţia optimă a problemei este $(u^* = 0, U^* = z^*)$ cu z^* dat de formula dinainte. Valoarea minimă a funcţiei obiectiv este atunci

(R4) $F(u^*, z^*) = cq + \sqrt{2Kqh}.$

Rezultă de aici şi timpul optim între două comenzi consecutive

(R5) $\theta^* = \dfrac{z^*}{q} = \sqrt{\dfrac{2K}{hq}}.$

Exemplu. Să presupunem că cererea zilnică a unui anumit material de construcţie este

$$q = 20000 \text{ bucăţi/zi.}$$

De asemenea, se mai cunosc

– costul de lansare a unei comenzi: K=34000 lei/lot

– costul de achiziţie: C=15 lei/buc

– costul de stocare: h=2 lei/(unitate de produs)×(zi).

Se cere să se determine lotul optim de marfă care trebuie comandat precum şi intervalul optim între două comenzi succesive, pentru ca costul total de aprovizionare şi stocare pe unitatea de timp să fie minim.
Pentru rezolvare,aplicăm formula precedentă:

$$u^* = 0,\ U^* = \sqrt{\frac{2Kq}{h}} = \sqrt{\frac{2 \cdot 34000 \cdot 20000}{2}} \approx 26000$$
$$\theta^* = \sqrt{\frac{2K}{h}} = \sqrt{\frac{2 \cdot 34000}{2 \cdot 20000}} \approx 1,3.$$

Cu alte cuvinte, cea mai economică metodă este de a face câte o comandă de 26000 unităţi din produsul respectiv, la fiecare 1,3 zile. Valoarea optimă a funcţiei obiectiv corespunzătoare acestor date este

$$F_{min} = cq + \sqrt{2Kqh} = 15 \cdot 20000 + \sqrt{2 \cdot 34000 \cdot 20000 \cdot 2} = 352000 \text{ lei.}$$

6.4.3 Model de stoc cu perioadă fixă, cerere constantă şi cu posibilitatea lipsei de stoc

Să considerăm acum un sistem de stoc în ipotezele următoare:

(C1) sistemul are perioadă orizont infinită, compusă din cicluri aprovizionare–consum identice, de lungime θ;

(C2) politica de stoc este de tip (u, U) unde $u < 0 < U$ (deci, se admit şi lipsuri din stoc);

(C3) comanda de aprovizionare de mărime $z = U - u$ este satisfăcută imediat $(p(t) = 0,\ 0 \leq t \leq \theta)$;

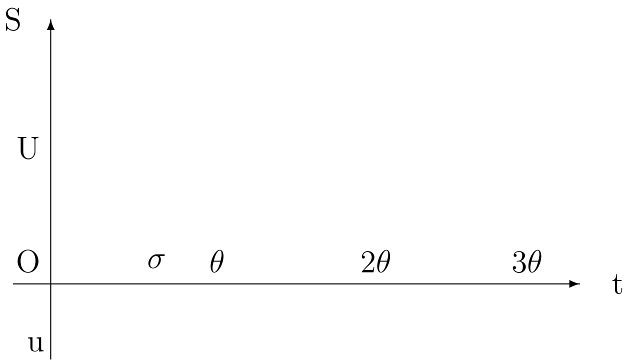
(C4) rata de consum din stoc este constantă $(q(t) = q,\ 0 \leq t \leq \theta)$.

Cunoaştem şi aici: costul de lansare a comenzii (K), costul de achiziţie (c), costul de stocare (h) ; şi,de asemenea,

– costul de penalizare (b).

Problema care se pune este următoarea: să se determine variabilele de decizie (u, U) astfel încât costul total mediu (de aprovizionare, stocare şi penalizare) pe unitate de timp să fie fie minim.

Pentru rezolvare vom porni şi aici de la graficul nivelului de stoc:



Din condiţiile problemei rezultă imediat formula stocului

(R1) $S(t) = u + z - qt,\ 0 \leq t \leq \theta.$

Întrucât $u < 0$, va exista un moment $\sigma \in (0, \theta)$ cu

$$S(\sigma) = 0; \quad \text{deci} \quad \sigma = \frac{u+z}{q} = \frac{U}{q}\cdot$$

În plus, evident

$$S(t) > 0;\ 0 \leq t < \sigma; \quad S(t) < 0,\ \sigma < t \leq \theta.$$

Sau, cu nişte notaţii anterior acceptate,

$$S^{(+)}(t) = \begin{cases} u + z - qt, & 0 \leq t \leq \sigma \\ 0, & \sigma \leq t \leq \theta \end{cases}$$

(formula de stoc disponibil pe perioada unui ciclu)

$$S^{(-)}(t) = \begin{cases} 0, & 0 \leq t \leq \sigma \\ qt - u - z, & \sigma \leq t \leq \theta \end{cases}$$

(formula de stoc indisponibil pe perioada unui ciclu).
Calculăm acum valoarea stocului disponibil pe perioada unui ciclu:

$$\int_0^\theta S^{(+)}(\tau)d\tau = \int_0^\sigma (u + z - q\tau)d\tau = (u + z)\sigma - q\frac{\sigma^2}{2}\cdot$$

Deci, costul de stocare pe unitatea de timp este

(R2) $C_1 = \dfrac{h}{\theta}\displaystyle\int_0^\theta S^{(+)}(\tau)d\tau = h\left[(u + z)\frac{\sigma}{\theta} - q\frac{\sigma^2}{2\theta}\right].$

Valoarea stocului indisponibil (adică, a cererii totale neonorate) pe perioada unui ciclu este

$$\int_0^\theta S^{(-)}(\tau)d\tau = \int_\sigma^\theta (q\tau - u - z)d\tau = q\frac{\theta^2 - \sigma^2}{2} - (u + z)(\theta - \sigma).$$

Şi, deci, costul de penalizare pe unitatea de timp va avea expresia:

(R3) $C_2 = \dfrac{b}{\theta}\displaystyle\int_0^\theta S^{(-)}(\tau)d\tau = b\left[\frac{q}{2}\,\frac{\theta^2 - \sigma^2}{2\theta} - (u + z)\frac{\theta - \sigma}{\theta}\right].$

În fine, costul de aprovizionare pe unitate de timp este

(R4) $C_3 = \dfrac{1}{\theta}(K + cz).$

Urmează de aici că, costul total mediu (de stocare, penalizare şi transport) pe unitatea de timp este

$$G(u, z) = C_1 + C_2 + C_3 = h\left[(u + z)\frac{\sigma}{\theta} - q\frac{\sigma^2}{2\theta}\right] +$$
$$+ b\left[\frac{q}{2}\,\frac{\theta^2 - \sigma^2}{2\theta} - (u + z)\frac{\theta - \sigma}{\theta}\right] + \frac{1}{\theta}(K + cz).$$

Cu formulele anterioare ($\theta = \frac{z}{q}$, $\sigma = \frac{U}{q}$), se obţine în final funcţia obiectiv (de două variabile)

$$F(z,U) = cq + \frac{Kq}{z} + h\frac{U^2}{2z} + b\frac{(z-U)^2}{2z},\ z > 0, U > 0.$$

Problema formulată iniţial se transcrie astfel

$$(P) \qquad \min\{F(z,U);\ z > 0, U > 0\}.$$

Aceasta este o problemă de extrem liber. Pentru rezolvare, vom anula derivatele parţiale

$$\frac{\partial F}{\partial z} = 0,\ \ \frac{\partial F}{\partial U} = 0$$

şi rezultă sistemul de ecuaţii

$$-\frac{Kq}{z^2} - \frac{h}{2}\left(\frac{U}{z}\right)^2 + \frac{b}{2}\left[1 - \left(\frac{U}{z}\right)^2\right] = 0;\ \frac{hU}{z} - \frac{b(z-U)}{z} = 0.$$

Soluţia acestui sistem este punctul staţionar

$$z^* = \sqrt{\frac{2Kq}{h}}\sqrt{\frac{b+h}{b}};\ \ U^* = \sqrt{\frac{2Kq}{h}}\sqrt{\frac{b}{b+h}}.$$

Pentru a studia natura acestui punct staţionar, vom trece la calculul derivatelor parţiale de ordinul doi în punctul $M = (z^*, U^*)$

$$\frac{\partial^2 F}{\partial z^2}(M) = (2Kqh + b)\frac{U^{*2}}{z^{*3}},\ \frac{\partial^2 F}{\partial U^2}(M) = (h+b)\frac{1}{z^{*2}}$$
$$\frac{\partial^2 F}{\partial z \partial U}(M) = \frac{\partial^2 F}{\partial U \partial z}(M) = -(h+b)\frac{U^*}{z^{*2}}.$$

Determinanţii principali ai hessianului corespunzător sunt

$$\Delta_1(M) = (2Kqh + b)\frac{U^{*2}}{z^{*3}} > 0;\ \ \Delta_2(M) = bh\frac{1}{z^{*2}} > 0.$$

Deci, punctul staţionar considerat este de minim. Mai obţinem în acelaşi timp nivelul minim de stoc

$$\text{(R4)} \ \ u^* = U^* - z^* = -\sqrt{\frac{2Kqh}{b(b+h)}}.$$

În fine,valoarea minimă corespunzătoare a funcţiei obiectiv este

(R5) $F(z^*, U^*) = cq + \sqrt{2Kqh} \cdot \sqrt{\dfrac{b}{b+h}}$

iar lungimea optimă a unui ciclu de funcţionare are expresia

(R6) $\theta^* = \dfrac{z^*}{q} = \sqrt{\dfrac{2K}{hq}} \cdot \sqrt{\dfrac{b+h}{b}}.$

Să notăm că, dacă preţul de penalizare (b) tinde către infinit, sistemul de stoc discutat se apropie de cel anterior (în care nu existau lipsuri în stoc).

Putem, de asemenea, calcula lungimea optimă a intervalului în care stocul este disponibil

(R7) $\sigma^* = \dfrac{U^*}{q} = \sqrt{\dfrac{2K}{hq}}\sqrt{\dfrac{b}{b+h}}.$

Mai obţinem de aici raportul

(R8) $\dfrac{\sigma^*}{\theta^*} = \dfrac{b}{b+h} \quad \left(\Longrightarrow 1 - \dfrac{\sigma^*}{\theta^*} = \dfrac{h}{b+h}\right).$

Aceasta serveşte beneficiarului la a impune un preţ de penalizare astfel încât produsul să lipsească un timp cât mai scurt din stoc.

Exemplu. Vom relua exemplul anterior, în care însă , admitem lipsă de stoc cu

– costul de penalizare: b=4 lei/(unitate de produs)×(zi).

Se cere să se stabilească lotul optim ce trebuie comandat şi intervalul optim între comenzi, astfel ca costul total de aprovizionare, stocare şi penalizare pe unitatea de timp să fie minim.
Vom aplica formulele precedente:

$$u^* = -\sqrt{\frac{2Kqh}{b(b+h)}} = -\sqrt{\frac{2 \cdot 34000 \cdot 20000 \cdot 2}{4(4+2)}} = -10600$$

$$U^* = -\sqrt{\frac{2Kqb}{h(b+h)}} = \sqrt{\frac{2 \cdot 34000 \cdot 20000 \cdot 4}{2(2+4)}} = 21000$$

$$\theta^* = \sqrt{\frac{2K(b+h)}{hqb}} = \sqrt{\frac{2 \cdot 34000 \cdot (4+2)}{2 \cdot 20000 \cdot 4}} = 1,5\ .$$

Cu alte cuvinte, cea mai economică metodă este următoarea: când lipsa de stoc a ajuns la 10600 bucăţi,se face o comandă de 21000–(-10600)=31600 bucăţi cu care se acoperă în primul rând lipsa respectivă şi apoi se consumă normal din stocul rămas. Comanda trebuie făcută la fiecare 1,5 zile.

6.4.4 Model de stoc cu perioadă orizont finită şi cerere variabilă

Considerăm un sistem de stoc în ipotezele următoare:

(C1) sistemul are perioadă orizont finită din n cicluri, cu caracteristici distincte, dar de aceeaşi lungime (θ);

(C2) politica de stoc pe fiecare ciclu (i) este de tip (u_i, U_i), cu $0 \leq u_i \leq U_i \leq U$; adică, nu se admite lipsă de stoc ($u_i \geq 0$) şi nici depăşirea unui stoc maxim ($U_i \leq U$);

(C3) comanda de aprovizionare pe fiecare ciclu (i), de mărime $z_i = U_i - u_i$ este satisfăcută imediat ($p_i(t) = 0$);

(C4) rata de consum pe fiecare ciclu (i) este constantă ($q_i(t) = q_i$) cu $q_i\theta \leq U,\ 0 \leq i \leq n-1$.

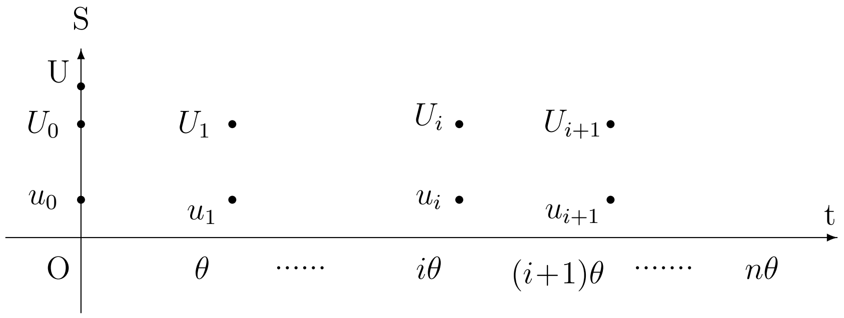
Se mai dau costul de lansare a comenzii (K) şi, pentru fiecare ciclu (i),

– costul de achiziţie (c_i),

– costul de stocare (h_i).

Problema care se pune este următoarea: să se determine variabilele de decizie (u_i, U_i) astfel încât costul total (de aprovizionare şi stocare) pe cele n perioade să fie minim.

Schema evoluţiei sistemului este dată de diagrama de mai jos:



După cum se vede, avem de-a face cu un proces secvenţial de decizie. Mulţimea stărilor sistemului este

(R1) $X_i = [0, U]$, pentru $i = 0, 1, ..., n$.

Să ne oprim la un moment oarecare $i\theta$, $0 \leq i \leq n-1$. Dată starea $u \in X_i$ mulţimea deciziilor (z) ce se pot lua asupra acesteia este caracterizată de condiţiile

$$u \leq u + z \leq U; \quad \text{adică} \quad 0 \leq z \leq U - u.$$

Fie z o asemenea decizie. Nivelul de stoc pe intervalul $[i\theta, (i+1)\theta]$ va avea forma

(R2) $S(t) = u + z - q_i(t - i\theta), \ i\theta \leq t \leq (i+1)\theta$.

Cum nu avem lipsă de stoc, este necesar ca

$$u + z - q_i\theta \geq 0; \quad \text{adică } z \geq q_i\theta - u.$$

Deducem de aici mulţimea stărilor posibile la care se poate ajunge din starea $u \in X_i$ prin aplicarea unei decizii z (de asemenea formă)

(R3) $X_{i+1}(i; u) = \{x \in [0, U]; \ \max(0, u - q_i\theta) \leq x \leq U - q_i\theta\}$,
$0 \leq i \leq n-1$.

Să notăm că această mulţime este nevidă, deoarece

$$u \leq U \Longrightarrow u - q_i\theta \leq U - q_i\theta; \quad \text{şi,în plus,}$$
$$0 \leq U - q_i\theta \quad \text{(conform ipotezei admise iniţial).}$$

Să calculăm acum costul unei tranziţii din starea $u \in X_i$ la starea $v = u + z - q_i\theta \in X_{i+1}(i; u)$. În primul rând, costul de aprovizionare (cu cantitatea z) pe intervalul $[i\theta, (i+1)\theta]$ este

(R4) $C_1^{(i)} = K + c_i z = K + c_i(v - u + q_i\theta)$.

Apoi, costul de stocare pe acelaşi interval este

(R5) $C_2^{(i)} = h_i \int_{i\theta}^{(i+1)\theta} S(\tau)d\tau = h_i(u + z - \frac{q_i}{2}\theta)\theta = h_i(v + \frac{q_i}{2}\theta)\theta.$

Pr4n urmare, costul total al tranziţiei este

$$W_i(u, v) = C_1^{(i)} + C_2^{(i)} = K + c_i q_i \theta + \frac{h_i}{2} q_i \theta^2 - c_i u + (h_i + c_i)v.$$

De acum încolo, rezolvarea problemei se obţine prin metodele programării dinamice. Mai exact, pentru fiecare stare $u \in X_i$ să notăm

(D1) $\begin{cases} F_i(u) = \text{costul minim al unei tranziţii de stare} \\ \text{de forma } (u, u_{i+1}, ..., u_n), \text{ unde } u_n \text{ este arbitrar în } X_n. \end{cases}$

În baza formulelor de recurenţă de tip retrospectiv ale programării dinamice, avem

$$F_i(u) = \min\{W_i(u, v) + F_{i+1}(v);\ v \in X_{i+1}(i; u)\},\ 0 \le i \le n - 2.$$

Aceasta ne permite ca, din aproape în aproape să determinăm funcţia de cost minim

(D2) $F_0(u)$, u = stare arbitrară din X_0.

Şi astfel, problema se poate considera rezolvată.

Exemplu. Să presupunem că, la fiecare decadă $\theta = 10$ zile, cererea zilnică a unui produs, costurile de achiziţie şi costurile de stocare sunt cele date în tabelul de mai jos:

Perioada (i)	0	1	2
rata de cerere (q_i) (bucăţi/zi)	10	30	20
costul de achiziţie (c_i) (lei/bucată)	15	25	20
costul de stocare (h_i) (lei/(bucată)×(zi))	2	4	2

Capacitatea de stocare a produsului este $U = 600$ bucăţi, iar costul de lansare

a unei comenzi este $K = 200$ lei/lot. Stocul iniţial este $u_0 = 100$ buc c3i. Se cere să se determine cantităţile ce trebuie comandate la începutul fiecărei perioade aşa încât cheltuielile lunare de aprovizionare şi stocare să fie minime.
Avem, deocamdată, mulţimile de stări

$$X_0 = X_1 = X_2 = X_3 = [0, 600].$$

Pentru starea arbitrară $u \in X_2$, mulţimea stărilor de tranziţie înspre X_3 este caracterizată de relaţiile

$$X_3(2; u) = \{x \in [0, 600];\ \max(0, u - 100) \le x \le 500\};\quad \text{adică}$$
$$X_3(2; u) = \begin{cases} [0, 500], & \text{dacă } 0 \le u \le 100 \\ [u - 100, 500], & \text{dacă } 100 \le u \le 600. \end{cases}$$

Costul total al tranziţiei de la starea $u \in X_2$ la starea $v \in X_3(2; u)$ este

$$W_2(u, v) = 200 + 1500 + 1000 - 15\,u + 17\,v = 2700 - 15\,u + 17\,v.$$

Deci, costul minim al unei tranziţii de stare de acest fel este

$$F_2(u) = \min\{W_2(u, v);\ v \in X_3(2; u)\},\quad \text{adică}$$
$$F_2(u) = \begin{cases} 2700 - 15\,u, & \text{dacă } u \in [0, 100] \\ 1000 + 2\,u, & \text{dacă } u \in [100, 600]. \end{cases}$$

Mai departe, pentru starea arbitrară $u \in X_1$, mulţimea stărilor de tranziţie înspre X_2 este definită de

$$X_2(1; u) = \{x \in [0, 600];\ \max(0, u - 300) \le x \le 500\};\quad \text{adică}$$
$$X_2(1; u) = \begin{cases} [0, 300], & \text{dacă } 0 \le u \le 300 \\ [u - 300, 300], & \text{dacă } 300 \le u \le 600. \end{cases}$$

Costul total al tranziţiei de la starea $u \in X_1$ la starea $v \in X_2(1; u)$ este

$$W_1(u, v) = 200 + 7500 + 6000 - 25\,u + 29\,v = 13700 - 25\,u + 29\,v.$$

Conform unei formule iterative anterioare, costul minim al unei tranziţii de stare (u, u_2, u_3) cu $u_2 \in X_2(1; u), u_3 \in X_3(2; u_2)$ este

$$F_1(u) = \min\{W_1(u, v) + F_2(v);\ v \in X_2(1; u)\}.$$

Din expresiile anterioare, combinate cu monotonia crescătoare a funcţiei $v \vdash 29v + F_2(v)$, găsim

$$F_1(u) = \begin{cases} 16400 - 25\,u, & u \in [0, 300] \\ \ \ 3500 + 18\,u, & u \in [300, 400] \\ -3300 + 35\,u & u \in [400, 600]. \end{cases}$$

În fine, pentru starea arbitrară $u \in X_0$, mulţimea stărilor de tranziţie înspre X_1 este caracterizată de relaţiile

$$X_1(0; u) = \{x \in [0, 600];\ \max(0, u - 200) \le x \le 400\}; \quad \text{adică}$$
$$X_1(0; u) = \begin{cases} [0, 400], & \text{dacă } 0 \le u \le 200 \\ [u - 200, 400], & \text{dacă } 200 \le u \le 600]. \end{cases}$$

Costul total al tranziţiei de la starea $u \in X_0$ la starea $v \in X_1(0; u)$ este

$$W_0(u, v) = 200 + 4000 + 2000 - 20\,u + 22\,v = 6200 - 20\,u + 22\,v.$$

Pe baza aceleiaşi formule iterative, costul minim al unei tranziţii de stare (u, u_1, u_2, u_3) cu $u_1 \in X_1(0; u), u_2 \in X_2(1; u_1), u_3 \in X_3(2; u_2)$ este

$$F_0(u) = \min\{W_0(u, v) + F_1(v);\ v \in X_1(0, u)\}.$$

Din expresiile anterioare, găsim în final

$$F_0(u) = \begin{cases} 21700 - 20\,u, & u \in [0, 500] \\ \ \ 1700 + 20\,u, & u \in [500, 600] \end{cases}$$

În particular, pentru starea iniţială $u_0 = 100$ avem

$$F_0(u_0) = 21700 - 20 \cdot 100 = 19700\ .$$

Aceasta este deci valoarea minimă a cheltuielilor lunare de aprovizionare şi stocare pentru sistemul nostru.
Să determinăm acum stările succesive (respectiv, variabilele de decizie) care o realizează. Notând cu u_1 starea din $X_1(0; u_0)$ care a dus la valoarea $F_0(u_0)$, avem

$F_0(u_0) = W_0(u_0, u_1) + F_1(u_1)$; sau, echivalent,

$19700 = 6200 - 2000 + 22\,u_1 + F_1(u_1)$; de unde, $u_1 = 272$.

Dar $u_1 = u_0 + z_0 - q_0\theta$; şi, deci,

$272 = 100 + z_0 - 100$; de unde $z_0 = 272$, $U_0 = u_0 + z_0 = 372$.
Avem acum
$F_1(u_1) = 16400 - 25 \cdot 272 = 9600$.
Notând cu u_2 starea din $X_2(1; u_1)$ care a dus la această valoare, avem
$F_1(u_1) = W_1(u_1, u_2) + F_2(u_2)$; sau echivalent,
$9600 = 13700 - 6800 + 29\, u_2 + F_2(u_2)$; de unde $u_2 = 0$.
Dar $u_2 = u_1 + z_1 - q_1\theta$; şi, deci,
$0 = 272 + z_1 - 300$; de unde, $z_1 = 28$, $U_1 = u_1 + z_1 = 300$.
În fine, avem
$F_2(u_2) = 2700 - 15 \cdot 0 = 2700$.
Notând cu u_3 starea din $X_3(2; u_2)$ care a condus la această valoare, avem egalitatea
$F_2(u_2) = W_2(u_2, u_3)$, sau, echivalent,
$2700 = 2700 - 0 + 17\, u_3$; de unde, $u_3 = 0$.
Cum $u_3 = u_2 + z_2 - q_2\theta$, deducem
$0 = 0 + z_2 - 200$; de unde $z_2 = 200$, $U_2 = u_2 + z_2 = 200$.

În concluzie, cea mai bună politică de urmat pentru minimizarea cheltuielilor de aprovizionare şi stocare este următoarea:

(a) La începutul primei decade, se face o comandă de $z_0 = 272$ bucăţi din produsul respectiv (având în stoc în acel moment $U_0 = 372$ bucăţi din acel produs) şi se consumă din stoc pe această perioadă până când nivelul de stoc va atinge $u_1 = 272$ bucăţi.

(b) La începutul celei de a doua decade, se face o comandă de $z_1 = 28$ bucăţi din produsul respectiv (având în stoc $U_1 = 300$ bucăţi în acel moment) şi se consumă din stoc pe această perioadă până la nivelul $u_2 = 0$ bucăţi.

(c) La începutul celei de a treia decade, se face o comandă de $z_2 = 200$ bucăţi din produsul respectiv (având în stoc, în acel moment, $U_2 = 200$ bucăţi din acel produs) şi se consumă din stoc pe această perioadă întregul stoc disponibil (deci, până la nivelul $u_3 = 0$ bucăţi).

Desigur, în legătură cu aceasta, se mai pot imagina şi situaţii în care pot apare cheltuieli de penalizare; dar, în esenţă, rezolvarea este cea indicată anterior.

Probleme propuse la §6.4

1. Necesităţile zilnice ale unei întreprinderi relativ la o anumită materie primă s-au evaluat la q=20 tone/zi, şi nu se admite lipsă de stoc. Se cunosc

– costul de lansare a unei comenzi K=15000 lei/lot,

– costul de achiziţie c=2000 lei/tonă,

– costul de stocare h=100 lei/(tonă)×(zi).

Presupunând că sistemul de stocare este cu perioadă fixă, iar politica de stoc de tip (u, U) se cere să se stabilească lotul optim care trebuie comandat precum şi intervalul optim între două comenzi succesive, astfel încât costul total de aproxizionare şi stocare pe unitate de timp să fie minim.

2. Rata de consum dintr-o anumită marfă, a unui magazin este q=200 bucăţi/zi şi se admite lipsă de stoc. Se cunosc:

– costul de lansare a unei comenzi K=5000 lei/lot

– costul de achiziţie c=400 lei/bucată

– costul de stocare h=20 lei/(bucată)×(zi)

– costul de penalizare b=40 lei/(bucată)×(zi).

Presupunând că sistemul de stocare este cu perioadă fixă, iar politica de stoc de tip (u, U), se cere să se stabilească lotul optim de marfă care trebuie comandat precum şi intervalul optim între două comenzi succesive, astfel încât costul total de aprovizionare, stocare şi penalizare pe unitatea de timp să fie minim.

3. Să presupunem că la fiecare trimestru (θ=90 zile) cererea zilnică a unui produs, costurile de achiziţie şi costurile de stocare sunt date de tabelul prezentat mai jos:

Perioada (i)	0	1	2	3
rata de cerere (bucăţi/zi)	20	40	50	10
costul de achiziţie (lei/bucată)	150	200	250	100
costul de stocare (lei/(bucată)×(zi))	10	20	25	10

Presupunem că nu se admit lipsuri de stoc, iar capacitatea maximă de stocare este de 3000 bucăţi. În plus, costul de lansare a unei comenzi este de 200 lei/lot. Se cere să se determine comenzile la începutul fiecărui trimestru astfel încât cheltuielile anuale de aprovizionare şi stocare să fie minime, în una din condiţiile:

(a) valoarea iniţială de stoc este 200 bucăţi, iar cea finală este nedeterminată;

(b) valoarea iniţială de stoc este nedeterminată, iar valoarea finală de stoc este de 100 bucăţi.

Bibliografie

A. Lucrări de sinteză

1. AMIHĂESEI, C., *Curs de cercetări operaţionale*, Universitatea "Al.I. Cuza" Iaşi, 1988.

2. BOLDUR–LĂŢESCU, G., SĂCUIU, I., ŢIGĂNESCU, E., *Cercetarea operaţională cu aplicaţii în economie,* Editura Didactică şi Pedagogică, Bucureşti, 1979.

3. DIACONIŢA, V., MANOLACHI, A., RUSU, GH., *Matematici aplicate în economie* (vol. II), Editura Fund. "Chemarea" Iaşi, 1993.

4. IACOB, C. ş.a., *Matematici clasice şi moderne* (Vol. I), Editura Tehnică, Bucureşti, 1978.

5. KAUFMAN, A., *Metode şi modele ale cercetării operaţionale* (Vol. I, II), Editura Ştiinţifică, Bucureşti, 1968.

6. MIHĂILĂ, N., POPESCU, O., *Matematici speciale aplicate în economie*, Editura Didactică şi Pedagogică, Bucureşti, 1978.

7. MIHOC, GH., ŞTEFĂNESCU, A., *Programare matematică,* Editura Didactică şi Pedagogică, Bucureşti, 1973.

8. POSTELNICU, T., DINESCU, C., SĂVULESCU, B., *Matematici speciale aplicate în economie,* Editura Didactică şi Pedagogică, Bucureşti, 1977.

9. SMADICI, C., *Cercetare operaţională* (partea II), Universitatea "Al.I. Cuza" Iaşi, 1985.

10. Stanciu, P. ş.a., *Matematici aplicate în economie* (Vol. I), Editura Facla, Timişoara, 1981.

11. Stavre, P., *Matematici speciale cu aplicaţii în economie*, Editura Scrisul Românesc, Craiova, 1982.

12. Şabac, I.Gh., *Matematici speciale* (Vol. I), Editura Didactică şi Pedagogică, Bucureşti, 1981.

13. Ştefănescu, A., Zidăroiu, C., *Cercetări operaţionale*, Editura Didactică şi Pedagogică, Bucureşti, 1981.

14. Tamaş,V. ş.a., *Matematici generale pentru economişti*, Editura Graphix, Iaşi, 1993.

15. Văduva, I., Dinescu, C., Săvulescu, B., *Modele matematice de organizare şi conducere a producţiei* (Vol. I, II), Editura Didactică şi Pedagogică, Bucureşti, 1975.

B. Lucrări specializate

1. Allen, R.G.D., *Analiză matematică pentru economişti*, Editura Ştiinţifică, Bucureşti, 1971.

2. Fihtenholţ, G.M., *Curs de calcul diferenţial şi integral* (Vol. I, II), Editura Tehnică, Bucureşti, 1964.

3. Ghelfond, A.O., *Calculul cu diferenţe finite,* Editura Tehnică, Bucureşti, 1956.

4. Marinescu, Gh., *Analiză matematică* (Vol. I, II), Editura Academiei R.S.R., Bucureşti, 1984.

5. Nicolescu, M., Dinculeanu, N., Marcus, S., *Manual de analiză matematică* (Vol. I, II), Editura Didactică şi Pedagogică, Bucureşti, 1962, 1964.

6. Roşculeţ, M., *Analiză matematică*, Editura Didactică şi Pedagogică, Bucureşti, 1979.

Volumul de faţă – care continuă expunerea de natură algebrică din prima parte – poate fi descris ca o introducere topologică în unele domenii de actualitate ale cercetării operaţionale. Capitolul patru este consacrat unor teme de bază ale Analizei : limită, continuitate, derivată (parţială), integrală (multiplă). Capitolul cinci tratează chestiuni de programare neliniară (extreme libere şi condiţionate). În fine, capitolul şase are ca obiect procesele dinamice discrete (şiruri recurente) sau continue (ecuaţii diferenţiale), programare dinamică şi teoria stocurilor.

Ca şi în prima parte,toate noţiunile cu care se lucrează sunt în prealabil definite; iar, rezultatele de bază sunt întregite cu exemple numerice concrete. S–a urmărit o permanentă corelare cu primele trei capitole cât şi posibilitatea aplicării ulterioare a acestor fapte în alte domenii ale cercetării operaţionale care n–au putut fi cuprinse aici din lipsă de spaţiu.

Volumul este util tuturor celor interesaţi de aspectele teoretice şi practice ale matematicilor economice.