

UNIVERZA V LJUBLJANI  
FAKULTETA ZA MATEMATIKO IN FIZIKO

# Štiri poglavja iz numerične analize

Gradivo za študente praktične matematike

Jan Grošelj

Ljubljana, 2024



**O gradivu.** Gradivo obravnava štiri temeljna poglavja numerične analize in je nastalo med izvajanjem predmeta Numerične metode 2 na študijskem programu Praktična matematika na Fakulteti za matematiko in fiziko v Ljubljani.

Namen gradiva je predstavitev osnovnih numeričnih metod za aproksimacijo funkcij, modeliranje s parametričnimi krivuljami, računanje približkov za odvode in integrale funkcij ter reševanje diferencialnih enačb. Snov je pospremljena s primeri in rešenimi nalogami. Teorija je večinoma dobro podprta z dokazi, a navedeni so tudi nekateri znameniti izreki, pri katerih bi utemeljevanje zahtevalo uvedbo poglobljenih analitičnih orodij, zato so dokazi izpuščeni. Zainteresiran bralec lahko te rezultate poišče v učbenikih (Plestenjak, 2015) in (Kozak, 2008), ki dobršen del tem, predstavljenih v tem gradivu, obravnavata obširneje in podrobneje.



# Kazalo

|          |                                                                 |           |
|----------|-----------------------------------------------------------------|-----------|
| <b>1</b> | <b>Aproksimacija funkcij</b>                                    | <b>1</b>  |
| 1.1      | Aproksimacijski problem . . . . .                               | 1         |
| 1.1.1    | Normirani prostori funkcij in aproksimacijski podprostor        | 1         |
| 1.1.2    | Polinomska aproksimacija in Weierstrassov izrek . . . . .       | 3         |
| 1.2      | Problem najboljše aproksimacije . . . . .                       | 5         |
| 1.2.1    | Element najboljše aproksimacije . . . . .                       | 6         |
| 1.2.2    | Najboljša enakomerna aproksimacija funkcij s polinomi . . . . . | 7         |
| 1.3      | Odsekoma polinomske funkcije . . . . .                          | 11        |
| 1.3.1    | Polinomska interpolacija . . . . .                              | 11        |
| 1.3.2    | Funkcijski zlepci . . . . .                                     | 13        |
| 1.3.3    | Odsekoma linearne funkcije . . . . .                            | 15        |
| 1.3.4    | Odsekoma kvadratične funkcije . . . . .                         | 18        |
| 1.3.5    | Odsekoma kubične funkcije . . . . .                             | 19        |
| <b>2</b> | <b>Računalniško podprto geometrijsko oblikovanje</b>            | <b>24</b> |
| 2.1      | Bézierjeve krivulje . . . . .                                   | 24        |
| 2.1.1    | Predstavitev z Bernsteinovimi baznimi polinomi . . . . .        | 24        |
| 2.1.2    | De Casteljauev postopek . . . . .                               | 27        |
| 2.1.3    | Afina invariantnost . . . . .                                   | 29        |
| 2.1.4    | Višanje stopnje Bézierjeve krivulje . . . . .                   | 31        |
| 2.2      | Zlepki Bézierjevih krivulj . . . . .                            | 32        |
| 2.2.1    | Odvodi Bézierjeve krivulje . . . . .                            | 32        |
| 2.2.2    | Zlepek dveh Bézierjevih krivulj . . . . .                       | 34        |
| 2.2.3    | Zlepek kvadratičnih Bézierjevih krivulj . . . . .               | 37        |
| 2.2.4    | Zlepek kubičnih Bézierjevih krivulj . . . . .                   | 39        |
| 2.2.5    | Opis krožnega loka z zlepkom Bézierjevih krivulj . . . . .      | 41        |
| <b>3</b> | <b>Numerično odvajanje in integriranje</b>                      | <b>44</b> |
| 3.1      | Pravila za odvajanje . . . . .                                  | 44        |
| 3.1.1    | Diferenčne formule za prve odvode . . . . .                     | 44        |
| 3.1.2    | Diferenčna formula za drugi odvod . . . . .                     | 49        |
| 3.2      | Newton–Cotesova pravila za integriranje . . . . .               | 49        |
| 3.2.1    | Osnovna pravila . . . . .                                       | 49        |
| 3.2.2    | Sestavljena pravila . . . . .                                   | 54        |
| 3.3      | Izboljšana pravila za integriranje . . . . .                    | 57        |
| 3.3.1    | Rombergova metoda . . . . .                                     | 58        |

---

|          |                                                               |           |
|----------|---------------------------------------------------------------|-----------|
| 3.3.2    | Adaptivna pravila . . . . .                                   | 60        |
| 3.3.3    | Gaussova pravila . . . . .                                    | 62        |
| <b>4</b> | <b>Numerično reševanje diferencialnih enačb</b>               | <b>65</b> |
| 4.1      | Reševanje začetnega problema . . . . .                        | 65        |
| 4.1.1    | Eulerjevi metodi . . . . .                                    | 65        |
| 4.1.2    | Runge–Kutta metode . . . . .                                  | 70        |
| 4.1.3    | Sistemi diferencialnih enačb in enačbe višjega reda . . . . . | 72        |
| 4.1.4    | Veččlenske metode . . . . .                                   | 75        |
| 4.2      | Reševanje robnega problema . . . . .                          | 77        |
| 4.2.1    | Diferenčna metoda . . . . .                                   | 78        |
| 4.2.2    | Streška metoda . . . . .                                      | 82        |



# Poglavje 1

## Aproksimacija funkcij

Pri aproksimaciji funkcije iščemo nadomestek za funkcijo, ki je zapletena ali pa je sploh ne poznamo v zaključeni obliki. Kot nadomestilo želimo uporabiti funkcijo, ki jo je enostavno predstaviti, analizirati in z njo računati. Pri tem pa moramo paziti, da funkcija v čim manjši meri odstopa od prvotne in da v čim večji meri ohranja njene lastnosti.

### 1.1 Aproksimacijski problem

Za izračun odstopanja približka oziroma aproksimacije funkcije potrebujemo mero, ki jo opredelimo s funkcijsko normo. Poleg tega je treba izbrati prostor, v katerem bomo iskali aproksimacijo. S temi pojmi definiramo splošen aproksimacijski problem.

#### 1.1.1 Normirani prostori funkcij in aproksimacijski podprostori

Naj  $\mathbb{X}$  označuje vektorski prostor funkcij nad realnimi števili. Najpogosteje bo to prostor zveznih funkcij  $C([a, b])$  na izbranem intervalu  $[a, b]$  ali pa prostor funkcij višjega reda zvezne odvedljivosti  $C^r([a, b])$  za  $r > 0$ . Vektorski prostor  $\mathbb{X}$  opremimo z normo.

**Definicija 1.1.** Funkcija  $\|\cdot\| : \mathbb{X} \rightarrow \mathbb{R}$  je norma realnega vektorskega prostora  $\mathbb{X}$ , če zadošča naslednjim lastnostim.

- *Nenegativnost:* Za vsak element  $f \in \mathbb{X}$  velja  $\|f\| \geq 0$ . Pri tem je  $\|f\| = 0$  natanko tedaj, ko je  $f = 0$ .
- *Homogenost:* Za vsak element  $f \in \mathbb{X}$  in vsak skalar  $\alpha \in \mathbb{R}$  velja  $\|\alpha f\| = |\alpha| \|f\|$ .
- *Trikotniška neenakost:* Za vsaka elementa  $f, g \in \mathbb{X}$  velja  $\|f + g\| \leq \|f\| + \|g\|$ .

Prostoru  $\mathbb{X}$ , skupaj z normo  $\|\cdot\|$ , pravimo normiran vektorski prostor.

**Primer 1.1.** Enakomerna (tudi neskončna ali supremum) norma, definirana za prostor zveznih funkcij  $C([a, b])$  na intervalu  $[a, b]$ , je za funkcijo  $f \in C([a, b])$  podana s predpisom

$$\|f\|_{\infty, [a, b]} = \max_{x \in [a, b]} |f(x)|.$$

Pri računanju na računskem stroju interval  $[a, b]$  pogosto zamenjamo s končno množico točk  $X \subset [a, b]$  in normo nadomestimo z diskretno izpeljanko

$$\|f\|_{\infty, X} = \max_{x \in X} |f(x)|,$$

ki pa ni norma, saj lahko za neničelno funkcijo  $f$  velja  $\|f\|_{\infty, X} = 0$ . Pravimo, da je  $\|\cdot\|_{\infty, X}$  semi norma.

**Naloga 1.1.** Izračunajte enakomerno normo funkcije

$$f(x) = \frac{1}{2} - \frac{1}{1 + (\pi x - 1)^2}$$

na intervalu  $[0, 1]$ . V Matlabu izračunajte vrednosti  $\|f\|_{\infty, X_n}$ ,  $n = 10, 100, 1000$ , za

$$X_n = \left\{ \frac{i}{n}; i = 0, 1, \dots, n \right\} \quad (1.1)$$

in jih primerjajte z dejansko normo.

*Rešitev.* Po definiciji enakomerne norme iščemo maksimum absolutne vrednosti funkcije  $f$ . Zato z odvajanjem funkcije najprej določimo njene ekstreme. Iz

$$f'(x) = \frac{2\pi(\pi x - 1)}{(1 + (\pi x - 1)^2)^2}$$

sledi, da je  $y = 1/\pi$  edina stacionarna točka funkcije  $f$ . Ker točka  $y$  leži na intervalu  $[0, 1]$  in je  $f(y) = -1/2$ , je vrednost  $1/2$  kandidat za vrednost  $\|f\|_{\infty, [0, 1]}$ . Preveriti moramo še vrednosti funkcije  $f$  na robu intervala. Vrednost v točki 0 je enaka 0, v točki 1 pa dobimo

$$f(1) = \frac{1}{2} - \frac{1}{1 + (\pi - 1)^2} \approx 0.3210.$$

To je manj od ekstremne vrednosti v točki  $y$ , zato je  $\|f\|_{\infty, [0, 1]} = 1/2$ .

V Matlabu seznam točk iz množice  $X_n$  zgeneriramo z ukazom `Xn = linspace(0, 1, n+1)`, vrednost  $\|f\|_{\infty, X_n}$  pa izračunamo z ukazom `max(abs(f(Xn)))`, pri čemer prej definiramo funkcijo `f` v obliki `f = @(x) 1/2 - 1./(1 + (pi*x - 1).^2)`. Pri  $n = 10, 100, 1000$  po vrsti dobimo 0.49670211, 0.49997181, 0.49999905.  $\square$

Pri iskanju aproksimacije za funkcijo  $f \in \mathbb{X}$  se omejimo na elemente aproksimacijskega podprostora  $\mathbb{Y} \subset \mathbb{X}$ . Ta mora biti dovolj preprost, da je primeren za numerično računanje. Po drugi strani si želimo, da ima čim večjo aproksimacijsko moč, kazalnik katere je razdalja

$$\text{dist}(f, \mathbb{Y}) = \inf_{g \in \mathbb{Y}} \|f - g\|.$$

Ker je aproksimacija za funkcijo  $f$  iz podprostora  $\mathbb{Y}$ , je to najmanjša napaka, ki jo lahko pričakujemo. Za  $\mathbb{Y}$  navadno izberemo prostor polinomov  $\mathbb{P}_n$  stopnje največ  $n \in \mathbb{N}_0$  ali pa katero izmed njihovih posplošitev, kot so odsekoma polinomske funkcije, racionalne funkcije ali trigonometrični polinomi.

V jeziku vpeljanih pojmov se aproksimacijski problem glasi takole: za dano funkcijo  $f$  iz prostora  $\mathbb{X}$ , opremljenega z normo  $\|\cdot\|$ , poišči funkcijo  $g$  iz podprostora  $\mathbb{Y} \subset \mathbb{X}$ , da je napaka, ki jo določa norma  $\|f - g\|$  residuala  $f - g$ , čim bližje vrednosti  $\text{dist}(f, \mathbb{Y})$ .

### 1.1.2 Polinomska aproksimacija in Weierstrassov izrek

V tem razdelku obravnavamo aproksimacijo funkcije  $f \in C([a, b])$  s polinomom iz prostora  $\mathbb{P}_n$  in navedemo izrek, ki zagotavlja, da lahko z višanjem stopnje  $n$  zagotovimo poljubno dobro aproksimacijo v smislu enakomerne norme  $\|\cdot\|_{\infty, [a, b]}$ .

Polinom  $p \in \mathbb{P}_n$  najpogosteje predstavimo v potenčni bazi kot

$$p(x) = a_n x^n + \dots + a_1 x + a_0$$

z realnimi koeficienti  $a_0, a_1, \dots, a_n$ , a obstajajo številni drugi načini, ki se jih poslužujemo pri specifičnih problemih. Z vidika interpolacije je na primer zanimiva predstavitev v bazi Lagrangeevih polinomov ali v bazi zamaknjenih potenc. Pri metodi najmanjših kvadratov je nepogrešljiva baza ortogonalnih polinomov. Z vidika stabilnosti in geometrijske predstavljivosti pa so posebej dobrodošli Bernsteinovi bazni polinomi, ki so na intervalu  $[0, 1]$  podani s predpisi

$$B_i^n(x) = \binom{n}{i} x^i (1-x)^{n-i}, \quad i = 0, 1, \dots, n. \quad (1.2)$$

**Naloga 1.2.** Zapišite in narišite Bernsteinove bazne polinome za  $n = 0, 1, 2$ . V Matlabu pripravite skripto, ki izriše Bernsteinove bazne polinome za poljuben  $n$ . Kakšne so vrednosti polinomov v točkah  $x = 0$  in  $x = 1$ ?

*Rešitev.* Pri stopnji  $n = 0$  imamo le en Bernsteinov bazni polinom, to je  $B_0^0(x) = 1$ . Pri stopnji  $n = 1$  sta Bernsteinova bazna polinoma dva,  $B_0^1(x) = 1 - x$  in  $B_1^1(x) = x$ , pri stopnji  $n = 2$  pa trije,  $B_0^2(x) = (1-x)^2$ ,  $B_1^2(x) = 2x(1-x)$  in  $B_2^2(x) = x^2$ . Polinomi  $B_i^2$ ,  $i = 0, 1, 2$ , so prikazani na sliki 1.1a, po vrsti v modri, rdeči in rumeni barvi.

V točki  $x = 0$  pri poljubni stopnji  $n$  velja  $B_i^n(0) = 0$  za vse indekse  $i$ , razen za indeks  $i = 0$ , pri katerem je  $B_0^n(0) = 1$ . Podobno velja v točki  $x = 1$ . Za vse indekse  $i$  velja  $B_i^n(1) = 0$ , razen za indeks  $i = n$ , pri katerem je  $B_n^n(1) = 1$ .  $\square$

**Trditev 1.1.** *Bernsteinovi bazni polinomi sestavljajo konveksno razčlenitev enote: za vsak  $x \in [0, 1]$  velja  $B_i^n(x) \geq 0$ ,  $i = 0, 1, \dots, n$ , in*

$$\sum_{i=0}^n B_i^n(x) = 1.$$

*Dokaz.* Nenegativnost na intervalu  $[0, 1]$  sledi direktno iz definicije Bernsteinovih baznih polinomov, razčlenitev enote

$$1 = (x + (1-x))^n = \sum_{i=0}^n \binom{n}{i} x^i (1-x)^{n-i}$$

pa iz binomskega izreka.  $\square$

S pomočjo Bernsteinovih baznih polinomov lahko osnujemo aproksimacijo za funkcijo  $f$  na intervalu  $[0, 1]$  v obliki

$$\mathcal{B}_n f(x) = \sum_{i=0}^n f\left(\frac{i}{n}\right) B_i^n(x).$$

To je polinom iz  $\mathbb{P}_n$ ,  $n \in \mathbb{N}$ , ki ga imenujemo kar Bernsteinov polinom. O njegovem aproksimacijskem potencialu priča naslednja trditev.

**Trditev 1.2.** Za vsak  $p \in \mathbb{P}_2$  in vsak  $x \in \mathbb{R}$  velja

$$\lim_{n \rightarrow \infty} \mathcal{B}_n p(x) = p(x).$$

*Dokaz.* Naj bo  $p_i(x) = x^i$ ,  $i = 0, 1, 2$ . Vsak polinom  $p \in \mathbb{P}_2$  lahko predstavimo kot  $p = a_2 p_2 + a_1 p_1 + a_0 p_0$  za neka realna števila  $a_2, a_1, a_0$ . Ker je

$$\mathcal{B}_n(a_2 p_2 + a_1 p_1 + a_0 p_0) = a_2 \mathcal{B}_n p_2 + a_1 \mathcal{B}_n p_1 + a_0 \mathcal{B}_n p_0,$$

je za dokaz trditve dovolj preveriti, da za vsak  $x \in \mathbb{R}$  velja

$$\lim_{n \rightarrow \infty} \mathcal{B}_n p_i(x) = p_i(x).$$

Za  $p_0$  velja kar  $\mathcal{B}_n p_0 = p_0$ , kot je razvidno iz dokaza trditve 1.1. Iz

$$\begin{aligned} \mathcal{B}_n p_1(x) &= \sum_{i=0}^n \frac{i}{n} \binom{n}{i} x^i (1-x)^{n-i} \\ &= \sum_{i=0}^{n-1} \frac{i+1}{n} \frac{n!}{(i+1)!(n-1-i)!} x^{i+1} (1-x)^{n-1-i} \\ &= x \mathcal{B}_{n-1} p_0(x) \\ &= x, \end{aligned}$$

razberemo, da je tudi  $\mathcal{B}_n p_1 = p_1$ , medtem ko lahko za  $p_2$  z malo več računanja izpeljemo

$$\mathcal{B}_n p_2(x) = x^2 + \frac{1}{n} x(1-x),$$

kar seveda pomeni, da je  $\lim_{n \rightarrow \infty} \mathcal{B}_n p_2 = p_2$ . □

**Naloga 1.3.** Zapišite Bernsteinov polinom  $\mathcal{B}_2 f$  za funkcijo  $f$ , podano s predpisom

$$f(x) = \frac{1}{3x+1},$$

in ga izrazite v potenčni bazi. V Matlabu izračunajte in narišite še Bernsteinove polinome  $\mathcal{B}_n f$  za stopnje  $n = 3, 4, \dots, 10$  ter si oglejte vrednosti  $\|f - \mathcal{B}_n f\|_{\infty, X_{1000}}$ , pri čemer je množica  $X_{1000}$  podana z (1.1).

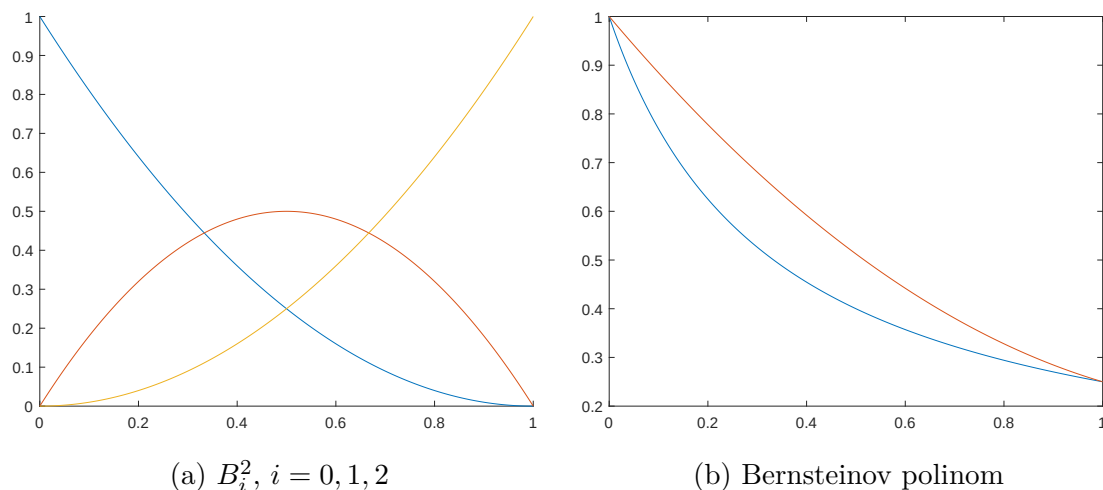
*Rešitev.* Bernsteinov polinom  $\mathcal{B}_2 f$  je podan s predpisom

$$\mathcal{B}_2 f(x) = f(0)B_0^2(x) + f(\tfrac{1}{2})B_1^2(x) + f(1)B_2^2(x) = (1-x)^2 + \tfrac{4}{5}x(1-x) + \tfrac{1}{4}x^2$$

in se v potenčni bazi glasi

$$\mathcal{B}_2 f(x) = \frac{9}{20}x^2 - \frac{6}{5}x + 1.$$

Skupaj s funkcijo  $f$  (v modri barvi) je prikazan na sliki 1.1b (v rdeči barvi). Napaka aproksimacije, izračunana na podlagi vrednosti  $\|f - \mathcal{B}_2 f\|_{\infty, X_{1000}}$  je približno 0.1567.



Slika 1.1: Bernsteinovi bazni polinomi in aproksimacija z Bernsteinovim polinomom.

Izračuni v Matlabu pokažejo, da z višanjem stopnje napaka monotono pada in je pri  $n = 10$  približno 0.0374.  $\square$

Polinom  $\mathcal{B}_n f$  je torej kandidat za aproksimacijo funkcije  $f$  s polinomom stopnje  $n$ . Resda smo se omejili na aproksimacijo funkcije  $f$  na interval  $[0, 1]$ , a to ni težava, saj lahko s preprosto reparametrizacijo  $x \mapsto (x - a)/(b - a)$ , ki interval  $[a, b]$  preslika na interval  $[0, 1]$ , polinom  $\mathcal{B}_n f$  prilagodimo intervalu  $[a, b]$ .

Trditev 1.2 ob zgornji opazki zagotavlja, da za poljuben polinom  $f \in \mathbb{P}_2$  na intervalu  $[a, b]$  razdalja

$$\text{dist}_\infty(f, \mathbb{P}_n) = \max_{p \in \mathbb{P}_n} \|f - p\|_{\infty, [a, b]}$$

pada proti 0, ko pošljemo  $n$  proti neskončno. Zagotovitev enakega rezultata za poljubno funkcijo  $f \in C([a, b])$  se zdi močno oddaljen cilj, a z nekaj spretnosti (glej Kozak, 2008, izrek 1.1) je prav z uporabo trditve 1.2 mogoče dokazati, da za vsak  $x \in [a, b]$  velja

$$\lim_{n \rightarrow \infty} \mathcal{B}_n f(x) = f(x).$$

Iz tega dejstva sledi naslednji znameniti aproksimacijski izrek.

**Izrek 1.1** (Weierstrass). *Za  $f \in C([a, b])$  velja*

$$\lim_{n \rightarrow \infty} \text{dist}_\infty(f, \mathbb{P}_n) = 0.$$

Variante izreka 1.1 obstajajo tudi za druge pomembne aproksimacijske podprostore, osnovno vodilo pa je definicija aproksimacijske funkcije na podlagi skrbno izbranih baznih funkcij podprostoru s podobnimi lastnostmi, kot jih imajo Bernsteinovi bazni polinomi.

## 1.2 Problem najboljše aproksimacije

Pri problemu najboljše aproksimacije iščemo element aproksimacijskega podprostoru, ki med vsemi najboljše aproksimira funkcijo v izbrani normi.

### 1.2.1 Element najboljše aproksimacije

Naj bo  $\mathbb{X}$  vektorski prostor, opremljen z normo  $\|\cdot\|$ ,  $f$  element prostora  $\mathbb{X}$  in  $\mathbb{Y}$  aproksimacijski podprostor za  $\mathbb{X}$ .

**Definicija 1.2.** Element  $g^* \in \mathbb{Y}$  je element najboljše aproksimacije za  $f \in \mathbb{X}$  v  $\mathbb{Y}$ , če je

$$\|f - g^*\| \leq \|f - g\|$$

za vsak  $g \in \mathbb{Y}$ . Ekvivalentno,  $\|f - g^*\| = \text{dist}(f, \mathbb{Y})$ .

Obstoj elementa najboljše aproksimacije ni samoumeven, a je zagotovljen vedno, kar-  
dar je prostor  $\mathbb{Y}$  končno razsežen. To pomeni, da aproksimacijsko funkcijo določimo z  
izbiro končnega števila parametrov, k čemur seveda vedno stremimo, saj bi bila sicer  
aproksimacijska metoda računsko neuporabna.

Več težav je lahko z zagotavljanjem enoličnosti elementa najboljše aproksimacije, kot  
namiguje naslednja naloga.

**Naloga 1.4.** Naj bo  $\mathbb{X} = \mathbb{R}^2$  in  $\mathbf{x} = (1/2, 1) \in \mathbb{X}$ . Poiščite element najboljše aproksima-  
cije za  $\mathbf{x}$  v podprostorih

$$\mathbb{Y}_1 = \{\alpha(1, 1) \in \mathbb{R}^2; \alpha \in \mathbb{R}\} \quad \text{in} \quad \mathbb{Y}_2 = \{\alpha(1, 0) \in \mathbb{R}^2; \alpha \in \mathbb{R}\}$$

glede na vektorski normi  $\|\cdot\|_2$  in  $\|\cdot\|_\infty$ .

*Rešitev.* Naj bo  $\alpha(1, 1)$  poljuben element iz  $\mathbb{Y}_1$ . Najprej poiščimo  $\alpha \in \mathbb{R}$ , pri katerem je  
dosežena minimalna vrednost izraza

$$\left\| \left( \frac{1}{2}, 1 \right) - \alpha(1, 1) \right\|_2 = \left\| \left( \frac{1}{2} - \alpha, 1 - \alpha \right) \right\|_2 = \sqrt{\left( \frac{1}{2} - \alpha \right)^2 + (1 - \alpha)^2} = \sqrt{2\alpha^2 - 3\alpha + \frac{5}{4}}.$$

To je ekvivalentno iskanju minimuma funkcije  $\alpha \mapsto 2\alpha^2 - 3\alpha + 5/4$ , od kjer sledi, da je  
 $\alpha = 3/4$  iskana vrednost in  $(3/4, 3/4)$  element najboljše aproksimacije za  $\mathbf{x}$  iz  $\mathbb{Y}_1$  glede  
na  $\|\cdot\|_2$ . Pri enakomerni normi  $\|\cdot\|_\infty$  iščemo minimalno vrednost izraza

$$\left\| \left( \frac{1}{2}, 1 \right) - \alpha(1, 1) \right\|_\infty = \left\| \left( \frac{1}{2} - \alpha, 1 - \alpha \right) \right\|_\infty = \max \left\{ \left| \frac{1}{2} - \alpha \right|, |1 - \alpha| \right\},$$

torej minimum funkcije  $\alpha \mapsto \max \{ |1/2 - \alpha|, |1 - \alpha| \}$ , ki je dosežen pri  $\alpha = 3/4$ . Torej  
je  $(3/4, 3/4)$  tudi element najboljše aproksimacije za  $\mathbf{x}$  v  $\mathbb{Y}_1$  glede na  $\|\cdot\|_\infty$ .

Elementi iz prostora  $\mathbb{Y}_2$  so oblike  $\alpha(1, 0)$  in po enakem razmisleku kot zgoraj iz

$$\left\| \left( \frac{1}{2}, 1 \right) - \alpha(1, 0) \right\|_2 = \left\| \left( \frac{1}{2} - \alpha, 1 \right) \right\|_2 = \sqrt{\left( \frac{1}{2} - \alpha \right)^2 + 1} = \sqrt{\alpha^2 - \alpha + \frac{5}{4}},$$

sledi, da je  $\alpha$ , ki določa element najboljše aproksimacije za  $\mathbf{x}$  v  $\mathbb{Y}_2$  glede na  $\|\cdot\|_2$ , minimum  
funkcije  $\alpha \mapsto \alpha^2 - \alpha + 5/4$ , to je  $\alpha = 1/2$ . Pri enakomerni normi pa lahko iz

$$\left\| \left( \frac{1}{2}, 1 \right) - \alpha(1, 0) \right\|_\infty = \left\| \left( \frac{1}{2} - \alpha, 1 \right) \right\|_\infty = \max \left\{ \left| \frac{1}{2} - \alpha \right|, 1 \right\}$$

zaključimo, da bo minimum funkcije  $\alpha \mapsto \max \{ |1/2 - \alpha|, 1 \}$  dosežen za poljuben  $\alpha$  z  
intervala  $[-1/2, 3/2]$ , saj je povsod tam  $|1/2 - \alpha| \geq 1$ . Torej je elementov najboljše  
aproksimacije za  $\mathbf{x}$  iz  $\mathbb{Y}_2$  glede na  $\|\cdot\|_\infty$  več, pravzaprav jih je neskončno mnogo.  $\square$

### 1.2.2 Najboljša enakomerna aproksimacija funkcij s polinomi

Pri problemu najboljše enakomerne aproksimacije funkcije  $f$  s polinomi na intervalu  $[a, b]$  iščemo  $p^* \in \mathbb{P}_n$ ,  $n \in \mathbb{N}_0$ , za katerega velja, da je

$$\|f - p^*\|_{\infty, [a, b]} \leq \|f - p\|_{\infty, [a, b]}$$

za vsak  $p \in \mathbb{P}_n$ . Rešitve problema ni enostavno poiskati, v veliki meri pa to poenostavi naslednji karakterizacijski izrek (Kozak, 2008, izreka 1.11 in 1.12).

**Izrek 1.2** (O alterniranju residuala). *Polinom najboljše enakomerne aproksimacije za funkcijo  $f \in C([a, b])$  iz  $\mathbb{P}_n$  na intervalu  $[a, b]$  obstaja in je enoličen. To je polinom  $p^* \in \mathbb{P}_n$ , za katerega obstaja množica  $\{x_0, x_1, \dots, x_{n+1}\} \subset [a, b]$ , sestavljena iz  $n + 2$  točk  $x_0 < x_1 < \dots < x_{n+1}$ , da je*

$$f(x_i) - p^*(x_i) = (-1)^i \varepsilon \|f - p^*\|_{\infty, [a, b]}, \quad i = 0, 1, \dots, n + 1,$$

kjer je  $\varepsilon$  bodisi 1 bodisi  $-1$ .

Če za funkcijo  $f \in C([a, b])$  najdemo polinom  $p^* \in \mathbb{P}_n$ , da residual  $f - p^*$  v  $n + 2$  točkah alternirajoče doseže svojo ekstremno vrednost (to je lastnost, ki jo opisuje izrek 1.2), smo rešili problem najboljše enakomerne aproksimacije.

**Naloga 1.5.** Zvezna funkcija  $f$  je podana s predpisom  $f(x) = \sin(3x)$ . Določite polinom najboljše enakomerne aproksimacije za  $f$  iz prostora  $\mathbb{P}_4$  na intervalu  $[0, 2\pi]$ .

*Rešitev.* Funkcija  $f$  ima na intervalu  $[0, 2\pi]$  sedem ničel, to so  $k\pi/3$ ,  $k = 0, 1, \dots, 6$ . Med temi ničlami ima  $f$  pet ekstremov, to so  $k\pi/3 + \pi/6$ ,  $k = 0, 1, \dots, 5$ . Opazimo, da funkcija  $f$  v teh točkah alternirajoče doseže vrednost  $\pm 1$ , kar pomeni, da za  $p^* = 0$  residual  $f - p^*$  alternira, njegova absolutna vrednost pa šestkrat zavzame maksimalno vrednost. Če torej iščemo polinom najboljše enakomerne aproksimacije za  $f$  na intervalu  $[0, 2\pi]$  v prostoru  $\mathbb{P}_4$ , je iskani polinom kar  $p^* = 0$ .  $\square$

Kadar je funkcija  $f$  polinom stopnje  $n + 1$ , lahko polinom  $p^* \in \mathbb{P}_n$  poiščemo s preprostim receptom, ki temelji na posebni bazi za  $\mathbb{P}_n$ , polinomih Čebiševa.

Polinom Čebiševa  $T_{n+1}$  stopnje  $n + 1$  lahko na intervalu  $[-1, 1]$  definiramo s predpisom

$$T_{n+1}(x) = \cos((n + 1) \arccos(x)), \quad (1.3)$$

splošneje pa je določen z rekurzivno zvezo

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geq 1, \quad T_0(x) = 1, \quad T_1(x) = x.$$

**Naloga 1.6.** Predstavite polinome  $T_n$ ,  $n = 0, 1, 2, 3, 4$ , v potenčni bazi.

*Rešitev.* Z uporabo rekurzivne formule, začeni s  $T_0(x) = 1$  in  $T_1(x) = x$ , izračunamo  $T_2(x) = 2x^2 - 1$ ,  $T_3(x) = 4x^3 - 3x$  in  $T_4(x) = 8x^4 - 8x^2 + 1$ . Polinomi  $T_i$ ,  $i = 0, 1, 2, 3, 4$ , so po vrsti v modri, rdeči, vijola, zeleni in rumeni barvi prikazani na sliki 1.2a.  $\square$

Iz predpisa (1.3) sledi, da so točke

$$\cos\left(\frac{\left(k - \frac{1}{2}\right)\pi}{n+1}\right), \quad k = 1, 2, \dots, n+1,$$

ničle polinoma  $T_{n+1}$  na intervalu  $[-1, 1]$ . Na robu in med njimi, to je v točkah

$$\cos\left(\frac{k\pi}{n+1}\right), \quad k = 0, 1, \dots, n+1,$$

pa  $T_{n+1}$  doseže ekstremne vrednosti  $\pm 1$ . Torej je po izreku 1.2 polinom najboljše enakomerne aproksimacije za  $T_{n+1}$  na intervalu  $[-1, 1]$  iz prostora  $\mathbb{P}_n$  kar polinom, ki je identično enak 0. Ta opazka ima naslednjo posledico.

**Trditev 1.3.** *Naj bo funkcija  $f \in \mathbb{P}_{n+1}$  izražena v bazi polinomov Čebiševa kot*

$$f = \sum_{i=0}^{n+1} \alpha_i T_i$$

*za neke realne koeficiente  $\alpha_i$ ,  $i = 0, 1, \dots, n$ . Potem je*

$$p^* = \sum_{i=0}^n \alpha_i T_i$$

*polinom najboljše enakomerne aproksimacije za  $f$  na intervalu  $[-1, 1]$  iz prostora  $\mathbb{P}_n$  in velja  $\text{dist}(f, \mathbb{P}_n) = |\alpha_n|$ .*

*Dokaz.* Residual  $f - p^*$  je enak  $\alpha_n T_n$  in zato po zgornjih ugotovitvah alternirajoče doseže ekstremne vrednosti v  $n+1$  točkah na intervalu  $[-1, 1]$ . Ker je  $\|T_n\|_{\infty, [-1, 1]} = 1$ , je  $\|f - p^*\|_{\infty, [-1, 1]} = |\alpha_n|$ .  $\square$

**Naloga 1.7.** Dan je polinom  $f(x) = 8x^4 + 4x^3 + 2x^2 + x$ . Predstavite  $f$  s pomočjo polinomov Čebiseva in na podlagi te predstavitve določite polinom iz  $\mathbb{P}_3$ , ki predstavlja element najboljše enakomerne aproksimacije za  $f$  na intervalu  $[-1, 1]$ .

*Rešitev.* Polinom  $f$  lahko s pomočjo polinomov Čebiševa, izpeljanih v nalogi 1.6, izrazimo kot

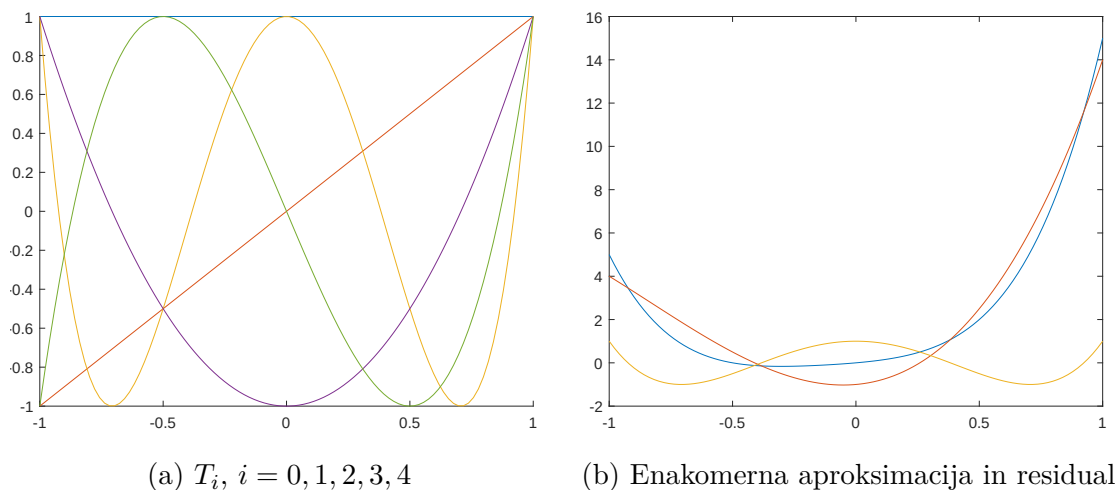
$$f = T_4 + T_3 + 5T_2 + 4T_1 + 4T_0.$$

Po trditvi 1.3 nato sledi, da je

$$p^*(x) = T_3(x) + 5T_2(x) + 4T_1(x) + 4T_0(x) = 4x^3 + 10x^2 + x - 1$$

polinom najboljše enakomerne aproksimacije za  $f$  na intervalu  $[-1, 1]$  iz  $\mathbb{P}_3$ . Za residual velja  $\|f - p^*\|_{\infty, [-1, 1]} = \|T_4\|_{\infty, [-1, 1]} = 1$ . Grafi polinomov  $f$ ,  $p^*$  in  $f - p^*$  so prikazani na sliki 1.2b, po vrsti v modri, rdeči in rumeni barvi.  $\square$

Za splošno zvezno funkcijo  $f$  lahko polinom najboljše enakomerne aproksimacije  $p^*$  iz  $\mathbb{P}_n$  na intervalu  $[a, b]$  (približno) poiščemo z iterativnim pristopom, ki se imenuje Remesov



Slika 1.2: Polinomi Čebiševa in njihova uporaba pri enakomerni aproksimaciji.

postopek. Pri tem postopku v koraku  $k$  določimo polinom  $p_k^*$ , ki je boljši približek za  $p^*$  kot polinom  $p_{k-1}^*$  iz koraka  $k - 1$ , saj velja

$$\|f - p_k^*\|_{\infty, [a, b]} < \|f - p_{k-1}^*\|_{\infty, [a, b]}.$$

Postopek začnemo s poljubno množico točk  $E_0 \subset [a, b]$  moči  $n + 2$ , na podlagi katere določimo polinom  $p_0^*$ . Ob koncu začetnega koraka množico  $E_0$  zamenjamo z množico  $E_1 \subset [a, b]$ , ki se od prve razlikuje le v eni točki, in postopek ponovimo.

V koraku  $k$  je polinom  $p_k^*$  določen z množico  $E_k = \{x_0, x_1, \dots, x_{n+1}\}$  kot polinom, ki zadošča lastnosti

$$f(x_i) - p_k^*(x_i) = (-1)^{-i}m, \quad i = 0, 1, \dots, n + 1, \quad (1.4)$$

za nek  $m \in \mathbb{R}$ . To je torej polinom, za katerega residual  $f - p_k^*$  alternirajoče doseže vrednost  $m$  v točkah iz množice  $E_k$ . Če polinom  $p_k^*$  predstavimo v potenčni bazi kot

$$p_k^*(x) = a_n x^n + \dots + a_1 x + a_0,$$

ga lahko glede na zahteve (1.4) izračunamo kot rešitev sistema linearnih enačb

$$\begin{bmatrix} 1 & 1 & x_0 & x_0^2 & \dots & x_0^n \\ -1 & 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (-1)^{n+1} & 1 & x_{n+1} & x_{n+1}^2 & \dots & x_{n+1}^n \end{bmatrix} \begin{bmatrix} m \\ a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_{n+1}) \end{bmatrix}. \quad (1.5)$$

Ko rešimo sistem (1.5), lahko obravnavamo residual  $r = f - p_k^*$ . Če se izkaže, da je  $\|r\|_{\infty, [a, b]} = |m|$ , polinom  $p_k^*$  po (1.4) zadošča lastnosti, opisani v izreku 1.2, in  $p^* = p_k^*$ !

**Naloga 1.8.** Napravite korak Remesovega postopka za funkcijo  $f$ , podano v nalogi 1.3, na začetni množici  $E_0 = \{0, 1/6, 2/3, 1\} \subset [0, 1]$ . Obravnavajte residual  $r = f - p_0^*$  na intervalu  $[0, 1]$ .

*Rešitev.* Sistem (1.5), ki določa polinom  $p_0^*(x) = a_2x^2 + a_1x + a_0$ , je v tem konkretnem primeru enak

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & \frac{1}{6} & \frac{1}{36} \\ 1 & 1 & \frac{2}{3} & \frac{4}{9} \\ -1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} m \\ a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{4} \end{bmatrix},$$

njegova rešitev pa ustreza  $m = 1/24$ ,  $a_0 = 23/24$ ,  $a_1 = -5/3$ ,  $a_2 = 1$ . Odvod residuala

$$r(x) = \frac{1}{3x+1} - \left(x^2 - \frac{5}{3}x + \frac{23}{24}\right)$$

je

$$r'(x) = -\frac{3}{(3x+1)^2} - 2x + \frac{5}{3} = -\frac{(3x+2)(6x-1)(3x-2)}{3(3x+1)^2},$$

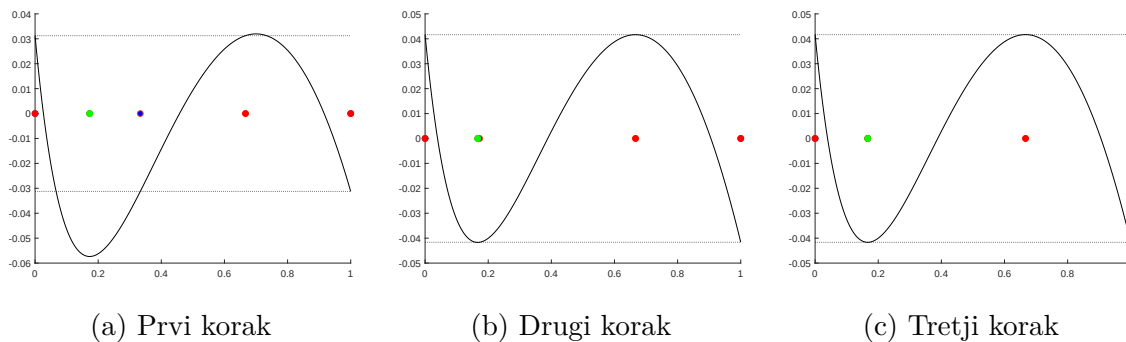
torej sta  $1/6$  in  $2/3$  stacionarni točki  $r$  na intervalu  $[0, 1]$ . Vrednosti  $r$  v teh dveh točkah sta  $-1/24$  in  $1/24$  ter predstavljata (lokalni) minimum in maksimum residuala. Na robu velja  $r(0) = 1/24$  in  $r(1) = -1/24$ , torej je  $\|r\|_{\infty,[0,1]} = 1/24 = m$ , kar pomeni, da je  $p_0^*$  polinom najboljše enakomerne aproksimacije za  $f$  iz  $\mathbb{P}_2$  na intervalu  $[0, 1]$ .  $\square$

Medtem ko pri eksaktnem računanju včasih naletimo na primer, ko polinom  $p_k^*$  kar ustreza polinomu  $p^*$ , se pri numeričnem računanju že zaradi numeričnih napak to ne zgodi. Kot v nalogi 1.8 postopek nadaljujemo tako, da poiščemo točko  $y \in [a, b]$ , v kateri residual  $r$  doseže globalni ekstrem na intervalu  $[a, b]$ . To je točka, za katero velja  $|r(y)| = \|r\|_{\infty,[a,b]}$ , zato je  $|r(y)| - |m| \geq 0$ . Če je razlika  $|r(y)| - |m|$  manjša od neke vnaprej predpisane (majhne) tolerance  $\varepsilon > 0$ , polinom  $p_k^*$  skoraj izpolnjuje lastnost iz izreka 1.2 in zato je  $p_k^* \approx p^*$ . V takem primeru z iteracijo zaključimo.

V primeru, da pogoj za konec iteracije ni izpolnjen, je treba določiti množico  $E_{k+1}$  za nadaljevanje postopka. Dobimo jo tako, da eno izmed točk v množici  $E_k$  nadomestimo s točko  $y$ . Za konvergenco postopka je bistvenega pomena, kako to storimo. Poiščemo indeks  $j \in \{1, 2, \dots, n+1\}$ , za katerega velja  $x_{j-1} < y < x_j$ . Če je predznak  $r(y)$  enak predznaku  $r(x_j)$ , definiramo  $E_{k+1} = E_k \setminus \{x_j\} \cup \{y\}$ , sicer pa  $E_{k+1} = E_k \setminus \{x_{j-1}\} \cup \{y\}$ . S tem zagotovimo, da residual  $r$  alternira tudi na točkah iz množice  $E_{k+1}$ . Lahko se primeri tudi  $a \leq y < x_0$  ali  $x_{n+1} < y \leq b$ . V tem primeru ravnamo podobno, le da pri zamenjavi izbiramo med točkama  $x_0$  in  $x_{n+1}$ .

**Naloga 1.9.** V Matlabu napravite nekaj korakov Remesovega postopka za primer iz naloge 1.8, le da začnete z množico  $E_0 = \{0, 1/3, 2/3, 1\}$ .

*Rešitev.* Rezultate prikazuje slika 1.3, na kateri so narisani residuali  $f - p_k^*$  za  $k = 0, 1, 2$ . Rdeče točke predstavljajo točke iz množic  $E_k$ . Zelena točka označuje točko iz  $E_{k+1}$ , ki nadomesti modro točko iz  $E_k$ . Pikčasti črti predstavljata konstanti  $-m$  in  $m$ . Razlika med  $|m|$  in  $\|f - p_3^*\|_{\infty,[0,1]}$  v četrtem koraku postopka je reda  $10^{-16}$ , enako pa velja tudi za koeficiente polinoma  $p_3^*$  v primerjavi s polinomom  $p^*$ , eksaktno izračunanim v nalogi 1.8.  $\square$



Slika 1.3: Residuali pri izvedbi Remesovega postopka za primer iz naloge 1.9.

## 1.3 Odsekoma polinomske funkcije

Iskanje najboljše aproksimacije za dano funkcijo je lahko računsko zahtevno, zato so za praktično rabo navadno privlačnejši enostavnejši pristopi za konstrukcijo aproksimantov. Med take prav gotovo sodi interpolacija s polinomi, ki pa postane nezanesljiva z višanjem stopnje polinomov. Mnogo boljše rezultate dobimo, če funkcijo interpoliramo s polinomi nizkih stopenj na manjših odsekih aproksimacijskega intervala, nato pa posamezne aproksimacije povežemo z določenim redom gladkosti.

### 1.3.1 Polinomska interpolacija

Funkciji  $f$  lahko priredimo polinom  $\mathcal{I}_n f \in \mathbb{P}_n$ , ki se s funkcijo ujema v izbranih točkah  $x_0, x_1, \dots, x_n$ . Tak polinom imenujemo interpolacijski polinom.

**Trditev 1.4.** *Naj bo  $f$  funkcija, ki je definirana v paroma različnih točkah  $x_0, x_1, \dots, x_n$ . V prostoru  $\mathbb{P}_n$  obstaja enoličen polinom  $\mathcal{I}_n f$ , ki zadošča pogojem*

$$\mathcal{I}_n f(x_i) = f(x_i), \quad i = 0, 1, \dots, n.$$

*Dokaz.* Glej (Plestenjak, 2015, Izrek 10.1). □

Za točke  $x_0, x_1, \dots, x_n$  lahko brez škode za splošnost predpostavimo, da so urejene po velikosti:

$$x_0 \leq x_1 \leq \dots \leq x_n.$$

Polinom  $\mathcal{I}_n f$  lahko izrazimo s pomočjo deljenih diferenc v bazi premaknjenih potenc  $x \mapsto (x - x_0) \dots (x - x_{i-1})$ ,  $i = 0, 1, \dots, n$ .

Deljena diferenca funkcije  $f$  na točkah iz množice  $\{x_i, x_{i+1}, \dots, x_{i+k}\} \subseteq \{x_0, x_1, \dots, x_n\}$  je vodilni koeficient polinoma iz prostora  $\mathbb{P}_k$ , ki interpolira vrednosti funkcije  $f$  v točkah  $x_i, x_{i+1}, \dots, x_{i+k}$ . Označimo jo z  $[x_i, x_{i+1}, \dots, x_{i+k}]f$ , izračunamo pa jo lahko s pomočjo rekurzivne formule

$$[x_i, x_{i+1}, \dots, x_{i+k}]f = \frac{[x_{i+1}, x_{i+2}, \dots, x_{i+k}]f - [x_i, x_{i+1}, \dots, x_{i+k-1}]f}{x_{i+k} - x_i}.$$

Za  $k = 0$  je  $[x_i]f = f(x_i)$ . Deljena diferenca je dobro definirana tudi v primeru, ko je  $x_i = x_{i+1} = \dots = x_{i+k}$ , in sicer kot

$$[x_i, x_{i+1}, \dots, x_{i+k}]f = \frac{f^{(k)}(x_i)}{k!}.$$

Naslednji izrek podaja polinom  $\mathcal{I}_n f$  v tako imenovani Newtonovi obliki.

**Izrek 1.3.** Naj bo  $f$  funkcija, ki je definirana v točkah  $x_0, x_1, \dots, x_k$  in  $r$ -krat odvedljiva v vsaki točki  $x_i$ , za katero velja  $x_i = x_{i+1} = \dots = x_{i+r}$ . Potem je

$$\mathcal{I}_n f(x) = \sum_{i=0}^n [x_0, x_1, \dots, x_i]f \cdot (x - x_0)(x - x_1) \dots (x - x_{i-1})$$

interpolacijski polinom za funkcijo  $f$ , ki za vsako točko  $x_i$  zadošča pogoju

$$(\mathcal{I}_n f)^{(j)}(x_i) = f^{(j)}(x_i), \quad j = 0, 1, \dots, r,$$

kjer je  $x_i = x_{i+1} = \dots = x_{i+r}$ .

**Naloga 1.10.** Naj bo  $f$  funkcija iz naloge 1.3. Zapišite polinom  $\mathcal{I}_3 f$ , ki interpolira vrednosti  $f$  v točkah  $x_0 = 0$ ,  $x_1 = 1/3$ ,  $x_2 = 2/3$ ,  $x_3 = 1$ , v Newtonovi obliki.

*Rešitev.* Polinom  $p$  določimo na podlagi naslednje sheme. Vrednosti deljenih diferenc so izračunane s pomočjo rekurzivne formule.

| $\cdot$       | $[ \cdot ]f$  | $[ \cdot, \cdot ]f$ | $[ \cdot, \cdot, \cdot ]f$ | $[ \cdot, \cdot, \cdot, \cdot ]f$ |
|---------------|---------------|---------------------|----------------------------|-----------------------------------|
| 0             | 1             |                     |                            |                                   |
| $\frac{1}{3}$ | $\frac{1}{2}$ | $-\frac{3}{2}$      | $\frac{3}{2}$              |                                   |
| $\frac{2}{3}$ | $\frac{1}{3}$ | $-\frac{1}{2}$      | $\frac{3}{8}$              | $-\frac{9}{8}$                    |
| 1             | $\frac{1}{4}$ | $-\frac{1}{4}$      |                            |                                   |

Po izreku 1.3 iz podatkov v zgornji diagonali sheme razberemo, da je interpolacijski polinom za funkcijo  $f$  podan s predpisom

$$\mathcal{I}_3 f(x) = 1 - \frac{3}{2}x + \frac{3}{2}x \left(x - \frac{1}{3}\right) - \frac{9}{8}x \left(x - \frac{1}{3}\right) \left(x - \frac{2}{3}\right).$$

Prikazan je na sliki 1.4a. □

Pri analizi napake aproksimacije funkcije z interpolacijskim polinomom si lahko pomagamo z naslednjim izrekom.

**Izrek 1.4.** Naj bo  $[a, b]$  interval, ki vsebuje točke  $x_0, x_1, \dots, x_n$ , in naj bo  $f$  funkcija iz prostora  $C^{n+1}([a, b])$ . Za vrednost residuala  $\mathcal{I}_n f - f$  v točki  $x \in [a, b]$  velja

$$\mathcal{I}_n f(x) - f(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega(x),$$

pri čemer je  $\xi \in (a, b)$  točka, odvisna od  $x$ , in  $\omega$  polinom iz  $\mathbb{P}_{n+1}$ , ki je podan s predpisom  $\omega(x) = (x - x_0)(x - x_1) \dots (x - x_n)$ .

**Naloga 1.11.** Naj bosta  $f$  in  $\mathcal{I}_3 f$  funkcija in interpolacijski polinom iz naloge 1.10. Ocenite napako  $\|f - \mathcal{I}_3 f\|_{\infty, [0,1]}$ .

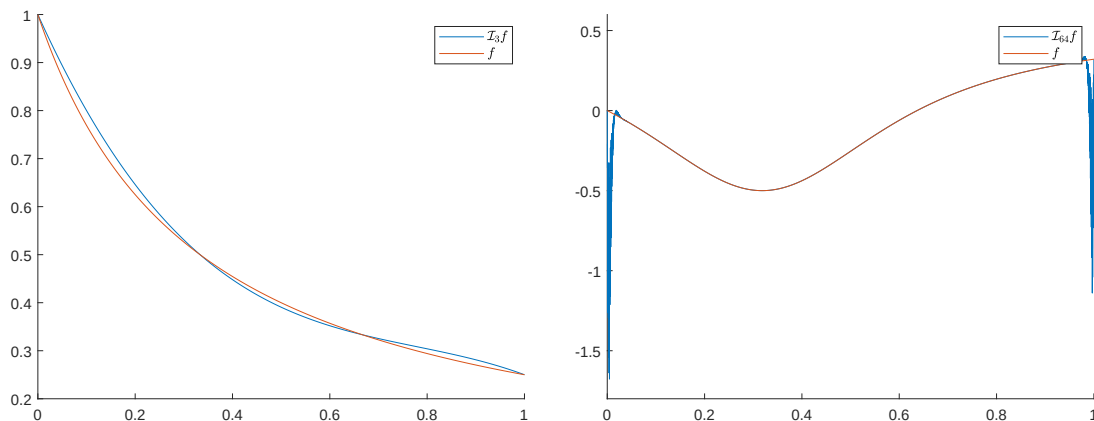
*Rešitev.* Na podlagi izreka 1.4 sklepamo

$$\max_{x \in [0,1]} |f(x) - \mathcal{I}_3 f(x)| \leq \frac{1}{4!} \max_{x \in [0,1]} |f^{(4)}(x)| \max_{x \in [0,1]} \left| x(x - \frac{1}{3})(x - \frac{2}{3})(x - 1) \right|.$$

Ker je  $f^{(4)}(x) = 3^4 \cdot 4!/(x+1)^5$ , je prvi del zgornje ocene navzgor omejen s  $3^4$ . Za poenostavitev drugega dela ocene posebej obravnavamo maksimalni absolutni vrednosti parabol  $x \mapsto x(x-1)$  in  $x \mapsto (x-1/3)(x-2/3)$  na intervalu  $[0,1]$ . Prva je omejena z  $1/4$ , druga pa z  $2/9$ . To pomeni, da je drugi maksimum v oceni gotovo manjši od  $1/18$ , in  $\|f - \mathcal{I}_3 f\|_{\infty, [0,1]} \leq 9/2$ .  $\square$

**Primer 1.2.** V Matlabu izračunamo vrednosti interpolacijskega polinoma  $\mathcal{I}_n f$  za ekvidistantne interpolacijske točke  $x_0, x_1, \dots, x_n$  na intervalu  $[0,1]$  in funkcijo  $f$  iz naloge 1.1. Na sliki 1.4b je prikazan polinom  $\mathcal{I}_{64} f$ , pri katerem je napaka v bližini robov intervala bistveno večja, kot bi jo pričakovali pri tako visoki stopnji. Izračun napake  $\|f - \mathcal{I}_n f\|_{\infty, X_{100000}}$  pokaže, da je pri  $n = 40$  napaka  $5.50 \cdot 10^{-5}$ , torej majhna, pri  $n = 64$  pa naraste na 1.67. Z višanjem stopnje se aproksimacija še bistveno poslabšuje, pri  $n = 100$  je napaka reda  $10^{-10}$ .

Obnašanje interpolacijskega polinoma iz primera 1.2 lahko omilimo z drugačno izbiro interpolacijskih točk. A v splošnem višanje stopnje polinoma ni dobra rešitev, za boljši pristop se izkaže deljenje intervala aproksimacije na manjše dele.



(a) Interpolacija funkcije iz naloge 1.10.

(b) Interpolacija funkcije iz primera 1.2.

Slika 1.4: Polinomska interpolacija funkcij.

### 1.3.2 Funkcijski zleпки

Interval  $[a, b]$  razdelimo na manjše intervale s točkami iz množice  $X = \{x_0, x_1, \dots, x_m\}$ . Predpostavljamo, da so točke urejene po velikosti in da določajo delitev

$$a = x_0 < x_1 < x_2 < \dots < x_m = b.$$

Funkciji, ki je odsekoma podana na podintervalih  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, m$ , rečemo zlepek. Posamezni odseki zleпка se stikajo v točkah  $x_0, x_1, \dots, x_m$ , zato te imenujemo stične točke. Red gladkosti zleпка je odvisen od ujemanja lokalnih predpisov v notranjih stičnih točkah  $x_1, x_2, \dots, x_{m-1}$ .

**Naloga 1.12.** Naj bo zlepek  $z : [-2, 2] \rightarrow \mathbb{R}$  podan s predpisom

$$z(x) = \begin{cases} \ln(x+2), & x \in [-2, -1) \\ (x+1)^2, & x \in [-1, 0) \\ \frac{2x-1}{x^2+1}, & x \in [0, 1) \\ \frac{x}{2}, & x \in [1, 2] \end{cases}.$$

Obravnavajte zveznost in zvezno odvedljivost zleпка  $z$  v stičnih točkah.

*Rešitev.* Najprej izračunamo, da je odvod zleпка  $z$  podan s predpisom

$$z'(x) = \begin{cases} \frac{1}{x+2}, & x \in [-2, -1) \\ 2(x+1), & x \in [-1, 0) \\ \frac{-2x^2+2x+2}{(x^2+1)^2}, & x \in [0, 1) \\ \frac{1}{2}, & x \in [1, 2] \end{cases}.$$

Sedaj lahko obravnavamo zveznost in zvezno odvedljivost v treh stičnih točkah. V točki  $x = -1$  imamo

$$\begin{aligned} \lim_{x \nearrow -1} z(x) &= \ln(-1+2) = 0, & \lim_{x \searrow -1} z(x) &= (-1+1)^2 = 0, \\ \lim_{x \nearrow -1} z'(x) &= \frac{1}{-1+2} = 1, & \lim_{x \searrow -1} z'(x) &= 2(-1+1) = 0, \end{aligned}$$

kar pomeni, da je funkcija  $z$  v točki  $-1$  zvezna, ni pa zvezno odvedljiva. Pri  $x = 0$  je

$$\lim_{x \nearrow 0} z(x) = (0+1)^2 = 1, \quad \lim_{x \searrow 0} z(x) = \frac{0-1}{0+1} = -1,$$

torej funkcija  $z$  v točki  $0$  ni niti zvezna. Pri  $x = 1$  pa je

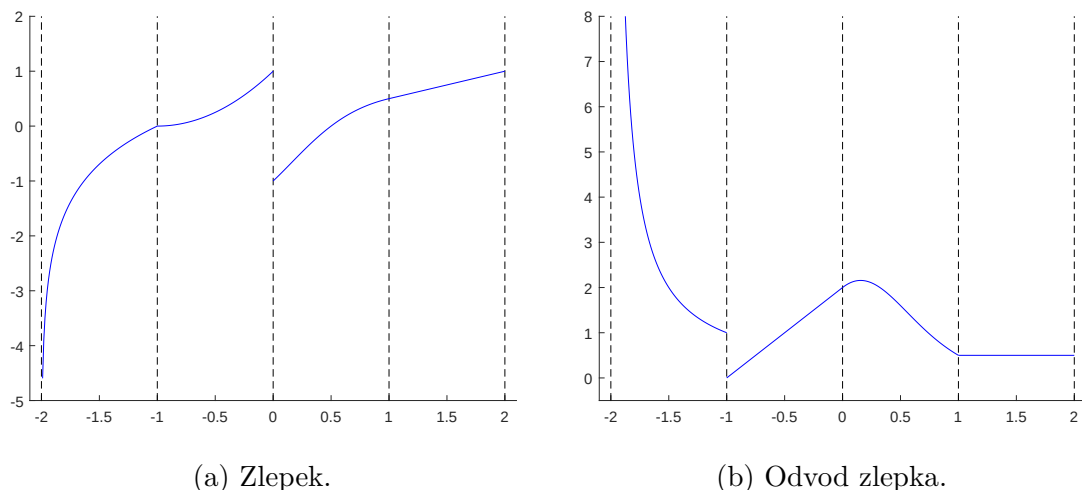
$$\begin{aligned} \lim_{x \nearrow 1} z(x) &= \frac{2-1}{1+1} = \frac{1}{2}, & \lim_{x \searrow 1} z(x) &= \frac{1}{2}, \\ \lim_{x \nearrow 1} z'(x) &= \frac{-2+2+2}{(1+1)^2} = \frac{1}{2}, & \lim_{x \searrow 1} z'(x) &= \frac{1}{2}, \end{aligned}$$

iz česar sledi, da je  $z$  v točki  $1$  zvezno odvedljiva. Zlepek  $z$  in njegov odvod  $z'$  sta prikazana na sliki 1.5.  $\square$

Najpogostejši primer zleпkov so odsekoma polinomske funkcije, ki sestavljajo prostor

$$\mathbb{S}_n^r(X) = \left\{ s \in C^r([a, b]); s|_{[x_{i-1}, x_i]} \in \mathbb{P}_n, i = 1, 2, \dots, m \right\}.$$

To so torej  $r$ -krat zvezno odvedljive funkcije z lastnostjo, da je zožitev na poljuben interval, določen z zaporednima stričnima točkama iz množice  $X$ , polinom iz  $\mathbb{P}_n$ . Število prostih parametrov tovrstne funkcije ustreza dimenziji prostora  $\mathbb{S}_n^r(X)$ .



Slika 1.5: Zlepek iz naloge 1.12.

**Trditev 1.5.** Dimenzija prostora  $\mathbb{S}_n^r(X)$  je  $mn - (m - 1)r + 1$ .

*Dokaz.* Ker je zožitev na posamezen podinterval polinom stopnje največ  $n$ , je na njem na voljo  $n + 1$  parametrov. Intervalov je  $m$ , torej je na razpolago  $m(n + 1)$  koeficientov. A ker moramo v notranjih stičnih točkah zagotoviti  $r$ -kratno zvezno odvedljivost, imamo v vsaki izmed  $m - 1$  točk  $r + 1$  dodatnih pogojev. Torej je

$$\dim(\mathbb{S}_n^r(X)) = m(n + 1) - (m - 1)(r + 1).$$

Krajši račun pokaže, da je zgornja vrednost enaka  $mn - (m - 1)r + 1$ . □

Pri konstrukciji odsekoma polinomskih funkcij si navadno pomagamo z interpolacijo vrednosti (in tudi odvodov) funkcije. Stremimo k najvišji možni gladkosti  $r = n - 1$ , pri kateri se dimenzija prostora poenostavi v  $m + n$ . Če bi v notranji delilni točki  $x_i \in X$  zahtevali red  $r = n$ , bi na podintervalih  $[x_{i-1}, x_i]$  in  $[x_i, x_{i+1}]$  dobili isti polinom, kar pravzaprav pomeni, da bi eliminirali delilno točko  $x_i$ .

### 1.3.3 Odsekoma linearne funkcije

Preprost primer odsekoma polinomskih funkcij so zvezne odsekoma linearne funkcije, ki sestavljajo prostor  $\mathbb{S}_1^0(X)$  z dimenzijo  $m + 1$ . Graf take funkcije ustreza lomljeni črti. Zvezna odsekoma linearna funkcija  $\mathcal{I}_1^0 f \in \mathbb{S}_1^0(X)$ , ki interpolira vrednosti funkcije  $f$  v točkah iz  $X$ , predstavlja aproksimacijo za funkcijo  $f$ , ki jo lahko v točki  $x \in [x_{i-1}, x_i]$  izrazimo s pomočjo interpolacijskega polinoma kot

$$\mathcal{I}_1^0 f(x) = [x_{i-1}]f + [x_{i-1}, x_i]f \cdot (x - x_{i-1}) = f(x_{i-1}) + \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}(x - x_{i-1}),$$

kar se poenostavi v

$$\mathcal{I}_1^0 f(x) = \frac{x_i - x}{x_i - x_{i-1}} f(x_{i-1}) + \frac{x - x_{i-1}}{x_i - x_{i-1}} f(x_i). \quad (1.6)$$

Iz tega predpisa je razvidno, da s parametri  $x$  med  $x_{i-1}$  in  $x_i$  dobimo premico med točkama  $(x_{i-1}, f(x_{i-1}))$  in  $(x_i, f(x_i))$ .

**Naloga 1.13.** Naj bo  $x_0 = 0$ ,  $x_1 = 1/4$ ,  $x_2 = 1/2$ ,  $x_3 = 3/4$ ,  $x_4 = 1$ . Zapišite predpise za funkcijo  $\mathcal{I}_1^0 f \in \mathbb{S}_1^0(X)$ , ki interpolira vrednosti funkcije  $f(x) = \sin(2\pi x)$  v točkah iz  $X$ . Dokažite, da je  $\|f - \mathcal{I}_1^0 f\|_{\infty, [0,1]} < \pi/2 - 1$ .

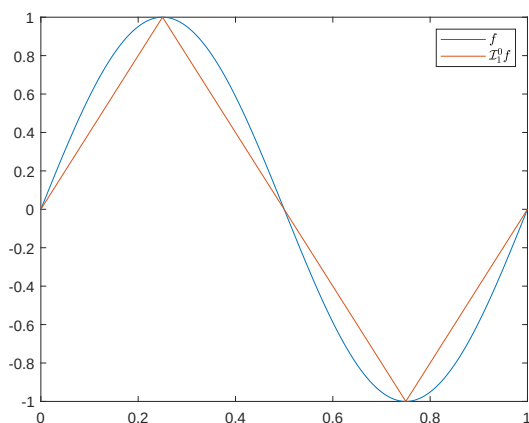
*Rešitev.* Iz (1.6) sledi, da je

$$\mathcal{I}_1^0 f(x) = \begin{cases} 4x, & x \in [x_0, x_1] \\ -4x + 2, & x \in [x_1, x_2] \\ -4x + 2, & x \in [x_2, x_3] \\ 4x - 4, & x \in [x_3, x_4] \end{cases}.$$

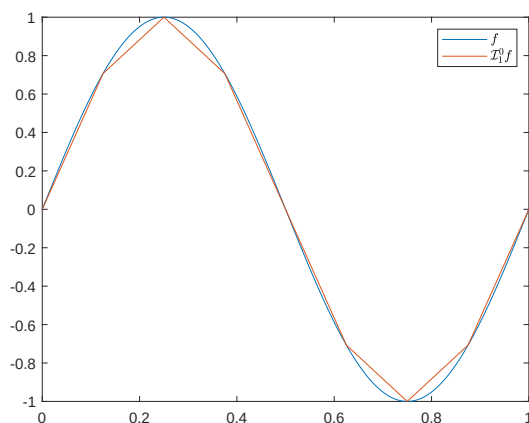
Graf funkcije  $\mathcal{I}_1^0 f$  je prikazan na sliki 1.6a. Obravnavajmo še vrednosti  $f(x) - \mathcal{I}_1^0 f(x)$  za  $x \in [0, 1]$ . Za začetek se omejimo na primer, ko je  $x \in [x_0, x_1]$ . Ker za vse take abscise  $x$  velja

$$f(x) - \mathcal{I}_1^0 f(x) = \sin(2\pi x) - 4x \leq 2\pi x - 4x,$$

sledi  $\|f - \mathcal{I}_1^0 f\|_{\infty, [x_0, x_1]} < \pi/2 - 1$ . Strogi enačaj je posledica dejstva, da v zgornji oceni vrednost  $\pi/2 - 1$  dobimo le pri  $x = 1/4$ , v tej točki pa vemo, da zlepek interpolira  $f$ . Podobne sklepe lahko po simetriji opravimo na intervalih  $[x_1, x_2]$ ,  $[x_2, x_3]$  in  $[x_3, x_4]$ , na podlagi česar zaključimo, da je  $\|f - \mathcal{I}_1^0 f\|_{\infty, [0,1]} < \pi/2 - 1$ .  $\square$



(a)  $m = 4$ .



(b)  $m = 8$ .

Slika 1.6: Interpolacija z zveznima odsekoma linearnima funkcijama (naloga 1.13).

Če je funkcija  $f$  dvakrat zvezno odvedljiva, lahko napako aproksimacije z interpolantom  $\mathcal{I}_1^0 f \in \mathbb{S}_1^0(X)$  ocenimo s pomočjo izreka 1.4.

**Trditev 1.6.** Za  $f \in C^2([a, b])$  velja

$$\|f - \mathcal{I}_1^0 f\|_{\infty, [a,b]} \leq \frac{1}{8} \left( \max_{i=1,2,\dots,m} \Delta x_{i-1} \right)^2 \|f^{(2)}\|_{\infty, [a,b]},$$

kjer je  $\Delta x_{i-1} = x_i - x_{i-1}$ .

*Dokaz.* Za vsak  $i \in \{1, 2, \dots, m\}$  in vsak  $x \in [x_{i-1}, x_i]$  velja

$$f(x) - \mathcal{I}_1^0 f(x) = \frac{f^{(2)}(\xi)}{2} (x - x_{i-1})(x - x_i),$$

kjer je  $\xi \in (x_{i-1}, x_i)$ . Ker je vrednost parabole  $x \mapsto (x - x_{i-1})(x - x_i)$  največja v središčni točki  $(x_{i-1} + x_i)/2$ , je

$$|f(x) - \mathcal{I}_1^0 f(x)| \leq \frac{\|f^{(2)}\|_{\infty, [a, b]}}{2} \frac{1}{4} (x_i - x_{i-1})^2,$$

iz česar po definiciji enakomerne norme sledi dokazovana ocena.  $\square$

**Naloga 1.14.** Naj bodo  $x_i = i/m$ ,  $i = 0, 1, \dots, m$ , ekvidistantne stične točke na intervalu  $[0, 1]$ . S pomočjo trditve 1.6 čimbolj natančno ocenite, katero je najmanjše število  $m$ , pri katerem za funkcijo  $f$  iz naloge 1.13 velja  $\|f - \mathcal{I}_1^0 f\|_{\infty, [0, 1]} < 10^{-6}$ .

*Rešitev.* Drugi odvod funkcije  $f$  lahko v poljubni točki  $z$  intervala  $[0, 1]$  absolutno omejimo z  $(2\pi)^2$ . Poleg tega je  $\Delta x_{i-1} = 1/m$  za vsak  $i \in \{1, 2, \dots, m\}$ , zato po trditvi 1.6 velja

$$\|f - \mathcal{I}_1^0 f\|_{\infty, [0, 1]} \leq \frac{1}{8} (2\pi)^2 \frac{1}{m^2} = \frac{\pi^2}{2m^2}.$$

Najmanjše naravno število  $m$ , ki zadošča lastnosti  $\pi^2 \leq 2m^2 \cdot 10^{-6}$ , je 2222.  $\square$

Prostor  $\mathbb{S}_1^0(X)$  lahko opremimo z bazo. Označimo bazne funkcije s  $H_i$ ,  $i = 0, 1, \dots, m$ , in jih definirajmo z zahtevo

$$H_i(x_j) = \begin{cases} 1, & i = j \\ 0, & \text{sicer} \end{cases}.$$

S tem smo zastavili interpolacijski problem in iz (1.6) sledi

$$H_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}}, & x \in [x_{i-1}, x_i] \\ \frac{x_{i+1} - x}{x_{i+1} - x_i}, & x \in [x_i, x_{i+1}] \\ 0, & \text{sicer} \end{cases}.$$

Prav tako iz (1.6) sledi tudi

$$\mathcal{I}_1^0 f = \sum_{i=0}^m f(x_i) H_i,$$

kar predstavlja ekonomičen zapis interpolanta za funkcijo  $f$ , ki spominja na aproksimacijo funkcije z Bernsteinovim polinomom. V resnici se da podobno kot v primeru Bernsteinovih polinomov dokazati, da za poljubno funkcijo  $f \in C([a, b])$  za vsak  $x \in [a, b]$  velja

$$\lim_{\Delta X \rightarrow 0} \mathcal{I}_1^0 f(x) = f(x).$$

Pri tem  $\Delta X \rightarrow 0$  pomeni, da gostimo število stičnih točk iz množice  $X$  na intervalu  $[a, b]$  tako, da gre razlika  $\Delta x_{i-1}$  za vsak  $i \in \{1, 2, \dots, m\}$  proti 0. To dokazuje aproksimacijski izrek, ki je alternativa Weierstrassovemu izreku 1.1 za prostor polinomov.

**Izrek 1.5.** Za  $f \in C([a, b])$  velja

$$\lim_{\Delta X \rightarrow 0} \text{dist}_{\infty}(f, \mathbb{S}_1^0(X)) = 0.$$

### 1.3.4 Odsekoma kvadratične funkcije

Z odsekoma kvadratičnimi funkcijami lahko konstruiramo zvezno odvedljivo aproksimacijo za funkcijo  $f$ . Naj bo  $\mathcal{I}_2^1 f \in \mathbb{S}_2^1(X)$  funkcija, ki zadošča interpolacijskim pogojem

$$\mathcal{I}_2^1 f(x_i) = f(x_i), \quad i = 0, 1, \dots, m. \quad (1.7)$$

Ker je dimenzija prostora  $\mathbb{S}_2^1(X)$  enaka  $m+2$ , je s temi pogoji aproksimacija  $\mathcal{I}_2^1 f$  določena do enega parametra natančno.

Ker je  $\mathcal{I}_2^1 f$  na intervalu  $[x_{i-1}, x_i]$ ,  $i \in \{1, 2, \dots, m\}$ , polinom iz  $\mathbb{P}_2$ , ki interpolira vrednosti funkcije  $f$  v točkah  $x_{i-1}$  in  $x_i$ , ga lahko za poljuben  $x \in [x_{i-1}, x_i]$  zapišemo v obliki

$$\mathcal{I}_2^1 f(x) = [x_{i-1}]f + [x_{i-1}, x_{i-1}]f \cdot (x - x_{i-1})^2 + [x_{i-1}, x_{i-1}, x_i]f \cdot (x - x_{i-1})^2.$$

Ta izraz se poenostavi v

$$\mathcal{I}_2^1 f(x) = f(x_{i-1}) + s_{i-1}(x - x_{i-1}) + \frac{f(x_i) - f(x_{i-1}) - s_{i-1}\Delta x_{i-1}}{(\Delta x_{i-1})^2}(x - x_{i-1})^2, \quad (1.8a)$$

če namesto odvoda  $f'(x_{i-1}) = [x_{i-1}, x_{i-1}]f$ , ki ni podan ali celo ni definiran, uvedemo parameter  $s_{i-1}$ . Ker zahtevamo, da je  $\mathcal{I}_2^1 f$  zvezno odvedljiva, mora v vsaki notranji delilni točki  $x_i$ ,  $i = 1, 2, \dots, m-1$ , veljati

$$\lim_{x \nearrow x_i} (\mathcal{I}_2^1 f)'(x) = \lim_{x \searrow x_i} (\mathcal{I}_2^1 f)'(x).$$

Na intervalu  $[x_{i-1}, x_i]$  je odvod  $\mathcal{I}_2^1 f$  podan s predpisom

$$x \mapsto s_{i-1} + 2 \frac{f(x_i) - f(x_{i-1}) - s_{i-1}\Delta x_{i-1}}{(\Delta x_{i-1})^2}(x - x_{i-1})$$

na intervalu  $[x_i, x_{i+1}]$  pa s predpisom

$$x \mapsto s_i + 2 \frac{f(x_{i+1}) - f(x_i) - s_i\Delta x_i}{(\Delta x_i)^2}(x - x_i)$$

Torej je  $\mathcal{I}_2^1 f$  zvezno odvedljiva natanko tedaj ko velja

$$s_{i-1} + 2 \frac{f(x_i) - f(x_{i-1}) - s_{i-1}\Delta x_{i-1}}{(\Delta x_{i-1})^2}\Delta x_{i-1} = s_i$$

oziroma

$$s_i = -s_{i-1} + 2 \frac{f(x_i) - f(x_{i-1})}{\Delta x_{i-1}} \quad (1.8b)$$

za vsak  $i \in \{1, 2, \dots, m-1\}$ . Če si izberemo  $s_0$ , so parametri  $s_1, s_2, \dots, s_{m-1}$  po rekurziji enolično določeni.

Parameter  $s_0$  lahko izberemo na primer tako, da bo  $\mathcal{I}_2^1 f$  na intervalu  $[x_0, x_1]$  premica. To velja, če je  $(\mathcal{I}_2^1 f)''(x_0) = 0$ , kar ustreza izbiri

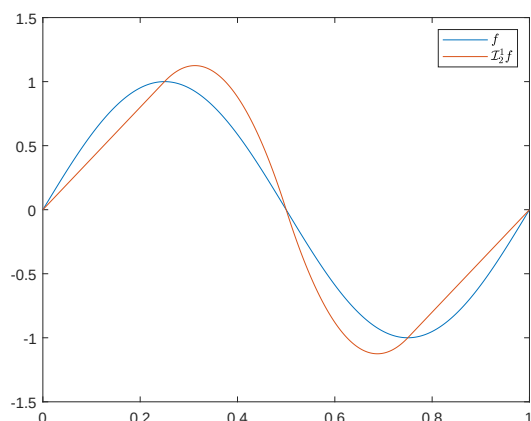
$$s_0 = \frac{f(x_1) - f(x_0)}{\Delta x_0}. \quad (1.8c)$$

**Naloga 1.15.** Naj bo  $f$  funkcija iz naloge 1.13. Zapišite formule za zožitve  $\mathcal{I}_2^1 f$  na intervale med točkami  $x_0 = 0$ ,  $x_1 = 1/4$ ,  $x_2 = 1/2$ ,  $x_3 = 3/4$ ,  $x_4 = 1$ .

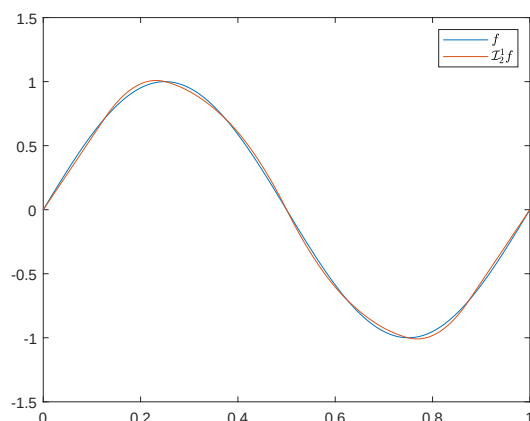
*Rešitev.* S pomočjo (1.8) izračunamo, da je

$$\mathcal{I}_2^1 f(x) = \begin{cases} 4x, & x \in [0, \frac{1}{4}] \\ -32(x - \frac{1}{4})^2 + 4(x - \frac{1}{4}) + 1, & x \in [\frac{1}{4}, \frac{1}{2}] \\ 32(x - \frac{1}{2})^2 - 12(x - \frac{1}{2}), & x \in [\frac{1}{2}, \frac{3}{4}] \\ 4(x - \frac{3}{4}) - 1, & x \in [\frac{3}{4}, 1] \end{cases}$$

funkcija iz  $\mathbb{S}_2^1(X)$ , ki zadošča pogojem (1.7). Graf  $\mathcal{I}_2^1 f$  je prikazan na sliki 1.7a.  $\square$



(a)  $m = 4$ .



(b)  $m = 8$ .

Slika 1.7: Interpolacija z zvezno odvedljivima odsekoma kvadratičnima funkcijama (naloga 1.15).

### 1.3.5 Odsekoma kubične funkcije

Dimenzija prostora  $\mathbb{S}_3^1(X)$  je po trditvi 1.5 enaka  $2(m+1)$ . To namiguje, da lahko funkciji  $f$ , ki je v točkah iz množice  $X$  odvedljiva, priredimo zvezno odvedljivo odsekoma kubično funkcijo  $\mathcal{I}_3^1 f \in \mathbb{S}_3^1(X)$ , za katero velja

$$\mathcal{I}_3^1 f(x_i) = f(x_i), \quad (\mathcal{I}_3^1 f)'(x_i) = f'(x_i), \quad i = 0, 1, \dots, m.$$

Na intervalu  $[x_{i-1}, x_i]$ ,  $i \in \{1, 2, \dots, m\}$ , je funkcija  $\mathcal{I}_3^1 f$  polinom iz  $\mathbb{P}_3$ , ki interpolira vrednosti in odvode funkcije  $f$  v točkah  $x_{i-1}$  in  $x_i$ . Zato za vsak  $x \in [x_{i-1}, x_i]$  velja

$$\begin{aligned} \mathcal{I}_3^1 f(x) = & [x_{i-1}]f + [x_{i-1}, x_{i-1}]f \cdot (x - x_{i-1}) + [x_{i-1}, x_{i-1}, x_i]f \cdot (x - x_{i-1})^2 \\ & + [x_{i-1}, x_{i-1}, x_i, x_i]f \cdot (x - x_{i-1})^2(x - x_i). \end{aligned}$$

Iz sheme deljenih diferenc

| $\cdot$   | $[\cdot]f$   | $[\cdot, \cdot]f$ | $[\cdot, \cdot, \cdot]f$                               | $[\cdot, \cdot, \cdot, \cdot]f$                                       |
|-----------|--------------|-------------------|--------------------------------------------------------|-----------------------------------------------------------------------|
| $x_{i-1}$ | $f(x_{i-1})$ | $f'(x_{i-1})$     |                                                        |                                                                       |
| $x_{i-1}$ | $f(x_{i-1})$ | $[x_{i-1}, x_i]f$ | $\frac{[x_{i-1}, x_i]f - f'(x_{i-1})}{\Delta x_{i-1}}$ | $\frac{f'(x_{i-1}) + f'(x_i) - 2[x_{i-1}, x_i]f}{(\Delta x_{i-1})^2}$ |
| $x_i$     | $f(x_i)$     | $f'(x_i)$         | $\frac{f'(x_i) - [x_{i-1}, x_i]f}{\Delta x_{i-1}}$     |                                                                       |
| $x_i$     | $f(x_i)$     |                   |                                                        |                                                                       |

sledi

$$\begin{aligned} \mathcal{I}_3^1 f(x) = & f(x_{i-1}) + f'(x_{i-1})(x - x_{i-1}) + \frac{[x_{i-1}, x_i]f - f'(x_{i-1})}{\Delta x_{i-1}}(x - x_{i-1})^2 \\ & + \frac{f'(x_{i-1}) + f'(x_i) - 2[x_{i-1}, x_i]f}{(\Delta x_{i-1})^2}(x - x_{i-1})^2(x - x_i). \end{aligned} \quad (1.9)$$

**Naloga 1.16.** Naj bo  $f$  funkcija iz naloge 1.13. Zapišite formule za zožitve  $\mathcal{I}_3^1 f$  na intervale med točkami  $x_0 = 0$ ,  $x_1 = 1/4$ ,  $x_2 = 1/2$ ,  $x_3 = 3/4$ ,  $x_4 = 1$ .

*Rešitev.* S pomočjo (1.11) izračunamo, da je

$$\mathcal{I}_3^1 f(x) = \begin{cases} 2\pi x + 8(2 - \pi)x^2 + 32(\pi - 4)x^2(x - \frac{1}{4}), & x \in [0, \frac{1}{4}] \\ 1 - 16(x - \frac{1}{4})^2 + 32(4 - \pi)(x - \frac{1}{4})^2(x - \frac{1}{2}), & x \in [\frac{1}{4}, \frac{1}{2}] \\ -2\pi(x - \frac{1}{2}) + 8(\pi - 2)(x - \frac{1}{2})^2 + 32(4 - \pi)(x - \frac{1}{2})^2(x - \frac{3}{4}), & x \in [\frac{1}{2}, \frac{3}{4}] \\ -1 + 16(x - \frac{3}{4})^2 + 32(\pi - 4)(x - \frac{3}{4})^2(x - 1), & x \in [\frac{3}{4}, 1] \end{cases}$$

funkcija iz  $\mathbb{S}_3^1(X)$ , ki interpolira vrednosti in odvode funkcije  $f$  v stičnih točkah. Graf funkcije  $\mathcal{I}_3^1 f$  je prikazan na sliki 1.8a.  $\square$

**Trditev 1.7.** Za  $f \in C^4([a, b])$  velja

$$\|f - \mathcal{I}_3^1 f\|_{\infty, [a, b]} \leq \frac{1}{384} \left( \max_{i=1, 2, \dots, m} \Delta x_{i-1} \right)^4 \|f^{(4)}\|_{\infty, [a, b]}.$$

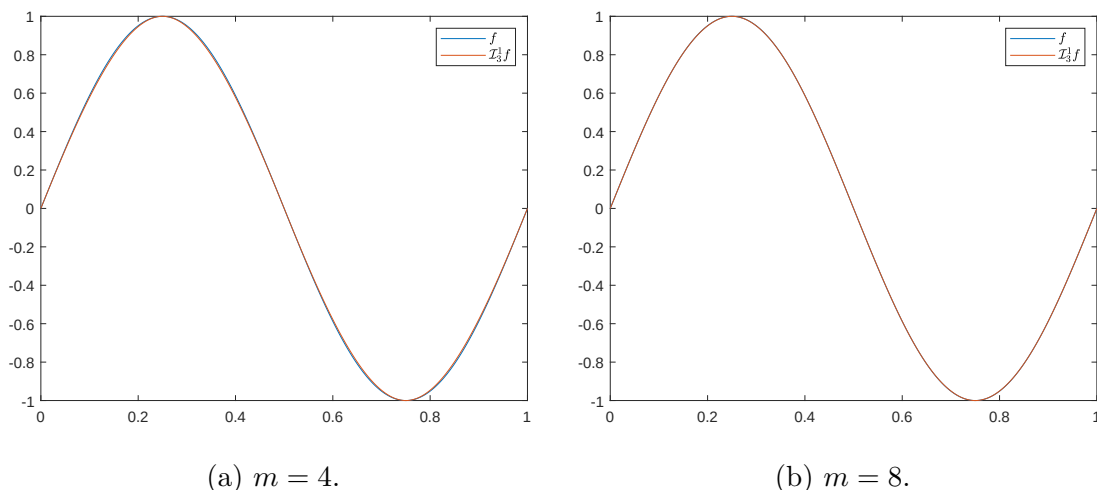
*Dokaz.* Dokaz trditve je podoben dokazu trditve 1.6. Za vsak  $i \in \{1, 2, \dots, m\}$  in vsak  $x \in [x_{i-1}, x_i]$  velja

$$f(x) - \mathcal{I}_3^1 f(x) = \frac{f^{(4)}(\xi)}{4!}(x - x_{i-1})^2(x - x_i)^2,$$

kjer je  $\xi \in (x_{i-1}, x_i)$ . Polinom  $x \mapsto ((x - x_{i-1})(x - x_i))^2$  doseže maksimalno absolutno vrednost v točki  $(x_{i-1} + x_i)/2$ , zato je

$$|f(x) - \mathcal{I}_3^1 f(x)| \leq \frac{\|f^{(4)}\|_{\infty, [a, b]}}{4!} \frac{1}{16} (\Delta x_{i-1})^4.$$

Po definiciji enakomerne norme to zaključuje dokaz.  $\square$



Slika 1.8: Interpolacija z zvezno odvedljivima odsekoma kubičnima funkcijama (naloge 1.16).

**Naloga 1.17.** Naj bodo  $x_i = i/m$ ,  $i = 0, 1, \dots, m$ , ekvidistantne stične točke na intervalu  $[0, 1]$ . S pomočjo trditve 1.7 čimbolj natančno ocenite, katero je najmanjše število  $m$ , pri katerem za funkcijo  $f$  iz naloge 1.13 velja  $\|f - \mathcal{I}_3^1 f\|_{\infty, [0,1]} < 10^{-6}$ .

*Rešitev.* Četrty odvod funkcije  $f$  je v poljubni točki z intervala  $[0, 1]$  omejen z  $(2\pi)^4$ . Ker je  $\Delta x_{i-1} = 1/m$  za vsak  $i \in \{1, 2, \dots, m\}$ , po trditvi 1.7 velja

$$\|f - \mathcal{I}_3^1 f\|_{\infty, [0,1]} \leq \frac{1}{384} (2\pi)^4 \frac{1}{m^4} = \frac{\pi^4}{24m^4}.$$

Najmanjše naravno število  $m$ , ki zadošča lastnosti  $\pi^4 \leq 24m^4 \cdot 10^{-6}$ , je 45. □

Prostor odsekoma kubičnih funkcij z maksimalno gladkostjo je  $\mathbb{S}_3^2(X)$  in ima dimenzijo  $m+3$ . Če torej želimo konstruirati interpolant za funkcijo  $f$  iz  $\mathbb{S}_3^2(X)$ , lahko predpišemo vrednosti funkcije  $f$  v točkah iz  $X$ , ne pa tudi vseh odvodov, saj ostaneta na voljo le še dva parametra svobode. Navadno ju določimo z interpolacijo odvodov v robnih točkah  $x_0$  in  $x_m$ .

Naj bo torej  $\mathcal{I}_3^2 f \in \mathbb{S}_3^2(X)$  funkcija, ki zadošča interpolacijskim pogojem

$$\mathcal{I}_3^2 f(x_i) = f(x_i), \quad i = 0, 1, \dots, m, \quad (1.10a)$$

in

$$(\mathcal{I}_3^2 f)'(x_0) = f'(x_0), \quad (\mathcal{I}_3^2 f)'(x_m) = f'(x_m). \quad (1.10b)$$

Za izpeljavo eksplicitne formule  $\mathcal{I}_3^2 f$  na intervalu  $[x_{i-1}, x_i]$ ,  $i \in \{1, 2, \dots, m\}$ , vrednosti  $f'(x_{i-1})$  in  $f'(x_i)$  v (1.11) nadomestimo s parametroma  $s_{i-1}$  in  $s_i$ . Dobimo polinom, ki je za  $x \in [x_{i-1}, x_i]$  podan s predpisom

$$\begin{aligned} \mathcal{I}_3^2 f(x) = & f(x_{i-1}) + s_{i-1}(x - x_{i-1}) + \frac{[x_{i-1}, x_i]f - s_{i-1}}{\Delta x_{i-1}}(x - x_{i-1})^2 \\ & + \frac{s_{i-1} + s_i - 2[x_{i-1}, x_i]f}{(\Delta x_{i-1})^2}(x - x_{i-1})^2(x - x_i). \end{aligned} \quad (1.11)$$

Če vzamemo  $s_0 = f'(x_0)$  in  $s_m = f'(x_m)$ , funkcija  $\mathcal{I}_3^2 f$  zadošča interpolacijskim pogojem (1.10). Parametri  $s_i$ ,  $i = 1, 2, \dots, m-1$ , so prosti in jih lahko določimo tako, da je v točkah  $x_1, x_2, \dots, x_{m-1}$  zadoščeno pogojem dvakratne zvezne odvedljivosti. Za vsak  $i \in \{1, 2, \dots, m-1\}$  mora veljati

$$\lim_{x \nearrow x_i} (\mathcal{I}_3^2 f)''(x) = \lim_{x \searrow x_i} (\mathcal{I}_3^2 f)''(x)$$

Z odvajanjem (1.11) izračunamo, da je funkcija  $(\mathcal{I}_3^2 f)''$  na intervalu  $[x_{i-1}, x_i]$  podana s predpisom

$$x \mapsto 2 \left( \frac{[x_{i-1}, x_i]f - s_{i-1}}{\Delta x_{i-1}} + (2(x - x_{i-1}) + (x - x_i)) \frac{s_{i-1} + s_i - 2[x_{i-1}, x_i]f}{(\Delta x_{i-1})^2} \right),$$

na intervalu  $[x_i, x_{i+1}]$  pa s predpisom

$$x \mapsto 2 \left( \frac{[x_i, x_{i+1}]f - s_i}{\Delta x_i} + (2(x - x_i) + (x - x_{i+1})) \frac{s_i + s_{i+1} - 2[x_i, x_{i+1}]f}{(\Delta x_i)^2} \right).$$

V točki  $x_i$  mora torej veljati

$$\frac{[x_{i-1}, x_i]f - s_{i-1}}{\Delta x_{i-1}} + 2 \frac{s_{i-1} + s_i - 2[x_{i-1}, x_i]f}{\Delta x_{i-1}} = \frac{[x_i, x_{i+1}]f - s_i}{\Delta x_i} - \frac{s_i + s_{i+1} - 2[x_i, x_{i+1}]f}{\Delta x_i},$$

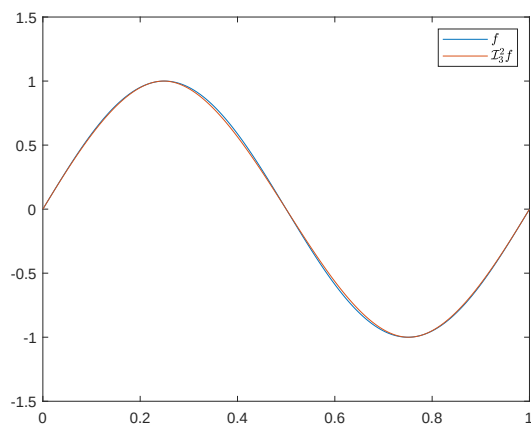
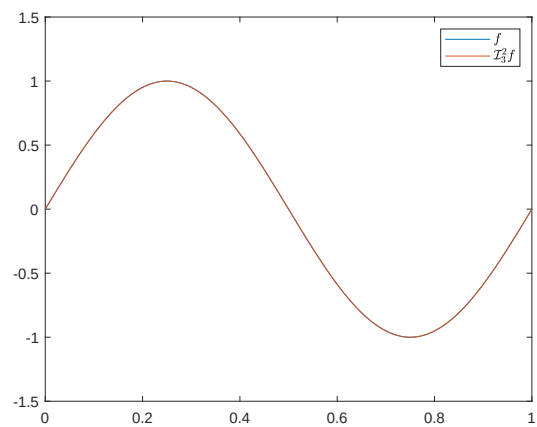
kar lahko zapišemo v obliki

$$\frac{1}{\Delta x_{i-1}} s_{i-1} + 2 \left( \frac{1}{\Delta x_{i-1}} + \frac{1}{\Delta x_i} \right) s_i + \frac{1}{\Delta x_i} s_{i+1} = 3 \left( \frac{[x_{i-1}, x_i]f}{\Delta x_{i-1}} + \frac{[x_i, x_{i+1}]f}{\Delta x_i} \right).$$

Za  $i = 1, 2, \dots, m-1$  to predstavlja sistem  $m-1$  enačb z  $m-1$  neznankami, ki ga lahko predstavimo v matrični obliki. Matrika sistema ima tridiagonalno strukturo, saj se v vrstici  $i$  neničelni le koeficienti, ki pripadajo spremenljivkam  $s_{i-1}$ ,  $s_i$  in  $s_{i+1}$  (za  $i = 1$  oziroma  $i = m-1$  je  $s_0$  oziroma  $s_m$  določen in ga upoštevamo v vektorju prostih vrednosti). Poleg tega velja, da je v vrstici  $i$  koeficient pred  $s_i$  strogo večji od vsote koeficientov pred  $s_{i-1}$  in  $s_{i+1}$ , kar pomeni, da je matrika strogo diagonalno dominantna in zato obrnljiva. To zagotavlja, da je sistem enolično rešljiv, in potrjuje obstoj  $\mathcal{I}_3^2 f$  za poljubno funkcijo  $f$ .

**Naloga 1.18.** V Matlabu implementirajte metodo, ki za dane vrednosti funkcije  $f$  v točkah iz množice  $X$  in vrednosti odvodov funkcije  $f$  v točkah  $x_0$  in  $x_m$  izračuna vrednost  $\mathcal{I}_3^2 f$  v poljubni točki z intervala  $[x_0, x_m]$ . Narišite graf  $\mathcal{I}_3^2 f$  za funkcijo  $f$  iz naloge 1.13. Oglejte si tudi obnašanje napake  $\|f - \mathcal{I}_3^2 f\|_{\infty, X_{1000}}$ , ko povečujete  $m$ .

*Rešitev.* Grafa funkcij  $\mathcal{I}_3^2 f$  za  $m = 4$  in  $m = 8$  sta prikazana na sliki 1.9. Napaka aproksimacije  $\|f - \mathcal{I}_3^2 f\|_{\infty, X_{1000}}$  je pri  $m = 4$  približno  $0.24 \cdot 10^{-1}$ , pri  $m = 8$  približno  $0.11 \cdot 10^{-2}$ , pri  $m = 50$  pa približno  $6.52 \cdot 10^{-7}$ .  $\square$

(a)  $m = 4$ .(b)  $m = 8$ .

Slika 1.9: Interpolacija z dvakrat zvezno odvedljivima odsekoma kubičnima funkcijama (naloge 1.18).

## Poglavje 2

# Računalniško podprto geometrijsko oblikovanje

Predstavitev objektov v računalniški grafiki temelji na končnih množicah točk, iz katerih po specifičnih postopkih generiramo krivulje in ploskve, ki sestavljajo objekte. Opis krivulj in ploskev pogosto temelji na polinomskih parametrizacijah.

### 2.1 Bézierjeve krivulje

Bézierjeve krivulje so podane s polinomsko parametrizacijo, ki temelji na Bernsteinovih baznih polinomih. Odlikujejo se po tem, da jih je enostavno generirati in upravljati z naborom točk v ravnini ali prostoru.

#### 2.1.1 Predstavitev z Bernsteinovimi baznimi polinomi

Krivulja  $\mathbf{k}$  v prostoru  $\mathbb{R}^d$ , kjer je  $d$  običajno 2 ali 3, je množica točk, ki jo pogosto opišemo v parametrični obliki

$$\mathbf{k} = \left\{ (k_1(t), k_2(t), \dots, k_d(t)) \in \mathbb{R}^d; t \in [0, 1] \right\},$$

kjer so  $k_1, k_2, \dots, k_d$  koordinatne funkcije z intervala  $[0, 1]$  v  $\mathbb{R}$ . Krivuljo zato pogosto kar enačimo s parametrizacijo  $\mathbf{k} : [0, 1] \rightarrow \mathbb{R}^d$ , ki vsakemu parametru  $t \in [0, 1]$  priredi točko  $\mathbf{k}(t)$  v  $\mathbb{R}^d$ .

Če so koordinatne funkcije  $k_1, k_2, \dots, k_d$  polinomi iz  $\mathbb{P}_n$ , krivuljo  $\mathbf{k}$  imenujemo polinomska krivulja stopnje  $n$ .

**Definicija 2.1.** Bézierjeva krivulja  $\mathbf{b} : [0, 1] \rightarrow \mathbb{R}^d$  stopnje  $n$  je polinomska krivulja, ki je podana s parametrizacijo

$$\mathbf{b}(t) = \sum_{i=0}^n \mathbf{b}_i B_i^n(t),$$

kjer so  $B_i^n$  Bernsteinovi bazni polinomi stopnje  $n$ , podani s predpisom (1.2), in  $\mathbf{b}_i$  točke v prostoru  $\mathbb{R}^d$ , ki jih imenujemo kontrolne točke Bézierjeve krivulje  $\mathbf{b}$ .

Ker Bernsteinovi bazni polinomi po trditvi 1.1 sestavljajo konveksno razčlenitev enote, lahko vsako točko  $\mathbf{b}(t)$  razumemo kot konveksno kombinacijo točk  $\mathbf{b}_i$ ,  $i = 0, 1, \dots, n$ . To pomeni, da celotna krivulja leži v konveksni ovojnici kontrolnih točk. Lomljeni črti, sestavljeni iz daljic med točkami  $\mathbf{b}_{i-1}$  in  $\mathbf{b}_i$ ,  $i = 1, 2, \dots, n$ , pravimo kontrolni poligon krivulje  $\mathbf{b}$ .

**Naloga 2.1.** Naj bo  $\mathbf{b}$  Bézierjeva krivulja stopnje 3, ki je podana s kontrolnimi točkami

$$\mathbf{b}_0 = (-1, 0), \quad \mathbf{b}_1 = (-a, 1), \quad \mathbf{b}_2 = (a, 1), \quad \mathbf{b}_3 = (1, 0)$$

za nek parameter  $a \in \mathbb{R}$ .

1. Pokažite, da je v primeru  $a = -1$  tangentni vektor krivulje  $\mathbf{b}$  pri parametru  $t = 1/2$  nezvezen.
2. Preverite, da ima krivulja  $\mathbf{b}$  pri izbiri  $a = -5/3$  samopresečišče. To pomeni, da obstajata različna parametra  $t_1$  in  $t_2$ , za katera velja  $\mathbf{b}(t_1) = \mathbf{b}(t_2)$ . Poskušajte ju poiskati z nastavkom  $t_{1,2} = 1/2 \mp \alpha$ ,  $\alpha \in \mathbb{R}$ .
3. V primeru, ko je  $a = 0$ , sta druga in tretja kontrolna točka enaki. Ali to pomeni, da je krivulja  $\mathbf{b}$  stopnje 2? Če ne, ali obstaja tak  $a$ , da to drži?

*Rešitev.* Zapišimo parametrizacijo podane ravninske Bézierjeve krivulje stopnje 3 po komponentah. Z upoštevanjem definicije Bernsteinovih baznih polinomov dobimo

$$\mathbf{b}(t) = \left( -(1-t)^3 - 3at(1-t)^2 + 3at^2(1-t) + t^3, 3t(1-t)^2 + 3t^2(1-t) \right),$$

kar lahko poenostavimo v

$$\mathbf{b}(t) = \left( (2-6a)t^3 + (-3+9a)t^2 + (3-3a)t - 1, 3t(1-t) \right)$$

1. Odvod  $\dot{\mathbf{b}}$  krivulje  $\mathbf{b}$  je podan z

$$\dot{\mathbf{b}}(t) = \left( 3(2-6a)t^2 + 2(-3+9a)t + 3-3a, -6t+3 \right).$$

Če vzamemo  $a = -1$ , je

$$\dot{\mathbf{b}}(t) = \left( 24 \left( t - \frac{1}{2} \right)^2, -6 \left( t - \frac{1}{2} \right) \right)$$

in pri  $t = 1/2$  dobimo  $\dot{\mathbf{b}}(t) = (0, 0)$ . Tangentni vektor krivulje  $\mathbf{b}$  pri parametru  $t$  je enotski vektor, podan z

$$\frac{\dot{\mathbf{b}}(t)}{\|\dot{\mathbf{b}}(t)\|} = \frac{\dot{\mathbf{b}}(t)}{\sqrt{576 \left( t - \frac{1}{2} \right)^4 + 36 \left( t - \frac{1}{2} \right)^2}} = \frac{\dot{\mathbf{b}}(t)}{6 \left| t - \frac{1}{2} \right| \sqrt{16 \left( t - \frac{1}{2} \right)^2 + 1}}.$$

Torej je

$$\lim_{t \nearrow \frac{1}{2}} \frac{\dot{\mathbf{b}}(t)}{\|\dot{\mathbf{b}}(t)\|} = (0, 1), \quad \lim_{t \searrow \frac{1}{2}} \frac{\dot{\mathbf{b}}(t)}{\|\dot{\mathbf{b}}(t)\|} = (0, -1),$$

kar pomeni, da tangentni vektor krivulje  $\mathbf{b}$  pri parametru  $t = 1/2$  obrne smer in zato ni zvezen.

2. Pri  $a = -5/3$  se krivulja poenostavi v

$$\mathbf{b}(t) = (12t^3 - 18t^2 + 8t - 1, 3t(1 - t))$$

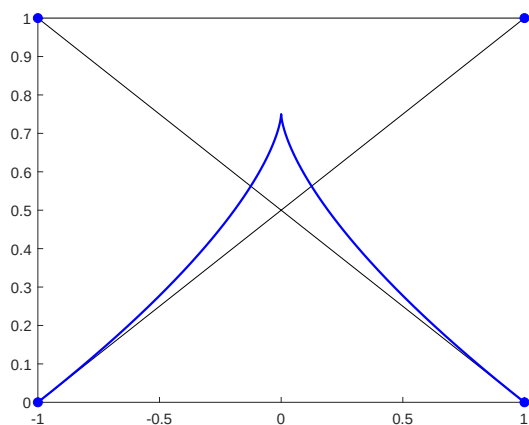
Če vstavimo  $t_{1,2} = 1/2 \mp \alpha$  in zahtevamo  $\mathbf{b}(t_1) = \mathbf{b}(t_2)$ , v prvi komponenti dobimo pogoj

$$12 \left( \left( \frac{1}{2} - \alpha \right)^3 - \left( \frac{1}{2} + \alpha \right)^3 \right) - 18 \left( \left( \frac{1}{2} - \alpha \right)^2 - \left( \frac{1}{2} + \alpha \right)^2 \right) + 8 \left( \left( \frac{1}{2} - \alpha \right) - \left( \frac{1}{2} + \alpha \right) \right) = 0,$$

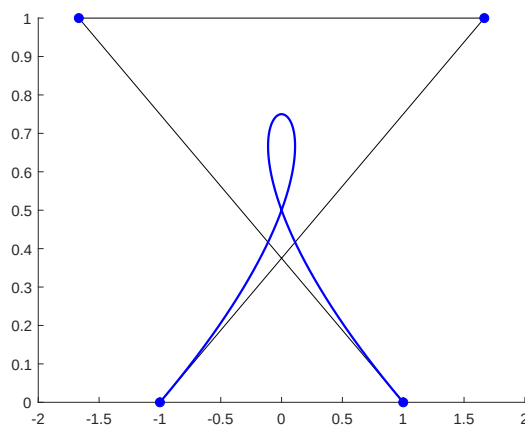
ki se poenostavi v  $\alpha(1 - 12\alpha^2) = 0$ . Nadalje opazimo, da v drugi komponenti dobimo  $-3\alpha^2 + 3/4$ , ne glede na to, ali vstavimo  $t_1$  ali  $t_2$ . Torej sta za  $\alpha = 1/\sqrt{12}$  parametra  $t_1$  in  $t_2$  različna, točki  $\mathbf{b}(t_1)$  in  $\mathbf{b}(t_2)$  pa ustrezata isti točki  $(0, 1/2)$ .

3. Pri  $a = 0$  v prvi komponenti parametrizacije krivulje  $\mathbf{b}$  nastopa kubični polinom, zato je krivulja stopnje 3, čeprav ima le 3 različne kontrolne točke. Če vzamemo  $a = 1/3$ , pa imamo v obeh komponentah parametrizacije kvadratična polinoma, zato je krivulja stopnje 2, čeprav je opisana s štirimi različnimi točkami.

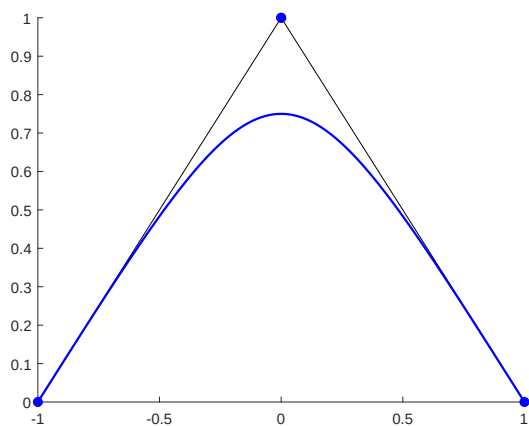
Krivulje iz zgornjih primerov so prikazane na sliki 2.1. □



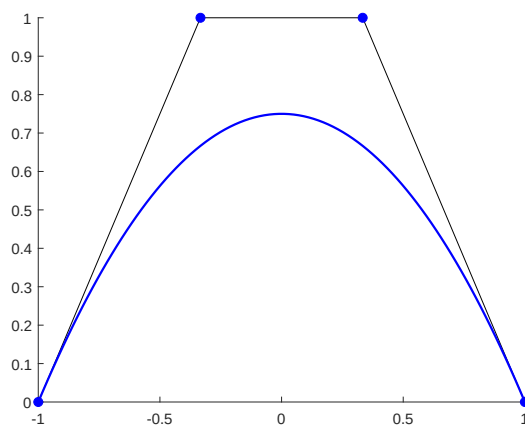
(a) Krivulja z nezvezno enotsko tangento.



(b) Krivulja s samopresečiščem.



(c) Krivulja s podvojeno kontrolno točko.



(d) Krivulja stopnje 2.

Slika 2.1: Primeri Bézierjevih krivulj iz naloge 2.1.

### 2.1.2 De Casteljaujev postopek

Stabilen postopek za izračun točke  $\mathbf{b}(t)$  na Bézierjevi krivulji  $\mathbf{b}$  stopnje  $n$  pri parametru  $t \in [0, 1]$  temelji na rekurzivni zvezi, ki ji zadoščajo Bernsteinovi bazni polinomi.

**Trditev 2.1.** *Za Bernsteinov bazni polinom  $B_i^n$ ,  $i \in \{0, 1, \dots, n\}$ , stopnje  $n \in \mathbb{N}$  velja*

$$B_i^n(t) = (1-t)B_i^{n-1}(t) + tB_{i-1}^{n-1}(t).$$

*Pri tem privzemamo, da je  $B_{-1}^{n-1} = 0$  in  $B_n^{n-1} = 0$ .*

*Dokaz.* Če upoštevamo, da za binomski simbol velja

$$\binom{n}{i} = \binom{n-1}{i} + \binom{n-1}{i-1},$$

po definiciji Bernsteinovih baznih polinomov sledi

$$B_i^n(t) = \binom{n-1}{i} t^i (1-t)^{n-i} + \binom{n-1}{i-1} t^{i-1} (1-t)^{n-i+1} = (1-t)B_i^{n-1}(t) + tB_{i-1}^{n-1}(t),$$

kar dokazuje trditev. □

De Casteljaujev postopek, s katerim računamo točko  $\mathbf{b}(t)$ , je postopek v  $n$  korakih. Začne se z zaporedjem kontrolnih točk  $\mathbf{b}_i$ ,  $i = 0, 1, \dots, n$ , nato pa se v vsakem koraku izračuna novo zaporedje točk, ki je za eno točko krajše od zaporedja na prejšnjem koraku. Tako na koncu ostane le ena točka, ki ustreza točki  $\mathbf{b}(t)$ .

Vpeljimo oznako  $\mathbf{b}_i^{(0)}(t) = \mathbf{b}_i$ ,  $i = 0, 1, \dots, n$ , za točke, s katerimi se de Casteljaujev postopek prične. Vmesne točke  $\mathbf{b}_i^{(r)}(t)$ , ki jih izračunamo v koraku  $r$ ,  $r \in \{1, 2, \dots, n\}$ , definiramo kot

$$\mathbf{b}_i^{(r)}(t) = (1-t)\mathbf{b}_i^{(r-1)}(t) + t\mathbf{b}_{i+1}^{(r-1)}(t), \quad i = 0, 1, \dots, n-r.$$

Ob zaključku postopka pri  $r = n$  dobimo točko  $\mathbf{b}_0^{(n)}(t)$ , ki po spodnjem izreku ustreza točki  $\mathbf{b}(t)$  na Bézierjevi krivulji.

**Izrek 2.1.** *Za vmesne točke de Casteljaujevega postopka velja*

$$\mathbf{b}_i^{(r)}(t) = \sum_{j=0}^r \mathbf{b}_{i+j} B_j^r(t).$$

*Dokaz.* Izrek dokažimo z indukcijo. Pri  $r = 0$  je Bernsteinov polinom  $B_0^0$  identično enak 1, zato je  $\mathbf{b}_i^{(0)}(t) = \mathbf{b}_i$ ,  $i = 0, 1, \dots, n$ , kar se ujema z definicijo točk  $\mathbf{b}_i^{(0)}$ . Predpostavimo, da izražava velja za  $r-1$  in dokažimo, da tedaj velja tudi za  $r$ . Po definiciji vmesnih točk v postopku je

$$\mathbf{b}_i^{(r)}(t) = (1-t)\mathbf{b}_i^{(r-1)}(t) + t\mathbf{b}_{i+1}^{(r-1)}(t) = (1-t) \sum_{j=0}^{r-1} \mathbf{b}_{i+j} B_j^{r-1}(t) + t \sum_{j=0}^{r-1} \mathbf{b}_{i+1+j} B_j^{r-1}(t).$$

Ob privzetku  $B_{-1}^{r-1} = B_r^{r-1} = 0$  lahko pišemo

$$\mathbf{b}_i^{(r)}(t) = (1-t) \sum_{j=0}^r \mathbf{b}_{i+j} B_j^{r-1}(t) + t \sum_{j=0}^r \mathbf{b}_{i+j} B_{j-1}^{r-1}(t) = \sum_{j=0}^r \mathbf{b}_{i+j} \left( (1-t) B_j^{r-1}(t) + t B_{j-1}^{r-1}(t) \right)$$

in po trditvi 2.1 ta preureditev zaključuje dokaz.  $\square$

Vsaka vmesna točka v de Casteljaujevem postopku je za  $t \in [0, 1]$  izračunana kot konveksna kombinacija dveh vmesnih točk iz prejšnjega koraka postopka. To daje izračunu točke na krivulji lepo geometrijsko interpretacijo, poleg tega pa zagotavlja stabilen izračun, saj pri konveksnih kombinacijah ne more priti do velikih numeričnih napak.

**Naloga 2.2.** Naj bo  $\mathbf{b}$  Bézierjeva krivulja stopnje 3 iz naloge 2.1 pri parametru  $a = 1/2$ . Z uporabo de Casteljaujevega postopka izračunajte točko  $\mathbf{b}(1/3)$ .

*Rešitev.* V prvem koraku postopka izračunamo točke

$$\begin{aligned} \mathbf{b}_0^{(1)}\left(\frac{1}{3}\right) &= \left(1 - \frac{1}{3}\right) \mathbf{b}_0 + \frac{1}{3} \mathbf{b}_1 = \frac{2}{3}(-1, 0) + \frac{1}{3}\left(-\frac{1}{2}, 1\right) = \left(-\frac{5}{6}, \frac{1}{3}\right), \\ \mathbf{b}_1^{(1)}\left(\frac{1}{3}\right) &= \left(1 - \frac{1}{3}\right) \mathbf{b}_1 + \frac{1}{3} \mathbf{b}_2 = \frac{2}{3}\left(-\frac{1}{2}, 1\right) + \frac{1}{3}\left(\frac{1}{2}, 1\right) = \left(-\frac{1}{6}, 1\right), \\ \mathbf{b}_2^{(2)}\left(\frac{1}{3}\right) &= \left(1 - \frac{1}{3}\right) \mathbf{b}_2 + \frac{1}{3} \mathbf{b}_3 = \frac{2}{3}\left(\frac{1}{2}, 1\right) + \frac{1}{3}(1, 0) = \left(\frac{2}{3}, \frac{2}{3}\right). \end{aligned}$$

V drugem koraku dobimo

$$\begin{aligned} \mathbf{b}_0^{(2)}\left(\frac{1}{3}\right) &= \left(1 - \frac{1}{3}\right) \mathbf{b}_0^{(1)}\left(\frac{1}{3}\right) + \frac{1}{3} \mathbf{b}_1^{(1)}\left(\frac{1}{3}\right) = \frac{2}{3}\left(-\frac{5}{6}, \frac{1}{3}\right) + \frac{1}{3}\left(-\frac{1}{6}, 1\right) = \left(-\frac{11}{18}, \frac{5}{9}\right), \\ \mathbf{b}_1^{(2)}\left(\frac{1}{3}\right) &= \left(1 - \frac{1}{3}\right) \mathbf{b}_1^{(1)}\left(\frac{1}{3}\right) + \frac{1}{3} \mathbf{b}_2^{(1)}\left(\frac{1}{3}\right) = \frac{2}{3}\left(-\frac{1}{6}, 1\right) + \frac{1}{3}\left(\frac{2}{3}, \frac{2}{3}\right) = \left(\frac{1}{9}, \frac{8}{9}\right), \end{aligned}$$

v tretjem koraku pa še

$$\mathbf{b}_0^{(3)}\left(\frac{1}{3}\right) = \left(1 - \frac{1}{3}\right) \mathbf{b}_0^{(2)}\left(\frac{1}{3}\right) + \frac{1}{3} \mathbf{b}_1^{(2)}\left(\frac{1}{3}\right) = \frac{2}{3}\left(-\frac{11}{18}, \frac{5}{9}\right) + \frac{1}{3}\left(\frac{1}{9}, \frac{8}{9}\right) = \left(-\frac{10}{27}, \frac{2}{3}\right).$$

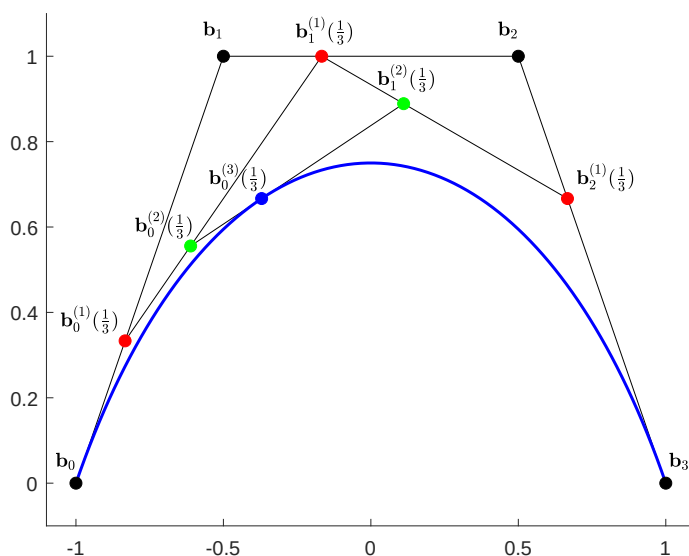
Torej je  $\mathbf{b}(1/3) = (-10/27, 2/3)$ . Postopek je povzet v shemi

|                                |                                          |                                            |                                            |
|--------------------------------|------------------------------------------|--------------------------------------------|--------------------------------------------|
| $\mathbf{b}_0^{(0)}$           | $\mathbf{b}_1^{(0)}$                     | $\mathbf{b}_2^{(0)}$                       | $\mathbf{b}_3^{(0)}$                       |
| $(-1, 0)$                      |                                          |                                            |                                            |
|                                | $\left(-\frac{5}{6}, \frac{1}{3}\right)$ |                                            |                                            |
| $\left(-\frac{1}{2}, 1\right)$ |                                          | $\left(-\frac{11}{18}, \frac{5}{9}\right)$ |                                            |
|                                | $\left(-\frac{1}{6}, 1\right)$           |                                            | $\left(-\frac{10}{27}, \frac{2}{3}\right)$ |
| $\left(\frac{1}{2}, 1\right)$  |                                          | $\left(\frac{1}{9}, \frac{8}{9}\right)$    |                                            |
|                                | $\left(\frac{2}{3}, \frac{2}{3}\right)$  |                                            |                                            |
| $(1, 0)$                       |                                          |                                            |                                            |

in na sliki 2.2.  $\square$

**Naloga 2.3.** V Matlabu pripravite naslednje metode za izračun in izris Bézierjeve krivulje s pomočjo de Casteljaujevega postopka.

1. Metoda `decasteljau` naj izračuna celotno de Casteljaujevo shemo za dano zaporedje kontrolnih točk oziroma njihovih koordinat pri danem parametru  $t \in [0, 1]$ .



Slika 2.2: Geometrijski prikaz de Casteljaujevega postopka iz naloge 2.2.

- Metoda `bezier` naj s pomočjo metode `decasteljau` izračuna točke na Bézierjevi krivulji pri danem seznamu parametrov z intervala  $[0, 1]$ . Bézierjeva krivulja naj bo podana v obliki seznama kontrolnih točk.
- Metoda `plotbezier` naj na podlagi kontrolnih točk in seznama parametrov, ki določajo točke na pripadajoči Bézierjevi krivulji, s pomočjo metode `bezier` nariše krivuljo ter lomljenko, sestavljeno iz zaporednih kontrolnih točk, ki ji pravimo kontrolni poligon.

Nazadnje s pomočjo ukazov `input` in `ginput` pripravite še skripto, z zagonom katere lahko uporabnik določi stopnjo Bézierjeve krivulje in s klikom vnese ustrezno število kontrolnih točk, nato pa se na zaslonu izriše vnešena Bézierjeva krivulja.

### 2.1.3 Afina invariantnost

Pomembna lastnost Bézierjevih krivulj je invariantnost na afine transformacije. Afina transformacija je preslikava, ki vektorju  $\mathbf{x} \in \mathbb{R}^d$  priredi vektor  $\mathbf{A}\mathbf{x} + \mathbf{v} \in \mathbb{R}^d$ , kjer je  $\mathbf{A} \in \mathbb{R}^{d \times d}$  matrika, ki predstavlja linearno preslikavo, in  $\mathbf{v} \in \mathbb{R}^d$  vektor, ki določa translacijo. Za Bézierjeve krivulje velja

$$\mathbf{A} \left( \sum_{i=0}^n \mathbf{b}_i B_i^n(t) \right) + \mathbf{v} = \sum_{i=0}^n (\mathbf{A}\mathbf{b}_i + \mathbf{v}) B_i^n(t),$$

kar pomeni, da je za afino transformacijo krivulje dovolj transformirati njene kontrolne točke. To je dobrodošla lastnost, saj nam pri transformaciji krivulje v računalniški grafiki ni treba preslikati vseh točk, ki opisujejo krivuljo.

**Naloga 2.4.** Naj bo  $\mathbf{b}$  Bézierjeva krivulja stopnje 3 iz naloge 2.1 pri parametru  $a = 2$ . Zapišite Bézierjeve krivulje, ki jih dobite, če na krivulji  $\mathbf{b}$  po vrsti izvedete naslednje transformacije.

- Zrcalite krivuljo čez  $x$  os.

2. Zarotirajte krivuljo za kot  $\pi/2$  v nasprotni smeri urinega kazalca.
3. Premaknite krivuljo za vektor  $(2, -3)$ .

Narišite dobljene krivulje v Matlabu in zapišite predpis za afino transformacijo, ki je kompozicija zgornjih transformacij.

*Rešitev.*

1. Zrcaljenje čez  $x$  os opišemo z afino transformacijo  $\Phi_1(\mathbf{x}) = \mathbf{A}_1\mathbf{x} + \mathbf{v}_1$  za matriko

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

in vektor  $\mathbf{v}_1 = \mathbf{0}$ . Kontrole točke  $\mathbf{c}_i = \Phi_1(\mathbf{b}_i)$ ,  $i = 0, 1, 2, 3$ , zrcaljenje krivulje so

$$\mathbf{c}_0 = (-1, 0), \quad \mathbf{c}_1 = (-2, -1), \quad \mathbf{c}_2 = (2, -1), \quad \mathbf{c}_3 = (1, 0).$$

2. Rotacijo krivulje za kot  $\pi/2$  opišemo z afino transformacijo  $\Phi_2(\mathbf{x}) = \mathbf{A}_2\mathbf{x} + \mathbf{v}_2$  za matriko

$$\mathbf{A}_2 = \begin{bmatrix} \cos(\frac{\pi}{2}) & -\sin(\frac{\pi}{2}) \\ \sin(\frac{\pi}{2}) & \cos(\frac{\pi}{2}) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

in vektor  $\mathbf{v}_2 = \mathbf{0}$ . Kontrole točke  $\mathbf{d}_i = \Phi_2(\mathbf{c}_i)$ ,  $i = 0, 1, 2, 3$ , rotirane krivulje so

$$\mathbf{d}_0 = (0, -1), \quad \mathbf{d}_1 = (1, -2), \quad \mathbf{d}_2 = (1, 2), \quad \mathbf{d}_3 = (0, 1).$$

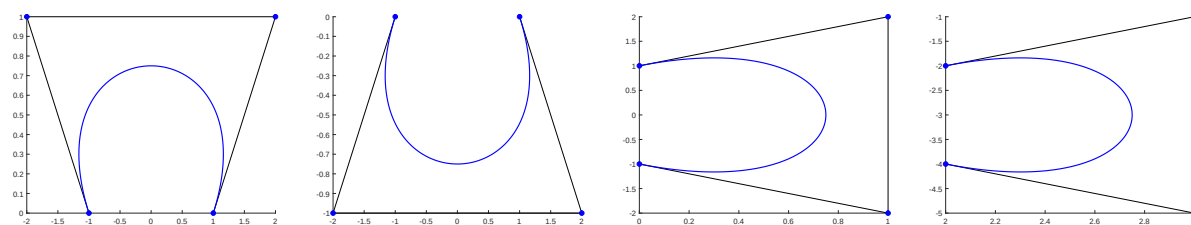
3. Premik krivulje opišemo z afino transformacijo  $\Phi_3(\mathbf{x}) = \mathbf{A}_3\mathbf{x} + \mathbf{v}_3$  za matriko  $\mathbf{A}_3 = \mathbf{I}$  in vektor  $\mathbf{v}_3 = (2, -3)$ . Kontrole točke  $\mathbf{e}_i = \Phi_3(\mathbf{d}_i)$ ,  $i = 0, 1, 2, 3$ , so

$$\mathbf{e}_0 = (2, -4), \quad \mathbf{e}_1 = (3, -5), \quad \mathbf{e}_2 = (3, -1), \quad \mathbf{e}_3 = (2, -2).$$

Kompozicija opisanih transformacij  $\Phi = \Phi_3 \circ \Phi_2 \circ \Phi_1$  je podana s predpisom

$$\Phi(\mathbf{x}) = \mathbf{A}_2\mathbf{A}_1\mathbf{x} + \mathbf{v}_3 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 2 \\ -3 \end{bmatrix}.$$

Začetna krivulja in njene transformacije so prikazane na sliki 2.3 po vrsti od leve proti desni.  $\square$



Slika 2.3: Primeri Bézierjevih krivulj iz naloge 2.1.

### 2.1.4 Višanje stopnje Bézierjeve krivulje

Naj bo  $\mathbf{b}$  Bézierjeva krivulja stopnje  $n$  s kontrolnimi točkami  $\mathbf{b}_i$ ,  $i = 0, 1, \dots, n$ . Ker so vse koordinatne funkcije, ki določajo parametrizacijo  $\mathbf{b}$ , polinomi stopnje  $n$ , lahko te polinome izrazimo tudi s pomočjo polinomov stopnje  $n+1$ . To pomeni, da lahko krivuljo  $\mathbf{b}$  za neke točke  $\mathbf{c}_i$ ,  $i = 0, 1, \dots, n+1$ , predstavimo kot Bézierjevo krivuljo

$$\mathbf{b}(t) = \sum_{i=0}^{n+1} \mathbf{c}_i B_i^{n+1}(t) \quad (2.1)$$

stopnje  $n+1$ .

**Trditev 2.2.** Naj bo  $\mathbf{b}$  Bézierjeva krivulja s kontrolnimi točkami  $\mathbf{b}_i$ ,  $i = 0, 1, \dots, n$ . Krivuljo  $\mathbf{b}$  lahko predstavimo v obliki (2.1) s kontrolnimi točkami

$$\mathbf{c}_0 = \mathbf{b}_0, \quad \mathbf{c}_i = \frac{i}{n+1} \mathbf{b}_{i-1} + \left(1 - \frac{i}{n+1}\right) \mathbf{b}_i, \quad i = 1, 2, \dots, n, \quad \mathbf{c}_{n+1} = \mathbf{b}_n. \quad (2.2)$$

*Dokaz.* Po definiciji Bernsteinovih baznih polinomov mora veljati

$$\sum_{i=0}^n \mathbf{b}_i \binom{n}{i} t^i (1-t)^{n-i} = \sum_{i=0}^{n+1} \mathbf{c}_i \binom{n+1}{i} t^i (1-t)^{n+1-i}.$$

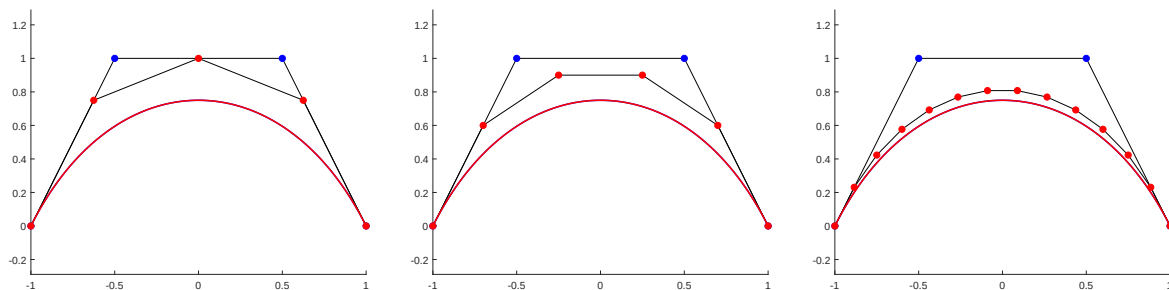
Če levo stran enačbe pomnožimo z  $1 = (1-t) + t$ , dobimo

$$\sum_{i=0}^n \mathbf{b}_i \binom{n}{i} \left( t^i (1-t)^{n+1-i} + t^{i+1} (1-t)^{n-i} \right).$$

Sedaj s primerjavo koeficientov pri funkcijah  $t \mapsto t^i (1-t)^{n-i+1}$ ,  $i = 0, 1, \dots, n+1$ , ki so linearno nedvisne, dobimo zveze

$$\mathbf{c}_i \binom{n+1}{i} = \mathbf{b}_i \binom{n}{i} + \mathbf{b}_{i-1} \binom{n}{i-1}, \quad i = 0, 1, \dots, n+1.$$

Te zveze se s korektnim upoštevanjem indeksov poenostavijo v (2.2). □



Slika 2.4: Višanje stopnje Bézierjeve krivulje iz naloge 2.5.

**Naloga 2.5.** V Matlabu sestavite metodo, ki zviša stopnjo Bézierjeve krivulje. Testirajte jo z Bézierjevo krivuljo iz naloge 2.2. Oglejte si, kaj se dogaja s kontrolnim polinomom krivulj, ki jih dobite z večkratnim zaporednim višanjem stopnje.

*Rešitev.* Slika 2.4 prikazuje Bézierjeve krivulje (v rdečem), ki jih dobimo tako, da originalni kubični krivulji (v modrem) enkrat, dvakrat in desetkrat zvišamo stopnjo. Razvidno je, da se kontrolni poligon z višanjem stopnje približuje krivulji.  $\square$

## 2.2 Zlepki Bézierjevih krivulj

Podobnega pristopa, kot je aproksimacija zapletenih funkcij z združevanjem polinomov v odsekoma polinomske funkcije, se lahko poslužujemo tudi pri parametrično podanih krivuljah, s čimer pridobimo orodje za opis kompleksnejših oblik. Lepljenje Bézierjevih krivulj je mogoče opisati s priročnimi geometrijskimi pogoji.

### 2.2.1 Odvodi Bézierjeve krivulje

Za zagotovitev vizualno ali parametrično gladkega prehoda med posameznimi polinomskimi krivuljami, ki so povezane v zlepek, si je treba najprej ogledati odvode Bézierjeve krivulje v robovih parametrizacije na intervalu  $[0, 1]$ .

Prvi korak k odvajanju Bézierjeve krivulje je izpeljava formule za odvod Bernsteinovega polinoma.

**Trditev 2.3.** Za Bernsteinov bazni polinom  $B_i^n$ ,  $i \in \{0, 1, \dots, n\}$ , stopnje  $n \in \mathbb{N}$  velja

$$\frac{d}{dt} B_i^n(t) = n \left( B_{i-1}^{n-1}(t) - B_i^{n-1}(t) \right),$$

kjer je  $B_{-1}^{n-1} = B_n^{n-1} = 0$ .

*Dokaz.* Po definiciji Bernsteinovega baznega polinoma velja

$$\frac{d}{dt} B_i^n(t) = \frac{d}{dt} \binom{n}{i} t^i (1-t)^{n-i} = \binom{n}{i} \left( i t^{i-1} (1-t)^{n-i} - (n-i) t^i (1-t)^{n-1-i} \right).$$

Binomski simbol lahko predstavimo v oblikah

$$\binom{n}{i} = \frac{n}{i} \frac{(n-1)!}{(i-1)!(n-i)!} = \frac{n}{i} \binom{n-1}{i-1}, \quad \binom{n}{i} = \frac{n}{n-i} \frac{(n-1)!}{i!(n-1-i)!} = \frac{n}{n-i} \binom{n-1}{i},$$

zato lahko odvod Bernsteinovega baznega polinoma zapišemo kot

$$\frac{d}{dt} B_i^n(t) = n \left( \binom{n-1}{i-1} t^{i-1} (1-t)^{(n-1)-(i-1)} - \binom{n-1}{i} t^i (1-t)^{n-1-i} \right).$$

Formula za odvod  $B_i^n$  tako sledi po definiciji  $B_{i-1}^{n-1}$  in  $B_i^{n-1}$ .  $\square$

Oglejmo si odvod Bézierjeve krivulje  $\mathbf{b}$  stopnje  $n$  s kontrolnimi točkami  $\mathbf{b}_i$ ,  $i = 0, 1, \dots, n$ . Po trditvi 2.3 je

$$\frac{d}{dt} \mathbf{b}(t) = \sum_{i=0}^n \mathbf{b}_i \frac{d}{dt} B_i^n(t) = n \sum_{i=0}^n \mathbf{b}_i \left( B_{i-1}^{n-1}(t) - B_i^{n-1}(t) \right),$$

kar lahko lepše zapišemo kot

$$\frac{d}{dt}\mathbf{b}(t) = n \left( \sum_{i=1}^n \mathbf{b}_i B_{i-1}^{n-1}(t) - \sum_{i=0}^{n-1} \mathbf{b}_i B_i^{n-1}(t) \right) = n \sum_{i=0}^{n-1} (\mathbf{b}_{i+1} - \mathbf{b}_i) B_i^{n-1}(t).$$

To pravzaprav pomeni, da smo izpeljali formulo za odvod Bézierjeve krivulje, ki je podan kot Bézierjeva krivulja stopnje  $n - 1$  s kontrolnimi točkami  $\Delta\mathbf{b}_i = \mathbf{b}_{i+1} - \mathbf{b}_i$ ,  $i = 0, 1, \dots, n - 1$ , pomnoženimi z  $n$ :

$$\frac{d}{dt}\mathbf{b}(t) = n \sum_{i=0}^{n-1} \Delta\mathbf{b}_i B_i^{n-1}(t). \quad (2.3)$$

Z indukcijo lahko na podoben način izrazimo tudi višje odvode.

**Izrek 2.2.** *Naj bo  $\mathbf{b}$  Bézierjeva krivulja s kontrolnimi točkami  $\mathbf{b}_i$ ,  $i = 0, 1, \dots, n$ . Odvod krivulje  $\mathbf{b}$  reda  $r \leq n$  je podan s predpisom*

$$\frac{d^r}{dt^r}\mathbf{b}(t) = \frac{n!}{(n-r)!} \sum_{i=0}^{n-r} \Delta^r \mathbf{b}_i B_i^{n-r}(t), \quad (2.4)$$

kjer je  $\Delta^r \mathbf{b}_i = \Delta^{r-1} \mathbf{b}_{i+1} - \Delta^{r-1} \mathbf{b}_i$  rekurzivna posplošitev difference  $\Delta^1 \mathbf{b}_i = \Delta \mathbf{b}_i = \mathbf{b}_{i+1} - \mathbf{b}_i$ .

**Naloga 2.6.** Določite kontrolne točke  $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4$  ravninske Bézierjeve krivulje  $\mathbf{b}$  stopnje štiri tako, da njeni tangenti pri parametrih  $t = 0$  in  $t = 1$  ustrezata vektorjema  $(4, \pm 4)$  in da ima krivulja pri parametru  $t = 1/2$  špico, s katero se dotakne kontrolne točke  $\mathbf{b}_2 = (0, 0)$ .

*Rešitev.* Z uporabo de Casteljaujevega postopka lahko izračunamo, da je

$$\mathbf{b}(\tfrac{1}{2}) = \tfrac{1}{16}\mathbf{b}_0 + \tfrac{1}{4}\mathbf{b}_1 + \tfrac{3}{8}\mathbf{b}_2 + \tfrac{1}{4}\mathbf{b}_3 + \tfrac{1}{16}\mathbf{b}_4 = (0, 0).$$

Izraz se še dodatno poenostavi, saj velja  $\mathbf{b}_2 = (0, 0)$ . Iz podatkov o tangentah v robnih točkah nadalje sledi  $4(\mathbf{b}_1 - \mathbf{b}_0) = (4, 4)$  in  $4(\mathbf{b}_4 - \mathbf{b}_3) = (4, -4)$ ; z upoštevanjem tega lahko zgornjo enačbo prepišemo v  $\mathbf{b}_1 + \mathbf{b}_3 = (0, 2/5)$ . Po drugi strani mora veljati

$$\dot{\mathbf{b}}(\tfrac{1}{2}) = \tfrac{1}{8}(\mathbf{b}_1 - \mathbf{b}_0) + \tfrac{3}{8}(\mathbf{b}_2 - \mathbf{b}_1) + \tfrac{3}{8}(\mathbf{b}_3 - \mathbf{b}_2) + \tfrac{1}{8}(\mathbf{b}_4 - \mathbf{b}_3) = (0, 0),$$

kar se hitro poenostavi v  $\mathbf{b}_3 - \mathbf{b}_1 = (-2/3, 0)$ . Na podlagi tega sledi, da so

$$\mathbf{b}_0 = (-\tfrac{2}{3}, -\tfrac{4}{5}), \quad \mathbf{b}_1 = (\tfrac{1}{3}, \tfrac{1}{5}), \quad \mathbf{b}_2 = (0, 0), \quad \mathbf{b}_3 = (-\tfrac{1}{3}, \tfrac{1}{5}), \quad \mathbf{b}_4 = (\tfrac{2}{3}, -\tfrac{4}{5})$$

kontrolne točke krivulje  $\mathbf{b}$ , ki je prikazana na sliki 2.5. □

**Naloga 2.7.** Zapišite prvi in drugi odvod krivulje iz naloge 2.6 ter ju skupaj s krivuljo narišite v Matlabu.

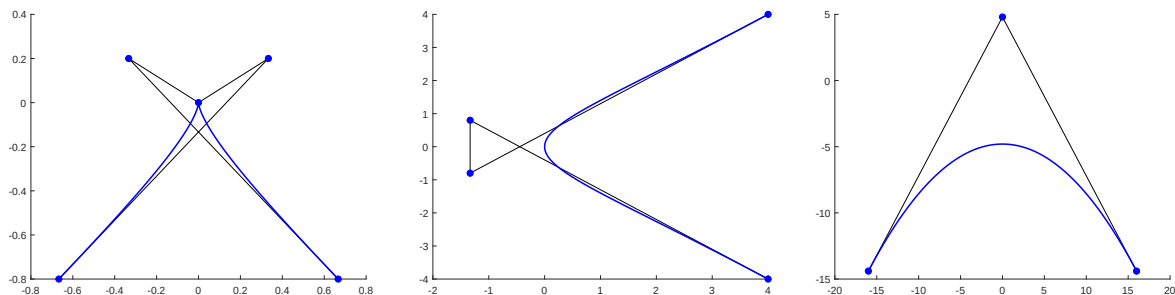
*Rešitev.* Z uporabo formule (2.3) izračunamo, da prvi odvod ustreza Bézierjevi krivulji

$$\dot{\mathbf{b}}(t) = (4, 4)B_0^3(t) + (-\tfrac{4}{3}, -\tfrac{4}{5})B_1^3(t) + (-\tfrac{4}{3}, \tfrac{4}{5})B_2^3(t) + (4, -4)B_3^3(t),$$

z uporabo formule (2.4) pa še, da drugi odvod ustreza Bézierjevi krivulji

$$\ddot{\mathbf{b}}(t) = (-16, -\tfrac{72}{5})B_0^2(t) + (0, \tfrac{24}{5})B_1^2(t) + (16, -\tfrac{72}{5})B_2^2(t).$$

Odvoda  $\dot{\mathbf{b}}$  in  $\ddot{\mathbf{b}}$ , ki ju lahko narišemo s pomočjo metod iz naloge 2.3, sta prikazana na sliki 2.5. □



Slika 2.5: Bézierjeva krivulja iz naloge 2.6 (levo) ter njen prvi in drugi odvod (v sredi in desno).

### 2.2.2 Zlepek dveh Bézierjevih krivulj

V nadaljevanju bi radi dve Bézierjevi krivulji povezali v eno. Ker želimo zlepek podati z enotno parametrizacijo, posamezna Bézierjeva krivulja pa je parametrizirana na intervalu  $[0, 1]$ , moramo poseči po reparametrizaciji.

Denimo, da za interval parametrizacije zleпка  $\mathbf{z}$ , ki je sestavljen iz Bézierjevih krivulj  $\mathbf{b}$  in  $\mathbf{c}$ , izberemo  $[u_0, u_2]$  za  $u_0 < u_2$  in da stik krivulj  $\mathbf{b}$  in  $\mathbf{c}$  predvidimo v točki  $u_1$ , ki leži med  $u_0$  in  $u_2$ . Krivuljo  $\mathbf{b}$  želimo torej parametrizirati na intervalu  $[u_0, u_1]$ , krivuljo  $\mathbf{c}$  pa na intervalu  $[u_1, u_2]$ , kar pomeni, da potrebujemo funkciji  $\varphi_1 : [u_0, u_1] \rightarrow [0, 1]$  in  $\varphi_2 : [u_1, u_2] \rightarrow [0, 1]$ , ki dana intervala preslikata na standardno definicijsko območje Bézierjeve krivulje. Uporabimo lahko kar linearni funkciji

$$\varphi_1(u) = \frac{u - u_0}{u_1 - u_0}, \quad \varphi_2(u) = \frac{u - u_1}{u_2 - u_1},$$

določeni z lastnostmi  $\varphi_1(u_0) = \varphi_2(u_1) = 0$  in  $\varphi_1(u_1) = \varphi_2(u_2) = 1$ . Zlepek  $\mathbf{z}$  lahko potem podamo s parametrizacijo

$$\mathbf{z}(u) = \begin{cases} \mathbf{b}(\varphi_1(u)), & u \in [u_0, u_1] \\ \mathbf{c}(\varphi_2(u)), & u \in [u_1, u_2] \end{cases} \quad (2.5)$$

na intervalu  $[u_0, u_2]$ .

Naj bosta Bézierjevi krivulji  $\mathbf{b}$  in  $\mathbf{c}$  enake stopnje  $n \in \mathbb{N}$  in naj bodo kontrolne točke teh dveh krivulj označene z  $\mathbf{b}_i$  in  $\mathbf{c}_i$ ,  $i = 0, 1, \dots, n$ . Če želimo, da je zlepek  $\mathbf{z}$  zvezen, se morata reparametrizirani Bézierjevi krivulji  $\mathbf{b}$  in  $\mathbf{c}$  pri parametru  $u = u_1$  stakniti. Zahtevati moramo, da je

$$\mathbf{b}(1) = \mathbf{b}(\varphi_1(u_1)) = \mathbf{c}(\varphi_2(u_1)) = \mathbf{c}(0)$$

oziroma

$$\mathbf{b}_n = \sum_{i=0}^n \mathbf{b}_i B_i^n(1) = \sum_{i=0}^n \mathbf{c}_i B_i^n(0) = \mathbf{c}_0.$$

Z drugimi besedami, zlepek  $\mathbf{z}$  je zvezen, če se kontrolni točki  $\mathbf{b}_n$  in  $\mathbf{c}_0$  ujemata, kar je geometrijski pogoj, neodvisen od parametrizacije.

Da bi dosegli bolj gladek prehod v stiku Bézierjevih krivulj, moramo obravnavati še odvedljivost zlepek. Zlepek  $\mathbf{z}$  je zvezno odvedljiv pri parametru  $u = u_1$  natanko tedaj, ko je

$$\frac{d}{du}\mathbf{b}(\varphi_1(u))|_{u=u_1} = \frac{d}{du}\mathbf{c}(\varphi_2(u))|_{u=u_1}.$$

Po pravilu za verižno odvajanje to pomeni

$$\frac{d}{dt}\mathbf{b}(t)|_{t=\varphi_1(u_1)} \cdot \frac{d}{du}\varphi_1(u)|_{u=u_1} = \frac{d}{dt}\mathbf{c}(t)|_{t=\varphi_2(u_1)} \cdot \frac{d}{du}\varphi_2(u)|_{u=u_1},$$

kar se ob upoštevanju definicije  $\varphi_1$  in  $\varphi_2$  ter formule (2.3) za odvod Bézierjevih krivulj  $\mathbf{b}$  in  $\mathbf{c}$  prepiše v

$$n \sum_{i=0}^{n-1} \Delta \mathbf{b}_i B_i^{n-1}(1) \cdot \frac{1}{u_1 - u_0} = n \sum_{i=0}^{n-1} \Delta \mathbf{c}_i B_i^{n-1}(0) \cdot \frac{1}{u_2 - u_1}$$

in poenostavi v

$$\frac{1}{u_1 - u_0} \Delta \mathbf{b}_{n-1} = \frac{1}{u_2 - u_1} \Delta \mathbf{c}_0.$$

Na podoben način lahko s pomočjo izreka 2.2 izpeljemo tudi pogoje za višji red gladkosti.

**Izrek 2.3.** *Naj bo  $\mathbf{z}$  zlepek Bézierjevih krivulj  $\mathbf{b}$  in  $\mathbf{c}$  stopenj  $n \in \mathbb{N}$ , ki je podan s predpisom (2.5). Zlepek  $\mathbf{z}$  je  $r$ -krat zvezno odvedljiv natanko tedaj, ko je*

$$\left(\frac{1}{\Delta u_0}\right)^j \Delta^j \mathbf{b}_{n-j} = \left(\frac{1}{\Delta u_1}\right)^j \Delta^j \mathbf{c}_0, \quad j = 0, 1, \dots, r,$$

kjer je  $\Delta u_i = u_{i+1} - u_i$ ,  $i = 0, 1$ .

Iz izreka 2.3 je razvidno, da na gladkost zlepek (pri  $r > 0$ ) poleg lege kontrolnih točk Bézierjevih krivulj vpliva tudi izbran interval parametrizacije.

**Naloga 2.8.** Naj bosta  $\mathbf{b}$  in  $\mathbf{c}$  Bézierjevi krivulji stopnje 3 s kontrolnimi točkami  $\mathbf{b}_i$  in  $\mathbf{c}_i$ ,  $i = 0, 1, 2, 3$ ; del njih je podanih:

$$\begin{array}{lll} \mathbf{b}_0 = (-3, 0), & \mathbf{b}_1 = (-2, 1), & \mathbf{b}_2 = (-1, 1), \\ \mathbf{c}_1 = (1, -1), & \mathbf{c}_2 = (2, -3), & \mathbf{c}_3 = (-1, -1). \end{array}$$

1. Zlepek Bézierjevih krivulj  $\mathbf{z}_1 : [-1, 1] \rightarrow \mathbb{R}^2$  s stičnimi točkami  $u_0 = -1$ ,  $u_1 = 0$ ,  $u_2 = 1$  je definiran s predpisom

$$\mathbf{z}_1(u) = \begin{cases} \mathbf{b}(u+1); & u \in [-1, 0) \\ \mathbf{c}(u); & u \in [0, 1] \end{cases}.$$

Določite  $\mathbf{b}_3$  in  $\mathbf{c}_0$  tako, da bo  $\mathbf{z}_1$  zvezno odvedljiv.

2. Zlepek Bézierjevih krivulj  $\mathbf{z}_2 : [-1, 1] \rightarrow \mathbb{R}^2$  s stičnimi točkami  $u_0 = -1$ ,  $u_1 = 1/2$ ,  $u_2 = 1$  je definiran s predpisom

$$\mathbf{z}_2(u) = \begin{cases} \mathbf{b}(\frac{2}{3}(u+1)); & u \in [-1, \frac{1}{2}) \\ \mathbf{c}(2u-1); & u \in [\frac{1}{2}, 1] \end{cases}.$$

Kakšni morata biti v tem primeru kontrolni točki  $\mathbf{b}_3$  in  $\mathbf{c}_0$ , da bo  $\mathbf{z}_2$  zvezno odvedljiv?

3. Ali je zlepek  $\mathbf{z}_1$  dvakrat zvezno odvedljiv? Kaj pa  $\mathbf{z}_2$ ?

*Rešitev.*

1. Pogoji za zveznost zlepka  $\mathbf{z}_1$  je  $\mathbf{b}_3 = \mathbf{b}(1) = \mathbf{c}(0) = \mathbf{c}_0$ , pogoj za zvezno odvedljivost pa

$$3(\mathbf{b}_3 - \mathbf{b}_2) = \dot{\mathbf{b}}(1) = \dot{\mathbf{c}}(0) = 3(\mathbf{c}_1 - \mathbf{c}_0),$$

od kjer sledi  $\mathbf{b}_3 = \mathbf{c}_0 = (0, 0)$ . Zlepek  $\mathbf{z}_1$  je prikazan na levi strani slike 2.6.

2. Pogoji za zveznost zlepka  $\mathbf{z}_2$  je tudi v tem primeru  $\mathbf{b}_3 = \mathbf{b}(1) = \mathbf{c}(0) = \mathbf{c}_0$ , medtem ko se pogoj za zvezno odvedljivost zaradi drugačne parametrizacije spremeni v

$$3 \cdot \frac{2}{3}(\mathbf{b}_3 - \mathbf{b}_2) = \dot{\mathbf{b}}(1) = \dot{\mathbf{c}}(0) = 3 \cdot 2(\mathbf{c}_1 - \mathbf{c}_0),$$

kar pomeni, da je  $\mathbf{b}_3 = \mathbf{c}_0 = (1/2, -1/2)$ . Zlepek  $\mathbf{z}_2$  je prikazan v sredini slike 2.6.

3. Oglejmo si druga odvoda zlepkov  $\mathbf{z}_1$  in  $\mathbf{z}_2$  pri stičnem parametru. Za  $\mathbf{z}_1$  pri  $u = 0$  po eni strani dobimo

$$\ddot{\mathbf{b}}(1) = 6(\mathbf{b}_3 - 2\mathbf{b}_2 + \mathbf{b}_1) = (0, -6),$$

po drugi pa

$$\ddot{\mathbf{c}}(0) = 6(\mathbf{c}_2 - 2\mathbf{c}_1 + \mathbf{c}_0) = (0, -6),$$

na podlagi česar sklepamo, da je  $\mathbf{z}_1$  dvakrat zvezno odvedljiv. Za  $\mathbf{z}_2$  pa pri  $u = 1/2$  dobimo

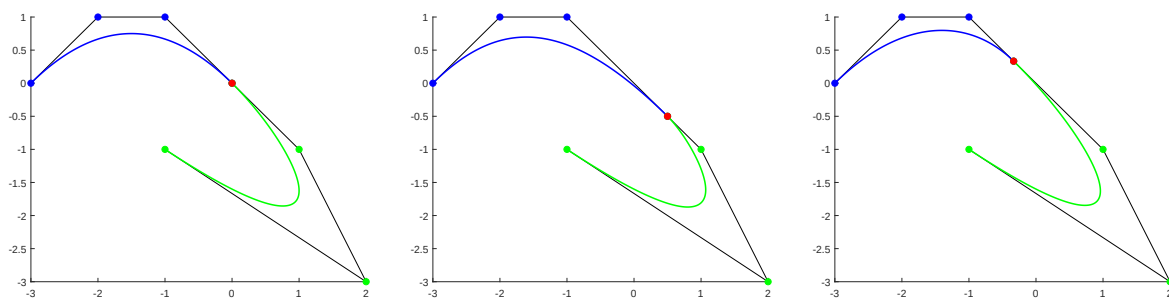
$$\ddot{\mathbf{b}}(1) = 6 \cdot \left(\frac{2}{3}\right)^2 (\mathbf{b}_3 - 2\mathbf{b}_2 + \mathbf{b}_1) = \left(\frac{4}{3}, -4\right)$$

in

$$\ddot{\mathbf{c}}(0) = 6 \cdot 2^2 (\mathbf{c}_2 - 2\mathbf{c}_1 + \mathbf{c}_0) = (12, -36),$$

kar kaže na to, da se druga odvoda v stiku ne ujemata in  $\mathbf{z}_2$  ni dvakrat zvezno odvedljiv.

□



Slika 2.6: Zlepki dveh Bézierjevih krivulj iz nalog 2.8 in 2.9. Z modro je narisana krivulja  $\mathbf{b}$ , z zeleno krivulja  $\mathbf{c}$ , rdeča točka pa označuje kontrolno točko  $\mathbf{b}_3 = \mathbf{c}_0$ .

Kadar z zlepkom Bézierjevih krivulj načrtujemo gibanje telesa, je pomembno, da dosežemo zadostno gladkost parametrizacije. Če zlepek ni dvakrat zvezno odvedljiv, pride v stiku do nezvezne spremembe pospeška, če ni zvezno odvedljiv, pa še do nezvezne spremembe hitrosti.

Kadar pa zlepek uporabljamo zgolj kot orodje za oblikovanje, gladkost parametrizacije ni tako pomembna, saj lahko vizualno gladkost dosežemo že z zagotavljanjem ujemanja smeri tangente v stiku (pri zvezni odvedljivosti poleg enakosti smeri zahtevamo tudi enakost dolžine tangentnega vektorja). Temu konceptu pravimo geometrijska gladkost. Zlepek  $\mathbf{z}$ , definiran v (2.5), je geometrijsko odvedljiv natanko tedaj, ko je  $\mathbf{b}_n = \mathbf{c}_0$  in obstaja tako pozitivno realno število  $\alpha$ , da je

$$\Delta \mathbf{b}_{n-1} = \alpha \Delta \mathbf{c}_0. \quad (2.6)$$

Ta pogoj je povsem neodvisen od parametrizacije zleпка. Podobno velja tudi za geometrijsko gladkost višjega reda, a so pogoji bolj zapleteni. Ravninski zlepek je na primer dvakrat geometrijsko odvedljiv natanko tedaj, ko ima zvezno enotsko tangento in ukripljenost.

**Naloga 2.9.** V Matlabu narišite zlepek Bézierjevih krivulj  $\mathbf{b}$  in  $\mathbf{c}$  iz naloge 2.8 z ujemačima kontrolnima točkama  $\mathbf{b}_3 = \mathbf{c}_0$  in opazujte, kaj se dogaja, če točko pomikate vzdolž daljice med kontrolnima točkama  $\mathbf{b}_2$  in  $\mathbf{c}_1$ .

*Rešitev.* Za poljubno število  $\lambda \in (0, 1)$  točka  $\mathbf{b}_3 = \mathbf{c}_0 = (1 - \lambda)\mathbf{b}_2 + \lambda\mathbf{c}_1$  leži na daljici med  $\mathbf{b}_2$  in  $\mathbf{c}_1$  in točke  $\mathbf{b}_2, \mathbf{b}_3, \mathbf{c}_0, \mathbf{c}_1$  zadoščajo pogoju (2.6), zato je zlepek Bézierjevih krivulj  $\mathbf{b}$  in  $\mathbf{c}$  geometrijsko odvedljiv in posledično vizualno gladek. Na sliki 2.6 je desno prikazan zlepek pri parametru  $\lambda = 1/3$ .  $\square$

### 2.2.3 Zlepek kvadratičnih Bézierjevih krivulj

Naj bo  $[u_0, u_m]$  interval parametrizacije, ki ga razdelimo z notranjimi stičnimi točkami  $u_0 < u_1 < \dots < u_m$ , da bi na njem definirali zvezno odvedljiv zlepek  $\mathbf{z}$ , sestavljen iz  $m$  Bézierjevih krivulj stopnje 2. Pri konstrukciji zvezno odvedljive odsekoma kvadratične funkcije nad delitvijo z  $m + 1$  stičnimi točkami v razdelku 1.3.4 je bilo na voljo  $m + 2$  interpolacijskih podatkov, zato sklepamo, da lahko  $\mathbf{z}$  določimo na podlagi  $m + 2$  točk.

Naj bodo  $\mathbf{b}^{(i)}$ ,  $i = 1, 2, \dots, m$ , Bézierjeve krivulje stopnje 2, ki sestavljajo zlepek  $\mathbf{z}$ , in naj bodo  $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_{m+1}$  dane točke v prostoru  $\mathbb{R}^d$ , na podlagi katerih bi radi določili kontrolne točke Bézierjevih krivulj  $\mathbf{b}^{(i)}$ , ki jih označimo z  $\mathbf{b}_0^{(i)}, \mathbf{b}_1^{(i)}, \mathbf{b}_2^{(i)}$ . Predpišimo

$$\mathbf{b}_1^{(i)} = \mathbf{p}_i, \quad i = 1, 2, \dots, m. \quad (2.7a)$$

Po izreku 2.3 mora za zveznost zleпка veljati

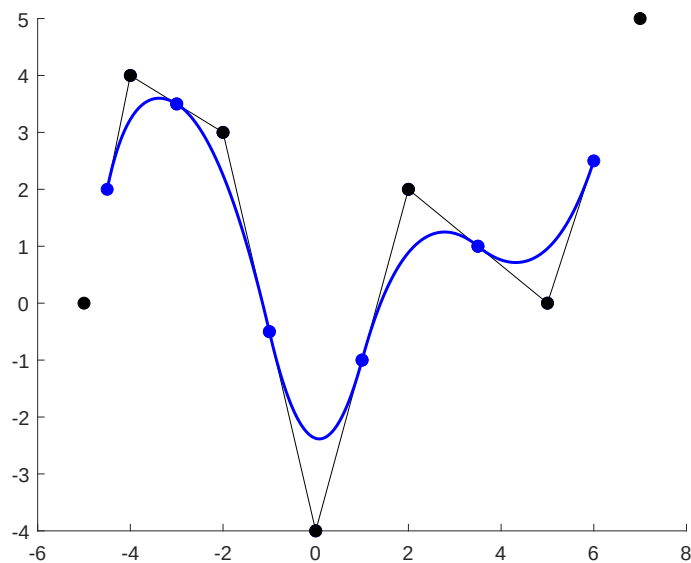
$$\mathbf{b}_2^{(i)} = \mathbf{b}_0^{(i+1)}, \quad i = 1, 2, \dots, m - 1,$$

za zvezno odvedljivost pa še

$$\frac{1}{\Delta u_{i-1}} \Delta \mathbf{b}_1^{(i)} = \frac{1}{\Delta u_i} \Delta \mathbf{b}_0^{(i+1)}, \quad i = 1, 2, \dots, m - 1.$$

Iz teh pogojev s krajšim računom izrazimo

$$\mathbf{b}_2^{(i)} = \mathbf{b}_0^{(i+1)} = \frac{\Delta u_i}{\Delta u_{i-1} + \Delta u_i} \mathbf{p}_i + \frac{\Delta u_{i-1}}{\Delta u_{i-1} + \Delta u_i} \mathbf{p}_{i+1}, \quad i = 1, 2, \dots, m - 1, \quad (2.7b)$$



Slika 2.7: Primer zvezno odvedljivega zlepk Bézierjevih krivulj stopnje 2. Črne točke ustrezajo izbranim točkam  $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_6$ , modre (in nekatere črne) točke pa določajo 5 Bézierjevih krivulj stopnje 2, ki sestavljajo zlepek.

kar kaže na to, da te kontrolne točke ležijo na daljicah med točkami  $\mathbf{p}_i$  in  $\mathbf{p}_{i+1}$ . Nedoločeni ostajata le še kontrolni točki  $\mathbf{b}_0^{(1)}$  in  $\mathbf{b}_2^{(m)}$ , za kateri vzamemo

$$\mathbf{b}_0^{(1)} = \frac{1}{2}\mathbf{p}_0 + \frac{1}{2}\mathbf{p}_1, \quad \mathbf{b}_2^{(m)} = \frac{1}{2}\mathbf{p}_m + \frac{1}{2}\mathbf{p}_{m+1}. \quad (2.7c)$$

Z izrazi v (2.7) je zlepek  $\mathbf{z}$ , ki je na intervalu  $[u_{i-1}, u_i]$ ,  $i \in \{1, 2, \dots, m\}$ , podan s predpisom

$$\mathbf{z}(u) = \mathbf{b}_0^{(i)} B_0^2 \left( \frac{u - u_{i-1}}{u_i - u_{i-1}} \right) + \mathbf{b}_1^{(i)} B_1^2 \left( \frac{u - u_{i-1}}{u_i - u_{i-1}} \right) + \mathbf{b}_2^{(i)} B_2^2 \left( \frac{u - u_{i-1}}{u_i - u_{i-1}} \right), \quad (2.8)$$

enolično določen z izbiro točk  $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_{m+1}$ . Primer zlepk za  $m = 5$  je prikazan na sliki 2.7.

**Naloga 2.10.** V Matlabu sestavite program, ki omogoča vnos točk  $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_{m+1}$  za izbran  $m \in \mathbb{N}$ . Na zaslon naj se izriše zvezno odvedljiv zlepek Bézierjevih krivulj stopnje 2, ki je določen s temi točkami in parametriziran na intervalu  $[0, m]$  s stičnimi točkami  $u_i = i$ ,  $i = 0, 1, \dots, m$ .

*Rešitev.* Pri pripravi programa si pomagamo z ukazoma `input` za pridobitev števila točk in `ginput` za pridobitev lokacije točk. Nato izračunamo kontrolne točke zlepk za vsak interval  $[u_{i-1}, u_i]$ ,  $i = 1, 2, \dots, m$ , posebej. Za izris Bézierjeve krivulje nam te ni treba reparametrizirati (saj reparametrizacija ne vpliva na obliko) in jo lahko narišemo kar s pomočjo funkcije `plotbezier` iz naloge 2.3 na intervalu  $[0, 1]$ .  $\square$

Zlepke Bézierjevih krivulj lahko generiramo tudi z iteracijskimi postopki, ki začetno zaporedje točk v vsakem koraku zgostijo, dokler ni zaporedje dovolj gosto, da predstavlja vizualno gladko predstavitev krivulje. Eden prvih znanih tovrstnih postopkov je Chaikinov postopek, ki iz zaporedja točk  $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_{m+1}$  generira zaporedja, za katera velja,

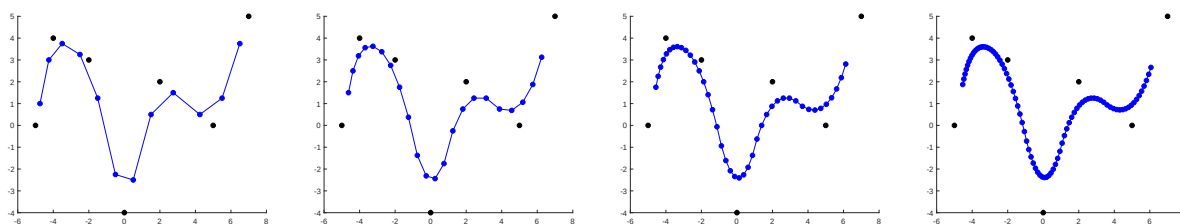
da konvergirajo k zvezno odvedljivemu kvadratičnemu zlepku (2.8), ki je podan s kontrolnimi točkami (2.7).

Chaikinov postopek temelji na izvajanju konveksnih kombinacij. Na začetku postavimo  $\mathbf{p}_i^{(0)} = \mathbf{p}_i$ ,  $i = 0, 1, \dots, m+1$ , nato pa v vsakem koraku  $k = 1, 2, \dots$  rekurzivno izračunamo novo zaporedje točk  $\mathbf{p}_i^{(k)}$ ,  $i = 0, 1, \dots, 2^k m + 1$  s formulama

$$\mathbf{p}_{2j}^{(k)} = \frac{3}{4}\mathbf{p}_j^{(k-1)} + \frac{1}{4}\mathbf{p}_{j+1}^{(k-1)}, \quad \mathbf{p}_{2j+1}^{(k)} = \frac{1}{4}\mathbf{p}_j^{(k-1)} + \frac{3}{4}\mathbf{p}_{j+1}^{(k-1)}, \quad j = 0, 1, \dots, 2^{k-1}m,$$

za sode in lihe indekse. Število točk se eksponentno povečuje, zato navadno za izris krivulje zadošča že štiri ali pet korakov postopka. Primer, generiran z metodo iz naslednje naloge, je prikazan na sliki 2.8.

**Naloga 2.11.** V Matlabu sestavite metodo, ki za dano naravno število  $n$  izvede  $n$  korakov Chaikinovega postopka na podanem zaporedju točk  $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_{m+1}$ . Implementirano metodo testirajte tako, da program iz naloge 2.10 dopolnite z izrisom lomljene črte, ki povezuje s Chaikinovim postopkom generirano zaporedje točk pri parametru  $n = 4$ .



Slika 2.8: Primer zvezno odvedljivega zlepka Bézierjevih krivulj stopnje 2 iz slike 2.7, ki je generiran s Chaikinovim postopkom. Črne točke ustrezajo začetnemu zaporedju točk, lomljena črta, sestavljena s povezovanjem modrih točk, pa ustreza generiranim zaporedjem po enem, dveh, treh in štirih korakih postopka.

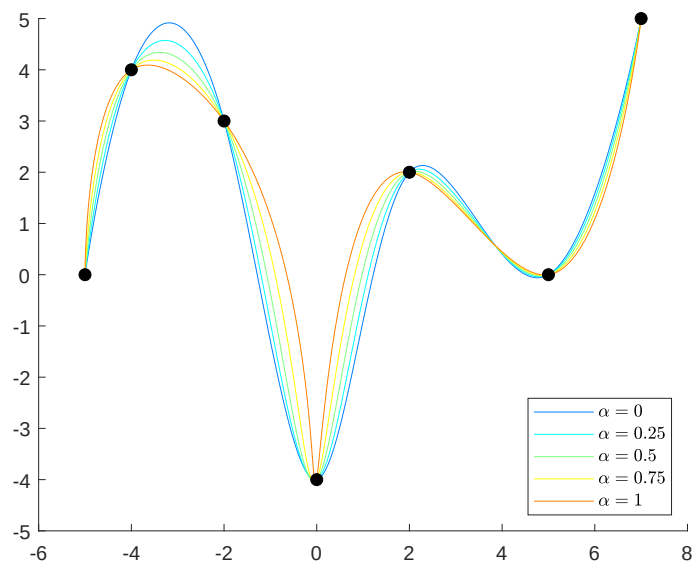
Postopki, podobni Chaikinovemu, so znani tudi za zlepke Bézierjevih krivulj višjih stopenj. Gre za postopke, ki spadajo v širši koncept teorije subdivizijskih shem. Ta se ukvarja s študijem lastnosti krivulj (predvsem gladkosti), ki jih lahko dobimo z zgoščanjem začetnega zaporedja točk. Pri splošnih postopkih tega tipa parametrizacije limitnih krivulj niso nujno polinomske in jih ne poznamo v zaključeni obliki.

### 2.2.4 Zlepek kubičnih Bézierjevih krivulj

V razdelku 1.3.5 smo ugotovili, da je odsekoma kubično funkcijo priročno opisati z interpolacijo vrednosti in odvodov v stičnih točkah definicijskega intervala. Zato je zlepek Bézierjevih krivulj stopnje 3 nad delitvijo intervala  $[u_0, u_m]$ ,  $m \in \mathbb{N}$ , s stičnimi točkami  $u_0, u_1, \dots, u_m$  smiselno izpeljati v odvisnosti od točk  $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_m$ , ki predpisujejo lego zlepka, in vektorjev  $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_m$ , ki predpisujejo smer tangentnega vektorja zlepka v stičnih točkah parametrizacije.

Natančneje, konstruirajmo zlepek  $\mathbf{z}$ , za katerega velja

$$\mathbf{z}(u_i) = \mathbf{p}_i, \quad \dot{\mathbf{z}}(u_i) = \mathbf{v}_i, \quad i = 0, 1, \dots, m.$$



Slika 2.9: Primeri dvakrat zvezno odvedljivih zlepkov Bézierjevih krivulj stopnje 3. Črne točke ustrezajo točkam  $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_6$ , ki določajo zlepke šestih Bézierjevih krivulj v različnih  $\alpha$ -parametrizacijah (naloga 2.12).

Na intervalu  $[u_{i-1}, u_i]$  ga zapišemo v obliki reparametrizirane Bézierjeve krivulje

$$\mathbf{z}(u) = \mathbf{b}_0^{(i)} B_0^3\left(\frac{u-u_{i-1}}{u_i-u_{i-1}}\right) + \mathbf{b}_1^{(i)} B_1^3\left(\frac{u-u_{i-1}}{u_i-u_{i-1}}\right) + \mathbf{b}_2^{(i)} B_2^3\left(\frac{u-u_{i-1}}{u_i-u_{i-1}}\right) + \mathbf{b}_3^{(i)} B_3^3\left(\frac{u-u_{i-1}}{u_i-u_{i-1}}\right)$$

s kontrolnimi točkami  $\mathbf{b}_0^{(i)}, \mathbf{b}_1^{(i)}, \mathbf{b}_2^{(i)}, \mathbf{b}_3^{(i)}$ . Iz formul za odvod Bézierjeve krivulje takoj sledi

$$\mathbf{b}_0^{(i)} = \mathbf{p}_{i-1}, \quad \mathbf{b}_1^{(i)} = \mathbf{p}_{i-1} + \frac{1}{3}\Delta u_{i-1}\mathbf{v}_{i-1}, \quad \mathbf{b}_2^{(i)} = \mathbf{p}_i - \frac{1}{3}\Delta u_{i-1}\mathbf{v}_i, \quad \mathbf{b}_3^{(i)} = \mathbf{p}_i.$$

Če torej podamo točke  $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_m$  in vektorje  $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_m$ , je zlepek enolično določen. Gre za zlepek, ki je po konstrukciji zvezno odvedljiv, višjim pogojem gladkosti pa v splošnem ne zadošča.

Če sprostimo izbiro vektorjev  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{m-1}$  in jih določimo s točkami  $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_m$ , lahko dosežemo, da je zlepek dvakrat zvezno odvedljiv. Izpeljava izrazov je podobna kot v funkcijskem primeru (razdelek 1.3.5) in vodi do sistema enačb

$$\Delta u_i \mathbf{v}_{i-1} + 2(\Delta u_{i-1} + \Delta u_i) \mathbf{v}_i + \Delta u_{i-1} \mathbf{v}_{i+1} = 3 \left( \frac{\Delta u_i}{\Delta u_{i-1}} \Delta \mathbf{p}_{i-1} + \frac{\Delta u_{i-1}}{\Delta u_i} \Delta \mathbf{p}_i \right)$$

za  $i = 1, 2, \dots, m-1$ . Sistem rešimo v vsaki komponenti prostora posebej in s tem določimo komponente vektorjev  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{m-1}$ . Vektorja  $\mathbf{v}_0$  in  $\mathbf{v}_m$  sta prosta in ju lahko izberemo na različne načine. Če so na voljo podatkih o tangentah v robnih točkah, jih izkoristimo, sicer pa lahko vzamemo  $\mathbf{v}_0 = \mathbf{v}_m = \mathbf{0}$ , kar ima za posledico, da je  $\mathbf{b}_0^{(1)} = \mathbf{b}_1^{(1)} = \mathbf{p}_0$  in  $\mathbf{b}_3^{(m)} = \mathbf{b}_2^{(m)} = \mathbf{p}_m$ .

**Naloga 2.12.** V Matlabu pripravite metodo, ki za dano zaporedje točk  $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_m$  izračuna kontrolne točke Bézierjevih krivulj, ki določajo dvakrat zvezno odvedljiv kubičen zlepek. Vektorja  $\mathbf{v}_0$  in  $\mathbf{v}_m$  naj bosta enaka  $\mathbf{0}$ . Metodo testirajte s programom, ki

uporabniku omogoči vnos interpolacijskih točk in izriše pripadajoči zlepek. Pri tem za delitev uporabite stične točke

$$u_0 = 0, \quad u_i = u_{i-1} + \|\Delta \mathbf{p}_{i-1}\|^\alpha, \quad i = 1, 2, \dots, m,$$

ki določajo tako imenovano  $\alpha$ -parametrizacijo zleпка. Eksperimentirajte z izbiro  $\alpha \in [0, 1]$  in opazujte, kako vrednost parametra vpliva na obliko zleпка.

*Rešitev.* Primeri zleпkov, dobljenih pri parametrih  $\alpha = i/4$ ,  $i = 0, 1, 2, 3, 4$ , so prikazani na sliki 2.9. Med običajne izbire sodijo  $\alpha = 0$  (enakomerna parametrizacija),  $\alpha = 1/2$  (centripetalna parametrizacija) in  $\alpha = 1$  (tetivna parametrizacija).  $\square$

## 2.2.5 Opis krožnega loka z zleпkom Bézierjevih krivulj

Ker s polinomskimi krivuljami ne moremo eksaktno opisati krožnice, je treba v sistemih, ki omogočajo le delo s polinomskimi krivuljami, poseči po aproksimaciji. Z Bézierjevo krivuljo nizke stopnje ni moč dobiti kvalitetne aproksimacije celotne krožnice, zato jo aproksimiramo po delih. Če pripravimo Bézierjevo krivuljo  $\mathbf{b} : [0, 1] \rightarrow \mathbb{R}^2$  nizke stopnje, ki dobro opisuje krožni lok za kote v območju  $(-\varphi, \varphi)$  za nek  $\varphi > 0$  ter v robovih interpolira krožnico, lahko celotno krožnico aproksimiramo z zleпkom Bézierjevih krivulj, dobljenim z rotacijami  $\mathbf{b}$ .

**Naloga 2.13.** Ugotovite, kakšnim pogojem morajo zadoščati kontrolne točke Bézierjeve krivulje  $\mathbf{b}$  stopnje  $n$ , da pri parametrih  $t = 0$  in  $t = 1$  interpolira krožnico ter smeri njenih tangent pri kotih  $-\varphi$  in  $\varphi$ .

1. Določite Bézierjevo krivuljo stopnje 2, ki porodi  $G^1$  interpolant krožnice.
2. Določite Bézierjevo krivuljo stopnje 3, ki porodi  $C^1$  interpolant krožnice.
3. Določite Bézierjevo krivuljo stopnje 3, ki porodi  $G^1$  interpolant krožnice in za katero velja  $\mathbf{b}(1/2) = (1, 0)$ .

V Matlabu narišite aproksimacije krožnic iz (a), (b) in (c) pri izbiri  $\varphi = \pi/6$ . Primerjajte napake  $\max_{t \in t} |1 - \|\mathbf{b}(t)\||$  za  $\mathbf{t} = (i/1000)_{i=0}^{1000}$ .

*Rešitev.* Naj bodo  $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_n$  kontrolne točke Bézierjeve krivulje  $\mathbf{b}$ . Iz interpolacije v robovih krožnega loka sledi

$$\mathbf{b}_0 = (\cos(\varphi), -\sin(\varphi)), \quad \mathbf{b}_n = (\cos(\varphi), \sin(\varphi)).$$

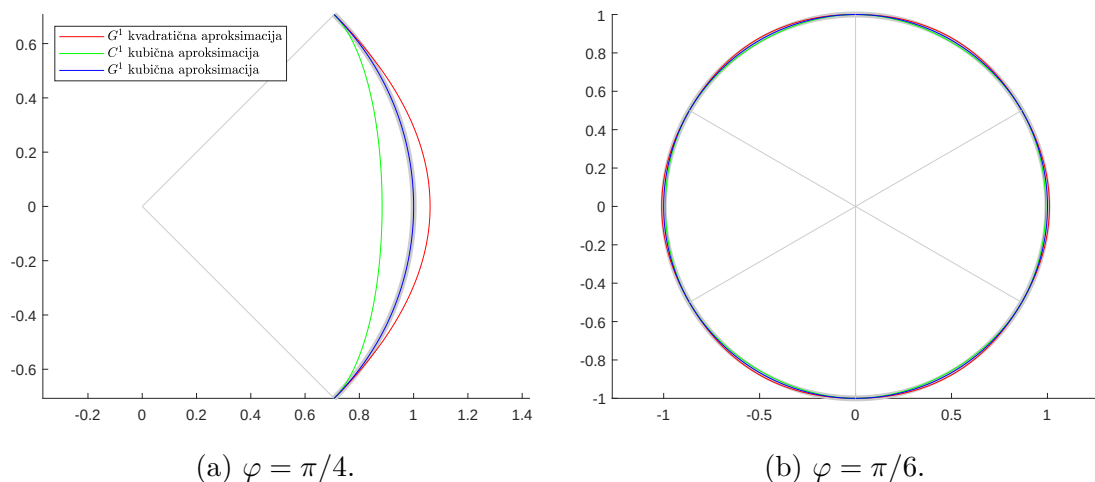
Smer tangente krivulje  $\mathbf{b}$  se v robovih ujema s krožnico, če za neki realni števili  $d_1$  in  $d_2$  velja  $\dot{\mathbf{b}}(0) = d_1(\sin(\varphi), \cos(\varphi))$  in  $\dot{\mathbf{b}}(1) = d_2(-\sin(\varphi), \cos(\varphi))$ , od kjer po formuli za odvod Bézierjeve krivulje sledi

$$\mathbf{b}_1 = \mathbf{b}_0 + \frac{1}{n}\dot{\mathbf{b}}(0) = \left(\cos(\varphi) + \frac{1}{n}d_1\sin(\varphi), -\sin(\varphi) + \frac{1}{n}d_1\cos(\varphi)\right),$$

$$\mathbf{b}_{n-1} = \mathbf{b}_n - \frac{1}{n}\dot{\mathbf{b}}(1) = \left(\cos(\varphi) + \frac{1}{n}d_2\sin(\varphi), \sin(\varphi) - \frac{1}{n}d_2\cos(\varphi)\right).$$

1. Bézierjeva krivulja stopnje 2 ima samo tri kontrolne točke:  $\mathbf{b}_0$ ,  $\mathbf{b}_1$  in  $\mathbf{b}_2$ . Iz zgornjih ugotovitev sledi, da sta  $\mathbf{b}_0$  in  $\mathbf{b}_2$  že določeni, za  $\mathbf{b}_1$  pa dobimo pogoja

$$\begin{aligned} \cos(\varphi) + \frac{1}{2}d_1\sin(\varphi) &= \cos(\varphi) + \frac{1}{2}d_2\sin(\varphi), \\ -\sin(\varphi) + \frac{1}{2}d_1\cos(\varphi) &= \sin(\varphi) - \frac{1}{2}d_2\cos(\varphi). \end{aligned}$$



Slika 2.10: Primeri aproksimacij krožnega loka (oziroma krožnice) iz naloge 2.13.

Iz prve enačbe sledi  $d_1 = d_2$ , iz druge pa  $d_1 = d_2 = 2 \tan(\varphi)$ . To pomeni, da je  $\mathbf{b}_0 = (1/\cos(\varphi), 0)$ .

2. Če za Bézierjevo krivuljo stopnje 3 zahtevamo, da je  $C^1$  interpolant krožnega loka, mora biti  $d_1 = d_2 = 1$ . S tem so določene vse štiri kontrolne točke  $\mathbf{b}_0$ ,  $\mathbf{b}_1$ ,  $\mathbf{b}_2$  in  $\mathbf{b}_3$ .
3. Oglejmo si Bézierjevo krivuljo stopnje 3, ki porodi  $G^1$  interpolant in za katero velja  $\mathbf{b}(1/2) = (1, 0)$ . Dodaten pogoj implicira

$$\frac{1}{8}\mathbf{b}_0 + \frac{3}{8}\mathbf{b}_1 + \frac{3}{8}\mathbf{b}_2 + \frac{1}{8}\mathbf{b}_3 = (1, 0).$$

V drugi komponenti zgornje enačbe dobimo

$$-\frac{1}{8}\sin(\varphi) + \frac{3}{8}\left(-\sin(\varphi) + \frac{1}{3}d_1\cos(\varphi)\right) + \frac{3}{8}\left(\sin(\varphi) - \frac{1}{3}d_2\cos(\varphi)\right) + \frac{1}{8}\sin(\varphi) = 0,$$

od koder je razvidno, da mora biti  $d = d_1 = d_2$ . Iz prve komponente enačbe, ki jo zapišemo v obliki

$$\frac{1}{8}\cos(\varphi) + \frac{3}{8}\left(\cos(\varphi) + \frac{1}{3}d_1\sin(\varphi)\right) + \frac{3}{8}\left(\cos(\varphi) + \frac{1}{3}d_2\sin(\varphi)\right) + \frac{1}{8}\cos(\varphi) = 1,$$

potem sledi

$$\cos(\varphi) + \frac{1}{4}d\sin(\varphi) = 1,$$

kar pomeni, da je  $d = 4(1/\sin(\varphi) - \cot(\varphi))$ . Z nekaj dodatnega računanja izluščimo, da sta kontrolni točki  $\mathbf{b}_1$  in  $\mathbf{b}_2$  dani z

$$\begin{aligned}\mathbf{b}_1 &= \frac{1}{3}\left(4 - \cos(\varphi), 4\cot(\varphi) - \frac{4}{\sin(\varphi)} + \sin(\varphi)\right), \\ \mathbf{b}_2 &= \frac{1}{3}\left(4 - \cos(\varphi), -4\cot(\varphi) + \frac{4}{\sin(\varphi)} - \sin(\varphi)\right).\end{aligned}$$

Primeri aproksimacije krožnega loka z izpeljanimi konstrukcijami so predstavljeni na sliki 2.10. Na levi je prikazana aproksimacija krožnega loka pri kotu  $\varphi = \pi/4$ , od

kjer je razvidno da najboljšo aproksimacijo nudi konstrukcija iz tretje točke. Napaka je približno  $2.73 \cdot 10^{-4}$ . Napaka konstrukcij iz prve in druge točke sta po vrsti približno enaki  $6.07 \cdot 10^{-2}$  in  $1.16 \cdot 10^{-1}$ . Pri  $\varphi = \pi/6$  je napaka pri vseh konstrukcijah vsaj navidezno zelo majhna, zato lahko z uporabo petih rotacij opišemo celotno krožnico, kot je prikazano na desni strani slike. Tudi pri tem kotu je najboljša konstrukcija iz tretje točke (napaka približno  $2.39 \cdot 10^{-5}$ ), konstrukcija iz druge točke (napaka je približno  $8.98 \cdot 10^{-3}$ ) pa je v tem primeru boljša od konstrukcije iz prve točke (napaka je približno  $1.04 \cdot 10^{-2}$ ).  $\square$

## Poglavje 3

# Numerično odvajanje in integriranje

Odvajanje ali integriranje funkcije ni vedno preprost problem, saj je funkcija lahko zapletena ali pa je celo ne poznamo v zaključeni obliki. Odvod in integral lahko približno izračunamo z numeričnimi pravili, ki temeljijo na funkcijskih vrednostih. Ta pravila pogosto izpeljemo tako, da namesto funkcije odvajamo oziroma integriramo njeno aproksimacijo.

### 3.1 Pravila za odvajanje

Osnovna pravila za numerično odvajanje niso zanimiva le z vidika aproksimacij odvodov funkcije, pomembno vlogo igrajo tudi pri reševanju diferencialnih enačb. Njihova izpeljava je tesno povezana s polinomske interpolacije.

#### 3.1.1 Diferenčne formule za prve odvode

Naj bo  $f$  odvedljiva funkcija v okolici točke  $x_0$ . Oglejmo si, kako lahko približek za njen odvod v točki  $x_0$  poiščemo z odvajanjem funkciji  $f$  prirejenega interpolacijskega polinoma.

Naj bo  $h$  pozitivno realno število in  $p_1 \in \mathbb{P}_1$  polinom, ki interpolira vrednosti funkcije  $f$  v točkah  $x_0$  in  $x_0 + h$ . Zapišemo ga lahko v obliki

$$p_1(x) = [x_0]f + [x_0, x_0 + h]f \cdot (x - x_0) = f(x_0) + \frac{f(x_0 + h) - f(x_0)}{h}(x - x_0). \quad (3.1)$$

Če polinom  $p_1$  odvajamo in odvod izvednotimo v  $x_0$ , dobimo

$$p'_1(x_0) = \frac{f(x_0 + h) - f(x_0)}{h}. \quad (3.2)$$

Ker je v okolici  $x_0$  polinom  $p_1$  aproksimacija za funkcijo  $f$ , pričakujemo, da se vrednost  $p'_1(x_0)$  približno ujema z vrednostjo  $f'(x_0)$ . To predvidevanje natančneje utemeljuje naslednja trditev.

**Trditev 3.1.** *Naj bo  $f$  dvakrat zvezno odvedljiva funkcija na intervalu  $[x_0, x_0 + h]$ ,  $h > 0$ . Potem obstaja število  $\xi \in (x_0, x_0 + h)$ , da velja*

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0)}{h} - \frac{1}{2}hf''(\xi). \quad (3.3)$$

*Dokaz.* Formulo (3.3) lahko izpeljemo na podlagi zveze med funkcijo  $f$  in interpolacijskim polinom  $p_1$ , ki je podan s predpisom (3.1). Za vsak  $x \in [x_0, x_0 + h]$  obstaja tako število  $\xi_x \in (x_0, x_0 + h)$ , odvisno od  $x$ , da velja

$$f(x) = p_1(x) + \frac{1}{2!}f''(\xi_x)(x - x_0)(x - x_0 - h).$$

Z odvajanjem te zveze dobimo

$$f'(x) = p_1'(x) + \frac{1}{2}f''(\xi_x)((x - x_0 - h) + (x - x_0)) + \frac{1}{2}\left(\frac{d}{dx}f''(\xi_x)\right)(x - x_0)(x - x_0 - h)$$

in če vstavimo  $x = x_0$ , se formula poenostavi v

$$f'(x_0) = p_1'(x_0) - \frac{1}{2}hf''(\xi_{x_0}),$$

kar po (3.2) dokazuje (3.3). □

Ulomek v formuli (3.3) ustreza enemu izmed klasičnih pravil za numerično odvajanje, ki se imenuje prema diferenca. Členi, ki vsebuje drugi odvod funkcije  $f$ , pravimo ostanek pravila. Na podoben način, z interpolacijo funkcije v točkah  $x_0 - h$  in  $x_0$ , lahko izpeljemo formulo, ki jo podaja sledeča trditev in določa numerično pravilo, imenovano obratna diferenca.

**Trditev 3.2.** *Naj bo  $f$  dvakrat zvezno odvedljiva funkcija na intervalu  $[x_0 - h, x_0]$ ,  $h > 0$ . Potem obstaja število  $\xi \in (x_0 - h, x_0)$ , da velja*

$$f'(x_0) = \frac{f(x_0) - f(x_0 - h)}{h} + \frac{1}{2}hf''(\xi). \quad (3.4)$$

Iz formul (3.3) in (3.4) je razvidno, da sta izpeljani pravili točni, kadar v vlogi funkcije  $f$  nastopa polinom iz  $\mathbb{P}_1$ , ne pa če v vlogi  $f$  nastopa polinom višje stopnje. Red napake pravil je  $\mathcal{O}(h)$ , kar pomeni, da napaka, ko zmanjšujemo korak  $h$ , pada relativno počasi.

Poiščimo boljšo aproksimacijo za odvod funkcije  $f$  v točki  $x_0$  s pomočjo polinoma  $p_2 \in \mathbb{P}_2$ , ki interpolira vrednosti funkcije  $f$  v točkah  $x_0 - h$ ,  $x_0$  in  $x_0 + h$ . Tak polinom lahko zapišemo v obliki

$$\begin{aligned} p_2(x) &= [x_0 - h]f + [x_0 - h, x_0]f \cdot (x - x_0 + h) \\ &\quad + [x_0 - h, x_0, x_0 + h]f \cdot (x - x_0 + h)(x - x_0). \end{aligned} \quad (3.5)$$

Z odvajanjem dobimo

$$p_2'(x) = [x_0 - h, x_0]f + [x_0 - h, x_0, x_0 + h]f \cdot ((x - x_0) + (x - x_0 + h)).$$

Če odvod izvednotimo v točki  $x = x_0$  in upoštevamo definicijo deljenih diferenc, dobimo približek za  $f'(x_0)$ , ki je podan z

$$p'_2(x_0) = \frac{f(x_0) - f(x_0 - h)}{h} + \frac{\frac{f(x_0+h)-f(x_0)}{h} - \frac{f(x_0)-f(x_0-h)}{h}}{2h}h$$

in se z nekaj dodatnega računanja poenostavi v

$$p'_2(x_0) = \frac{f(x_0 + h) - f(x_0 - h)}{2h}. \quad (3.6)$$

Nekoliko presenetljivo v izrazu (3.6) ne nastopa vrednost  $f(x_0)$ , kar pomeni, da lahko približek za  $f'(x_0)$  izračunamo samo na podlagi dveh funkcijskih vrednosti, kar je enako kot pri premi in obratni diferenci. Ker pa smo formulo dobili z odvajanjem kvadratičnega polinoma, je pričakovati, da bo pravilo natančnejše.

**Trditev 3.3.** *Naj bo  $f$  trikrat zvezno odvedljiva funkcija na intervalu  $[x_0 - h, x_0 + h]$ ,  $h > 0$ . Potem obstaja število  $\xi \in (x_0 - h, x_0 + h)$ , da velja*

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0 - h)}{2h} - \frac{1}{6}h^2 f'''(\xi). \quad (3.7)$$

*Dokaz.* Trditev se dokaže na podoben način kot trditev 3.1. Za vsak  $x \in [x_0 - h, x_0 + h]$  obstaja število  $\xi_x \in (x_0 - h, x_0 + h)$ , da velja

$$f(x) = p_2(x) + \frac{1}{3!}f'''(\xi_x)(x - x_0 + h)(x - x_0)(x - x_0 - h), \quad (3.8)$$

kjer je  $p_2$  polinom, ki je podan s predpisom (3.5). Ko zgornji izraz odvajamo in vstavimo  $x = x_0$ , dobimo

$$f'(x_0) = p'_2(x_0) - \frac{1}{6}h^2 f'''(\xi_{x_0}).$$

Od tod iz (3.6) sledi (3.7). □

Pravilo, ki ga predstavlja ulomek v formuli (3.7), se imenuje simetrična diferenca. Iz formule je razvidno, da je ostanek pravila reda  $\mathcal{O}(h^2)$ . Pravilo je točno za vse polinome iz  $\mathbb{P}_2$ .

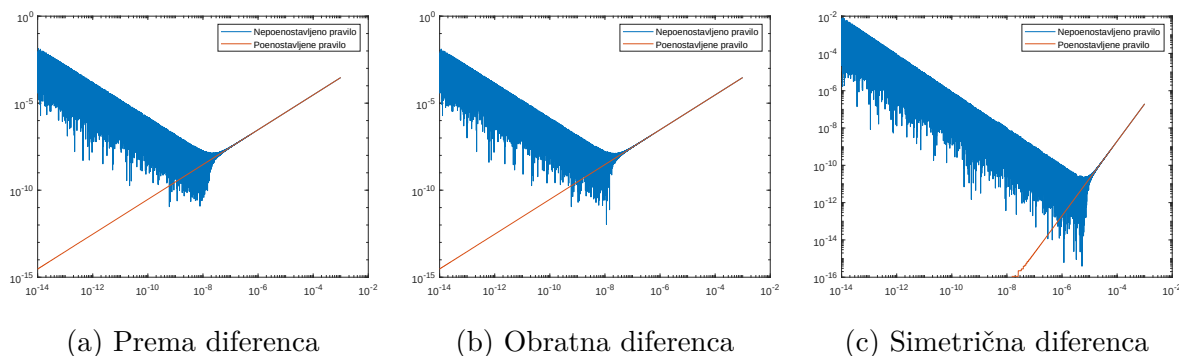
**Naloga 3.1.** Funkcija  $f$  je podana s predpisom  $f(x) = 1/(x + 1)$ . Izračunajte premo, obratno in simetrično diferenco za funkcijo  $f$  v točki  $x_0 = 1/2$  za poljuben  $h \in (0, 1)$  ter rezultate primerjajte z vrednostjo  $f'(x_0)$  v limiti, ko gre  $h$  proti 0. Nato v Matlabu implementirajte izračun diferenc za poljubno funkcijo in izračunajte vrednosti diferenc za funkcijo  $f$  pri majhnih vrednostih  $h$ . Kaj opazite? Poskušajte pojasniti rezultate.

*Rešitev.* V pravila (3.3), (3.4) in (3.7) po vrsti vstavimo predpis za funkcijo  $f$  in izraze poenostavimo v

$$\begin{aligned} f'\left(\frac{1}{2}\right) &\approx \frac{1}{h} \left( \frac{1}{\frac{1}{2} + h + 1} - \frac{1}{\frac{1}{2} + 1} \right) = -\frac{4}{9 + 6h}, \\ f'\left(\frac{1}{2}\right) &\approx \frac{1}{h} \left( \frac{1}{\frac{1}{2} + 1} - \frac{1}{\frac{1}{2} + h + 1} \right) = -\frac{4}{9 - 6h}, \\ f'\left(\frac{1}{2}\right) &\approx \frac{1}{2h} \left( \frac{1}{\frac{1}{2} + h + 1} - \frac{1}{\frac{1}{2} - h + 1} \right) = -\frac{4}{9 - 4h^2}. \end{aligned}$$

Odvod funkcije  $f$  je podan s predpisom  $f'(x) = -1/(x+1)^2$ , torej je  $f'(1/2) = -4/9$ . Slednje je tudi vrednost limite vseh zgornjih izrazov, ko pošljemo  $h$  proti 0.

V Matlabu se lahko prepričamo, da kadar približke za odvod funkcije  $f$  računamo na podlagi poenostavljenih formul, napaka pada, ko gre  $h$  proti 0. Pri direktni uporabi numeričnega pravila pa napaka z zmanjševanjem  $h$  na neki stopnji začne naraščati. Krive so napake pri numeričnem računanju vrednosti funkcij, ki jih v formuli množimo z  $1/h$ , kar je veliko število, ko je  $h$  majhen. Obnašanje napake je prikazano na sliki 3.1. Razvidno je, da pri simetrični diferenci napaka pada hitreje kot pri premi in obratni diferenci, vendar je pravilo tudi bolj občutljivo na numerične napake. Pri uporabi pravil za numerično odvajanje je torej treba biti zelo previden.  $\square$



Slika 3.1: Prikaz napake pri računanju preme, obratne in simetrične difference iz naloge 3.1 pri uporabi nepoenostavljenega pravila (v modri barvi) in poenostavljenega pravila (v rdeči barvi). Grafi so narisani v logaritemski skali.

Še natančnejša pravila za numerično odvajanje lahko dobimo tako, da približek določimo na podlagi večjega števila funkcijskih vrednosti. Tovrstna pravila lahko podobno kot premo, obratno in simetrično diferenco izpeljemo z odvajanjem interpolacijskih polinomov višjih stopenj, drug pristop pa ponuja metoda nedoločenih koeficientov. Pri tej metodi pravilo zastavimo kot uteženo kombinacijo funkcijskih vrednosti v različnih točkah, nato pa uteži oziroma koeficiente določimo z zahtevo, da je pravilo točno za polinome čim višjih stopenj. Opisana je v naslednji nalogi.

**Naloga 3.2.** Za dano funkcijo  $f$  bi radi izpeljali formulo za odvod funkcije  $f$  v točki  $x_0$  oblike

$$F(x_0) = Af(x_0 - 2h) + Bf(x_0 - h) + Cf(x_0) + Df(x_0 + h) + Ef(x_0 + 2h),$$

kjer je  $h > 0$  izbran korak.

1. Določite konstante  $A, B, C, D$  in  $E$  tako, da bo formula točna za vse polinome stopnje manjše ali enake 4. Pomagajte si s sistemom enačb, ki ga dobite, če za funkcijo  $f$  vzamete funkcije  $x \mapsto (x - x_0)^i$ ,  $i = 0, 1, 2, 3, 4$ .
2. Predpostavite, da je funkcija  $f$  dovoljkrat zvezno odvedljiva in poiščite izražavo ostanka  $f'(x_0) - F(x_0)$  v obliki  $Kf^{(r)}(\xi)$ , kjer je  $K$  konstanta,  $\xi$  število z intervala  $(x_0 - 2h, x_0 + 2h)$  in  $r$  najmanjše naravno število, za katero velja, da je formula točna za vse polinome stopnje manjše ali enake  $r - 1$ .

3. Z izpeljano formulo bi radi poiskali čimbolj natančen približek za odvod funkcije  $f(x) = 1/(1+x)$  v točki  $x_0 = 1/2$ . Napaka pri računanju odvoda je odvisna od  $h$  in je sestavljena iz dveh delov, napake metode, ki jo določa ostanek, in zaokrožitvene napake pri računanju funkcije  $f$ . Predpostavite, da je napaka pri izračunu vrednosti funkcije  $f$  manjša od  $\varepsilon$ , in določite  $h$ , pri katerem je napaka pri izračunu odvoda najmanjša.

*Rešitev.*

1. Vsak polinom stopnje manjše ali enake 4 lahko predstavimo v predlagani zamaknjeni potenčni bazi. Torej bo formula za odvajanje zanje točna, če konstante  $A, B, C, D$  in  $E$  zadoščajo enačbam

$$\begin{aligned} 0 &= A + B + C + D + E, \\ 1 &= -2hA - hB + hD + 2hE, \\ 0 &= 4h^2A + h^2B + h^2D + 4h^2E, \\ 0 &= -8h^3A - h^3B + h^3D + 8h^3E, \\ 0 &= 16h^4A + h^4B + h^4D + 16h^4E, \end{aligned}$$

ki jih dobimo, če v formuli za  $f$  po vrsti vzamemo bazne funkcije  $x \mapsto (x - x_0)^i$ ,  $i = 0, 1, 2, 3, 4$ . Od tod izračunamo

$$A = \frac{1}{12h}, \quad B = -\frac{2}{3h}, \quad C = 0, \quad D = \frac{2}{3h}, \quad E = -\frac{1}{12h},$$

torej je

$$F(x_0) = \frac{1}{h} \left( \frac{1}{12}f(x_0 - 2h) - \frac{2}{3}f(x_0 - h) + \frac{2}{3}f(x_0 + h) - \frac{1}{12}f(x_0 + 2h) \right).$$

2. Če vzamemo  $f(x) = (x - x_0)^5$ , je

$$f'(x_0) - F(x_0) = 0 - \frac{1}{h} \left( \frac{1}{12}(-2h)^5 - \frac{2}{3}(-h)^5 + \frac{2}{3}h^5 - \frac{1}{12}(2h)^5 \right) = 4h^4$$

Ker je razlika neničelna in oblike  $Kf^{(r)}(\xi)$ , sledi  $r = 5$  in  $K5! = 4h^4$ , torej se za dovolj lepo funkcijo  $f$  napaka izpeljane formule za numerično odvajanje izraža kot

$$f'(x_0) - F(x_0) = \frac{h^4}{30}f^{(5)}(\xi).$$

3. Napaka metode za dano funkcijo  $f$  ustreza absolutni vrednosti ostanka. Za majhen  $h$  bo približno enaka

$$\frac{h^4}{30} \frac{5!}{(1+1/2)^6} = \frac{2^8}{3^6}h^4.$$

Zaokrožitveno napako lahko na podlagi formule za odvajanje ocenimo z

$$\frac{1}{h} \left| \frac{1}{12}\varepsilon_1 - \frac{2}{3}\varepsilon_2 + \frac{2}{3h}\varepsilon_3 - \frac{1}{12}\varepsilon_4 \right| \leq \frac{3\varepsilon}{2h}, \quad |\varepsilon_i| \leq \varepsilon, \quad i = 1, 2, 3, 4.$$

Celotna napaka je torej oblike

$$\frac{2^8}{3^6}h^4 + \frac{3\varepsilon}{2h}$$

in bo najmanjša pri  $h \approx \sqrt[5]{3^7/2^{11}}\varepsilon$ .

□

### 3.1.2 Diferenčna formula za drugi odvod

Podobno kot pravila za aproksimacijo prvih odvodov lahko izpeljemo tudi formule za aproksimacijo drugih odvodov.

Oglejmo si pravilo za drugi odvod funkcije  $f$  v točki  $x_0$ , ki temelji na vrednosti funkcije v točkah  $x_0 - h$ ,  $x_0$ ,  $x_0 + h$  za  $h > 0$ . Polinom  $p_2 \in \mathbb{P}_2$ , podan s predpisom (3.5), interpolira funkcijo v treh izbranih točkah, njegov drugi odvod pa je konstanten in podan s predpisom

$$p_2''(x) = 2[x_0 - h, x_0, x_0 + h]f = \frac{f(x_0 - h) - 2f(x_0) + f(x_0 + h)}{h^2}.$$

Pravilo, skupaj z ostankom, podaja naslednja trditev.

**Trditev 3.4.** *Naj bo  $f$  štirikrat zvezno odvedljiva funkcija na intervalu  $[x_0 - h, x_0 + h]$ ,  $h > 0$ . Potem obstaja število  $\xi \in (x_0 - h, x_0 + h)$ , da velja*

$$f''(x_0) = \frac{f(x_0 - h) - 2f(x_0) + f(x_0 + h)}{h^2} - \frac{1}{12}h^2 f^{(4)}(\xi). \quad (3.9)$$

*Dokaz.* Osnovna ideja dokaza je, da zvezo 3.8 odvajamo dvakrat in nato vstavimo  $x_0$ . Z nekaj računanja dobimo formulo (3.9). □

**Naloga 3.3.** Naj bo  $f$  funkcija iz naloge 3.1. Poenostavite numerično pravilo za drugi odvod funkcije  $f$  v točki  $1/2$ , ki je podano s formulo (3.9), in preverite, ali se ujema z vrednostjo  $f''(1/2)$  v limiti, ko gre  $h$  proti 0.

*Rešitev.* Na podlagi pravila za drugi odvod izračunamo, da je

$$f''(x_0) \approx \frac{1}{h^2} \left( \frac{1}{\frac{1}{2} + h + 1} - \frac{2}{\frac{1}{2} + 1} + \frac{1}{\frac{1}{2} - h + 1} \right) = \frac{16}{27 - 12h^2}.$$

Drugi odvod funkcije  $f$  je podan s predpisom  $f''(x) = 2/(x+1)^3$ , zato je  $f''(1/2) = 16/27$ . Torej se približek za drugi odvod približuje točni vrednosti v limiti, ko gre  $h$  proti 0. □

## 3.2 Newton–Cotesova pravila za integriranje

Podobno kot pravila za odvajanje lahko tudi pravila za integriranje izpeljemo tako, da funkcijo nadomestimo z interpolacijskim polinomom in ga integriramo. Newton–Cotesova pravila temeljijo na integraciji interpolacijskih polinomov nad ekvidistantno delitvijo intervala.

### 3.2.1 Osnovna pravila

Naj bo  $f$  integrabilna funkcija na intervalu  $[a, b]$ . Osnovno Newton–Cotesovo pravilo za integral

$$\int_a^b f(x) dx$$

dobimo tako, da interval  $[a, b]$  za izbran  $n \in \mathbb{N}$  razdelimo z delilnimi točkami  $x_i = a + ih$ ,  $i = 0, 1, \dots, n$ , z ekvidistantnim razmikom  $h = (b - a)/n$ , funkcijo  $f$  nadomestimo s polinomom  $p_n \in \mathbb{P}_n$ , ki interpolira vrednosti  $f$  v točkah  $x_i$ , in polinom  $p_n$  integriramo na intervalu  $[a, b]$ .

Pri  $n = 1$  sta delilni točki  $x_0 = a$  in  $x_1 = b$  le dve in ustrezata robu intervala  $[a, b]$ . Razmik  $h = b - a$  med njima ustreza dolžini intervala. Polinom  $p_1 \in \mathbb{P}_1$ , ki interpolira vrednosti funkcije  $f$  v točkah  $x_0$  in  $x_1$ , je podan s predpisom

$$p_1(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0). \quad (3.10)$$

Z njegovo integracijo dobimo

$$\int_a^b p_1(x) dx = hf(x_0) + \frac{h^2}{2} \frac{f(x_1) - f(x_0)}{h} = \frac{h}{2} (f(x_0) + f(x_1)),$$

kar predstavlja približek za integral funkcije  $f$  na intervalu  $[a, b]$ , ki je izračunan samo na podlagi dveh funkcijskih vrednosti. Integral interpolacijskega polinoma  $p_1$  ustreza ploščini trapeza, ki je določen s točkami  $(a, f(a))$ ,  $(a, 0)$ ,  $(b, 0)$ ,  $(b, f(b))$ , zato takemu približku za integral pravimo trapezno pravilo. Shematično je prikazano na sliki 3.2 (levo), napako približka pa podaja naslednja trditev.

**Trditev 3.5.** *Naj bo  $f$  dvakrat zvezno odvedljiva funkcija na intervalu  $[a, b]$ . Potem obstaja število  $\xi \in (a, b)$ , da velja*

$$\int_a^b f(x) dx = \frac{b-a}{2} (f(a) + f(b)) - \frac{(b-a)^3}{12} f''(\xi). \quad (3.11)$$

*Dokaz.* Formula (3.11) dokažemo s pomočjo zveze med funkcijo  $f$  in interpolacijskim polinomom  $p_1$ , ki je podan z (3.10). Za vsak  $x \in [a, b]$  obstaja število  $\xi_x \in (a, b)$ , da velja

$$f(x) = p_1(x) + \frac{1}{2!} f''(\xi_x)(x-a)(x-b).$$

Z integracijo te enačbe dobimo

$$\int_a^b f(x) dx = \int_a^b p_1(x) dx + \frac{1}{2} \int_a^b f''(\xi_x)(x-a)(x-b) dx. \quad (3.12)$$

Ker je funkcija  $f''$  zvezna na intervalu  $[a, b]$ , je omejena, zato obstajata realni konstanti  $k$  in  $K$ , da velja  $k \leq f''(\xi_x) \leq K$  za vsak  $x \in [a, b]$ . Nadalje velja

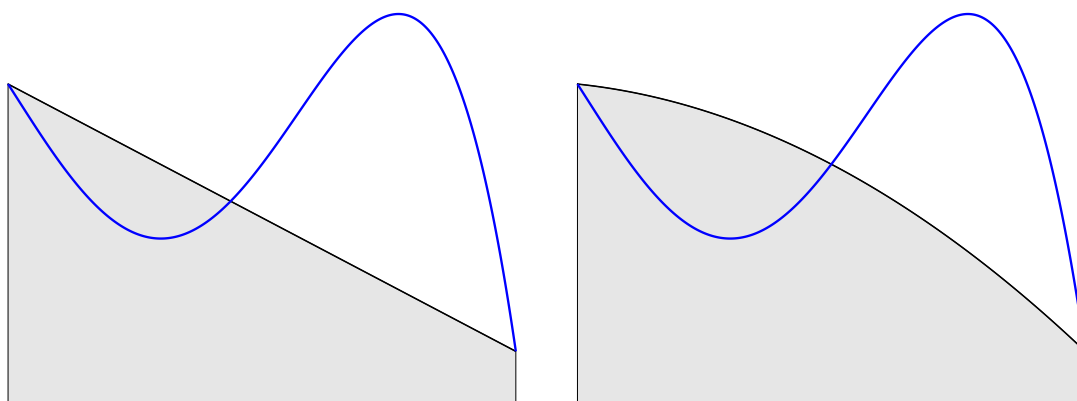
$$K(x-a)(x-b) \leq f''(\xi_x)(x-a)(x-b) \leq k(x-a)(x-b),$$

saj je parabola  $x \mapsto (x-a)(x-b)$  na intervalu  $[a, b]$  negativnega predznaka. Po integraciji zgornjih neenačb dobimo

$$K \int_a^b (x-a)(x-b) dx \leq \int_a^b f''(\xi_x)(x-a)(x-b) dx \leq k \int_a^b (x-a)(x-b) dx$$

in iz

$$\int_a^b (x-a)(x-b) dx = -\frac{1}{6}(b-a)^3$$



(a) Trapezno pravilo.

(b) Simpsonovo pravilo.

Slika 3.2: Prikaz približkov (sivo območje) za integral funkcije (označene z modro) z dvema osnovnima Newton–Cotesovima praviloma.

potem sledi

$$k \leq -\frac{6}{(b-a)^3} \int_a^b f''(\xi_x)(x-a)(x-b) dx \leq K.$$

Ponovno upoštevajmo, da je  $f''$  zvezna na intervalu  $[a, b]$ , kar pomeni, da zavzame vse vrednosti med  $k$  in  $K$ . Torej obstaja tako število  $\xi \in (a, b)$ , da je

$$f''(\xi) = -\frac{6}{(b-a)^3} \int_a^b f''(\xi_x)(x-a)(x-b) dx.$$

Formula (3.11) sedaj sledi iz (3.12) in (3.10).  $\square$

Iz trditve 3.5 sledi, da je trapezno pravilo točno, če je  $f$  linearna funkcija. A ostanek je v splošnem neničeln že, če za funkcijo  $f$  vzamemo polinom iz  $\mathbb{P}_2$ . Poskušajmo izboljšati numerično pravilo z integracijo interpolacijskega polinoma višje stopnje.

Osnovno Newton–Cotesovo pravilo za  $n = 2$  se imenuje Simpsonovo pravilo in ga izpeljemo na podoben način kot trapezno pravilo. Delilne točke so v tem primeru  $x_0 = a$ ,  $x_1 = (a+b)/2$ ,  $x_2 = b$  in so razmaknjene za  $h = (b-a)/2$ . Polinom  $p_2 \in \mathbb{P}_2$ , ki interpolira vrednosti funkcije  $f$  v točkah  $x_0, x_1, x_2$ , je podan s predpisom

$$p_2(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0) + \frac{\frac{f(x_2) - f(x_1)}{x_2 - x_1} - \frac{f(x_1) - f(x_0)}{x_1 - x_0}}{x_2 - x_0}(x - x_0)(x - x_1).$$

Z integriranjem dobimo

$$\int_a^b p_2(x) dx = 2hf(x_0) + \frac{(2h)^2}{2} \frac{f(x_1) - f(x_0)}{h} + \frac{(2h)^3}{12} \frac{f(x_0) - 2f(x_1) + f(x_2)}{2h^2},$$

kar lahko poenostavimo v

$$\int_a^b p_2(x) dx = \frac{h}{3} (f(x_0) + 4f(x_1) + f(x_2)).$$

Pričakovano smo dobili pravilo, ki temelji na treh vrednostih funkcije  $f$ . Shematično je prikazano na sliki 3.2 (desno), ostanek pa opisuje sledeča trditev.

**Trditev 3.6.** Naj bo  $f$  štirikrat zvezno odvedljiva funkcija na intervalu  $[a, b]$ . Potem obstaja število  $\xi \in (a, b)$ , da velja

$$\int_a^b f(x) dx = \frac{b-a}{6} \left( f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right) - \frac{(b-a)^5}{2880} f^{(4)}(\xi). \quad (3.13)$$

Dokaz trditve 3.6 ni preprost in zahteva uporabo še bolj prefinjenih analitičnih orodij, kot smo jih uporabili v dokazu trditve 3.5. Rezultat je tudi nekoliko presenetljiv, saj zagotavlja, da je Simpsonovo pravilo točno ne le za polinome iz  $\mathbb{P}_2$  temveč tudi za kubične polinome.

**Naloga 3.4.** Dan je integral  $\int_0^{\frac{\pi}{4}} \sin(x) dx$ . Numerično izračunajte vrednost integrala z uporabo trapeznega in Simpsonovega pravila ter primerjajte rezultata s točno vrednostjo.

*Rešitev.* Glede na formulo (3.11) je približek za integral po trapeznem pravilu enak

$$\frac{\pi}{8} \left( 0 + \frac{\sqrt{2}}{2} \right) = \frac{\pi\sqrt{2}}{16} \doteq 0.27768.$$

Če računamo po Simpsonovem pravilu, pa s pomočjo formule (3.13) dobimo približek

$$\frac{\pi}{24} \left( 0 + 4 \frac{\sqrt{2-\sqrt{2}}}{2} + \frac{\sqrt{2}}{2} \right) = \frac{\pi}{48} \left( \sqrt{2} + 4\sqrt{2-\sqrt{2}} \right) \doteq 0.29293.$$

Točna vrednost integrala je  $1 - \sqrt{2}/2 \doteq 0.29289$ . □

Tudi integracijska pravila lahko, podobno kot pravila za numerično odvajanje, izpeljemo s pomočjo metode nedoločenih koeficientov. Spodnja naloga zahteva izpeljavo Newton–Cotesovega pravila za  $n = 3$ , ki mu pravimo Simpsonovo drugo pravilo ali pa kar (3/8)-sko pravilo.

**Naloga 3.5.** Naj bo  $f$  označuje  $r$ -krat zvezno odvedljivo funkcijo na intervalu  $[a, b]$ , ki je razdeljen z ekvidistantnimi točkami  $x_i = a + i(b-a)/3$ ,  $i = 0, 1, 2, 3$ . Izpeljite integracijsko formulo oblike

$$\int_a^b f(x) dx = Af(x_0) + Bf(x_1) + Cf(x_2) + Df(x_3) + Ef^{(r)}(\xi),$$

tako da bo točno za polinome čim višjih stopenj. Koeficiente  $A$ ,  $B$ ,  $C$  in  $D$  izračunajte z reševanjem sistema linearnih enačb, ki ga dobite, če pravilo uporabite za funkcije  $x \mapsto (x - x_0)^i$ ,  $i = 0, 1, 2, 3$ . Nato na podoben način določite še konstanti  $E$  in  $r$ , medtem ko je  $\xi$  neznano število z intervala  $(a, b)$ .

*Rešitev.* V nastavku za integracijsko pravilo uporabimo bazne funkcije  $x \mapsto (x - x_0)^i$ ,  $i = 0, 1, 2, 3$ , in na ta način dobimo enačbe

$$\begin{aligned} 3h &= A + B + C + D, \\ (3h)^2/2 &= hB + 2hC + 3hD, \\ (3h)^3/3 &= h^2B + 4h^2C + 9h^2D, \\ (3h)^4/4 &= h^3B + 8h^3C + 27h^3D, \end{aligned}$$

kjer je  $h = (b - a)/3$ . Konstante, ki zadoščajo zgornjemu sistemu enačb, so

$$A = \frac{3}{8}h, \quad B = \frac{9}{8}h, \quad C = \frac{9}{8}h, \quad D = \frac{3}{8}h.$$

Oglejmo si še ostanek integracijskega pravila za funkcijo  $x \mapsto (x - x_0)^4$ . Ker je

$$\frac{(3h)^5}{5} - \frac{3h}{8} (0 + 3h^4 + 3(2h)^4 + (3h)^4) = -\frac{9}{10}h^5,$$

je  $r = 4$  in  $E = -3/80h^5$ . Integracijsko pravilo se torej glasi

$$\int_a^b f(x) dx = \frac{3}{8} (f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)) - \frac{3}{80}h^5 f^{(4)}(\xi)$$

in je (tako kot Simpsonovo pravilo) točno za polinome stopnje 4.  $\square$

Poleg pravil, ki temeljijo na ekvidistantnih točkah  $x_0, x_1, \dots, x_n$ , so pomembna tudi pravila, ki temeljijo le na točkah  $x_1, x_2, \dots, x_{n-1}$ . Da ta dva tipa pravil ločimo, prva imenujemo Newton–Cotesova pravila zaprtega tipa, druga pa Newton–Cotesova pravila odprtega tipa. Slednja so uporabna predvsem v primerih, kadar funkcija  $f$  na robovih intervala integracije  $[a, b]$  ni omejena ali njeni vrednosti nista znani, kar onemogoča uporabo pravil zaprtega tipa.

Najpreprostejše pravilo odprtega tipa je sredinsko pravilo, ki je za nek  $\xi \in (a, b)$  podano s formulo

$$\int_a^b f(x) dx = (b - a)f\left(\frac{a+b}{2}\right) + \frac{8(b-a)^3}{3}h^3 f''(\xi),$$

najpogosteje pa je v rabi Milneovo pravilo na treh točkah, ki ga določa formula

$$\int_a^b f(x) dx = \frac{b-a}{3} \left( 2f\left(\frac{3a+b}{4}\right) - f\left(\frac{a+b}{2}\right) + 2f\left(\frac{a+3b}{4}\right) \right) + \frac{14336}{45}(b-a)^5 f^{(4)}(\xi). \quad (3.14)$$

Pravila odprtega tipa lahko izpeljemo na enake načine kot pravila zaprtega tipa.

**Naloga 3.6.** Izpeljite Milneovo pravilo, ki je podano s formulo (3.14). Pri izpeljavi ostanka predpostavite, da je funkcija  $f$  štirikrat zvezno odvedljiva na intervalu  $(a, b)$  in da je oblike  $Df^{(r)}(\xi)$  za neko konstanto  $D$ , neko naravno število  $r \leq 4$  in  $\xi \in (a, b)$ .

*Rešitev.* Uporabimo metodo nedoločenih koeficientov in začnimo z nastavkom

$$\int_a^b f(x) dx = Af(x_1) + Bf(x_2) + Cf(x_3) + Df^{(r)}(\xi),$$

kjer so točke  $x_i$ ,  $i = 1, 2, 3$ , podane z  $x_i = a + ih$  za  $h = (b - a)/4$ . Koeficiente  $A$ ,  $B$ ,  $C$  lahko določimo tako, da je pravilo točno za vse polinome iz  $\mathbb{P}_2$ , kar pomeni, da gotovo velja  $r > 2$ . V pripravljen nastavek vstavimo funkcije  $x \mapsto (x - a)^i$ ,  $i = 0, 1, 2$ , in tako dobimo sistem treh linearnih enačb

$$4h = A + B + C, \quad 16h^2/2 = hA + 2hB + 3hC, \quad 64h^3/3 = h^2A + 4h^2B + 9h^2C,$$

z rešitvijo  $A = 8h/3$ ,  $B = -4h/3$ ,  $C = 8h/3$ . Torej je

$$\int_a^b f(x) dx = \frac{(b-a)}{3} (2f(x_1) - Bf(x_2) + 2f(x_3)) + Df^{(r)}(\xi).$$

Konstanto  $D$  in red  $r$  poskusimo določiti tako, da v deloma izpeljano formulo vstavimo funkcijo  $x \mapsto (x - a)^3$ . Izračunamo, da je

$$\int_a^b (x - a)^3 dx - \frac{(b - a)}{3} (2h^3 - (2h)^3 + 2(3h)^3) = 0,$$

kar pomeni, da je pravilo točno za vse polinome iz  $\mathbb{P}_3$ . Torej je  $r > 3$ . Nato v pravilo vstavimo še funkcijo  $x \mapsto (x - a)^4$  in z nekaj računanja dobimo

$$\int_a^b (x - a)^4 dx - \frac{(b - a)}{3} (2h^4 - (2h)^4 + 2(3h)^4) = \frac{112}{15} h^5.$$

Ostane je neničeln, torej je  $r = 4$  in  $4!D = 112h^5/15$  oziroma  $D = 14/45h^5$ . Te ugotovimo nas pripeljejo do formule (3.14).  $\square$

V primerjavi s pravili za odvajanje so Newton–Cotesova integracijska pravila numerično stabilnejša, saj utežene vsote funkcijskih vrednosti ne množimo z majhnimi števili. Vseeno pa pravila z večjim številom delilnih točk ne zagotavljajo nujno natančnejših približkov. Zato za izboljšanje rezultatov uberemo podobno pot kot pri aproksimaciji funkcij, kjer smo interval aproksimacije razdelili na več manjših delov.

### 3.2.2 Sestavljena pravila

Pri sestavljenih pravilih integral funkcije  $f$  na intervalu  $[a, b]$  razdelimo ekvidistantno z delilnimi točkami  $x_i = a + ih$ ,  $i = 0, 1, \dots, m$ , kjer je  $h = (b - a)/m$  in je  $m$  lahko veliko naravno število. Osnovna ideja je, da na posameznih podintervalih, ki jih določajo nekatere izmed delilnih točk, uporabimo osnovno Newton–Cotesovo pravilo.

Pri izpeljavi sestavljenih pravil se opremo na aditivnost integrala. Velja na primer

$$\int_a^b f(x) dx = \int_{x_0}^{x_1} f(x) dx + \int_{x_1}^{x_2} f(x) dx + \dots + \int_{x_{m-1}}^{x_m} f(x) dx,$$

kar pomeni, da lahko približek na vsakem podintervalu  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, m$ , poiščemo posebej in neodvisno od sosednjih intervalov. Zato je izpeljava sestavljenih integracijskih pravil manjši zalogaj kot konstrukcija odsekoma polinomskih funkcij, kjer je treba za smiselno aproksimacijo v obzir vzeti tudi zadosten red gladкости v stičnih točkah.

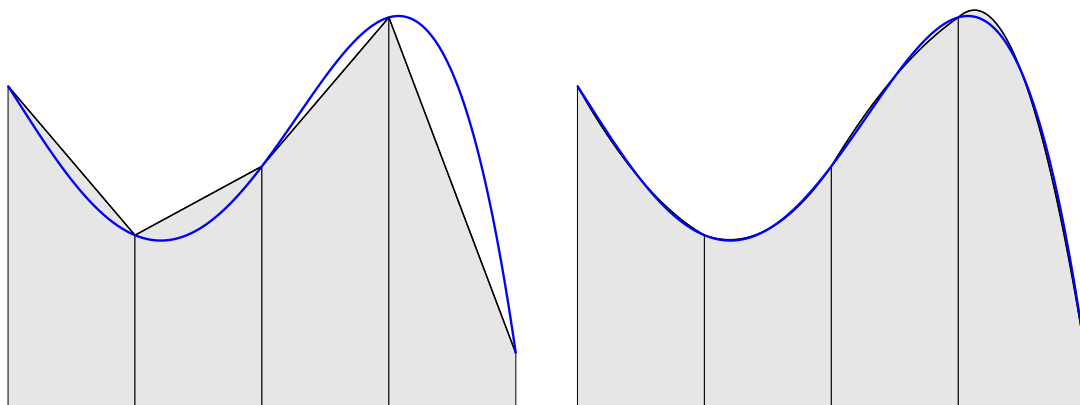
Če na vsakem izmed intervalov delitve uporabimo trapezno pravilo, po formuli (3.11) dobimo približek

$$\int_a^b f(x) dx \approx \frac{h}{2} (f(x_0) + f(x_1)) + \frac{h}{2} (f(x_1) + f(x_2)) + \dots + \frac{h}{2} (f(x_{m-1}) + f(x_m)),$$

ki ga lahko v bolj kompaktni obliki zapišemo kot

$$\int_a^b f(x) dx \approx \frac{h}{2} \left( f(x_0) + 2 \sum_{i=1}^{m-1} f(x_i) + f(x_m) \right).$$

To je ekonomično pravilo, ki temelji le na vrednostih funkcije  $f$  v točkah delitve osnovnega intervala in ga imenujemo sestavljeno trapezno pravilo. Prikazano je na sliki 3.3 (levo). Ostanek, iz katerega je razvidno, da pravilo konvergira, ko  $m$  večamo, podaja naslednja trditev.



(a) Trapezno pravilo.

(b) Simpsonovo pravilo.

Slika 3.3: Prikaz približkov (sivo območje) za integral funkcije (označene z modro) z dvema sestavljenima Newton–Cotesovima praviloma.

**Trditev 3.7.** Naj bo  $f$  dvakrat zvezno odvedljiva funkcija na intervalu  $[a, b]$  in naj bodo  $x_i = a + ih$ ,  $i = 0, 1, \dots, m$ , delilne točke intervala z razmikom  $h = (b - a)/m$ , kjer je  $m$  poljubno naravno število. Potem obstaja število  $\xi \in (a, b)$ , da velja

$$\int_a^b f(x) dx = \frac{h}{2} \left( f(x_0) + 2 \sum_{i=1}^{m-1} f(x_i) + f(x_m) \right) - \frac{h^2(b-a)}{12} f''(\xi).$$

*Dokaz.* Pri izpeljavi ostanka se opremo na trditev 3.5. Ostanki se na posameznih intervalih integracije seštevajo, zato je po formuli (3.11) za celotno pravilo enak

$$\int_a^b f(x) dx - \frac{h}{2} \left( f(x_0) + 2 \sum_{i=1}^{m-1} f(x_i) + f(x_m) \right) = -\frac{h^3}{12} \sum_{i=1}^m f''(\xi_i),$$

kjer  $\xi_i$  označuje število z intervala  $(x_{i-1}, x_i)$ . Ker je

$$\min_{i=1,2,\dots,m} f''(\xi_i) \leq \frac{1}{m} \sum_{i=1}^m f''(\xi_i) \leq \max_{i=1,2,\dots,m} f''(\xi_i)$$

in funkcija  $f''$  zaradi zveznosti zavzame vse vrednosti med zgornjo in spodnjo mejo v zgornjem izrazu, obstaja število  $\xi \in (a, b)$ , pri katerem velja

$$f''(\xi) = \frac{1}{m} \sum_{i=1}^m f''(\xi_i).$$

Torej lahko ostanek zapišemo kot

$$-\frac{h^3}{12} \sum_{i=1}^m f''(\xi_i) = -\frac{h^3}{12} m f''(\xi) = -\frac{h^2(hm)}{12} f''(\xi) = -\frac{h^2(b-a)}{12} f''(\xi),$$

kar dokazuje trditev. □

Sestavljeno Simpsonovo pravilo izpeljemo na podoben način kot trapezno, le da za  $m$  pridejo v poštev le soda števila, saj osnovno pravilo poleg vrednosti funkcije v robnih točkah upošteva tudi vrednost v sredinski točki intervala. Integral zapišemo kot

$$\int_a^b f(x) dx = \int_{x_0}^{x_2} f(x) dx + \int_{x_2}^{x_4} f(x) dx + \dots + \int_{x_{m-2}}^{x_m} f(x) dx$$

in ga z uporabo osnovnega pravila ocenimo z

$$\begin{aligned} \int_a^b f(x) dx \approx \frac{h}{3} (f(x_0) + 4f(x_1) + f(x_2)) + \frac{h}{3} (f(x_2) + 4f(x_3) + f(x_4)) + \dots \\ \dots + \frac{h}{3} (f(x_{m-2}) + 4f(x_{m-1}) + f(x_m)). \end{aligned}$$

Po poenostavitvi in izpeljavi ostanka, ki jo opravimo na podoben način kot v trditvi 3.7, dobimo pravilo, ki je shematično prikazano na sliki 3.3 (desno) in opisano v spodnji trditvi. Tudi v tem primeru imamo jasno zagotovilo, da približki za integral konvergirajo k točni vrednosti, ko gostimo število delilnih točk na osnovnem intervalu.

**Trditev 3.8.** Naj bo  $f$  štirikrat zvezno odvedljiva funkcija na intervalu  $[a, b]$  in naj bodo  $x_i = a + ih$ ,  $i = 0, 1, \dots, m$ , delilne točke intervala z razmikom  $h = (b - a)/m$ , kjer je  $m$  sodo naravno število. Potem obstaja število  $\xi \in (a, b)$ , da velja

$$\int_a^b f(x) dx = \frac{h}{3} \left( f(x_0) + 4 \sum_{i=1}^{\frac{m}{2}} f(x_{2i-1}) + 2 \sum_{i=1}^{\frac{m}{2}-1} f(x_{2i}) + f(x_m) \right) - \frac{h^4(b-a)}{180} f^{(4)}(\xi).$$

**Naloga 3.7.** Izračunajte približke za integral funkcije  $f(x) = \sin(x)$  na intervalu  $[0, \pi]$  z uporabo sestavljenega trapeznega in Simpsonovega pravila pri  $m = 2$  in  $m = 4$ . Primerjajte približke s točno vrednostjo.

*Rešitev.* Za izračun približkov integrala potrebujemo vrednosti funkcije  $f$  v točkah  $\pi j/4$ ,  $j = 0, 1, \dots, 4$ :

$$f(0) = 0, \quad f(\pi/4) = \frac{\sqrt{2}}{2}, \quad f(\pi/2) = 1, \quad f(3\pi/4) = \frac{\sqrt{2}}{2}, \quad f(\pi) = 0.$$

Z uporabo trapeznega pravila pri  $m = 2$  dobimo približek

$$\frac{\pi}{4} (0 + 2 + 0) = \frac{\pi}{2} \doteq 1.5708,$$

pri  $m = 4$  pa približek

$$\frac{\pi}{8} (0 + \sqrt{2} + 2 + \sqrt{2} + 0) = \frac{\pi}{4} (\sqrt{2} + 1) \doteq 1.8961.$$

Pri Simpsonovem pravilu sta približka dana z

$$\frac{\pi}{6} (0 + 4 + 0) = \frac{2\pi}{3} \doteq 2.0944$$

in

$$\frac{\pi}{12} (0 + 2\sqrt{2} + 2 + 2\sqrt{2} + 0) = \frac{\pi}{6} (2\sqrt{2} + 1) \doteq 2.0046.$$

Točna vrednost integrala  $\int_0^\pi \sin(x) dx$  je 2. □

**Naloga 3.8.** V Matlabu pripravite metodi za izračun približka integrala funkcije po sestavljenem trapeznem in Simpsonovem pravilu, ki sta podana v trditvah 3.7 in 3.8. Metodi naj sprejmeta funkcijske vrednosti v delilnih točkah  $x_0, x_1, \dots, x_m$  in razmik  $h$  med njimi. Preizkusite ju na primeru iz naloge 3.7. Za obe pravili določite najmanjše število  $m$ , pri katerem je napaka manjša od  $10^{-2}$ . Nato za izračun približka uporabite še vgrajeni ukaz `integral`.

*Rešitev.* Približki za integral funkcije  $f(x) = \sin(x)$  na intervalu  $[0, \pi]$  izračunani s sestavljenim trapeznim in Simpsonovim pravilom so podani v tabeli 3.1. Simpsonovo pravilo lahko uporabimo le za soda števila  $m$ . Iz rezultatov je razvidno, da za želeno natančno približka pri sestavljenem Simpsonovem pravilu zadošča že  $m = 4$ , medtem ko je napaka pri sestavljenem trapeznem pravilu manjša od  $10^{-2}$  šele pri  $m = 13$ . Z uporabo vgrajene funkcije `integral` dobimo približek, pri katerem je napaka reda  $10^{-16}$ . Pri sestavljenem Simpsonovem pravilu je za tako dober približek treba vzeti  $m \approx 3500$ , pri trapeznem pravilu pa  $m \approx 10^7$ .  $\square$

| $m$ | sestavljeno trapezno pravilo |                         | sestavljeno Simpsonovo pravilo |                         |
|-----|------------------------------|-------------------------|--------------------------------|-------------------------|
|     | približek                    | napaka                  | približek                      | napaka                  |
| 1   | $1.9236707 \cdot 10^{-16}$   | $2.00000 \cdot 10^0$    |                                |                         |
| 2   | 1.5707963                    | $4.29204 \cdot 10^{-1}$ | 2.0943951                      | $9.43951 \cdot 10^{-2}$ |
| 3   | 1.8137994                    | $1.86201 \cdot 10^{-1}$ |                                |                         |
| 4   | 1.8961189                    | $1.03881 \cdot 10^{-1}$ | 2.0045598                      | $4.55975 \cdot 10^{-3}$ |
| 5   | 1.9337656                    | $6.62344 \cdot 10^{-2}$ |                                |                         |
| 6   | 1.9540972                    | $4.59028 \cdot 10^{-2}$ | 2.0008632                      | $8.63190 \cdot 10^{-4}$ |
| 7   | 1.9663167                    | $3.36833 \cdot 10^{-2}$ |                                |                         |
| 8   | 1.9742316                    | $2.57684 \cdot 10^{-2}$ | 2.0002692                      | $2.69170 \cdot 10^{-4}$ |
| 9   | 1.9796508                    | $2.03492 \cdot 10^{-2}$ |                                |                         |
| 10  | 1.9835235                    | $1.64765 \cdot 10^{-2}$ | 2.0001095                      | $1.09517 \cdot 10^{-4}$ |
| 11  | 1.9863870                    | $1.36130 \cdot 10^{-2}$ |                                |                         |
| 12  | 1.9885638                    | $1.14362 \cdot 10^{-2}$ | 2.0000526                      | $5.26243 \cdot 10^{-5}$ |
| 13  | 1.9902572                    | $9.74282 \cdot 10^{-3}$ |                                |                         |
| 14  | 1.9916004                    | $8.39957 \cdot 10^{-3}$ | 2.0000283                      | $2.83436 \cdot 10^{-5}$ |
| 15  | 1.9926838                    | $7.31617 \cdot 10^{-3}$ |                                |                         |
| 16  | 1.9935703                    | $6.42966 \cdot 10^{-3}$ | 2.0000166                      | $1.65910 \cdot 10^{-5}$ |

Tabela 3.1: Primerjava sestavljenega trapeznega in Simpsonovega pravila iz naloge 3.8.

### 3.3 Izboljšana pravila za integriranje

Sestavljena pravila so robustno orodje za izračun natančnega približka za integral funkcije, a število funkcijskih vrednostih, ki jih pri tem uporabijo, je lahko veliko. Na voljo je več različnih načinov, ki nam omogočajo, da do enako dobrih približkov pridemo z manjšim številom izračunov.

### 3.3.1 Rombergova metoda

Numerično integriranje funkcije  $f$  na intervalu  $[a, b]$  s sestavljenim pravilom si lahko predstavljamo kot iterativen proces, pri katerem za izbran  $m \in \mathbb{N}$  izračunamo približek pri razmiku  $h = (b - a)/m$ , nato pa v primeru, da z rezultatom nismo zadovoljni, razmik razpolovimo in izračun ponovimo. Pričakujemo, da je približek pri koraku  $h/2$  boljši kot pri  $h$ , a vprašamo se, ali lahko s kombinacijo teh dveh približkov pridobimo še boljši rezultat.

Naj  $P_h f$  označuje numerično pravilo za integracijo funkcije  $f$  pri razmiku  $h$ . Predpostavimo, da zanj velja

$$\int_a^b f(x) dx = P_h f + K h^p + \mathcal{O}(h^r), \quad (3.15a)$$

kjer je  $r > p$  in  $K$  konstanta, ki je neodvisna od  $h$ . Potem velja tudi

$$\int_a^b f(x) dx = P_{h/2} f + K \left(\frac{h}{2}\right)^p + \mathcal{O}\left(\left(\frac{h}{2}\right)^r\right) = P_{h/2} f + \frac{1}{2^p} K h^p + \mathcal{O}(h^r) \quad (3.15b)$$

in če (3.15b) pomnožimo z  $2^p$  ter odštejemo (3.15a), dobimo

$$\int_a^b f(x) dx = \frac{2^p P_{h/2} f - P_h f}{2^p - 1} + \mathcal{O}(h^r).$$

To pa pomeni, da je vrednost  $(2^p P_{h/2} f - P_h f)/(2^p - 1)$  boljši približek za integral funkcije kot  $P_h f$  ali  $P_{h/2} f$ , saj je napaka reda  $r > p$ . Za izvedbo tovrstnega popravka ne potrebujemo nobene dodatne funkcijske vrednosti, treba je vedeti le, kako se obnaša napaka sestavljenega pravila. To zamisel imenujemo Richardsonova ekstrapolacija.

Naj  $T_h f$  označuje sestavljeno trapezno pravilo pri razmiku  $h$ . Trditev 3.7 namiguje, da bi lahko obstajala konstanta  $K$ , da bi veljalo

$$\int_a^b f(x) dx \approx T_h f + K h^2 + \mathcal{O}(h^3).$$

Izkaže se, da se ostanek v odvisnosti od razmika  $h$  obnaša še bistveno bolj zanimivo in da ustreza vrsti s sodimi potencami. Obstajajo torej konstante  $K_2, K_4, K_6, \dots$ , da velja

$$\int_a^b f(x) dx = T_h f + K_2 h^2 + K_4 h^4 + K_6 h^6 + \dots$$

Približek, dobljen z Richardsonovo ekstrapolacijo, označimo s

$$T_{h/2}^{(1)} f = \frac{4T_{h/2} f - T_h f}{3}. \quad (3.16)$$

Obstajajo konstante  $K_4^{(1)}, K_6^{(1)}, \dots$ , za katere velja

$$\int_a^b f(x) dx = T_{h/2}^{(1)} f + K_4^{(1)} h^4 + K_6^{(1)} h^6 + \dots$$

Zato lahko postopek induktivno ponavljamo in po rekurzivni formuli

$$T_{h/2^r}^{(r)} f = \frac{2^{2r} T_{h/2^r}^{(r-1)} f - T_{h/2^{r-1}}^{(r-1)} f}{2^{2r} - 1}, \quad r = 2, 3, \dots, k, \quad (3.17)$$

dobimo približek z lastnostjo

$$\int_a^b f(x) dx = T_{h/2^k}^{(k)} f + \mathcal{O}(h^{2k+2}).$$

Ta posebna uporaba Richardsonove ekstrapolacije je znana pod imenom Rombergova metoda.

**Naloga 3.9.** Z uporabo Rombergove metode izračunajte približek  $T_{\pi/4}^{(2)} f$  za integral funkcije  $f$  iz naloge 3.7 na intervalu  $[0, \pi]$ .

*Rešitev.* Za izračun  $T_{\pi/4}^{(2)} f$  je treba najprej izračunati približka  $T_{\pi/2}^{(1)} f$  in  $T_{\pi/4}^{(1)} f$ . Vrednost osnovnega trapeznega pravila, ki ustreza približku  $T_{\pi} f$ , je

$$T_{\pi} f = \frac{\pi}{2} (f(0) + f(\pi)) = \frac{\pi}{2} (0 + 0) = 0,$$

vrednosti

$$T_{\pi/2} f = \frac{\pi}{2}, \quad T_{\pi/4} f = \frac{\pi}{4}(\sqrt{2} + 1)$$

pa sta izračunani že v nalogi 3.7. Po formuli (3.16) je potem

$$\begin{aligned} T_{\pi/2}^{(1)} f &= \frac{4T_{\pi/2} f - T_{\pi} f}{3} = \frac{4 \cdot \frac{\pi}{2} - 0}{3} = \frac{2\pi}{3}, \\ T_{\pi/4}^{(1)} f &= \frac{4T_{\pi/4} f - T_{\pi/2} f}{3} = \frac{4 \cdot \frac{\pi}{4}(\sqrt{2} + 1) - \frac{\pi}{2}}{3} = \frac{\pi}{6}(2\sqrt{2} + 1). \end{aligned}$$

Opazimo, da ta dva približka ravno ustrezata približkoma po (sestavljene) Simpsonovem pravilu iz naloge 3.7. Po formuli (3.17) dobimo še

$$T_{\pi/4}^{(2)} f = \frac{16T_{\pi/4}^{(1)} f - T_{\pi/2}^{(1)} f}{15} = \frac{16 \cdot \frac{\pi}{6}(2\sqrt{2} + 1) - \frac{2\pi}{3}}{15} = \frac{2\pi}{45} (8\sqrt{2} + 3) \doteq 1.9986,$$

kar je boljši približek za točno vrednost integrala, ki je enak 2, kot približki, ki smo jih izračunali poprej.  $\square$

| $r$ | $T_{\pi/2^r} f$              | $T_{\pi/2^r}^{(1)} f$ | $T_{\pi/2^r}^{(2)} f$ | $T_{\pi/2^r}^{(3)} f$ | $T_{\pi/2^r}^{(4)} f$ |
|-----|------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 0   | $1.923670694 \cdot 10^{-16}$ |                       |                       |                       |                       |
| 1   | 1.570796327                  | 2.094395102           |                       |                       |                       |
| 2   | 1.896118898                  | 2.004559755           | 1.998570732           |                       |                       |
| 3   | 1.974231602                  | 2.000269170           | 1.999983131           | 2.000005550           |                       |
| 4   | 1.993570344                  | 2.000016591           | 1.999999752           | 2.000000016           | 1.999999995           |

Tabela 3.2: Rezultati izvedbe Rombergove metode iz naloge 3.10.

**Naloga 3.10.** V Matlabu implementirajte Rombergovo metodo, ki kot vhodne podatke sprejme funkcijo  $f$ , interval integracije  $[a, b]$  in podatek o številu korakov  $k \in \mathbb{N}$ . Poleg približka  $T_{(b-a)/2^{k-1}}^{(k-1)} f$  naj metoda vrne še celotno shemo približkov v tabeli velikosti  $k \times k$ . Metodo preizkusite na primeru iz naloge 3.7 in rezultate primerjate z rezultati iz naloge 3.8.

*Rešitev.* Rezultati izvedbe metode na primeru iz naloge 3.7 za  $k = 5$  so prikazana v tabeli 3.2.  $\square$

### 3.3.2 Adaptivna pravila

Izračun približka za integral funkcije  $f$  na intervalu  $[a, b]$  s sestavljenim Newton–Cotesovim pravilom temelji na vrednostih  $f$  v ekvidistatnih delilnih točkah. A funkcija se lahko na različnih delih intervala  $[a, b]$  obnaša zelo različno. Nekje je denimo položna, drugje pa strmo pada ali narašča, morda celo skokovito niha. Temu velja prilagoditi oziroma adaptirati tudi integracijsko pravilo in tam, kjer je funkcija položna, delilne točke zredčiti, tam, kjer se hitro spreminja, pa zgostiti.

Raje kot da analiziramo obnašanje funkcije, ki je pogosto niti ne poznamo v zaključeni obliki, pri snovanju integracijskega pravila opazujmo, kako se na določenih delih intervala spreminja ostanek. Za izhodišče izpeljave adaptivnega pravila vzemimo Simpsonovo pravilo.

Na intervalu  $[a, b]$  najprej uporabimo osnovno Simpsonovo pravilo

$$S_1 f = \frac{b-a}{6} \left( f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right),$$

nato pa še sestavljeno Simpsonovo pravilo

$$S_2 f = \frac{b-a}{12} \left( f(a) + 4f\left(\frac{3a+b}{4}\right) + 2f\left(\frac{a+b}{2}\right) + 4f\left(\frac{a+3b}{4}\right) + f(b) \right)$$

na dveh podintervalih. Po trditvi 3.8 obstajata števili  $\xi_1, \xi_2 \in (a, b)$ , da velja

$$\int_a^b f(x) dx = S_1 f - \frac{h^4(b-a)}{180} f^{(4)}(\xi_1) = S_2 f - \frac{h^4(b-a)}{16 \cdot 180} f^{(4)}(\xi_2).$$

Če predpostavimo, da je  $f^{(4)}(\xi_1) \approx f^{(4)}(\xi_2)$ , po konceptu Richardsonove ekstrapolacije sledi

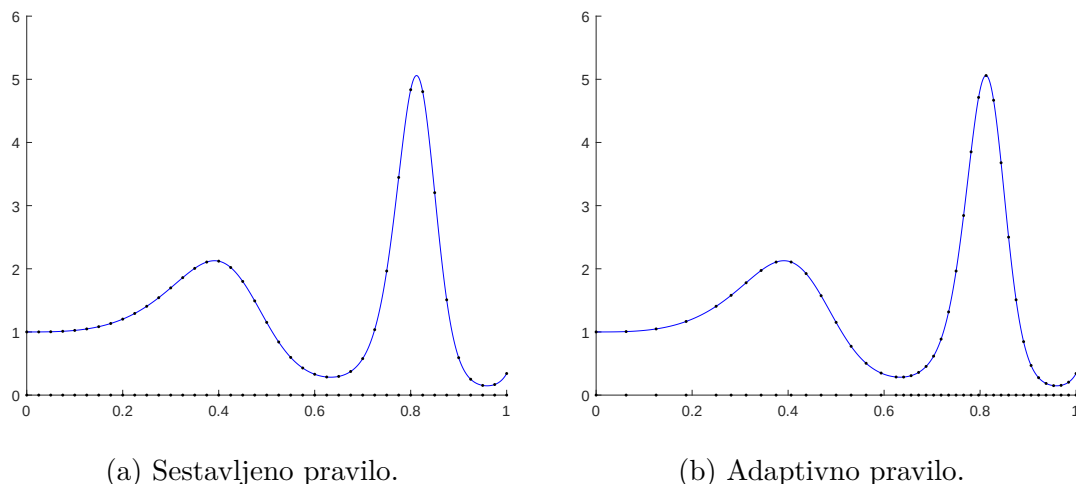
$$\int_a^b f(x) dx \approx \frac{16S_2 f - S_1 f}{15},$$

hkrati pa lahko na podlagi te predpostavke ocenimo tudi, da je

$$-\frac{h^4(b-a)}{16 \cdot 180} f^{(4)}(\xi_2) \approx \frac{S_2 f - S_1 f}{15}.$$

Zgornji opazki lahko izkoristimo takole. Vrednost  $|S_2 f - S_1 f|/15$  nam pove približno kako dober približek smo izračunali s  $S_2 f$ . Če je vrednost majhna (denimo manjša od predpisane tolerance  $\delta$ ), z računanjem približka zaključimo, a namesto  $S_2 f$  kot končni približek vzamemo  $(16S_2 f - S_1 f)/15$ , saj je ta vrednost še bližje točni vrednosti integrala.

V primeru, da je vrednost ocene za ostanek velika, lahko z računanjem nadaljujemo tako, da enak postopek s toleranco  $\delta/2$  ponovimo na podintervalih  $[a, \frac{a+b}{2}]$  in  $[\frac{a+b}{2}, b]$  ter dobljena rezultata seštejemo. S tem integracijo razdelimo na dva manjša podintervala oziroma rekurzivno na številne odseke začetnega intervala, računanje približka na njih pa izvajamo le tako natančno, kot je glede na oceno ostanka smiselno. Postopek zagotavlja, da napaka približka ob koncu računanja ni večja od  $\delta$ .



Slika 3.4: Prikaz uporabljenih funkcijskih vrednosti pri računanju približka za integral iz primera 3.1.

**Primer 3.1.** Oglejmo si primerjavo med adaptivnim in sestavljenim Simpsonovim pravilom pri računanju integrala funkcije

$$f(x) = e^{2x \sin(12x^2)}$$

na intervalu  $[0, 1]$ . Pri adaptivnem pravilu uporabimo tolerance  $10^{-e}$ ,  $e = 3, 4, \dots, 12$ , izračunamo približek in preštejemo število funkcijskih izračunov v različnih točkah, do katerih pride tekom izvedbe metode. Nato izračunamo še približek po sestavljenem Simpsonovem pravilu, ki temelji na enakem številu funkcijskih vrednosti (v ekvidistantnih točkah z intervala  $[0, 1]$ ). Rezultati so podani v tabeli 3.3. Računanje približkov z upodobitvijo 41 točk, na katerih temeljita adaptivno in sestavljeno Simpsonovo pravilo, je prikazano na sliki 3.4.

| funkcijske vrednosti | sestavljenem pravilo     | adaptivno pravilo        |
|----------------------|--------------------------|--------------------------|
| 41                   | $1.21648 \cdot 10^{-4}$  | $2.71801 \cdot 10^{-6}$  |
| 73                   | $8.55455 \cdot 10^{-6}$  | $2.09027 \cdot 10^{-5}$  |
| 121                  | $1.15580 \cdot 10^{-6}$  | $4.97486 \cdot 10^{-7}$  |
| 225                  | $9.66732 \cdot 10^{-8}$  | $1.09655 \cdot 10^{-10}$ |
| 401                  | $9.54627 \cdot 10^{-9}$  | $9.10448 \cdot 10^{-10}$ |
| 709                  | $9.73835 \cdot 10^{-10}$ | $1.80951 \cdot 10^{-11}$ |
| 1273                 | $9.35048 \cdot 10^{-11}$ | $4.48086 \cdot 10^{-13}$ |
| 2213                 | $1.02227 \cdot 10^{-11}$ | $2.37588 \cdot 10^{-14}$ |
| 3793                 | $1.18194 \cdot 10^{-12}$ | $1.73195 \cdot 10^{-14}$ |
| 6937                 | $1.05471 \cdot 10^{-13}$ | $2.88658 \cdot 10^{-15}$ |

Tabela 3.3: Primerjava sestavljenega in adaptivnega Simpsonovega pravila pri računanju približka za integral funkcije iz primera 3.1. Prvi stolpec podaja števila uporabljenih funkcijskih vrednosti, drugi in tretji stolpec pa napake približkov.

### 3.3.3 Gaussova pravila

Izpeljava osnovnih Newton–Cotesovih pravil za integracijo funkcije  $f$  temelji na izbiri ekvidistantnih točk na intervalu  $[a, b]$ . Na mestu je vprašanje, ali lahko natančnost približka, ki ga za izbrano naravno število  $n$  dobimo na podlagi  $n+1$  funkcijskih vrednosti, izboljšamo, če uporabimo vrednosti  $f$  v  $n+1$  točkah, ki jih izberemo na kak drug način.

Integracijsko pravilo zapišemo z nastavkom

$$\int_a^b f(x) dx \approx A_0 f(x_0) + A_1 f(x_1) + \dots + A_n f(x_n). \quad (3.18)$$

Če točke  $x_i$ ,  $i = 0, 1, \dots, n$ , z intervala  $[a, b]$  razumemo kot proste parametre, je v pravilu  $2n+2$  neznank. Pričakovati je, da lahko te neznanke določimo tako, da bo pravilo točno za vse polinome iz  $\mathbb{P}_{2n+1}$ . Takemu pravilo pravimo Gaussovo integracijsko pravilo. Izpeljavo lahko opravimo z metodo nedoločenih koeficientov.

**Naloga 3.11.** Naj bo  $f$  integrabilna funkcija na intervalu  $[-1, 1]$ . Določite točki  $x_0, x_1$  z intervala  $[-1, 1]$  in koeficienta  $A_0, A_1 \in \mathbb{R}$ , pri katerih je integracijsko pravilo

$$\int_{-1}^1 f(x) dx \approx A_0 f(x_0) + A_1 f(x_1)$$

točno za vse polinome stopnje manjše ali enake 3.

*Rešitev.* Integracijsko pravilo je točno za vsak polinom iz  $\mathbb{P}_3$ , če je točno za funkcije  $x \mapsto x^i$ ,  $i = 0, 1, 2, 3$ . Če pravilo uporabimo na teh funkcijah, po vrsti dobimo

$$\begin{aligned} 2 &= A_0 + A_1, \\ 0 &= A_0 x_0 + A_1 x_1, \\ \frac{2}{3} &= A_0 x_0^2 + A_1 x_1^2, \\ 0 &= A_0 x_0^3 + A_1 x_1^3. \end{aligned}$$

To je sistem nelinearnih enačb za neznanke  $x_0, x_1$  in  $A_0, A_1$ . Hitro se lahko prepričamo, da nobena izmed neznank ne more biti enaka 0. Če drugo enačbo pomnožimo z  $x_0^2$  in od nje odštejemo četrto, dobimo

$$A_1 x_1 (x_0^2 - x_1^2) = 0,$$

torej mora biti  $x_0^2 = x_1^2$ . Iz prve in tretje enačbe potem sledi

$$\frac{2}{3} = A_0 x_0^2 + A_1 x_1^2 = (A_0 + A_1) x_0^2 = 2x_0^2.$$

Brez škode za splošnost privzamemo, da je  $x_0 < x_1$ , torej je  $x_0 = -\sqrt{3}/3$  in  $x_1 = \sqrt{3}/3$ . Iz druge enačbe potem sledi  $A_0 = A_1$  in iz prve  $A_0 = A_1 = 1$ . S tem smo izpeljali, da je pravilo

$$\int_{-1}^1 f(x) dx \approx f\left(-\frac{\sqrt{3}}{3}\right) + f\left(\frac{\sqrt{3}}{3}\right)$$

točno za polinome stopnje manjše ali enake 3. □

Priprava Gaussovega pravila z metodo nedoločenih koeficientov je v splošnem zahtevna naloga, saj je treba za izračun prostih parametrov v integracijskem pravilu rešiti sistem nelinearnih enačb. Zato je pomembna naslednja ugotovitev.

**Trditev 3.9.** Naj bo skalarni produkt s kvadratom integrabilnih funkcij  $g$  in  $h$  podan s predpisom

$$\langle g, h \rangle = \int_a^b g(x)h(x) \, dx.$$

Denimo, da so  $x_0, x_1, \dots, x_n$  take paroma različne točke z intervala  $[a, b]$ , da je polinom

$$\omega(x) = (x - x_0)(x - x_1) \dots (x - x_n)$$

pravokoten na vse polinome iz  $\mathbb{P}_n$ . Potem je pravilo (3.18), ki je določeno s koeficienti

$$A_i = \int_a^b \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j} \, dx, \quad (3.19)$$

točno za vse polinome iz  $\mathbb{P}_{2n+1}$ .

*Dokaz.* Ker je  $\omega$  polinom stopnje  $n+1$ , za poljuben polinom  $p \in \mathbb{P}_{2n+1}$  obstajata polinoma  $s, r \in \mathbb{P}_n$ , da velja  $p = s \cdot \omega + r$ . Potem je

$$\int_a^b p(x) \, dx = \int_a^b s(x)\omega(x) \, dx + \int_a^b r(x) \, dx.$$

Po predpostavki je

$$\int_a^b s(x)\omega(x) \, dx = \langle s, \omega \rangle = 0,$$

polinom  $r$  pa lahko zapišemo v bazi Lagrangeevih baznih polinomov v obliki

$$r(x) = \sum_{i=0}^n p(x_i) \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j},$$

saj  $r$  interpolira vrednosti polinoma  $p$  v točkah  $x_0, x_1, \dots, x_n$ . To pomeni, da je

$$\int_a^b p(x) \, dx = \int_a^b \left( \sum_{i=0}^n p(x_i) \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j} \right) \, dx = \sum_{i=0}^n \left( \int_a^b \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j} \, dx \right) p(x_i),$$

kar zaključuje dokaz. □

Trditev 3.9 ponuja alternativni način za izpeljavo Gaussovega pravila (3.18) za izbrano število  $n \in \mathbb{N}$ . Najprej poiščemo polinom stopnje  $n+1$ , ki je pravokoten na vse polinome iz  $\mathbb{P}_n$ . To lahko napravimo s pomočjo Gram–Schmidtovega postopka. Ničle tega polinoma (za katere se izkaže, da so vedno realne in enostavne ter ležijo intervalu  $(a, b)$ ) določajo točke  $x_0, x_1, \dots, x_n$ . Koeficiente  $A_0, A_1, \dots, A_n$  nato izračunamo po formuli (3.19) ali z metodo nedoločenih koeficientov, pri kateri dobimo sistem linearnih enačb, saj so točke  $x_0, x_1, \dots, x_n$  že fiksirane.

**Naloga 3.12.** Preverite, da je polinom  $x \mapsto x^3 - \frac{3}{5}x$  pravokoten na vse polinome iz  $\mathbb{P}_2$  glede na skalarni produkt, ki je za funkciji  $g$  in  $h$  podan s predpisom

$$\langle g, h \rangle = \int_{-1}^1 g(x)h(x) \, dx.$$

Nato določite Gaussovo integracijsko pravilo oblike

$$\int_{-1}^1 f(x) dx \approx A_0 f(x_0) + A_1 f(x_1) + A_2 f(x_2),$$

ki je točno za vse polinome iz  $\mathbb{P}_5$ .

*Rešitev.* Pišimo

$$\omega(x) = x^3 - \frac{3}{5}x = \left(x + \frac{\sqrt{3}}{\sqrt{5}}\right) x \left(x - \frac{\sqrt{3}}{\sqrt{5}}\right).$$

Ker je polinom  $\omega$  lih, je  $\langle \omega, 1 \rangle = \langle \omega, x^2 \rangle = 0$ . Poleg tega je

$$\langle \omega, x \rangle = \int_{-1}^1 \omega(x)x dx = \left( \frac{x^5}{5} - \frac{x^4}{5} \right) \Big|_{x=-1}^{x=1} = 0,$$

na podlagi česar lahko sklepamo, da je  $\langle \omega, p \rangle = 0$  za vsak  $p \in \mathbb{P}_2$ . Po trditvi 3.9 lahko torej integracijsko pravilo zastavimo v obliki

$$\int_{-1}^1 f(x) dx \approx A_0 f\left(-\frac{\sqrt{3}}{\sqrt{5}}\right) + A_1 f(0) + A_2 f\left(\frac{\sqrt{3}}{\sqrt{5}}\right)$$

in poiščemo take koeficiente  $A_0, A_1, A_2$ , da bo pravilo točno za polinome iz  $\mathbb{P}_5$ . Enačbe, ki jih dobimo, če v pravilo vstavimo funkcije  $x \mapsto x^i$ ,  $i = 0, 1, 2$ , so

$$A_0 + A_1 + A_2 = 2, \quad -\frac{\sqrt{3}}{\sqrt{5}}A_0 + \frac{\sqrt{3}}{\sqrt{5}}A_2 = 0, \quad \frac{3}{5}A_0 + \frac{3}{5}A_2 = \frac{2}{3}$$

in enolično določajo koeficiente:  $A_0 = 5/9$ ,  $A_1 = 8/9$  in  $A_2 = 5/9$ . Izpeljali smo torej pravilo

$$\int_{-1}^1 f(x) dx \approx \frac{1}{9} \left( 5f\left(-\frac{\sqrt{3}}{\sqrt{5}}\right) + 8f(0) + 5f\left(\frac{\sqrt{3}}{\sqrt{5}}\right) \right),$$

ki je po trditvi 3.9 točno za vse polinome iz  $\mathbb{P}_5$ . □

**Naloga 3.13.** Izračunajte približka za integral funkcije  $f(x) = \cos(\pi x)$  na intervalu  $[-1, 1]$  z Gaussovima praviloma iz nalog 3.11 in 3.12 ter ju primerjajte s približkoma, izračunanima z osnovnim trapeznim in Simpsonovim pravilom, ki temeljita na enakem številu točk.

*Rešitev.* Točna vrednost integrala je 0. S kalkulatorjem izračunamo, da je približek po pravilu iz naloge 3.11 enak  $-0.4812$ , medtem ko je pri trapeznem pravilu, ki prav tako temelji na dveh točkah, enak  $-2$ . Približek, izračunan po pravilu iz naloge 3.12, je  $0.0449$  in je prav tako bistveno boljši od približka  $2/3$ , ki ga dobimo z uporabo Simpsonovega pravila na treh točkah. □

## Poglavje 4

# Numerično reševanje diferencialnih enačb

Diferencialna enačba je enačba, ki povezuje funkcijo z njenimi odvodi. S tovrstnimi enačbami lahko opišemo številne naravne in gospodarske procese in če uspemo enačbo rešiti, lahko tak proces simuliramo pod različnimi pogoji. Reševanje splošnih diferencialnih enačb je težak matematični problem, le redke enačbe imajo eksplicitne rešitve, zato je ključnega pomena numerično reševanje. Pri tem se lahko zatečemo tudi k aproksimacijskim shemam za odvajanje in integriranje.

### 4.1 Reševanje začetnega problema

Klasičen matematični model, povezan z reševanjem navadnih diferencialnih enačb, je začetni problem, imenovan tudi Cauchyjeva naloga. Za dano funkcijo  $f$  dveh spremenljivk iščemo funkcijo  $y$  ene spremenljivke, ki na intervalu  $(x_0, x_0 + \varepsilon)$ ,  $x_0 \in \mathbb{R}$ ,  $\varepsilon > 0$ , zadošča enačbi

$$y'(x) = f(x, y(x)), \quad (4.1)$$

ob tem pa je izpolnjen še začetni pogoj  $y(x_0) = y_0$ . V nadaljevanju bomo predpostavljali, da je funkcija  $f$  v prvi spremenljivki zvezna, v drugi pa Lipschitzovo zvezna (zadosten pogoj za to je, da je zvezno odvedljiva), kar zagotavlja, da funkcija  $y$  obstaja in je enolična, vsaj pri izbiri dovolj majhnega zamika  $\varepsilon$ .

#### 4.1.1 Eulerjevi metodi

Numerično reševanje robnega problema lahko zastavimo tako, da poiščemo približke za vrednosti rešitve  $y$  v izbranih točkah, desno od  $x_0$ .

Največkrat izberemo kar fiksen razmik  $h > 0$  in nato iz začetnega pogoja  $y_0$ , ki je predpisan v točki  $x_0$ , rekursivno izračunamo približke  $y_n$  za vrednosti  $y(x_n)$  v točkah

$$x_n = x_{n-1} + h, \quad n = 1, 2, \dots$$

Kadar  $y_n$  izračunamo le na podlagi prejšnjega približek  $y_{n-1}$  po formuli oblike

$$y_n = y_{n-1} + h\psi(x_{n-1}, y_{n-1}, h), \quad n = 1, 2, \dots, \quad (4.2)$$

govorimo o enočlenski metodi. Funkciji  $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}$  pravimo numerični odvod.

Denimo, da za funkcijo  $\psi$  velja

$$\psi(x_{n-1}, y_{n-1}, 0) = f(x_{n-1}, y_{n-1}). \quad (4.3)$$

Po nastavku (4.2) je

$$\frac{y_n - y_{n-1}}{h} = \psi(x_{n-1}, y_{n-1}, h)$$

in če ta izraz obravnavamo v limiti, ko gre  $h$  proti 0, dobimo ravno enačbo (4.1). Če je  $h$  dovolj majhen, lahko torej upravičeno pričakujemo, da zvezna odsekoma linearna funkcija, ki interpolira vrednosti  $y_n$  v točkah  $x_n$ , dobro aproksimira točno rešitev  $y$ .

Snovanje enočlenske metode za reševanje začetnega problema ni torej nič drugega kot izpeljava ustreznega izraza za numerični odvod  $\psi$ . Za začetek si oglejmo preprosto metodo, ki temelji na uporabi preme difference (3.3).

Enačbo (4.1) iz vrednotimo pri  $x_{n-1}$ . Ob predpostavki, da je rešitev  $y$  dvakrat zvezno odvedljiva funkcija, po trditvi 3.1 obstaja  $\xi \in (x_{n-1}, x_n)$ , da velja

$$y'(x_{n-1}) = \frac{y(x_n) - y(x_{n-1})}{h} - \frac{1}{2}hy''(\xi) = f(x_{n-1}, y(x_{n-1})).$$

Torej je

$$y(x_n) = y(x_{n-1}) + hf(x_{n-1}, y(x_{n-1})) + \mathcal{O}(h^2). \quad (4.4)$$

Denimo, da že poznamo približek  $y_{n-1}$  za vrednost  $y(x_{n-1})$ . Potem lahko, podprti z zgornjo relacijo, približek  $y_n$  za  $y(x_n)$  izračunamo po formuli

$$y_n = y_{n-1} + hf(x_{n-1}, y_{n-1}). \quad (4.5)$$

To je enačba oblike (4.2) pri izbiri  $\psi(x, y, h) = f(x, y)$ . Gre za metodo, ki je znana pod imenom eksplisitna Eulerjeva metoda.

Izpeljano metodo lahko interpretiramo geometrijsko. Če privzamemo, da je približek  $y_{n-1}$  enak točni vrednosti  $y(x_{n-1})$ , je vrednost  $f(x_{n-1}, y_{n-1})$  v (4.5) po enačbi (4.1) enaka odvodu  $y'(x_{n-1})$  oziroma koeficientu tangente na graf rešitve  $y$  v točki  $x_{n-1}$ . Torej novi približek  $y_n$  izračunamo tako, da se iz točke  $(x_{n-1}, y_{n-1})$  pomaknemo vzdolž tangente do točke  $(x_n, y_n)$ . Napaki  $y(x_n) - y_n$ , ki jo pri tem napravimo, rečemo lokalna napaka. Iz (4.4) je razvidno, da je enaka  $\mathcal{O}(h^2)$ .

Dejanska napaka  $y(x_n) - y_n$  brez privzetka  $y_{n-1} = y(x_{n-1})$  se imenuje globalna napaka. V globalni napaki se lokalne napake tekom rekurzivnega računanja približkov akumulirajo, zato je praviloma večja od lokalne napake. Na srečo pa analiza lokalne napake zadošča za določitev globalne napake. Velja zveza, da je globalna napaka za red večja od lokalne, v tem primeru torej  $\mathcal{O}(h)$ .

**Izrek 4.1.** *Naj bo metoda za reševanje začetnega problema, ki ga določa enačba (4.1), podana v obliki (4.2). Denimo, da je funkcija  $\psi$  zvezna v prvi in tretji spremenljivki ter Lipshitzovo zvezna v drugi spremenljivki, poleg tega pa zadošča pogoju (4.3). Če je lokalna napaka metode  $\mathcal{O}(h^{r+1})$ ,  $r > 1$ , je globalna napaka  $\mathcal{O}(h^r)$ .*

Na podoben način kot eksplisitno Eulerjevo metodo lahko izpeljemo še metodo, ki temelji na obratni diferenci (3.4). V tem primeru enačbo (4.1) izvednotimo v  $x_n$ , po trditvi 3.2 sklepamo

$$y'(x_n) = \frac{y(x_n) - y(x_{n-1}))}{h} + \frac{1}{2}hy''(\xi) = f(x_n, y(x_n))$$

in na podlagi tega zastavimo formulo

$$y_n = y_{n-1} + hf(x_n, y_n), \quad (4.6)$$

ki predstavlja implicitno Eulerjevo metodo. Tudi tu je lokalna napaka  $\mathcal{O}(h^2)$ , globalna pa zato  $\mathcal{O}(h)$ . Geometrijsko gledano približek  $y_n$  izračunamo tako, da velja, da se iz točke  $(x_n, y_n)$  vzdolž tangente lahko pomaknemo nazaj v točko  $(x_{n-1}, y_{n-1})$ .

**Naloga 4.1.** Naj bo začetni problem podan s funkcijo  $f(x, y) = x + y - 1$  in začetnim pogojem  $y_0 = 1$  v točki  $x_0 = 0$ . Poiščite numerično rešitev začetnega problema na intervalu  $[0, 1]$  s pomočjo eksplisitne in implicitne Eulerjeve metode pri koraku  $h = 1/4$ . Rezultate primerjajte s točno rešitvijo.

*Rešitev.* Z eksplisitno Eulerjevo metodo (4.5) dobimo približke

$$\begin{aligned} y_1 &= y_0 + hf(x_0, y_0) = 1 + \frac{1}{4}(0 + 1 - 1) = 1, \\ y_2 &= y_1 + hf(x_1, y_1) = 1 + \frac{1}{4}\left(\frac{1}{4} + 1 - 1\right) = \frac{17}{16} \doteq 1.0625, \\ y_3 &= y_2 + hf(x_2, y_2) = \frac{17}{16} + \frac{1}{4}\left(\frac{1}{2} + \frac{17}{16} - 1\right) = \frac{77}{64} \doteq 1.2031, \\ y_4 &= y_3 + hf(x_3, y_3) = \frac{77}{64} + \frac{1}{4}\left(\frac{3}{4} + \frac{77}{64} - 1\right) = \frac{77}{64} + \frac{61}{256} = \frac{369}{256} \doteq 1.4414. \end{aligned}$$

Ker je predpis za funkcijo  $f$  enostaven, lahko približek pri implicitni metodi (4.6) eksplisitno izrazimo. Iz formule

$$y_n = y_{n-1} + h(x_n + y_n - 1)$$

sledi

$$y_n = \frac{1}{1-h}(y_{n-1} + h(x_n - 1)),$$

torej so približki implicitne Eulerjeve metode dani z

$$\begin{aligned} y_1 &= \frac{4}{3}\left(1 + \frac{1}{4}\left(\frac{1}{4} - 1\right)\right) = \frac{13}{12} \doteq 1.0833, \\ y_2 &= \frac{4}{3}\left(\frac{13}{12} + \frac{1}{4}\left(\frac{1}{2} - 1\right)\right) = \frac{23}{18} \doteq 1.2778, \\ y_3 &= \frac{4}{3}\left(\frac{23}{18} + \frac{1}{4}\left(\frac{3}{4} - 1\right)\right) = \frac{175}{108} \doteq 1.6204 \\ y_4 &= \frac{4}{3}\left(\frac{175}{108} + \frac{1}{4}(1 - 1)\right) = \frac{175}{81} \doteq 2.1605. \end{aligned}$$

Točna rešitev enačbe je  $y(x) = Ce^x - x$  za neko konstanto  $C$ . Če upoštevamo začetni pogoj, dobimo  $y(x) = e^x - x$ . Vrednosti v točkah  $1/4, 1/2, 3/4, 1$  so

$$y(\tfrac{1}{4}) \doteq 1.0340, \quad y(\tfrac{1}{2}) \doteq 1.1487, \quad y(\tfrac{3}{4}) \doteq 1.3670, \quad y(1) \doteq 1.7183.$$

Opazimo, da so približki, dobljeni z eksplisitno metodo manjši, približki, dobljeni z implicitno metodo pa večji od točnih vrednosti.  $\square$

Računanje približkov po formuli (4.6) je očitno bolj zapleteno kot po formuli (4.5), saj  $y_n$  nastopa tako na levi kot desni strani enačbe. Funkcija  $f$  v enačbi (4.1) običajno ni linearna, zato je treba v vsakem koraku metode rešiti nelinearno enačbo. Pogosto lahko posežemo kar po navadni iteraciji: približek  $y_n$  izračunamo tako, da za začetni približek  $y_n^{(0)} = y_{n-1} + hf(x_{n-1}, y_{n-1})$  vzamemo vrednost, izračunano po eksplisitni Eulerjevi metodi, nato pa v nekaj iteracijah

$$y_n^{(r)} = y_{n-1} + hf(x_n, y_n^{(r-1)}), \quad r = 1, 2, \dots,$$

izračunamo rešitev  $y_n \approx y_n^{(k)}$  nelinearne enačbe. Pri tem je  $k$  najmanjše naravno število, za katerega velja, da je vrednost  $|y_n^{(k)} - y_n^{(k-1)}|$  manjša od neke predpisane tolerance. Ta način ni vedno zanesljiv. Kadar eksplisitna Eulerjeva metoda ni konvergentna, je treba uporabiti kako bolj robustno metodo za reševanje nelinearnih enačb.

**Naloga 4.2.** V Matlabu implementirajte eksplisitno in implicitno Eulerjevo metodo.

1. Metodi naj za vhodne podatke sprejmeta seznam točk  $(x_0, x_1, \dots)$ , v katerih računamo približke, funkcijo  $f$ , ki določa diferencialno enačbo, in vrednost  $y_0$ , ki podaja začetni pogoj v točki  $x_0$ . Implicitna metoda naj dodatno sprejme še parameter, ki določa toleranco pri reševanju nelinearne enačbe z navadno iteracijo na vsakem koraku metode. Obe metodi naj vrneta zaporedje izračunanih približkov.
2. S pomočjo implementiranih metod izračunajte približke za rešitev začetnega problema, ki je podan s funkcijo  $f(x, y) = y + 15e^x \cos(15x)$  in začetnim pogojem  $y_0 = 0$  v točki  $x = 0$ . Točna rešitev problema je  $y(x) = e^x \sin(15x)$ . Obravnavajte globalno napako v točkah delitve intervala  $[0, 1]$  pri razmikih  $h = 0.1 \cdot 2^{-r}$ ,  $r \in \{0, 1, \dots, 9\}$ .

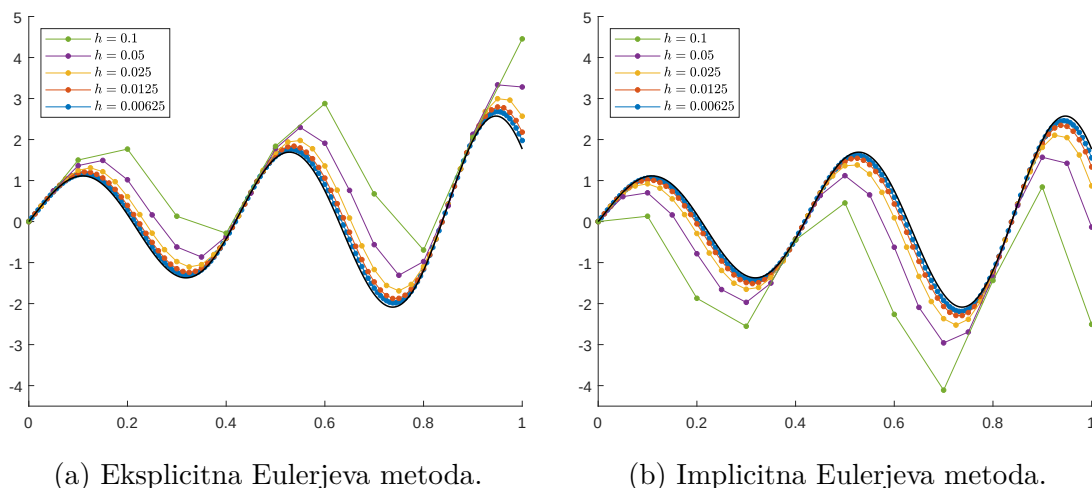
*Rešitev.* Izračunani približki začetnega problema z Eulerjevima metodama pri razmikih  $h$  za  $r = 0, 1, 2, 3, 4$  so prikazani na sliki 4.1, globalne napake v točkah delitve intervala  $[0, 1]$  pa vsebuje tabela 4.1.  $\square$

Izračun približka po formuli (4.6) je zaradi implicitne narave metode računsko zahtevnejši, kar se zdi z obzirom na to, da imata obe metodi globalno napako  $\mathcal{O}(h)$ , neupravičeno. A izkaže se, da ima implicitna posebno lastnost, ki ji pravimo A-stabilnost.

A-stabilnost metod za reševanje diferencialnih enačb pogosto testiramo tako, da preverimo, ali se numerični približki  $y_n$  za rešitev posebnega začetnega problema

$$y'(x) = \lambda y(x), \quad y(0) = y_0, \quad y_0 \neq 0, \quad \lambda < 0, \quad (4.7)$$

ko gre  $n$  proti neskončno, ujemajo z vrednostmi točne rešitve  $y(x)$ , ko gre  $x$  proti neskončno.



Slika 4.1: Približki za rešitev robnega problema iz naloge 4.2, izračunani z Eulerjevima metodama. Graf točne rešitve je narisan s črno barvo.

| $r$ | eksplicitna Eulerjeva metoda | implicitna Eulerjeva metoda |
|-----|------------------------------|-----------------------------|
| 0   | $2.68646 \cdot 10^0$         | $4.27361 \cdot 10^0$        |
| 1   | $1.51368 \cdot 10^0$         | $1.90205 \cdot 10^0$        |
| 2   | $8.01707 \cdot 10^{-1}$      | $8.98294 \cdot 10^{-1}$     |
| 3   | $4.12453 \cdot 10^{-1}$      | $4.36569 \cdot 10^{-1}$     |
| 4   | $2.09181 \cdot 10^{-1}$      | $2.15208 \cdot 10^{-1}$     |
| 5   | $1.05336 \cdot 10^{-1}$      | $1.06843 \cdot 10^{-1}$     |
| 6   | $5.28555 \cdot 10^{-2}$      | $5.32322 \cdot 10^{-2}$     |
| 7   | $2.64747 \cdot 10^{-2}$      | $2.65689 \cdot 10^{-2}$     |
| 8   | $1.32491 \cdot 10^{-2}$      | $1.32727 \cdot 10^{-2}$     |
| 9   | $6.62750 \cdot 10^{-3}$      | $6.63338 \cdot 10^{-3}$     |

Tabela 4.1: Napake Eulerjevih metod pri reševanju robnega problema iz naloge 4.2.

**Naloga 4.3.** Izpeljite splošni formuli za numerični približek  $y_n$  pri reševanju začetnega problema (4.7) po eksplisitni in implicitni Eulerjevi metodi. Ugotovite, pri katerih razmikih  $h$  velja  $\lim_{n \rightarrow \infty} y_n = \lim_{x \rightarrow \infty} y(x)$ .

*Rešitev.* Točna rešitev začetnega problema je  $y(x) = y_0 e^{\lambda x}$ . Če pošljemo  $x$  proti neskončno, je vrednost  $y(x)$  v limiti enaka 0.

Pri eksplisitni metodi za približek  $y_n$  sklepamo

$$y_n = y_{n-1} + \lambda h y_{n-1} = (1 + \lambda h) y_{n-1} = (1 + \lambda h)^2 y_{n-2} = \dots = (1 + \lambda h)^n y_0,$$

kar pomeni, da je  $\lim_{n \rightarrow \infty} y_n = 0$  le v primeru, ko je  $|1 + \lambda h| < 1$ . To omejuje izbiro koraka  $h$  na interval  $(0, -2/\lambda)$ . Metoda zato ni stabilna.

Pri implicitni metodi je

$$y_n = y_{n-1} + \lambda h y_n,$$

od koder izrazimo približek  $y_n$  v odvisnosti od  $y_0$  kot

$$y_n = \frac{1}{1 - \lambda h} y_{n-1} = \frac{1}{(1 - \lambda h)^2} y_{n-2} = \dots = \frac{1}{(1 - \lambda h)^n} y_0.$$

Sedaj je  $\lim_{n \rightarrow \infty} y_n = 0$  neodvisno od  $h$  in metodo smatramo za stabilnejšo.  $\square$

### 4.1.2 Runge–Kutta metode

Eulerjevi metodi (4.5) in (4.6) slonita na ideji, da novi približek izračunamo tako, da že izračunanemu približku prištejemo premik v smeri, ki ga določa diferencialna enačba. Če ta koncept ohranimo, a namesto enega premika hkrati uporabimo več premikov, dobimo širok razred enočlenskih metod, ki se imenujejo Runge–Kutta metode.

Pri Runge–Kutta metodah približek  $y_n$  za točno vrednost  $y(x_n)$  rešitve začetnega problema  $y$  v točki  $x_n$  iščemo na podlagi formule (4.2), pri kateri numerični odvod  $\psi$  podamo s predpisom

$$\psi(x_{n-1}, y_{n-1}, h) = \sum_{i=1}^s \gamma_i k_i(x_{n-1}, y_{n-1}, h). \quad (4.8a)$$

Pri tem je  $s$  stopnja metode, to je število pomožnih odvodov  $k_i$ ,  $i = 1, 2, \dots, s$ , ki jih uporabimo za izračun približka. Pomožni odvodi so uteženi s koeficienti  $\gamma_i$ , za katere navadno velja, da se seštejejo v 1. To pomeni, da vsota  $\sum_{i=1}^s \gamma_i k_i$  predstavlja povprečje odvodov  $k_i$ .

Odvodi  $k_i$  so v splošnem definirani s

$$k_i(x_{n-1}, y_{n-1}, h) = f \left( x_{n-1} + \alpha_i h, y_{n-1} + h \sum_{j=1}^s \beta_{i,j} k_j(x_{n-1}, y_{n-1}, h) \right) \quad (4.8b)$$

in so odvisni od izbire parametrov  $\alpha_i$  in  $\beta_{i,j}$ . Parametri  $\alpha_i$  določajo delež premika od  $x_{n-1}$  proti  $x_n$ , parametri  $\beta_{i,j}$  pa delež premika od  $y_{n-1}$  proti  $y_n$ .

Parametre Runge–Kutta metode ekonomično predstavimo z Butcherjevo shemo

|            |               |               |          |               |
|------------|---------------|---------------|----------|---------------|
| $\alpha_1$ | $\beta_{1,1}$ | $\beta_{1,2}$ | $\dots$  | $\beta_{1,s}$ |
| $\alpha_2$ | $\beta_{2,1}$ | $\beta_{2,2}$ | $\dots$  | $\beta_{2,s}$ |
| $\vdots$   | $\vdots$      | $\vdots$      | $\ddots$ | $\vdots$      |
| $\alpha_s$ | $\beta_{s,1}$ | $\beta_{s,2}$ | $\dots$  | $\beta_{s,s}$ |
|            | $\gamma_1$    | $\gamma_2$    | $\dots$  | $\gamma_s$    |

Ob nastavku (4.8) je s to shemo metoda enolično podana.

**Primer 4.1.** EksPLICITNA in IMPLICITNA Eulerjeva metoda sta primera Runge–Kutta metod stopnje  $s = 1$ . Iz (4.5) je razvidno, da eksplcitna Eulerjeva metoda ustreza izbiri  $\alpha_1 = 0$ ,  $\beta_{1,1} = 0$  in  $\gamma_1 = 1$ . Implicitno Eulerjevo metodo (4.6) pa dobimo, če vzamemo  $\alpha_1 = 1$ ,  $\beta_{1,1} = 1$  in  $\gamma_1 = 1$ , saj je

$$y_n = y_{n-1} + hf(x_n, y_n) = y_{n-1} + hf(x_n, y_{n-1} + hf(x_n, y_n)).$$

Pripadajoči Butcherjevi shemi sta

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array}, \quad \begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array}.$$

Izpeljava obeh Eulerjevih metod je slonela na uporabi pravil za odvajanje. Določimo na podoben način še metodo, ki temelji na simetrični diferenci, in premislimo, da spada pod okrilje Runge–Kutta metod.

**Primer 4.2** (Izboljšana Eulerjeva metoda). Diferencialno enačbo (4.1) izvednotimo v točki  $x_{n-1} + h/2$ , nato pa odvod  $y'(x_{n-1} + h/2)$  ob predpostavki, da je rešitev  $y$  trikrat zvezno odvedljiva, nadomestimo s pomočjo formule (3.7) z uporabo razmika  $h/2$ . Po trditvi 3.3 obstaja  $\xi \in (x_{n-1}, x_n)$ , da velja

$$y'(x_{n-1} + \tfrac{1}{2}h) = \frac{y(x_n) - y(x_{n-1}))}{h} - \frac{1}{24}h^2 y'''(\xi) = f(x_{n-1} + \tfrac{1}{2}h, y(x_{n-1} + \tfrac{1}{2}h)).$$

To še ne zadošča za vpeljavo metode, ki temelji na približkih  $y_{n-1}$  in  $y_n$ , a težavo lahko prebrodimo z razvojem vrednosti

$$y(x_{n-1} + \tfrac{1}{2}h) = y(x_{n-1}) + \frac{1}{2}hy'(x_{n-1}) + \mathcal{O}(h^2) = y(x_{n-1}) + \frac{1}{2}hf(x_{n-1}, y(x_{n-1})) + \mathcal{O}(h^2)$$

v Taylorjevo vrsto okoli točke  $x_{n-1}$ . Sedaj z vpeljavo premikov

$$\begin{aligned} k_1(x_{n-1}, y_{n-1}, h) &= f(x_{n-1}, y_{n-1}), \\ k_2(x_{n-1}, y_{n-1}, h) &= f(x_{n-1} + \tfrac{1}{2}h, y_{n-1} + \tfrac{1}{2}hk_1) \end{aligned}$$

definiramo tako imenovano izboljšano Eulerjevo metodo

$$y_n = y_{n-1} + hk_2(x_{n-1}, y_{n-1}, h),$$

ki se glede na zgornjo izpeljavo ponaša z lokalno napako  $\mathcal{O}(h^3)$ . Gre za dvostopenjsko Runge–Kutta metodo, podano z Butcherjevo shemo

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ \hline & 0 & 1 \end{array}.$$

Poleg pravil za odvajanje lahko za izpeljavo nekaterih metod za začetne probleme izkoristimo tudi integracijska pravila.

**Primer 4.3** (Trapezna metoda). Integrirajmo enačbo (4.1) v mejah od  $x_{n-1}$  od  $x_n$ . Dobimo

$$y(x_n) - y(x_{n-1}) = \int_{x_{n-1}}^{x_n} y'(x) dx = \int_{x_{n-1}}^{x_n} f(x, y(x)) dx.$$

Na desni strani uporabimo trapezno pravilo (3.11). Če je funkcija  $x \mapsto f(x, y(x))$  dvakrat zvezno odvedljiva, po trditvi 3.5 obstaja  $\xi \in (x_{n-1}, x_n)$ , da velja

$$y(x_n) - y(x_{n-1}) = \frac{1}{2}h(f(x_{n-1}, y(x_{n-1})) + f(x_n, y(x_n))) - \frac{1}{12}h^3 f''(\xi, y(\xi)).$$

Očitno smo spet na sledi metodi z lokalno napako  $\mathcal{O}(h^3)$ . Definiramo jo s formulo

$$y_n = y_{n-1} + \frac{1}{2}h(f(x_{n-1}, y_{n-1}) + f(x_n, y_n)),$$

ki kaže, da gre za implicitno metodo. V kontekstu Runge–Kutta metode jo podamo kot

$$y_n = y_{n-1} + h \left( \frac{1}{2} k_1(x_{n-1}, y_{n-1}, h) + \frac{1}{2} k_2(x_{n-1}, y_{n-1}, h) \right)$$

s premikoma

$$k_1(x_{n-1}, y_{n-1}, h) = f(x_{n-1}, y_{n-1}),$$

$$k_2(x_{n-1}, y_{n-1}, h) = f\left(x_{n-1} + h, y_{n-1} + h \left( \frac{1}{2} k_1(x_{n-1}, y_{n-1}, h) + \frac{1}{2} k_2(x_{n-1}, y_{n-1}, h) \right)\right).$$

Pripadajoča Butcherjeva shema je

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1/2 & 1/2 \\ \hline & 1/2 & 1/2 \end{array}.$$

Globalni napaki metod iz primerov (4.2) in (4.3) sta po izreku 4.1 enaki  $\mathcal{O}(h^2)$ , kar je za red bolje kot pri eksplicitni in implicitni Eulerjevi metodi. Z višanjem stopnje Runge–Kutta metod lahko red globalne napake še izboljšamo. Morda najbolj znamenit primerek iz razreda Runge–Kutta metod je štiristopenjska metoda RK4 z globalno napako  $\mathcal{O}(h^4)$ .

**Primer 4.4** (RK4). Metoda RK4 je eksplicitna metoda, ki je podana z Butcherjevo shemo

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline & 1/6 & 2/6 & 2/6 & 1/6 \end{array}.$$

Za lokalno napako se izkaže, da je enaka  $\mathcal{O}(h^5)$ , kar je mogoče dokazati s potrpežljivim razvojem približka v Taylorjevo vrsto.

**Naloga 4.4.** V Matlabu implementirajte Runge–Kutta metode iz primerov 4.2, 4.3 in 4.4 ter jih preizkusite na primeru iz naloge 4.2.

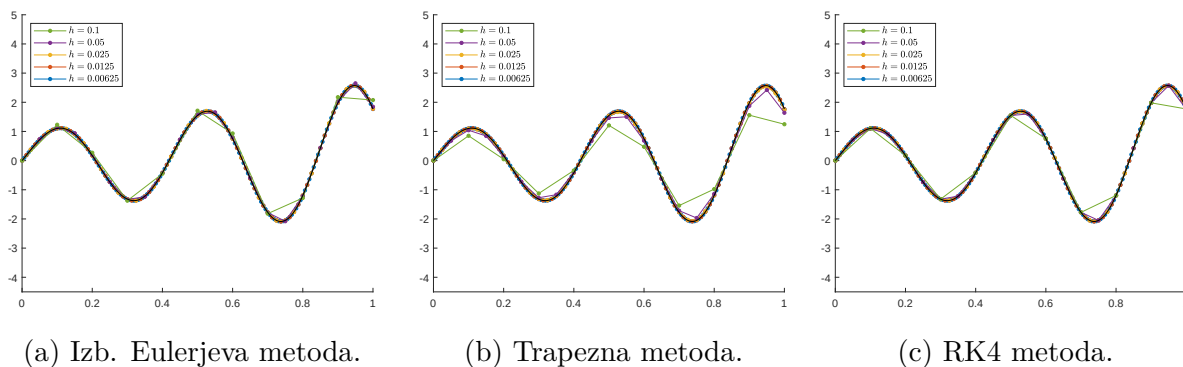
*Rešitev.* Izračunani približki začetnega problema s tremi Runge–Kutta metodami pri razmikih  $h = 0.1 \cdot 2^{-r}$  za  $r \in \{0, 1, 2, 3, 4\}$  so prikazani na sliki 4.2, globalne napake v točkah delitve intervala  $[0, 1]$  za  $r \in \{0, 1, \dots, 9\}$  pa vsebuje tabela 4.2.  $\square$

### 4.1.3 Sistemi diferencialnih enačb in enačbe višjega reda

Vse metode, izpeljane v prejšnjih dveh razdelkih, je mogoče brez težav uporabiti tudi za robne probleme, v katerih nastopajo sistemi navadnih diferencialnih enačb.

Enačbo (4.1) posplošimo v sistem

$$\mathbf{y}'(x) = \mathbf{f}(x, \mathbf{y}(x)), \quad (4.9)$$



Slika 4.2: Približki za rešitev robnega problema iz naloge 4.4, izračunani z Runge–Kutta metodami iz primerov 4.2, 4.3 in 4.4. Graf točne rešitve je narisana s črno barvo.

| $r$ | izb. Eulerjeva metoda   | trapezna metoda         | RK4 metoda               |
|-----|-------------------------|-------------------------|--------------------------|
| 0   | $3.08704 \cdot 10^{-1}$ | $5.20592 \cdot 10^{-1}$ | $3.72130 \cdot 10^{-3}$  |
| 1   | $8.00794 \cdot 10^{-2}$ | $1.46998 \cdot 10^{-1}$ | $2.88253 \cdot 10^{-4}$  |
| 2   | $2.16084 \cdot 10^{-2}$ | $3.65165 \cdot 10^{-2}$ | $1.78031 \cdot 10^{-5}$  |
| 3   | $5.48095 \cdot 10^{-3}$ | $9.21925 \cdot 10^{-3}$ | $1.10964 \cdot 10^{-6}$  |
| 4   | $1.38087 \cdot 10^{-3}$ | $2.30394 \cdot 10^{-3}$ | $6.93113 \cdot 10^{-8}$  |
| 5   | $3.46594 \cdot 10^{-4}$ | $5.75930 \cdot 10^{-4}$ | $4.33490 \cdot 10^{-9}$  |
| 6   | $8.68238 \cdot 10^{-5}$ | $1.44005 \cdot 10^{-4}$ | $2.70927 \cdot 10^{-10}$ |
| 7   | $2.17281 \cdot 10^{-5}$ | $3.60010 \cdot 10^{-5}$ | $1.69336 \cdot 10^{-11}$ |
| 8   | $5.43479 \cdot 10^{-6}$ | $9.00024 \cdot 10^{-6}$ | $1.06448 \cdot 10^{-12}$ |
| 9   | $1.35905 \cdot 10^{-6}$ | $2.25007 \cdot 10^{-6}$ | $7.10543 \cdot 10^{-14}$ |

Tabela 4.2: Napake Runge–Kutta metod pri reševanju robnega problema iz naloge 4.4.

ki je določen s preslikavo  $\mathbf{f} : \mathbb{R}^{1+k} \rightarrow \mathbb{R}^k$ ,  $k \in \mathbb{N}$ . Rešitev sistema enačb je preslikava  $\mathbf{y} : \mathbb{R} \rightarrow \mathbb{R}^k$ . Če preslikavi predstavimo po koordinatah kot  $\mathbf{f} = (f_1, f_2, \dots, f_k)$  in  $\mathbf{y} = (y_1, y_2, \dots, y_k)$ , lahko sistem (4.9) zapišemo bolj nazorno v obliki

$$\begin{aligned}
 y_1'(x) &= f_1(x, y_1(x), y_2(x), \dots, y_k(x)), \\
 y_2'(x) &= f_2(x, y_1(x), y_2(x), \dots, y_k(x)), \\
 &\vdots \\
 y_k'(x) &= f_k(x, y_1(x), y_2(x), \dots, y_k(x)).
 \end{aligned}$$

Začetni problem dobimo, če sistem enačb opremimo še s pogoji  $\mathbf{y}(x_0) = \mathbf{y}_0 \in \mathbb{R}^k$  v začetni točki  $x_0$ .

Splošno enočlensko metodo za tovrstne začetne probleme zapišemo v obliki

$$\mathbf{y}_n = \mathbf{y}_{n-1} + h\boldsymbol{\psi}(x_{n-1}, \mathbf{y}_{n-1}, h), \quad n = 1, 2, \dots$$

Ta rekurzivna formula opisuje, kako iz približka  $\mathbf{y}_{n-1} \in \mathbb{R}^k$  za vrednost  $\mathbf{y}(x_{n-1})$  izračunamo približek  $\mathbf{y}_n \in \mathbb{R}^k$  za vrednost  $\mathbf{y}(x_n)$  v točki  $x_n = x_{n-1} + h$  pri izbranem razmiku  $h > 0$ . Bistvene lastnosti metode odraža numerični odvod  $\boldsymbol{\psi} : \mathbb{R}^{1+k+1} \rightarrow \mathbb{R}^n$ . Vse Runge–Kutta metode lahko torej enostavno posplošimo na ta model tako, da operacije izvajamo vektorsko.

Tako posplošene metode omogočajo tudi reševanje začetnega problema z enačbo višjega reda oblike

$$y^{(k)}(x) = f(x, y(x), y'(x), \dots, y^{(k-1)}(x)),$$

ob kateri so predpisani še robni pogoji

$$y(x_0) = y_0, \quad y'(x_0) = y'_0, \quad \dots, \quad y^{(k-1)}(x_0) = y_0^{(k-1)}.$$

Navedeno diferencialno enačbo višjega reda lahko namreč z uvedbo funkcijskih spremenljivk  $y_1 = y, y_2 = y', \dots, y_k = y^{(k-1)}$  prepisemo v sistem navadnih diferencialnih enačb prvega reda

$$\begin{aligned} y_1'(x) &= y_2(x), \\ y_2'(x) &= y_3(x), \\ &\vdots \\ y_k'(x) &= f(x, y_1(x), y_2(x), \dots, y_k(x)) \end{aligned}$$

in s tem problem prevedemo na reševanje robnega problema pri pogojih

$$y_1(x_0) = y_0, \quad y_1(x_0) = y'_0, \quad \dots, \quad y_k(x_0) = y_0^{(k-1)}.$$

Približek za iskano rešitev  $y = y_1$  potemtakem dobimo z vrednostmi, ki ustrezajo prvim komponentam izračunanih približkov  $\mathbf{y}_0, \mathbf{y}_1, \mathbf{y}_2, \dots$

**Naloga 4.5.** Dana je navadna diferencialna enačba drugega reda

$$y''(x) = y'(x)y(x)^2 - y(x), \quad x > 0,$$

pri začetnih pogojih  $y(0) = 1$  in  $y'(0) = 0$ . Prevedite problem na sistem dveh navadnih diferencialnih enačb prvega reda in z uporabo eksplicitne Eulerjeve metode pri razmiku 0.1 izračunajte približka za  $y(0.1)$  in  $y(0.2)$ .

*Rešitev.* Če uvedemo funkcijo  $z = y'$ , lahko enačbo zapišemo kot

$$z'(x) = z(x)y(x)^2 - y(x).$$

To je sistem dveh navadnih diferencialnih enačb

$$\begin{bmatrix} y'(x) \\ z'(x) \end{bmatrix} = \begin{bmatrix} z(x) \\ z(x)y(x)^2 - y(x) \end{bmatrix}$$

pri robnih pogojih  $y(0) = 1$  in  $z(0) = 0$ . Začnemo z  $y_0 = y(0)$  in  $z_0 = z(0)$  ter z uporabo eksplicitne Eulerjeve metode v obeh komponentah dobimo

$$\begin{bmatrix} y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} y_0 \\ z_0 \end{bmatrix} + 0.1 \begin{bmatrix} z_0 \\ z_0 y_0^2 - y_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0.1 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.1 \end{bmatrix}.$$

Nato napravimo še en korak metode in izračunamo

$$\begin{bmatrix} y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ z_1 \end{bmatrix} + 0.1 \begin{bmatrix} z_1 \\ z_1 y_1^2 - y_1 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.1 \end{bmatrix} + 0.1 \begin{bmatrix} -0.1 \\ -1.1 \end{bmatrix} = \begin{bmatrix} 0.99 \\ -0.21 \end{bmatrix}.$$

Približka za vrednosti  $y(0.1)$  in  $y(0.2)$  sta  $y_1 = 1$  in  $y_2 = 0.99$ . □

### 4.1.4 Veččlenske metode

Izračun novega približka pri Runge–Kutta metodah temelji le na uporabi neposrednega prejšnjega približka, a za premik navadno uporabimo več vrednosti funkcije  $f$ . Veččlenske metode se poskušajo temu izogniti z uporabo večjega števila prejšnjih približkov.

Veččlenska metoda s  $k$  členi za reševanje začetnega problema z diferencialno enačbo (4.1) je metoda, ki je v splošni obliki podana z

$$\sum_{i=0}^k \alpha_i y_{n-i} + h \sum_{i=0}^k \beta_i f(x_{n-i}, y_{n-i}) = 0, \quad n = k, k+1, \dots, \quad (4.10)$$

za neke parametre  $\alpha_i, \beta_i$ ,  $i = 0, 1, \dots, k$ . Vrednosti  $y_{n-i}$ ,  $i = 1, 2, \dots, k$ , predstavljajo že izračunane približke za vrednosti točne rešitve  $y$  v točkah  $x_{n-i} = x_n - ih$ ,  $h > 0$ .

Iz (4.10) je razvidno, da za izračun novega približka  $y_n$  uporabimo le vrednosti funkcije  $f$  v točkah, za katere računamo približke, ne pa tudi vrednosti v vmesnih točkah. To je pglavitna prednost veččlenskih metod. Po drugi strani je pomanjkljivost tega pristopa, da moramo za začetek izvajanja tovrstnih metod priskrbeti približke  $y_1, y_2, \dots, y_{k-1}$ , ki jih navadno izračunamo z uporabo kake enočlenske metode.

Kadar je  $\beta_0 = 0$ , lahko iz enačbe (4.10) približek  $y_n$  eksplicitno izrazimo, torej gre za eksplicitno veččlensko metodo. V nasprotnem primeru imamo opravka z implicitno metodo in približek izračunamo z reševanjem nelinearne enačbe. Parametre veččlenskih metod navadno določimo z integracijo navadne diferencialne enačbe (4.1), pri tem pa namesto funkcije  $f$  integriramo interpolacijski polinom.

Konkretno, Adamsove metode izpeljemo tako, da enačbo (4.1) integriramo v mejah od  $x_{n-1}$  do  $x_n$ . Iz

$$\int_{x_{n-1}}^{x_n} y'(x) dx = \int_{x_{n-1}}^{x_n} f(x, y(x)) dx$$

izrazimo

$$y(x_n) = y(x_{n-1}) + \int_{x_{n-1}}^{x_n} f(x, y(x)) dx \quad (4.11)$$

in funkcijo  $x \mapsto f(x, y(x))$  zamenjamo s polinomom  $p \in \mathbb{P}_{k-1}$ , ki interpolira vrednosti funkcije  $f$  v  $k$  točkah. Z interpolacijo v točkah  $x_{n-k}, x_{n-k+1}, \dots, x_{n-1}$  dobimo eksplicitne metode, ki se imenujejo Adams–Bashforthove metode:

- $k = 1$ :  $y_n = y_{n-1} + hf(x_{n-1}, y_{n-1})$ ,
- $k = 2$ :  $y_n = y_{n-1} + \frac{1}{2}h(-f(x_{n-2}, y_{n-2}) + 3f(x_{n-1}, y_{n-1}))$ ,
- $k = 3$ :  $y_n = y_{n-1} + \frac{1}{12}h(5f(x_{n-3}, y_{n-3}) - 16f(x_{n-2}, y_{n-2}) + 23f(x_{n-1}, y_{n-1}))$ .

Interpolacija v točkah  $x_{n-k+1}, x_{n-k+2}, \dots, x_n$  pa vodi k implicitnim metodam, tako imenovanim Adams–Moultonovim metodam:

- $k = 1$ :  $y_n = y_{n-1} + hf(x_n, y_n)$ ,
- $k = 2$ :  $y_n = y_{n-1} + \frac{1}{2}h(f(x_{n-1}, y_{n-1}) + f(x_n, y_n))$ ,
- $k = 3$ :  $y_n = y_{n-1} + \frac{1}{12}h(-f(x_{n-2}, y_{n-2}) + 8f(x_{n-1}, y_{n-1}) + 5f(x_n, y_n))$ .

Metodi za  $k = 1$  ustrežata enočlenskima Eulerjevima metodama. Implicitna metoda za  $k = 2$  je trapezna metoda, obravnavana v primeru 4.3, torej je tudi to pravzaprav enočlenska metoda. Ostale tri metode so prave veččlenske metode. Za vse navedene metode velja, da imajo lokalno napako  $\mathcal{O}(h^{k+1})$ .

**Primer 4.5** (Adams–Bashforthova metoda za  $k = 2$ ). Oglejmo si izpeljavo veččlenske metode, ki temelji na interpolaciji funkcije  $x \mapsto f(x, y(x))$  v točkah  $x_{n-2}$  in  $x_{n-1}$ . Interpolacijski polinom  $p \in \mathbb{P}_1$  je podan s predpisom

$$p(x) = f(x_{n-2}, y(x_{n-2})) + \frac{f(x_{n-1}, y(x_{n-1})) - f(x_{n-2}, y(x_{n-2}))}{h}(x - x_{n-2}).$$

Z integriranjem v mejah od  $x_{n-1}$  do  $x_n$  dobimo

$$\int_{x_{n-1}}^{x_n} p(x) dx = hf(x_{n-2}, y(x_{n-2})) + \frac{3}{2}h^2 \frac{f(x_{n-1}, y(x_{n-1})) - f(x_{n-2}, y(x_{n-2}))}{h}.$$

Če vrednosti  $y(x_{n-2})$  in  $y(x_{n-1})$  nadomestimo s približkoma  $y_{n-2}$  in  $y_{n-1}$  po formuli (4.11) izpeljemo metodo

$$y_n = y_{n-1} + \frac{1}{2}h(-f(x_{n-2}, y_{n-2}) + 3f(x_{n-1}, y_{n-1})).$$

Še en znan razred veččlenskih metod so Milneove metode, ki jih dobimo z integracijo enačbe (4.1) v mejah od  $x_{n-k}$  do  $x_n$ . Iz

$$\int_{x_{n-k}}^{x_n} y'(x) dx = \int_{x_{n-k}}^{x_n} f(x, y(x)) dx$$

izrazimo

$$y(x_n) = y(x_{n-k}) + \int_{x_{n-k}}^{x_n} f(x, y(x)) dx \quad (4.12)$$

in integral aproksimiramo z osnovnim Newton–Cotesovim integracijskim pravilom. Če uporabimo pravilo zaprtega tipa, je rezultat implicitna metoda, če posežemo po pravilu odprtega tipa, pa dobimo eksplicitno metodo. Tako izpeljani metodi, osnovani na istem številu točk v integracijskem pravilu, pogosto uporabljamo skupaj, kot opisuje naslednji primer.

**Primer 4.6** (Milneova metoda prediktor-korektor). V enačbi (4.12) najprej vzamemo  $k = 2$  in vrednosti  $y(x_n)$  in  $y(x_{n-2})$  nadomestimo s približkoma  $y_n$  in  $y_{n-2}$ . Po aproksimaciji integrala s Simpsonovim pravilom (3.13) dobimo implicitno metodo

$$y_n = y_{n-2} + \frac{h}{3}(f(x_{n-2}, y_{n-2}) + 4f(x_{n-1}, y_{n-1}) + f(x_n, y_n)). \quad (4.13a)$$

Za izpeljavo eksplicitne metode, osnovane na integracijskem pravilu s tremi funkcijskimi vrednostmi, v enačbi (4.12) vzamemo  $k = 4$ , vrednosti  $y(x_n)$  in  $y(x_{n-4})$  nadomestimo s približkoma  $y_n$  in  $y_{n-4}$  ter približek za integral določimo po Milneovem pravilu (3.14). Dobimo metodo

$$y_n = y_{n-4} + \frac{4h}{3}(2f(x_{n-3}, y_{n-3}) - f(x_{n-2}, y_{n-2}) + 2f(x_{n-1}, y_{n-1})). \quad (4.13b)$$

Izpeljani metodi lahko uporabimo skupaj tako, da približek

$$y_n^{(P)} = y_{n-4} + \frac{4h}{3} (2f(x_{n-3}, y_{n-3}) - f(x_{n-2}, y_{n-2}) + 2f(x_{n-1}, y_{n-1}))$$

za vrednost rešitve  $y$  v točki  $x_n$  najprej izračunamo z eksplicitno metodo (4.13b), nato pa to vrednost uporabimo na desni strani implicitne metode (4.13a) in izračunamo

$$y_n^{(K_1)} = y_{n-2} + \frac{h}{3} (f(x_{n-2}, y_{n-2}) + 4f(x_{n-1}, y_{n-1}) + f(x_n, y_n^{(P)})).$$

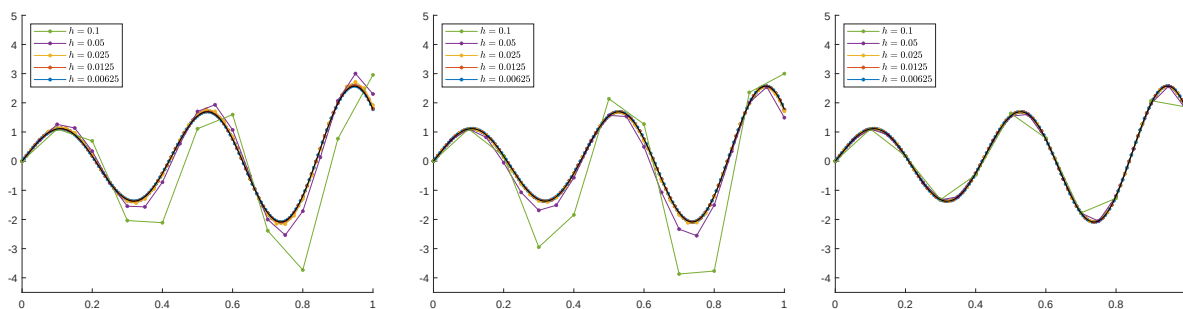
Nato izvajamo navadno iteracijo

$$y_n^{(K_r)} = y_{n-2} + \frac{h}{3} (f(x_{n-2}, y_{n-2}) + 4f(x_{n-1}, y_{n-1}) + f(x_n, y_n^{(K_{r-1})})), \quad r = 2, 3, \dots,$$

dokler za  $k \in \mathbb{N}$  absolutna vrednost razlike  $y_n^{(K_k)} - y_n^{(K_{k-1})}$  ni manjša od neke vnaprej predpisane tolerance. Takrat končamo in za približek  $y_n$  v točki  $x_n$  proglasimo vrednost  $y_n^{(K_k)}$ . Taki metodi rečemo metoda prediktor-korektor, saj z eksplicitno metodo najprej predvidimo, kakšen je dober približek v točki  $x_n$ , nato pa ga z implicitno metodo popravimo.

**Naloga 4.6.** V Matlabu implementirajte dvočlensko in tričlensko Adams-Bashforthovo metodo ter Milneovo metodo prediktor-korektor iz primera 4.6. Testirajte jih na primeru iz naloge 4.2. Za začetne približke uporabite točne vrednosti rešitve robnega problema.

*Rešitev.* Izračunani približki začetnega problema s tremi veččlenskimi metodami pri razmikih  $h = 0.1 \cdot 2^{-r}$  za  $r \in \{0, 1, 2, 3, 4\}$  so prikazani na sliki 4.3, napake v točkah delitve intervala  $[0, 1]$  za  $r \in \{0, 1, \dots, 9\}$  pa vsebuje tabela 4.3.  $\square$



(a) Dvočlenska A.-B. metoda. (b) Tričlenska A.-B. metoda. (c) Milneova metoda.

Slika 4.3: Približki za rešitev robnega problema iz naloge 4.6, izračunani z veččlenskimi metodami. Graf točne rešitve je narisana s črno barvo.

## 4.2 Reševanje robnega problema

Kadar diferencialna enačba (prvega ali višjega reda) ni opremljena (le) s pogojem v začetni točki, pač pa (tudi) s pogoji v drugih delih roba območja, govorimo o robnem problemu. Reševanje takega problema je navadno zahtevnejše od reševanja začetnega problema. Nenazadnje je vprašljiv že obstoj rešitve, ki je običajno zagotovljen le ob zelo specifičnih predpostavkah.

| $r$ | A.–B. dvočlenska metoda | A.–B. tričlenska metoda | Milneova metoda          |
|-----|-------------------------|-------------------------|--------------------------|
| 0   | $2.53800 \cdot 10^0$    | $2.57364 \cdot 10^0$    | $9.54304 \cdot 10^{-2}$  |
| 1   | $5.32452 \cdot 10^{-1}$ | $5.60803 \cdot 10^{-1}$ | $7.78604 \cdot 10^{-3}$  |
| 2   | $1.60498 \cdot 10^{-1}$ | $7.54836 \cdot 10^{-2}$ | $3.86353 \cdot 10^{-4}$  |
| 3   | $4.30269 \cdot 10^{-2}$ | $9.49984 \cdot 10^{-3}$ | $1.73998 \cdot 10^{-5}$  |
| 4   | $1.11562 \cdot 10^{-2}$ | $1.22842 \cdot 10^{-3}$ | $1.26583 \cdot 10^{-6}$  |
| 5   | $2.83547 \cdot 10^{-3}$ | $1.57800 \cdot 10^{-4}$ | $8.70828 \cdot 10^{-8}$  |
| 6   | $7.14451 \cdot 10^{-4}$ | $1.99744 \cdot 10^{-5}$ | $5.68120 \cdot 10^{-9}$  |
| 7   | $1.79312 \cdot 10^{-4}$ | $2.51187 \cdot 10^{-6}$ | $3.62550 \cdot 10^{-10}$ |
| 8   | $4.49148 \cdot 10^{-5}$ | $3.14909 \cdot 10^{-7}$ | $2.28910 \cdot 10^{-11}$ |
| 9   | $1.12395 \cdot 10^{-5}$ | $3.94210 \cdot 10^{-8}$ | $1.44595 \cdot 10^{-12}$ |

Tabela 4.3: Napake veččlenskih metod pri reševanju robnega problema iz naloge 4.6.

### 4.2.1 Diferenčna metoda

Najpogosteje uporabljana metoda za reševanje robnih problemov je diferenčna metoda. Osnovana je na diskretizaciji območja, na katerem je predpisana diferencialna enačba. Enačbo obravnavamo v posameznih točkah diskretizacije in odvode v njej s pomočjo pravil za numerično odvajanje nadomestimo z vrednostmi v več sosednjih točkah diskretizacije. To vzpostavi relacije med približki za vrednosti v točkah diskretizacije in predpisanimi vrednostmi na robu ter naposled pripelje do sistema linearnih ali nelinearnih enačb, z reševanjem katerega dobimo približek za robni problem.

Za zgled si oglejmo diferenčno metodo za robni problem, ki ga določa diferencialna enačba drugega reda oblike

$$-y''(x) = f(x) \quad (4.14)$$

na intervalu  $(a, b)$ . Na robovih intervala predpišemo vrednosti  $y(a)$  in  $y(b)$ .

Interval  $[a, b]$  razdelimo na  $m \in \mathbb{N}$  delov z uvedbo delilnih točk  $x_i = a + ih$ ,  $i = 0, 1, \dots, m$ , ki so ekvidistantno razmaknjene z razmikom  $h = (b - a)/m$ . Enačbo (4.14) izvednotimo v notrajni delilni točki  $x_i$ ,  $i \in \{1, 2, \dots, m - 1\}$ , in drugi odvod v enačbi nadomestimo s simetrično diferenco (3.9). Po trditvi 3.4 je

$$-\frac{y(x_{i-1}) - 2y(x_i) + y(x_{i+1}))}{h^2} + \mathcal{O}(h^2) = f(x_i). \quad (4.15)$$

Vsaki notranji delilni točki  $x_i$ ,  $i = 1, 2, \dots, m - 1$ , priredimo neznanko  $y_i$ , ki predstavlja približek za vrednost  $y(x_i)$ . Vrednosti  $y$  v robnih delilnih točkah  $x_0$  in  $x_m$  sta predpisani z robnimi pogoji, zato lahko vpeljemo še  $y_0 = y(a)$  in  $y_m = y(b)$ . Na podlagi zveze (4.15) vpeljemo  $m - 1$  enačb

$$-y_{i-1} + 2y_i - y_{i+1} = h^2 f(x_i)$$

za neznanke  $y_1, y_2, \dots, y_{m-1}$ . To je sistem linearnih enačb, ki je v matrični obliki podan

z

$$\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m-2} \\ y_{m-1} \end{bmatrix} = h^2 \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_{m-2}) \\ f(x_{m-1}) \end{bmatrix} + \begin{bmatrix} y_0 \\ 0 \\ \vdots \\ 0 \\ y_m \end{bmatrix}. \quad (4.16)$$

Matrika sistema je obrnljiva, zato je sistem enolično rešljiv. Zaradi tridiagonalne strukture matrike je reševanje sistema enostavno in časovne zahtevnosti  $\mathcal{O}(m)$ .

**Naloga 4.7.** Naj bo robni problem na intervalu  $(-1, 1)$  podan z enačbo (4.14), ki jo določa funkcija  $f(x) = 6x$ , ter pogoje  $y(-1) = 1$  in  $y(1) = -1$ . Poiščite približke za rešitev robnega problema z diferenčno metodo v točkah  $-0.5, 0, 0.5$ . Preverite, da se približki ujemajo s točno rešitvijo robnega problema in pojasnite, zakaj je tako.

*Rešitev.* Pri razmiku  $h = 0.5$  uvedemo neznanke  $y_1, y_2, y_3$ , ki predstavljajo približke za vrednosti  $y$  v točkah  $x_1 = -0.5, x_2 = 0, x_3 = 0.5$ . V robnih točkah  $x_0 = -1$  in  $x_4 = 1$  na podlagi robnih pogojev predpišemo vrednosti  $y_0 = 1$  in  $y_4 = -1$ . Glede na (4.16) so neznanke določene z rešitvijo sistema

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} -3 \\ 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Po kratkem izračunu dobimo  $y_1 = 1/8, y_2 = 0, y_3 = -1/8$ . To so vrednosti, ki se ujemajo s točno rešitvijo  $y(x) = -x^3$  v točkah  $x_1, x_2, x_3$ . Razlog za to se skriva v opazki, da je točna rešitev robnega problema kubičen polinom, simetrična diferenca (3.9), na kateri sloni izpeljava diferenčne metode za ta primer, pa po trditvi 3.4 predstavlja točno vrednost drugega odvoda za vse polinome iz  $\mathbb{P}_3$ .  $\square$

Če enačbo (4.14) posplošimo v dve dimenziji, dobimo znamenito Poissonovo enačbo

$$-\frac{\partial^2 u}{\partial x^2}(x, y) - \frac{\partial^2 u}{\partial y^2}(x, y) = f(x, y) \quad (4.17a)$$

na pravokotniku  $(a, b) \times (c, d)$ . Navadno jo zapišemo v kompaktni obliki  $-\Delta u = f$ , kjer  $\Delta$  označuje Laplacevo operator

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}.$$

Enačbo opremimo z robnimi pogoji vzdolž stranic pravokotnika:

$$\begin{aligned} u(x, c) &= g_c(x), & u(x, d) &= g_d(x), & x &\in (a, b), \\ u(a, y) &= g_a(y), & u(b, y) &= g_b(y), & y &\in (c, d), \end{aligned} \quad (4.17b)$$

kjer sta  $g_c, g_d : (a, b) \rightarrow \mathbb{R}$  dani funkciji v spremenljivki  $x$  in  $g_a, g_b : (c, d) \rightarrow \mathbb{R}$  dani funkciji v spremenljivki  $y$ . Naloga je poiskati funkcijo  $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ , ki zadošča (4.17).

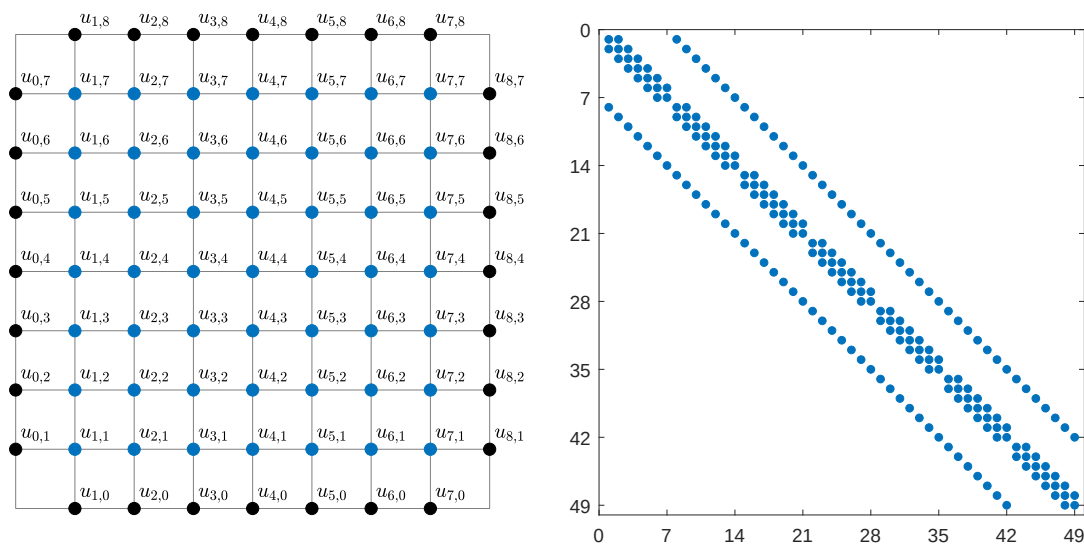
Numeričnega reševanja robnega problema (4.17) z diferenčno metodo se lahko lotimo na povsem enak način kot reševanja robnega problema, ki ga določa enačba (4.14) v eni dimenziji. Delilne točke je treba sedaj predpisati v dveh smereh. Točke

$$x_i = a + ih_x, \quad i = 0, 1, \dots, m_x,$$

delijo interval  $[a, b]$  ekvidistantno na  $m_x \in \mathbb{N}$  delov z razmikom  $h_x = (b - a)/m_x$ , točke

$$y_j = c + jh_y, \quad j = 0, 1, \dots, m_y,$$

pa interval  $[c, d]$  ekvidistantno na  $m_y \in \mathbb{N}$  delov z razmikom  $h_y = (d - c)/m_y$ . Pari  $(x_i, y_j)$  predstavljajo točke v ravnini, ki določajo diskretizacijo pravokotnika  $[a, b] \times [c, d]$  z mrežo.



(a) Diskretizacija kvadrata.

(b) Matrika sistema enačb.

Slika 4.4: Strukturni prikaz diferenčne metode za robni problem iz primera 4.7.

Enačbo (4.17a) iz vrednotimo v notranjih točkah mreže in po zgledu enodimenzionalnega primera uporabimo simetrično diferenco (3.9) za parcialna odvoda v  $x$  in  $y$  smeri. Za  $(i, j) \in \{1, 2, \dots, m_x - 1\} \times \{1, 2, \dots, m_y - 1\}$  vrednosti  $u(x_i, y_j)$  nadomestimo z neznankami  $u_{i,j}$ , vrednosti  $u_{0,j}$  in  $u_{m_x,j}$ ,  $j = 1, 2, \dots, m_y - 1$ , ter  $u_{i,0}$  in  $u_{i,m_y}$ ,  $i = 1, 2, \dots, m_x - 1$ , pa določimo na podlagi robnih pogojev (4.17b). Vpeljane neznanke in vrednosti povežemo z  $(m_x - 1)(m_y - 1)$  zvezami

$$-\frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h_x^2} - \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h_y^2} = f(x_i, y_j),$$

ki določajo razpršen sistem linearnih enačb, saj je vsaka neznanka  $u_{i,j}$  neposredno povezana le s štirimi sosednjimi vrednostmi  $u_{i-1,j}$ ,  $u_{i+1,j}$ ,  $u_{i,j-1}$ ,  $u_{i,j+1}$ . Če izberemo  $m_x$  in  $m_y$  tako, da je razmik  $h$  v  $x$  in  $y$  smeri enak ( $h = h_x = h_y$ ), se relacije poenostavijo v

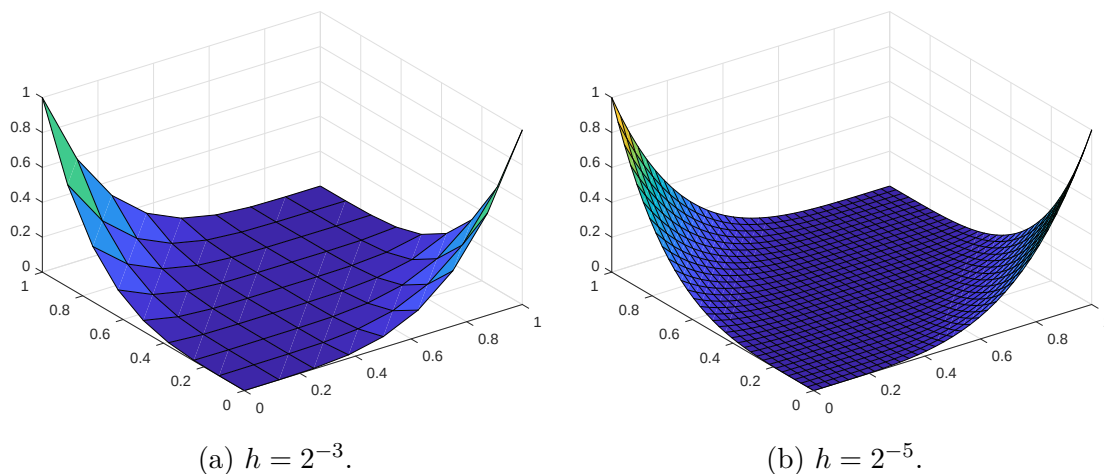
$$-u_{i-1,j} - u_{i,j-1} + 4u_{i,j} - u_{i+1,j} - u_{i,j+1} = h^2 f(x_i, y_j).$$

Neznanke lahko uredimo in vsako opišemo z enojnim indeksom tako, da relacije med njimi določajo matrični sistem s simetrično matriko, ki ima poleg neničelne glavne diagonale le še po dve dodatni neničelni diagonalni pod in nad glavno diagonalo. To ponovno omogoča sorazmerno učinkovito reševanje sistema enačb, tudi kadar sta  $m_x$  in  $m_y$  veliki števili in je mreža, nad katero računamo približke, zelo gosta.

**Primer 4.7.** Naj bo robni problem nad kvadratom  $(0, 1) \times (0, 1)$  podan s Poissonovo enačbo (4.17a), ki jo določa funkcija  $f(x, y) = -24(x - y)^2$ , ter robnimi pogoji oblike (4.17b), ki so podani z

$$\begin{aligned} u(x, 0) &= x^4, & u(x, 1) &= (x - 1)^4, & x &\in (0, 1), \\ u(0, y) &= y^4, & u(1, y) &= (1 - y)^4, & y &\in (0, 1). \end{aligned}$$

Hitro se lahko prepričamo, da je točna rešitev robnega problema podana s predpisom  $u(x, y) = (x - y)^4$ . Z diferenčno metodo lahko z večanjem števila delilnih točk v obeh koordinatnih smereh izračunamo natančne približke za vrednosti  $u$ . Na sliki 4.4 sta za  $h = 2^{-3}$  (to je pri  $m_x = m_y = 8$ ) prikazani mreža, ki določa diskretizacijo kvadrata, skupaj z neznankami, in struktura matrice velikosti  $49 \times 49$ , ki določa sistem, z reševanjem katerega izračunamo približke v točkah mreže. Če izračunane približke povežemo v prostorsko mrežo, dobimo aproksimacijo za graf točne rešitve  $u$ . Taka mreža je za razmika  $h = 2^{-3}$  in  $h = 2^{-5}$  prikazana na sliki 4.5. Tabela 4.4 pa vsebuje maksimalne absolutne napake v točkah mreže pri razmikih  $h = 2^{-(r+1)}$ ,  $r = 0, 1, \dots, 9$ .



Slika 4.5: Približka za rešitev robnega problema iz primera 4.7 po diferenčni metodi.

| $r$ | napaka                  |
|-----|-------------------------|
| 0   | $6.25000 \cdot 10^{-2}$ |
| 1   | $1.75781 \cdot 10^{-2}$ |
| 2   | $4.54891 \cdot 10^{-3}$ |
| 3   | $1.14759 \cdot 10^{-3}$ |
| 4   | $2.87558 \cdot 10^{-4}$ |
| 5   | $7.19308 \cdot 10^{-5}$ |
| 6   | $1.79853 \cdot 10^{-5}$ |
| 7   | $4.49649 \cdot 10^{-6}$ |
| 8   | $1.12413 \cdot 10^{-6}$ |
| 9   | $2.81034 \cdot 10^{-7}$ |

Tabela 4.4: Maksimalne absolutne napake približkov za rešitev robnega problema iz primera 4.7 v točkah diskretnih mrež pri razmikih  $h = 0.5 \cdot 2^{-r}$ .

### 4.2.2 Strelska metoda

Včasih je mogoče robni problem opisati s pomočjo začetnega problema. Če razvijemo metodo, ki poskrbi za prevedbo enega problema na drugega, lahko za reševanje robnega problema s pridom izkoristimo metode, razvite za reševanje začetnih problemov.

Naj bo preslikava  $\mathbf{y} : \mathbb{R} \rightarrow \mathbb{R}^k$ ,  $k \in \mathbb{N}$ , rešitev sistema navadnih diferencialnih enačb (4.9) na intervalu  $(a, b)$  pri začetnem pogoju  $\mathbf{y}(a) = \mathbf{y}_a$ . Predpostavimo, da je preslikava  $\mathbf{f} : \mathbb{R}^{1+k} \rightarrow \mathbb{R}^k$ , ki določa sistem, taka, da rešitev  $\mathbf{y}$  obstaja za poljuben vektor  $\mathbf{y}_a \in \mathbb{R}^k$ . Odvisnost  $\mathbf{y}$  od začetnega pogoja  $\mathbf{y}_a$  ponazorimo tako, da rešitev zapišemo kot funkcijo  $x \mapsto \mathbf{y}(x; \mathbf{y}_a)$ .

Če namesto rešitve sistema diferencialnih enačb (4.9) pri začetnem pogoju iščemo rešitev pri končnem pogoju  $\mathbf{y}(b) = \mathbf{y}_b$ , lahko problem interpretiramo kot iskanje začetnega vektorja  $\mathbf{y}_a \in \mathbb{R}^k$ , da velja

$$\mathbf{y}(b; \mathbf{y}_a) = \mathbf{y}_b.$$

To v splošnem predstavlja sistem nelinearnih enačb, ki ga lahko rešujemo iterativno. Pri približku  $\mathbf{y}_a^{(r)}$  v koraku  $r$  izračunamo rešitev začetnega problema  $x \mapsto \mathbf{y}(x, \mathbf{y}_a^{(r)})$ . Definiramo

$$\mathbf{y}_b^{(r)} = \mathbf{y}(b, \mathbf{y}_a^{(r)})$$

in si ogledamo napako  $\|\mathbf{y}_b - \mathbf{y}_b^{(r)}\|$ . Če je napaka manjša od predpisane tolerance,  $\mathbf{y}$  proglasimo za rešitev robnega problema, sicer na podlagi napake določimo približek  $\mathbf{y}_a^{(r+1)}$  za naslednji korak.

Zgoraj opisani postopek lahko slikovito opišemo kot streljanje na tarčo. Glede na to, kje je izstrelek končal, prilagodimo začetne nastavitve in ustrelimo ponovno, dokler ne zadene cilja. Zato to metodo reševanja robnega problema imenujemo strelska metoda.

Za iskanje rešitve začetnega problema v vsakem koraku strelske metode uporabimo zanesljivo metodo za reševanje začetnega problema. Strelska metoda je torej kompozicija numerične metode za reševanje začetnega problema in iterativne metode za reševanje sistemov nelinearnih enačb. Analiza konvergence in napak pri taki metodi je zato vse prej kot lahek zalogaj.

**Primer 4.8.** V priljubljeni računalniški igri Gorillas iz devetdesetih let se gorili na nebotičnikih metropolisa izmenično obmetujeta z bananami, dokler ena izmed njiju banane ne zaluča pod ustreznim kotom  $\alpha$  in z ustrežno hitrostjo  $v$ , da z njo zadene drugo. Po drugem Newtonovem zakonu za hitrost banane  $\mathbf{v} = (v_x, v_y)$  v vsakem trenutku leta velja

$$\mathbf{F}_u + \mathbf{F}_g = m\dot{\mathbf{v}}, \quad (4.18)$$

kjer  $\mathbf{F}_u$  označuje silo upora zraka na banano,  $\mathbf{F}_g$  gravitacijsko silo na banano,  $m$  maso banane,  $\dot{\mathbf{v}}$  pa odvod hitrosti po času (pospešek banane). Sila upora zraka  $\mathbf{F}_u$  je odvisna od hitrosti  $\mathbf{v}$  in deluje v nasprotni smeri  $\mathbf{v}$ . Za njeno velikost predpostavimo, da velja  $\|\mathbf{F}_u\| = c \|\mathbf{v}\|^2$ , kjer je  $c$  koeficient upora zraka. V kartezičnem koordinatnem sistemu jo lahko torej zapišemo kot

$$\mathbf{F}_u = -\|\mathbf{F}_u\|(\cos(\varphi), \sin(\varphi)) = -\|\mathbf{F}_u\|\left(\frac{v_x}{\|\mathbf{v}\|}, \frac{v_y}{\|\mathbf{v}\|}\right) = -c\|\mathbf{v}\|(v_x, v_y),$$

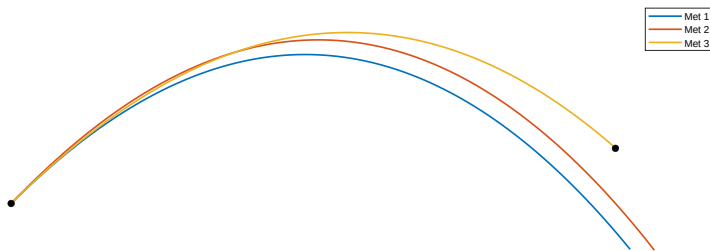
kjer  $\varphi$  označuje kot vektorja hitrosti. Gravitacijska sila  $\mathbf{F}_g$  deluje le v navpični smeri, njena velikost je  $mg$ , kjer je  $g$  gravitacijski pospešek. Enačbo (4.18) lahko potemtakem opišemo v vodoravni in navpični komponenti z enačbama

$$-c v_x \|\mathbf{v}\| = m\dot{v}_x, \quad -c v_y \|\mathbf{v}\| - mg = m\dot{v}_y.$$

Ker gorila poskuša s svojega položaja  $(x_1, y_1)$  zadeti položaj  $(x_2, y_2)$  druge gorile, je soočena z diferencialno enačbo drugega reda v dveh dimenzijah

$$m \cdot (\ddot{x}, \ddot{y}) = -c \sqrt{\dot{x}^2 + \dot{y}^2} \cdot (\dot{x}, \dot{y}) - (0, mg), \quad (4.19)$$

ki opisuje trajektorijo banane  $(x(t), y(t))$  v odvisnosti od časa  $t$  pri robnih pogojih  $(x(0), y(0)) = (x_1, y_1)$  in  $(x(T), y(T)) = (x_2, y_2)$  ob nekem času  $T > 0$ . Igra ni torej nič drugega kot reševanje robnega problema s strelsko metodo (slika 4.6): treba je uganiti, pri katerem začetnem pogoju  $(\dot{x}(0), \dot{y}(0)) = (v \cos(\alpha), v \sin(\alpha))$  je izpolnjen končni pogoj  $(x(T), y(T)) = (x_2, y_2)$ .



Slika 4.6: Trije meti (primer 4.8) z istega položaja, a pri različnih začetnih hitrostih.

**Naloga 4.8.** Zapišite sistem navadnih diferencialnih enačb oblike  $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y})$ , ki določa trajektorijo banane iz primera 4.8 pri danem začetnem položaju in začetni hitrosti.

1. V Matlabu sestavite program, ki uporabniku omogoči vnos hitrosti  $v$  in kota  $\alpha$ , nato pa s pomočjo eksplcitne Eulerjeve metode z razmikom  $h = 0.02$  simulira let banane v območju  $[0, 10] \times [0, 10]$  pri podatkih  $m = 10$ ,  $g = 9.81$ ,  $c = 0.1$  iz koordinatnega izhodišča.
2. V implementaciji programa iz prejšnje točke zamenjajte enočlensko Eulerjevo metodo s tričlensko Adams–Bashforthovo metodo, pri kateri približek  $\mathbf{y}_n$  na koraku  $n$  računamo na podlagi približkov  $\mathbf{y}_{n-1}$ ,  $\mathbf{y}_{n-2}$  in  $\mathbf{y}_{n-3}$  iz prejšnjih treh korakov po formuli

$$\mathbf{y}_n = \mathbf{y}_{n-1} + \frac{1}{12}h (5\mathbf{f}(\mathbf{y}_{n-3}) - 16\mathbf{f}(\mathbf{y}_{n-2}) + 23\mathbf{f}(\mathbf{y}_{n-1})).$$

Približka  $\mathbf{y}_1$  in  $\mathbf{y}_2$  naj bosta (še vedno) izračunana z eksplcitno Eulerjevo metodo.

3. Sestavite metodo, ki z avtomatskim poskušanjem brez uporabnikovega vmešavanja zadane naključno točko v območju  $[0, 10] \times [0, 10]$ . Del pripravljene kode za računanje približkov točk na trajektoriji meta preoblikujte v samostojno funkcijo, ki pri dani začetni hitrosti vrne minimalno razdaljo med izračunanimi približki in ciljno točko. Nato z vgrajeno metodo `fminsearch` poiščite ničlo te funkcije, ki določa začetno hitrost, pri kateri zadenemo ciljno točko. Metoda `fminsearch` izvaja iterativen postopek in vsak korak te iteracije predstavlja en poskus.

4. S prijemi iz prejšnjih točk implementirajte priredbo opisane igre, v kateri uporabnik tekmuje v obmetavanju proti drugemu igralcu ali računalniku.

*Rešitev.* Sistem navadnih diferencialnih enačb drugega reda (4.19) lahko s pomočjo spremenljivk  $v_x = \dot{x}$  in  $v_y = \dot{y}$  zapišemo kot

$$\begin{bmatrix} \dot{x} \\ \dot{v}_x \\ \dot{y} \\ \dot{v}_y \end{bmatrix} = \begin{bmatrix} v_x \\ -\frac{1}{m} \left( c v_x \sqrt{v_x^2 + v_y^2} \right) \\ v_y \\ -\frac{1}{m} \left( c v_y \sqrt{v_x^2 + v_y^2} \right) - g \end{bmatrix}.$$

To je sistem oblike  $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y})$ , če definiramo  $\mathbf{y} = (x, v_x, y, v_y)$  in

$$\mathbf{f}(\mathbf{y}) = \mathbf{f}(x, v_x, y, v_y) = \left( v_x, -\frac{1}{m} \left( c v_x \sqrt{v_x^2 + v_y^2} \right), v_y, -\frac{1}{m} \left( c v_y \sqrt{v_x^2 + v_y^2} \right) - g \right).$$

Začetni pogoj  $\mathbf{y}(0) = \mathbf{y}_0$ , ki enolično določa trajektorijo  $(x(t), y(t))$ ,  $t > 0$ , je podan z

$$\mathbf{y}_0 = (x_1, v \cos(\alpha), y_1, v \sin(\alpha)),$$

kjer sta  $v$  in  $\alpha$  izbrana začetna hitrost in kot lučaja banane. □

# Literatura

J. Kozak. *Numerična analiza*. DMFA - založništvo, 2008.

B. Plestenjak. *Razširjen uvod v numerične metode*. DMFA - založništvo, 2015.