



Matematické základy počítačové grafiky

Eduard Sojka, Martin Němec,
Tomáš Fabián

Text byl vytvořen v rámci realizace projektu *Matematika pro inženýry 21. století* (reg. č. CZ.1.07/2.2.00/07.0332), na kterém se společně podílela Vysoká škola báňská – Technická univerzita Ostrava a Západočeská univerzita v Plzni



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Eduard Sojka, Martin Němec, Tomáš Fabián
Matematické základy počítačové grafiky

© Eduard Sojka, Martin Němec, Tomáš Fabián, 2011
ISBN

Předmluva

Vážený čtenáři,

tento text by měl sloužit studentům jako výukový materiál pro studium předmětů „Základy počítačové grafiky“, „Počítačová grafika“ a „Geometrie pro počítačovou grafiku“. Celý výukový materiál je rozdělen do třech částí, jejichž obsah je volen tak, aby posluchač získal základní přehled probíraných témat v uvedených předmětech. První část se zabývá teorií afinních prostorů a afinních transformací. Následně je text zobecněn na projektivní prostory a projektivní transformace. Studium této kapitoly získáte informace o transformacích v počítačové grafice a jejich možnostech. Druhá kapitola se zabývá teorií křivek a ploch. Nejdříve jsou uvedeny základní obecné vlastnosti a následně jsou popsány jednotlivé typy křivek a ploch až po v dnešní době nejpoužívanější NURBS křivky a plochy, kterým je věnováno nejvíce prostoru. Třetí a poslední část tohoto textu je věnována popisu a možnostem modelování těles. V této kapitole najdete základní typy reprezentace a popisu těles používané v počítačové grafice. Prosíme čtenáře o schovávavost při studiu těchto materiálů a v případě nalezení nedostatků či chyb laskavě žádáme o jejich sdělení na adresy autorů¹

V Ostravě 31. 8. 2011

Autoři

¹eduard.sojka@vsb.cz martin.nemec@vsb.cz tomas.fabian@vsb.cz

Obsah

Předmluva	iii
1 Afinní a projektivní prostory	1
1.1 Afinní prostor a afinní transformace	1
1.2 Změna souřadné soustavy	3
1.3 Ortonormalita afinní transformace	4
1.4 Projektivní prostor, projektivní transformace	7
1.5 Stanovení matice zobrazovací transformace – příklad A	16
1.6 Stanovení matice zobrazovací transformace – příklad B	18
1.7 Stanovení matice zobrazovací transformace – příklad C	24
1.8 Transformace na výstupní zařízení	26
1.9 Inverze zobrazovací transformace – příklad D	27
1.10 Dodatky ke kapitole o afinních prostorech a afinních transformacích	29
2 Křivky a plochy v počítačové grafice	31
2.1 Křivka a její vlastnosti	32
2.1.1 Parametrická a geometrická spojitost	35
2.1.2 Vlastnosti křivek v počítačové grafice	36
2.1.3 Tečný vektor, tečna	36
2.1.4 Délka křivky	37
2.1.5 Parametrizace křivky	38
2.1.6 Parametrizace obloukem	39
2.1.7 Oskulační rovina	39
2.1.8 Křivost a torze křivky	41
2.1.9 Oskulační kružnice	43
2.2 Plochy v počítačové grafice	44
2.2.1 Vlastnosti ploch	45
2.2.2 u-křivky, v-křivky	45
2.2.3 Křivost plochy	46
2.3 Polynomiální křivky	47
2.3.1 Interpolace algebraickým polynomem	47
2.3.2 Interpolace po částech	49
2.3.3 Hermitův interpolační polynom	50

2.4	Metoda nejmenších čtverců	51
2.5	Fergusonova křivka	54
2.5.1	Navazování Fergusonových kubik	57
2.6	Spline křivky	57
2.7	Bézierova křivka	59
2.7.1	Základní vlastnosti Bézierovy křivky	60
2.7.2	Konverze Fergusonovy kubiky na Bézierovu kubiku	62
2.7.3	Bézierův spline	63
2.7.4	Geometrická konstrukce bodů Bézierových křivek	65
2.7.5	Racionální Bézierova křivka	66
2.7.6	Použití Bézierových křivek	66
2.8	Bézierova plocha	67
2.9	Coonsova B-spline křivka	69
2.9.1	Coonsův kubický B-spline	71
2.10	Coonsova B-spline plocha	72
2.11	NURBS křivky a plochy	74
2.11.1	B-spline bazové funkce	75
2.11.2	Derivace B-spline bazových funkcí	79
2.11.3	Neracionální B-spline křivky	81
2.11.4	Konvexní obal B-spline křivek	84
2.11.5	Plocha jako tenzorový součin dvou křivek	84
2.11.6	Neracionální B-spline plochy	85
2.11.7	Racionální B-spline křivky (NURBS)	87
2.11.8	Racionální B-spline plochy (NURBS)	90
3	Modelování těles	99
3.1	Topologické prostory a topologická ekvivalence	100
3.2	Soubor simplexů	101
3.3	Tělesa s hranicí 2-manifold v E^3	103
3.4	Eulerova formule	105
3.5	Regularizované booleovské operace	106
3.6	Reprezentace tělesa pomocí hranice	108
3.6.1	Implementace hraniční reprezentace	109
3.6.2	Eulerovy operátory	111
3.6.3	Provádění booleovských operací	114
3.6.4	Zobrazení tělesa popsaného hraniční reprezentací	116
3.7	CSG reprezentace tělesa	116
3.7.1	Realizace CSG reprezentace	118
3.7.2	Provádění booleovských operací v reprezentaci CSG	119
3.7.3	Určení vzájemné polohy bodu a CSG tělesa	119
3.7.4	Zobrazení tělesa reprezentovaného pomocí CSG	120

Kapitola 1

Afinní a projektivní prostory

Průvodce studiem



Po prostudování této kapitoly porozumíte vybraným částem z teorie afinních a projektivních prostorů a transformací, a to v míře potřebné pro realizaci grafických systémů, pro které je tato teorie základním východiskem. Bez znalostí alespoň základů z této oblasti byste v žádném případě nemohli grafické systémy konstruovat a mnohdy ani efektivně používat. Po prostudování kapitoly budete schopni realizovat jakoukoli afinní a projektivní transformaci. Zejména transformace, které v grafických systémech realizují rovnoběžná nebo perspektivní zobrazení, a také transformace modelovací, které se používají při vytváření scény z jednotlivých objektů.

Všechny jevy, které jsou zde prezentovány, byly vybrány velmi pečlivě na základě toho, že je jejich znalost při praktické realizaci systémů nevyhnutelná. Totéž plně platí i pro všechny zde uváděné řešené příklady, které jsou nedílnou součástí výkladu (často v nich naleznete důležité informace). Naopak všechny jevy, z nichž byste praktický užitek neměli, jsou potlačeny. Věřím proto, že překonáte případnou nedočkavost, nenecháte se odradit zdánlivě složitými matematickými vztahy a s chutí se do studia kapitoly pustíte.

1.1 Afinní prostor a afinní transformace

Afinní prostor a afinní transformace jsou základními pojmy, se kterými se lze při konstrukci grafických systémů setkat. V této podkapitole se s nimi seznámíte. Při studiu celého textu je pak budete téměř neustále potřebovat. Bylo by jistě možné začít výklad přesnými definicemi. S ohledem na účel tohoto textu však od takového postupu upustíme a spokojíme se s poněkud méně únavným (i když současně bohužel také poněkud méně přesným) úvodem vycházejícím z intuitivní představy a zkušenosti.

Intuitivně za afinní prostor považujeme prostor obsahující body. Navíc musí být současně dán (nebo zvolen) nějaký přidružený vektorový prostor (tzv. vektorové zaměření) a nějaké zobrazení, které ke každé dvojici z prostoru bodů přiřazuje nějaký prvek onoho přidruženého prostoru vektorového. Místo dalšího zpřesňování této in-

formace se zde spokojíme s konstatováním, že konstrukce afinního prostoru sleduje velmi praktický cíl, a to, aby bylo možné body v afinním prostoru jednoznačně specifikovat pomocí souřadnic (tedy pomocí prvků ze zmíněného přidruženého prostoru vektorového) a aby k manipulacím s body bylo možné využívat prostředků lineární algebry. Za afinní prostor je nutné, přísně vzato, považovat nejen samotný prostor bodů, ale i onen zmíněný přidružený prostor vektorový a také zobrazení, které dvojicím bodů přiřazuje vektory. Dimenze přidruženého vektorového prostoru (který je využit k měření souřadnic) určuje také dimenzi afinního prostoru. Pro n -rozměrný afinní prostor budeme používat označení A_n . Nejčastěji budeme pracovat s afinními prostory trojrozměrnými. Řekněme, že v trojrozměrném afinním prostoru máme bod X . Jeho souřadnice vzhledem ke zvolené souřadné soustavě (přesněji je budeme nazývat afinními souřadnicemi) jsou popsány vektorem o třech prvcích $\mathbf{x} = (x_1, x_2, x_3)$. Vektor \mathbf{x} je prvkem zmíněného přidruženého vektorového prostoru.

Afinní prostor, v němž je zaveden skalární součin a norma jako odmocnina ze skalárního součinu, se nazývá *euklidovským prostorem*. Pro euklidovský n -rozměrný prostor budeme používat značení E_n . Zavedení skalárního součinu a normy je důležité, protože umožňuje měřit délku vektorů a úhly mezi vektory (to určitě znáte z matematiky). Nyní už asi shledáváte, že s euklidovskými a tedy i s afinními prostory máte již vlastně bohaté zkušenosti, protože jste s nimi v minulosti nepochybně nejednou při řešení geometrických úloh pracovali. Možná jste jen nebyli zvyklí zdůrazňovat, o jaké prostory se jedná.

Afinní transformace je zobrazení bodů jednoho afinního prostoru do jiného afinního prostoru. Speciálním případem je zobrazení bodů afinního prostoru do téhož afinního prostoru, které je bijekcí. Tento případ se nazývá afinitou. (Bijekce je zobrazení, kdy dvěma různým vzorům vždy odpovídají dva různé obrazy. Zde tak konkrétně vylučujeme, aby se dva různé body zobrazily na bod jediný). Velmi často bude předmětem našeho zájmu afinní transformace mezi dvěma afinními prostory dimenze 3, případně afinita na prostoru dimenze 3. K matematickému popisu afinní transformace zavedeme následující vektory a matici (omezíme se zde na trojrozměrné prostory)

$$\mathbf{y} = (y_1, y_2, y_3), \quad \mathbf{x} = (x_1, x_2, x_3), \quad \mathbf{t} = (t_1, z_2, z_3),$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}. \quad (1.1)$$

Afinní transformaci lze pak popsat vztahem

$$\mathbf{y} = \mathbf{x}\mathbf{A} + \mathbf{t}. \quad (1.2)$$

Poznamenejme, že uvedený vztah je *důsledkem* toho, jak je afinní transformace definována (podkapitola 1.10). Pro naše potřeby je ale platnost vztahu (1.2) dostatečnou informací. V uvedeném vztahu matice \mathbf{A} a vektor \mathbf{t} popisují vlastnosti transformace.

Vektor \mathbf{y} reprezentuje bod, který je afinním obrazem bodu, který je reprezentován vektorem \mathbf{x} . Jestliže přímo známe matici \mathbf{A} a vektor \mathbf{t} , pak lze transformaci podle předpisu (1.2) již jednoduše provést. Někdy je ale situace komplikovanější v tom smyslu, že je transformace zadána jinak než přímo hodnotami \mathbf{A} , \mathbf{t} . V takovém případě je obvyklé, že se hodnoty \mathbf{A} , \mathbf{t} na základě zadání určí a pak se opět využije vztahu (1.2). Typickými postupy, které vedou k určení potřebných hodnot, se budeme v dalším textu zabývat podrobněji.

1.2 Změna souřadné soustavy

Velmi častým případem afinní transformace mezi dvěma prostory stejné dimenze je změna souřadné soustavy a odpovídající transformace souřadnic bodů. Opět se zde omezíme na afinní (případně euklidovské) prostory trojrozměrné, protože ty budeme dále potřebovat v převážné většině případů. Zobecnění závěrů zde uvedených pro prostory s libovolným počtem dimenzí je ale snadné (máte-li chuť, můžete se o zobecnění po přečtení textu pokusit, není to ale nezbytně nutné).

V grafických systémech se změna souřadné soustavy provádí dosti často. Někdy se jedná o vnitřní záležitost systému (pro uživatele neviditelnou), kdy je změna souřadné soustavy součástí nějakého výpočetního postupu nebo jeho odvození. Jindy může být manipulace se souřadnou soustavou patrná i pro uživatele. Jako příklad lze uvést tzv. modelovací transformace, s jejichž pomocí uživatel systému vytváří scénu (může pomocí ní zadávat např. umístění objektů ve scéně, jejich velikost atd.). Po prostudování této podkapitoly budete umět transformace souřadných soustav realizovat.

Transformace souřadnic způsobená pouhým posunutím počátku souřadné soustavy je triviální. Nechtě (p_x, p_y, p_z) jsou souřadnice nového počátku souřadné soustavy. Transformaci souřadnic z původní do nové souřadné soustavy pak popisuje vztah (1.2), v němž \mathbf{A} je jednotková matice a vektor \mathbf{t} je $\mathbf{t} = (-p_x, -p_y, -p_z)$.

Věnujme se nyní poněkud komplikovanějšímu případu. Počátek nyní při změně souřadné soustavy ponecháme beze změny. Změníme ale osy soustavy. Předpokládejme, že původní souřadná soustava měla *bázové vektory* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ (bázové vektory jsou vektory jednotkové délky udávající směry souřadných os). Bázové vektory nové souřadné soustavy jsou $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$. Můžeme si myslet, že bázové vektory obou soustav vyjadřujeme vzhledem k nějaké další (už třetí) referenční souřadné soustavě. Řekněme, že máme bod X , který je v původní souřadné soustavě reprezentován vektorem $\mathbf{x} = (x_1, x_2, x_3)$. Úkolem je zjistit souřadnice bodu X vzhledem k nové souřadné soustavě. Tyto souřadnice budou vyjádřeny vektorem, který označíme $\tilde{\mathbf{x}}$. Polohový vektor bodu X (vzhledem k oné myšlené třetí referenční souřadné soustavě) můžeme vyjádřit jako součet

$$\sum_{i=1}^3 x_i \mathbf{v}_i. \quad (1.3)$$

Vyjádříme bázové vektory \mathbf{v}_i souřadné soustavy původní pomocí bázových vektorů souřadné soustavy nové. Nepochybně musí být možné vyjádřit vektor \mathbf{v}_i jako lineární kombinaci vektorů $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ (tvoří-li trojice $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ bázi, pak lze každý vektor uvažovaného prostoru vyjádřit jako lineární kombinaci vektorů uvedené trojice, tedy i každý z vektorů \mathbf{v}_i). Máme proto

$$\mathbf{v}_i = \sum_{j=1}^3 a_{ij} \mathbf{w}_j. \quad (1.4)$$

kde a_{ij} jsou koeficienty lineární kombinace. Dosazením vztahu (1.4) do vztahu (1.3) a dostaneme

$$\begin{aligned} \sum_{i=1}^3 x_i \mathbf{v}_i &= \sum_{i=1}^3 x_i \sum_{j=1}^3 a_{ij} \mathbf{w}_j = \sum_{i=1}^3 \sum_{j=1}^3 x_i a_{ij} \mathbf{w}_j \\ &= (x_1 a_{11} + x_2 a_{21} + x_3 a_{31}) \mathbf{w}_1 \\ &\quad + (x_1 a_{12} + x_2 a_{22} + x_3 a_{32}) \mathbf{w}_2 \\ &\quad + (x_1 a_{13} + x_2 a_{23} + x_3 a_{33}) \mathbf{w}_3. \end{aligned} \quad (1.5)$$

Členy v závorkách v posledním ze vztahů (1.5) udávají souřadnice bodu X v nové souřadné soustavě. Uvažme, že podle našeho čekání by měl transformaci popisovat maticový vztah $\tilde{\mathbf{x}} = \mathbf{x} \mathbf{A}$ (protože se počátek při změně soustavy nezměnil, je ve vztahu (1.2) nutně $\mathbf{t} = \mathbf{0}$). Zjišťujeme, že když položíme

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad (1.6)$$

bude součin $\tilde{\mathbf{x}} = \mathbf{x} \mathbf{A}$ (jednotlivé složky vektoru $\tilde{\mathbf{x}}$) vyjadřovat členy v závorkách v posledním ze vztahů (1.5), a tedy souřadnice bodu X v nové souřadné soustavě. To je důležitý závěr. Vidíme, že prvky v i -tém řádku hledané matice \mathbf{A} jsou souřadnice i -tého bázového vektoru původní souřadné soustavy vzhledem k souřadné soustavě nové (říká to vztah (1.4)). Tohoto závěru později mnohokrát využijeme v situaci, kdy bude zapotřebí matici \mathbf{A} stanovit. Stojí za to zapamatovat si toto zjištění.

1.3 Ortonormalita afinní transformace

Mezi afinními transformacemi zaujímají významné místo transformace ortonormální. *Ortonormální transformace* jsou ty, při nichž zůstávají délky a úhly před a po transformaci nezměněny. Asi si vybavíte, že velmi obvyklé transformace posunutí a pootočení jsou příklady transformací ortonormálních. Často se při řešení praktických úloh vyplatí vědět o existenci této třídy transformací. Předem obvykle víte, zda zamýšlená transformace je či není ortonormální. Je-li tomu tak, pak matice \mathbf{A}

musí mít jisté speciální vlastnosti. Této znalosti můžete využít jednak při stanovení matice \mathbf{A} a také při manipulaci s ní. Po přečtení této podkapitoly budete umět ortonormality transformací využívat.

Jak délky, tak úhly (které mají být při ortonormální transformaci zachovány) souvisí se skalárním součinem. Uvažujme vektory \mathbf{x} , \mathbf{y} . Zopakujme, že *skalární součin* vektorů \mathbf{x} , \mathbf{y} (značíme jej zde $\langle \mathbf{x}, \mathbf{y} \rangle$) zavádíme vztahem

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}\mathbf{y}^\top. \quad (1.7)$$

Připomeňme dále dobře známé vztahy pro délku ($|\mathbf{x}|$) vektoru \mathbf{x} a pro úhel (φ) dvou vektorů \mathbf{x} , \mathbf{y} . Víte, že platí

$$|\mathbf{x}| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\mathbf{x}\mathbf{x}^\top}, \quad (1.8)$$

$$\cos \varphi = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{|\mathbf{x}||\mathbf{y}|} = \frac{\mathbf{x}\mathbf{y}^\top}{|\mathbf{x}||\mathbf{y}|}. \quad (1.9)$$

Je zřejmé, že jestliže transformace nemění hodnotu skalárního součinu, pak také úhly a délky nejsou transformací změněny. Nechť \mathbf{x}_1 , \mathbf{x}_2 jsou vektory reprezentující počáteční a koncový bod vektoru \mathbf{x} . Je tedy $\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1$. Afinní transformací (1.2) přejde vektor \mathbf{x} ve vektor $(\mathbf{x}_2\mathbf{A} + \mathbf{t}) - (\mathbf{x}_1\mathbf{A} + \mathbf{t}) = (\mathbf{x}_2 - \mathbf{x}_1)\mathbf{A} = \mathbf{x}\mathbf{A}$. Při transformaci vektoru se tedy translační složka transformace reprezentovaná vektorem \mathbf{t} vůbec neuplatní (což jste ale asi intuitivně očekávali). Významná je naopak ta část transformace, která je reprezentována maticí \mathbf{A} . Podmínku zachování skalárního součinu vyjádříme rovnicí

$$\langle \mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{x}\mathbf{A}, \mathbf{x}\mathbf{A} \rangle. \quad (1.10)$$

Odtud pak plyne požadavek

$$\mathbf{x}\mathbf{x}^\top = (\mathbf{x}\mathbf{A})(\mathbf{x}\mathbf{A})^\top = \mathbf{x}\mathbf{A}\mathbf{A}^\top\mathbf{x}^\top. \quad (1.11)$$

Z předchozího vztahu je již zřejmé, že afinní transformace bude zachovávat hodnotu skalárního součinu tehdy, jestliže bude splněna podmínka $\mathbf{A}\mathbf{A}^\top = \mathbf{I}$, kde \mathbf{I} je *jednotková matice*. Uvedená podmínka je nutnou a postačující podmínkou ortonormality transformace. Z definice inverzní matice máme $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. Vidíme tedy, že dále musí platit

$$\mathbf{A}^{-1} = \mathbf{A}^\top. \quad (1.12)$$

Uvedené zjištění má praktický význam. Ukazuje, jak lze jednoduše vypočítat inverzní matici k matici ortonormální transformace. Inverzní matice se jednoduše rovná matici transponované. Ukažme konečně ještě jednu vlastnost matice ortonormální transformace, a to, že její determinant musí být roven hodnotě ± 1 . Zdůvodnění

tohoto faktu není obtížné. Protože platí $\det(\mathbf{A}\mathbf{A}^\top) = \det(\mathbf{A})\det(\mathbf{A}^\top) = \det^2(\mathbf{A})$, a protože $\det(\mathbf{I}) = 1$, musí být také

$$\det^2(\mathbf{A}) = 1. \quad (1.13)$$



Příklad 1.1. V tomto příkladě shrneme dosud získané poznatky o afinní transformaci. Příklad je záměrně volen tak jednoduchý, abyste mohli porovnat výsledek získaný teoretickým řešením s výsledkem, který očekáváte. Uvažujte trojrozměrný afinní prostor. Nová soustava souřadnic je pootočená vůči původní soustavě o úhel φ kolem osy z . Nalezněte převodní vztah pro přepočítání souřadnic bodů z původní souřadné soustavy do nové.

Řešení. Se záměrem připravit se na řešení pozdějších úloh budeme postupovat takto: Bázové vektory nové soustavy označíme $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$. Jejich souřadnice vyjádříme vzhledem k původní souřadné soustavě. (Často je tento postup příjemnější, než vyjadřovat souřadnice bázových vektorů původní souřadné soustavy vzhledem k nové soustavě, jak bychom podle předcházející podkapitoly měli učinit.) Dostaneme

$$\mathbf{w}_1 = (\cos \varphi, \sin \varphi, 0), \quad \mathbf{w}_2 = (-\sin \varphi, \cos \varphi, 0), \quad \mathbf{w}_3 = (0, 0, 1).$$

Na základě výsledku z podkapitoly 1.2 nyní již můžeme zapsat předpis, pomocí kterého je možné přepočítat souřadnice bodů z nové do původní souřadné soustavy (což je zatím ale naopak, než bylo požadováno v zadání). Máme

$$\mathbf{x}_{\text{old}} = \mathbf{x}_{\text{new}}\mathbf{A}^\top, \quad \text{kde je } \mathbf{A} = \begin{bmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Z předchozího vztahu a s přihlédnutím ke skutečnosti, že rotace je ortonormální transformací (protože nemění délky ani úhly), snadno dostaneme

$$\mathbf{x}_{\text{new}} = \mathbf{x}_{\text{old}}\mathbf{A}^{-1} = \mathbf{x}_{\text{old}}\mathbf{A}^\top.$$



Příklad 1.2. Stanovte matici \mathbf{A} pro transformaci spočívající v rotaci souřadné soustavy postupně kolem souřadných os x, y, z o úhly α, β, γ . (Rotace o úhel β se myslí okolo osy y soustavy, která vznikla dříve provedenou rotací původní soustavy o úhel α kolem její osy x . Analogicky se předpokládá, že před rotací okolo osy z o úhel γ byly již obě předchozí rotace kolem os x a y již provedeny.)

Řešení. Při řešení budeme postupovat tak, že nejprve převedeme souřadnice ze souřadné soustavy původní do souřadnic v souřadné soustavě pootočené kolem osy x , ty pak dále převedeme do souřadné soustavy pootočené kolem osy y a ty opět nakonec i do souřadné soustavy pootočené kolem osy z . Tuto posloupnost postupně prováděných transformací lze matematicky popsat vztahem

$$\mathbf{x}_{\text{new}} = ((\mathbf{x}_{\text{old}}\mathbf{A}_x)\mathbf{A}_y)\mathbf{A}_z = \mathbf{x}_{\text{old}}\mathbf{A}_x\mathbf{A}_y\mathbf{A}_z.$$

Matice $\mathbf{A}_x, \mathbf{A}_y, \mathbf{A}_z$ odpovídají rotaci kolem jednotlivých souřadných os. S ohledem na výsledek předchozího příkladu, v němž jsme odvodili předpis pro matici \mathbf{A}_z^\top , je můžeme ihned zapsat. Máme

$$\mathbf{A}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}, \mathbf{A}_y = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix}, \mathbf{A}_z = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

▲

Příklad 1.3. Promyslete, kolik nezávislých hodnot obsahuje matice \mathbf{A} afinní ortonormální transformace. Uvažujte trojrozměrný prostor.



Řešení. Třebaže matice \mathbf{A} obsahuje celkem 9 prvků, jsou v případě ortonormální transformace jen tři z nich nezávislé. Ostatní prvky je totiž možné dopočítat ze vztahu $\mathbf{A}\mathbf{A}^\top = \mathbf{I}$, který poskytuje šest rovnic pro šest zbývajících prvků. Situaci již také ilustroval předchozí příklad. Rotace je ortonormální transformací. Nebylo proto divu, že k jejímu popisu stačily jen tři nezávislé hodnoty. V předchozím příkladě to byly úhly α, β, γ .

▲

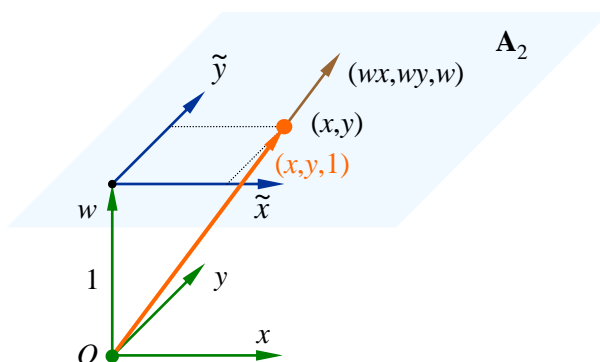
Shrnutí

V tuto chvíli byste měli zejména znát, jak lze matematicky popsat afinní transformaci. Měli byste být schopni stanovit matici a vektor, které transformaci popisují. Měli byste vědět, co je ortonormální transformace a také byste měli umět těžit ze speciálních vlastností matice, která ji popisuje.

1.4 Projektivní prostor, projektivní transformace

Při použití afinních prostorů v grafických systémech lze v některých situacích narazit na jisté problémy. Významným problémem je například to, že není jednoduše možné pracovat s body v nekonečnu (s tzv. nevlastními body). I když jsou praktické scény zpravidla konečné, mohou nevlastní body vzniknout při zobrazování scén (např. při velmi obvyklém středovém promítání na rovinou průmětnu může průmět některých bodů scény padnout do nekonečna). Problém spočívá v tom, že vektor reprezentující nevlastní bod v afinním prostoru by měl mít jednu nebo více složek rovných hodnotě $\pm\infty$. Hodnotu ∞ , jak víte, nelze ale v počítači jednoduše reprezentovat. Navíc by tato hodnota měla být i výsledkem aritmetických operací, což opět není jednoduše možné. Uvedené potíže lze překonat zavedením *projektivního prostoru*. *Projektivní transformace*, která se nad projektivními prostory zavádí, je také obecnější než transformace afinní. Například: Zatímco pomocí afinní transformace lze realizovat zobrazení pouze rovnoběžným promítáním, umožňuje projektivní transformace provádět promítání rovnoběžné i středové. V této podkapitole se projektivní prostor a projektivní transformaci naučíte používat.

Zavedení projektivního prostoru a tzv. homogenních souřadnic, které v projektivním prostoru jednotlivé body reprezentují, lze názorně vysvětlit ve dvojrozměrném



Obr. 1.1: Zavedení homogenních souřadnic

případě (obr. 1.1). Uvažujme dvojrozměrný afinní prostor \mathbf{A}_2 a zavedme nyní přidružený trojrozměrný prostor \mathbf{V}_3 tak, že všechny body prostoru \mathbf{A}_2 leží v rovině $w = 1$ tohoto přidruženého prostoru (obr. 1.1). Dvojrozměrný projektivní prostor \mathbf{P}_2 zavádíme jako množinu směrů ve \mathbf{V}_3 . Vysvětleme blíže smysl takového počínání. V afinním prostoru \mathbf{A}_2 uvažujme bod, jehož afinní souřadnice jsou (x, y) (obr. 1.1). Homogenními souřadnicemi uvažovaného bodu, které jej reprezentují v právě zavedeném prostoru \mathbf{P}_2 , je trojice (wx, wy, w) , kde w je libovolné reálné číslo různé od nuly. V \mathbf{P}_2 je tedy uvažovaný bod reprezentován nekonečným počtem vektorů přidruženého vektorového prostoru \mathbf{V}_3 , které jsou ovšem všechny navzájem kolineární. Speciálně můžeme také volit $w = 1$ (zvláštní význam této volby je patrný z obr. 1.1). Dostaneme tak trojici $(x, y, 1)$. Je-li naopak trojice (x, y, w) homogenními souřadnicemi nějakého bodu, pak také trojice $(\frac{x}{w}, \frac{y}{w}, 1)$ je homogenními souřadnicemi téhož bodu a dvojice $(\frac{x}{w}, \frac{y}{w})$ jeho souřadnicemi afinními. Přechod od souřadnic (x, y, w) k souřadnicím $(\frac{x}{w}, \frac{y}{w}, 1)$ bývá nazýván normalizací.

Shrneme-li to, co bylo dříve uvedeno, můžeme říci, že v projektivním prostoru je každý bod reprezentován nějakým vektorem, který k bodu směřuje. Je při tom jedno, jak dlouhý vektor je. Nyní je již zřejmé, proč nebudou s body v nekonečnu v projektivních prostorech potíže. I ony totiž mohou být reprezentovány vektory konečné délky. Uvažujme například bod ležící v obr. 1.1 v nekonečnu na souřadnicové ose \tilde{x} . Tento bod může být reprezentován např. trojicí $(1, 0, 0)$.

Jestliže jsme se až doposud zabývali pouze dvojrozměrnými projektivními prostory, pak to bylo proto, že bylo snadné s pomocí obrázku ukázat zavedení takového prostoru. Grafické systémy ale obvykle pracují s prostory trojrozměrnými. Vektor homogenních souřadnic má v trojrozměrném projektivním prostoru čtyři složky. Je typické, že poloha bodů ve scéně bývá uživatelem prvotně popsána pomocí souřadnic afinních. Řekněme, že (x, y, z) jsou afinní souřadnice bodu X ve scéně, který neleží v nekonečnu. Grafické systémy při své činnosti obvykle velmi brzy převádějí

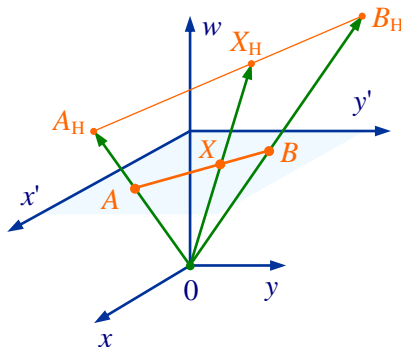
souřadnice afinní na homogenní. Nejčastěji jednoduše tak, že jako homogenní souřadnice bodu berou čtveřici $(x, y, z, 1)$. Tím je z nekonečně mnoha možných vektorů reprezentujících bod X v \mathbf{P}_3 vybrán jediný, a to ten, pro který platí $w = 1$. Zobrazení středovým promítáním se prakticky vždy řeší v homogenních souřadnicích. Před vykreslením výsledného obrazu (např. na obrazovku nebo do souboru) je ale zapotřebí vrátit se k souřadnicím afinním. To se provádí (jak již víte) normalizací. Před provedením normalizace bývá provedeno ořezání výsledného obrazu do zvoleného konečného zorného objemu, takže obraz již neobsahuje nevlastní body, které by jinak při provádění normalizace vedly k dělení nulou. Také rovnoběžné promítání a modelovací transformace se v grafických systémech obvykle řeší s využitím homogenních souřadnic, a to přestože by to v tomto případě nebylo nezbytně nutné. Zde se jednoduše i pro řešení speciální úlohy využívá obecnějšího nástroje, protože jej systém zpravidla stejně tak či tak již obsahuje.

Příklad 1.4. Cílem tohoto příkladu je ukázat, že úlohy, které se čtenář ve středoškolské a vysokoškolské matematice již naučil řešit v prostorech afinních, lze stejně dobře a jednoduše řešit také v prostorech projektivních. Uvažujme následující úlohu: Určete v \mathbf{P}_2 rovnici přímky, která prochází body A, B o homogenních souřadnicích $(x_A, y_A, w_A), (x_B, y_B, w_B)$. Řešte nejprve obecně, pak pro $A = (0, 0, 1), B = (1, 0, 1)$ a pro $A = (1, 0, 1), B = (0, 1, 1)$.



Řešení. Abychom co nejvíce ilustrovali příbuznost s již známými řešeními v afinním prostoru, ukážeme dvě řešení zadané úlohy.

1. Necht (\tilde{x}, \tilde{y}) jsou afinní souřadnice bodu. Uvažujme dobře známou rovnici přímky ve tvaru $a\tilde{x} + b\tilde{y} + c = 0$, která platí pro souřadnice afinní. Necht (x, y, w) jsou homogenní souřadnice téhož bodu. Víme, že afinní souřadnice lze získat normalizací. Máme $\tilde{x} = x/w, \tilde{y} = y/w$. Dosadíme-li hodnoty \tilde{x}, \tilde{y} do dříve uvedené rovnice přímky, dostaneme po úpravě rovnici ve tvaru $ax + by + cw = 0$, která platí pro body zadané v homogenních souřadnicích. Předpokládejme na okamžik, že hodnoty a, b, c jsou známy. Násobíme-li rovnici reálným číslem $\lambda \neq 0$, pak dostaneme rovnici $\lambda ax + \lambda by + \lambda cw = 0$. Je zřejmé, že také rovnice $\hat{a}x + \hat{b}y + \hat{c}z = 0$, kde $\hat{a} = \lambda a, \hat{b} = \lambda b, \hat{c} = \lambda c$, je rovnicí téže přímky. Odtud vyplývá, že hodnoty a, b, c nebude možné jednoznačně stanovit pouze ze znalosti souřadnic bodů, jimiž přímka prochází. Bude zapotřebí použít vhodné doplňující podmínky. (Na stejný problém bychom ovšem narazili i při řešení v prostoru afinním.) Pro stanovení koeficientů a, b, c tak máme rovnice $ax_A + by_A + cw_A = 0, ax_B + by_B + cw_B = 0$. „Univerzální“ doplňující podmínkou je např. podmínka ve tvaru $a^2 + b^2 + c^2 = 1$. Použití uvedené podmínky ovšem bohužel vede na systém, který je nelineární. Tuto potíž lze obejít např. použitím podmínky $a + b + c = 1$ nebo podmínky $a + b + c = 0$. Použije se ta varianta, pro níž má systém řešení. Uvedeným postupem získáme pro první a druhou dvojici zadaných bodů A, B rovnice přímek ve tvaru $x + y - w = 0, x - y = 0$.



Obr. 1.2: K rovnici přímky ve tvaru $X = (1-\lambda)A + \lambda B$ v homogenních souřadnicích

2. Velmi jednoduše je možné rovnici přímky procházející body A, B zapsat ve tvaru $X = (1-\lambda)A + \lambda B$, kde λ je reálný parametr ($-\infty < \lambda < \infty$). Je známo, že vztah $X = (1-\lambda)A + \lambda B$ je parametrickou rovnicí přímky v afinních souřadnicích. Že je rovnicí přímky i při použití souřadnic homogenních ukazuje obr. 1.2. Body A, B jsou na obr. 1.2 reprezentovány vektory, jejichž koncové body jsou označeny jako A_H, B_H . Rovnice $X_H = (1-\lambda)A_H + \lambda B_H$ je rovnicí přímky ve \mathbf{V}_3 . Protože body O, A, B, A_H, B_H leží v rovině, reprezentuje vektor s koncovým bodem v X_H bod X ležící na přímce spojující body A, B .



Zásadní úlohu v grafických systémech sehrává projektivní transformace (kolineace). Kolineace je zobrazením bodů jednoho projektivního prostoru na body téhož nebo jiného prostoru. Pro přesnou definici čtenáře odkazujeme na podkapitulu 1.10 (není však nezbytné ji studovat). Z hlediska praktické potřeby nám zde postačí konstatovat, že kolineaci lze matematicky popsat vztahem

$$\mathbf{y} = \mathbf{x}\mathbf{T}. \quad (1.14)$$

Vektory \mathbf{x} a \mathbf{y} reprezentují body před resp. po transformaci. Matice \mathbf{T} charakterizuje kolineaci. Grafické systémy pracují nejčastěji s trojrozměrným projektivním prostorem. Pak jsou \mathbf{x}, \mathbf{y} čtyřprvkové vektory homogenních souřadnic a \mathbf{T} je matice rozměru 4×4 . Podle tvaru matice \mathbf{T} lze usuzovat na transformaci, kterou matice provádí. Je naopak také možné k požadované transformaci stanovit matici \mathbf{T} . V následujícím přehledu uvedeme několik základních případů. Je užitečné se s nimi seznámit, a to mimo jiné také proto, že z nich lze sestavovat transformace složitější. Omezíme se na transformace v prostoru \mathbf{P}_3 .

Změna měřítka na jedné ose

$$\mathbf{T} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Transformace popsaná uvedenou maticí provádí změnu měřítka na ose x . Můžete se o tom snadno přesvědčit, když transformaci maticí \mathbf{T} aplikujete na bod o souřadnicích $(x, y, z, 1)$. Dostanete $(x, y, z, 1)\mathbf{T} = (s_x x, y, z, 1)$.

Změna měřítka na všech osách

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{s} \end{bmatrix}$$

Transformace provádí změnu měřítka na všech souřadnicových osách, což je patrné ze vztahu $(x, y, z, 1)\mathbf{T} = (x, y, z, \frac{1}{s}) \rightarrow (sx, sy, sz, 1)$ (symbolem \rightarrow jsme označili operaci normalizace homogenních souřadnic).

Zkosení

$$\mathbf{T} = \begin{bmatrix} 1 & a & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Transformace, která je popsána touto maticí, bývá obvykle nazývána zkosením. Je opět jednoduché se přesvědčit o oprávněnosti tohoto názvu. Provedením násobení dostáváme $(x, y, z, 1)\mathbf{T} = (x, y + ax, z, 1)$. Vidíme, že rovina $y = \text{const}$ transformací přejde v rovinu $y = \text{const} + ax$ (tedy „zkosí se“; a je směrnice úhlu zkosení). Roviny $x = \text{const}$, $z = \text{const}$ zůstanou uvedenou transformací nedotčeny. Matice popisující transformaci, která by „zkosila“ také tyto roviny, lze snadno zkonstruovat analogicky, což ponecháváme čtenáři jako drobné samostatné cvičení.

Posunutí

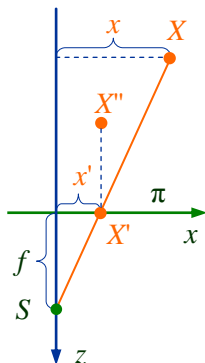
$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ p_x & 0 & 0 & 1 \end{bmatrix}$$

Jak vyplývá ze vztahu $(x, y, z, 1)\mathbf{T} = (x + p_x, y, z, 1)$, popisuje tato matice posunutí bodu o hodnotu p_x ve směru souřadné osy x . Matice realizující translace ve směru souřadných os y, z jsou triviální modifikací matice zde uvedené.

Rotace

$$\mathbf{T} = \begin{bmatrix} c & s & 0 & 0 \\ -s & c & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Transformace popsaná uvedenou maticí je rotací kolem souřadnicové osy z . Symboly s, c značí hodnoty sinu, resp. kosinu úhlu rotace. Musí být $s^2 + c^2 = 1$. Provedením



Obr. 1.3: K ověření činnosti matice středového promítání

maticového násobení dostáváme $(x, y, z, 1)\mathbf{T} = (xc - ys, xs + yc, z, 1)$. Matice realizující rotaci kolem os x, y je možné zkonstruovat analogicky. Je užitečné, abyste se opět pokusili odpovídající matice zapsat.

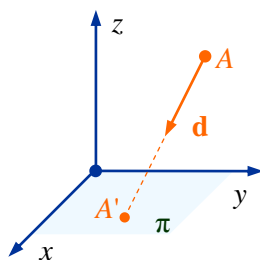
Středové promítání

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{-1}{f} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Matice popisuje projekci ze středu o souřadnicích $(0, 0, f)$ na rovinu $z = 0$ (tedy na rovinu xy). Ukážeme, jak je možné se o tom přesvědčit. Provedením transformace pro bod $(x, y, z, 1)$ podle předpisu (1.14) máme

$$(x, y, z, 1)\mathbf{T} = (x, y, z, \frac{f-z}{f}) \rightarrow (\frac{f}{f-z}x, \frac{f}{f-z}y, \frac{f}{f-z}z, 1).$$

Porovnejme tento výsledek s tím, co od dané projekce intuitivně očekáváme. Řekněme, že má být promítnut bod X o afinních souřadnicích x, y, z (obr. 1.3). Jeho průmět označme X' . Afinní souřadnice průmětu necht' jsou x', y', z' . Z podobnosti trojúhelníků (obr. 1.3) máme $\frac{x}{(f-z)} = \frac{x'}{f}$ (vezměte v úvahu, že souřadnice z bodu X na obr. 1.3 je záporná). Odtud plyne $x' = \frac{xf}{(f-z)}$. Podobně můžeme získat i souřadnici y' . Vyjde $y' = \frac{yf}{(f-z)}$. Souřadnici z' očekáváme $z' = 0$. Porovnáním uvedených hodnot s hodnotami získanými podle předpisu (1.14) zjišťujeme shodu u souřadnic x', y' . V souřadnici z' zjišťujeme rozdíl. Zatímco intuitivně očekávaná hodnota je $z' = 0$, projektivní transformace dává hodnotu $z' = \frac{zf}{(f-z)}$. Výsledkem projektivní transformace tedy není bod X' , ale bod X'' (obr. 1.3). To ale nevadí. Spíše se jedná o přednost. K vykreslení průmětu (např. na obrazovku) se totiž použijí souřadnice x', y' . Nenulová souřadnice z' může být využita k řešení viditelnosti.



Obr. 1.4: K ověření činnosti matice rovnoběžného promítání

Rovnoběžné promítání

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{d_x}{d_z} & -\frac{d_y}{d_z} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Matice popisuje rovnoběžné promítání na rovinu xy . Směr promítacího paprsku je $\mathbf{d} = (d_x, d_y, d_z)$ (obr. 1.4). O správnosti matice se opět snadno přesvědčíme. Nechť A je bod, který má být promítán, a necht' $\mathbf{a} = (a_x, a_y, a_z)$ je vektor jeho afinních souřadnic. Určíme polohový vektor $\mathbf{a}' = (a'_x, a'_y, a'_z)$ průmětu bodu A . Rovnice promítacího paprsku procházejícího bodem A je (t je parametr)

$$\mathbf{x} = \mathbf{a} + t\mathbf{d}.$$

Z podmínky $a'_z = 0$ dostáváme

$$t = -\frac{a_z}{d_z}.$$

Odtud dále máme

$$a'_x = a_x - \frac{a_z}{d_z}d_x, \quad a'_y = a_y - \frac{a_z}{d_z}d_y.$$

Porovnáním právě získaných výsledků s výsledky, které poskytuje transformace popsaná maticí \mathbf{T} , zjišťujeme shodu v prvních dvou souřadnicích. Jednička na pozici 3, 3 v matici \mathbf{T} zajišťuje, že souřadnice z průmětu je stejná jako u původního bodu. To je opět výhodné pro následné řešení viditelnosti.

Složitější transformace bývají běžně vytvářeny skládáním transformací jednodušších. Budeme se proto zabývat nyní i touto možností. Řekněme například, že jsme postupně provedli dvě transformace. Nejprve transformaci popsanou maticí \mathbf{T}_1 , pak transformaci \mathbf{T}_2 . Necht' \mathbf{x} je vektor souřadnic nějakého bodu před provedením obou transformací. Přepočít souřadnic odpovídající první transformaci provedeme podle vztahu $\mathbf{x}\mathbf{T}_1$. Skutečnost, že dále pak následovala ještě transformace \mathbf{T}_2 , zohledníme výpočtem podle vztahu $(\mathbf{x}\mathbf{T}_1)\mathbf{T}_2$. Protože platí $(\mathbf{x}\mathbf{T}_1)\mathbf{T}_2 = \mathbf{x}(\mathbf{T}_1\mathbf{T}_2)$, vidíme, že na uvedené dvě postupně prováděné transformace můžeme pohlížet jako na transformaci jedinou, která je popsána maticí $\mathbf{T}_1\mathbf{T}_2$. Právě popsáný postup skládání transformací násobením matic, které je popisují, lze pochopitelně zobecnit na libovolný počet dílčích transformací, jak ostatně v následujících příkladech velmi brzy uvidíte.

Z předchozího přehledu základních transformací vyplývá, že projektivní transformace je mocnější než transformace afinní. Afinní transformaci ze vztahu (1.2) lze realizovat jako speciální případ transformace projektivní (1.14) a tvůrci systémů to tak také velmi často dělají. Matice \mathbf{T} má pak v takovém případě tvar (viz též vztah (1.2))

$$\mathbf{T} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{t} & 1 \end{bmatrix},$$

kde $\mathbf{0}$ je sloupcový vektor obsahující 3 nuly. Rozdíl mezi afinní a projektivní transformací lze slovně charakterizovat takto. Obě transformace jsou lineární. To znamená, že lineární objekty (přímky, roviny) jsou oběma transformacemi zobrazeny opět na objekty lineární. Na rozdíl od projektivní transformace zachovává transformace afinní rovnoběžnost. Přímky nebo roviny, které jsou rovnoběžné před transformací, zůstanou rovnoběžné i po afinní transformaci.



Příklad 1.5. Zapište matici \mathbf{T}_a projektivní transformace, která realizuje posunutí objektu o vektor (x, y, z) , a dále matici \mathbf{T}_b , která realizuje rotaci objektu kolem osy x o úhel α a pak ještě rotaci kolem osy y o úhel β .

Řešení. Obě hledané matice snadno sestavíme s využitím předchozího přehledu, s využitím výsledku příkladu 1.2 a s uvážením, že, na rozdíl od příkladu 1.2, zde provádíme transformaci objektu a nikoli transformaci souřadné soustavy (obě varianty se navzájem liší znaménkem složek translace i úhlů rotace). Bude proto

$$\mathbf{T}_a = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ x & y & z & 1 \end{bmatrix}, \quad \mathbf{T}_b = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \beta & 0 & -\sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$



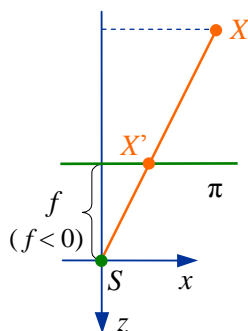
Příklad 1.6. Ukažte, že matice

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{f} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

realizuje středové promítání ze středu $(0, 0, 0)$ na rovinu $z = f$ (obr. 1.5).

Řešení. Snadno se o tom opět můžeme přesvědčit. Realizací transformace podle předpisu (1.14) máme

$$(x, y, z, 1)\mathbf{T} = \left(x, y, z, \frac{z}{f}\right) \rightarrow \left(\frac{f}{z}x, \frac{f}{z}y, f, 1\right).$$



Obr. 1.5: K ověření činnosti matice středového promítání z příkladu

Hodnoty, které od uvedené projekce očekáváme intuicí, lze opět snadno odvodit pomocí obr. 1.5, podobně jako jsme to provedli u středového promítání z předchozího přehledu. Tentokrát zjistíme shodu u všech tří souřadnic (x, y, z) , což prokazuje správnost dokazovaného tvrzení. Za povšimnutí stojí, že ke shodě nyní došlo i ve třetí souřadnici. Jak intuice, tak transformace popsaná maticí dávají shodně třetí souřadnici průmětu $z = f$. To ale činí uvedenou variantu poněkud méně praktickou, jestliže by měl následovat výpočet viditelnosti.

▲

Příklad 1.7. Ukažte, že matic realizujících nějakou jedinou konkrétní projektivní transformaci je nekonečně mnoho.



Řešení. Předpokládejme, že \mathbf{T} je matice popisující uvažovanou transformaci. Necht X značí bod, který je transformací zobrazen na bod Y . Body X, Y necht jsou reprezentovány takovými vektory \mathbf{x}, \mathbf{y} svých homogenních souřadnic, že platí $\mathbf{y} = \mathbf{x}\mathbf{T}$. Vezměme nyní libovolné reálné $\lambda \neq 0$ a násobme jím uvedenou rovnici. Dostaneme vztah $\lambda\mathbf{y} = \mathbf{x}(\lambda\mathbf{T})$. Jestliže vektor \mathbf{y} souřadnic reprezentoval bod Y (což jsme předpokládali), pak vektor $\lambda\mathbf{y}$ reprezentuje bod Y také (víme totiž, že v projektivním prostoru je každý bod reprezentován nekonečně mnoha kolineárními vektory). Shledáváme tedy, že jestliže \mathbf{T} byla matice realizující požadovanou projekci, pak lze tutéž projekci realizovat i maticí $\lambda\mathbf{T}$.

▲

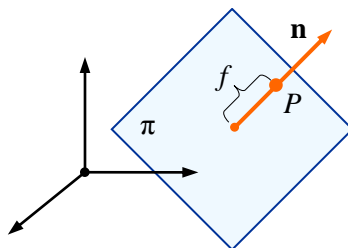
Příklad 1.8. Odhadněte, jakou transformaci realizuje matice



$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & \frac{-1}{f} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Řešení. Na základě předchozího přehledu lze snadno odhadnout, že se jedná o matici realizující středové promítání na rovinu yz . Afinní souřadnice středu projekce jsou $(f, 0, 0)$.

▲



Obr. 1.6: Prvky zadávající promítání

Shrnutí: V tuto chvíli byste měli zejména znát, jak lze matematicky popsat projektivní transformaci. Měli byste být schopni stanovit matici popisující jednoduché transformace nebo naopak rozhodnout, jakou transformaci zadaná (jednoduchá) matice provádí. Těchto znalostí využijeme v následujících příkladech, v nichž budeme z jednoduchých transformací skládat transformace složitější, které již budou velmi praktické.

1.5 Stanovení matice zobrazovací transformace – příklad A

V tomto příkladě ukážeme postup, jak lze dříve uvedených základních transformací a jim odpovídajících matic využít při sestavení matice realizující poněkud komplikovanější a prakticky již velmi užitečnou transformaci. Jedná se ve své podstatě o příklad, který je ale tak významný, že jej zařazujeme do textu jako samostatnou podkapitolu (v následujících podkapitolách budou prezentovány ještě další podobné příklady).

Úkolem je nalézt matici realizující středové promítání. Promítání je dáno středem $P = (p_x, p_y, p_z)$, normálou \mathbf{n} zobrazovací roviny, $\mathbf{n} = (n_x, n_y, n_z)$, a ohniskovou vzdáleností f , tj. vzdáleností středu projekce od zobrazovací roviny (obr. 1.6). Předpokládáme, že délka vektoru \mathbf{n} je $|\mathbf{n}|=1$ (tento předpoklad zjednoduší vztahy, které budeme později odvozovat).

Řešení. Hledanou transformaci rozdělíme na čtyři transformace dílčí. Cílem prvních tří kroků bude provést takovou transformaci souřadnic, abychom obdrželi jeden ze speciálních případů projekce z předchozí podkapitoly, pro který již odpovídající matice známe. Provedeme následující:

1. Počátek souřadné soustavy, kterou máme ve scéně zavedenu, posuneme do středu projekce P .
2. Provedeme rotaci této posunuté souřadné soustavy tak, aby její osa z' splýnula s normálou \mathbf{n} zobrazovací roviny.

3. Posuneme počátek potočené souřadné soustavy proti směru její osy z' o délku f tak, aby její počátek padl do zobrazovací roviny. (Po provedení dosud vyjmenovaných tří kroků leží rovina xy souřadné soustavy v rovině projekce. Střed projekce leží na ose z soustavy, a to v její kladné části ve vzdálenosti f od počátku.)
4. Provedeme projekci (nyní se již jedná o speciální případ uvedený v přehledu transformací v předchozí podkapitole, kdy střed projekce leží na ose z a zobrazovací rovina je $z = 0$).

Označíme-li matice uvedených transformací postupně jako $\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3, \mathbf{T}_4$, pak hledaná výsledná matice \mathbf{T} bude jejich součinem $\mathbf{T} = \mathbf{T}_1\mathbf{T}_2\mathbf{T}_3\mathbf{T}_4$. S využitím výsledků z předchozí podkapitoly můžeme ihned psát vztahy pro matice $\mathbf{T}_1, \mathbf{T}_3, \mathbf{T}_4$. Máme

$$\mathbf{T}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -p_x & -p_y & -p_z & 1 \end{bmatrix}, \quad \mathbf{T}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & f & 1 \end{bmatrix}, \quad \mathbf{T}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{f} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Ke stanovení zbývající matice rotace \mathbf{T}_2 použijeme postupu popsaného v podkapitole 1.2 a v příkladě 1.1. K tomu určíme souřadnice bázevých vektorů souřadné soustavy po rotaci (nová souřadná soustava) vzhledem k souřadné soustavě platné před rotací (stará souřadná soustava). Bázevé vektory staré souřadné soustavy označme $\mathbf{x}, \mathbf{y}, \mathbf{z}$. Bázevé vektory nové souřadné soustavy jsou $\mathbf{x}', \mathbf{y}', \mathbf{z}'$. (Poznamenejme, že i když je souřadná soustava platná před rotací vůči původní soustavě ve scéně již posunuta, jsou její bázevé vektory $\mathbf{x}, \mathbf{y}, \mathbf{z}$ stále tytéž jako u původní souřadné soustavy ve scéně.) Pro polohu souřadné soustavy po rotaci máme podmínku, aby osa z' byla kolmá k zobrazovací rovině (aby tedy měla směr zadaného vektoru \mathbf{n}). Podmínka však nespécifikuje polohu nové souřadné soustavy jednoznačně (soustava může zatím rotovat kolem své osy z'). Zvolíme proto další doplňující podmínku. Požadujeme, aby směr osy x' soustavy po rotaci byl kolmý nejen k ose z' , ale také k ose z souřadné soustavy před rotací, tedy k vektoru $\mathbf{z} = (0, 0, 1)$. Předpokládáme při tom, že vektory \mathbf{n} a \mathbf{z} nejsou kolineární. Z uvedených podmínek máme

$$\begin{aligned} \mathbf{z}' &= \mathbf{n}, \\ \mathbf{x}' &= \frac{\mathbf{z} \times \mathbf{n}}{|\mathbf{z} \times \mathbf{n}|} = \frac{(-n_y, n_x, 0)}{\sqrt{n_x^2 + n_y^2}}, \\ \mathbf{y}' &= \mathbf{z}' \times \mathbf{x}' = (n_x, n_y, n_z) \times \frac{(-n_y, n_x, 0)}{\sqrt{n_x^2 + n_y^2}} = \frac{(-n_x n_z, -n_y n_z, n_x^2 + n_y^2)}{\sqrt{n_x^2 + n_y^2}}. \end{aligned}$$

Zavedeme-li $\mathbf{x}' = (x'_x, x'_y, x'_z)$, $\mathbf{y}' = (y'_x, y'_y, y'_z)$, $\mathbf{z}' = (z'_x, z'_y, z'_z)$, pak s ohledem na ortonormalitu vyšetřované transformace (podkapitola 1.3, příklad 1.1) máme

$$\mathbf{T}_2 = \begin{bmatrix} x'_x & y'_x & z'_x & 0 \\ x'_y & y'_y & z'_y & 0 \\ x'_z & y'_z & z'_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$



Příklad 1.9. Spustte si program "úkol1_9.exe". Program předpokládá zadání zobrazení tak, jak bylo popsáno v této podkapitole, které čte ze souboru "úkol1_9.txt". Program počítá a vypisuje dílčí matice \mathbf{T}_1 až \mathbf{T}_4 a také výslednou transformační matici, která je jejich součinem. Na konkrétních hodnotách si tak můžete ověřit, zda jste vše správně pochopili. V případě potřeby můžete zadání modifikovat (jedná se o textový soubor).

Shrnutí: V této poněkud kratší podkapitole jste se naučili konstruovat matici realizující středové promítání, a to na základě takových zadávacích prvků, které jsou v grafických systémech v praxi obvyklé. V následujícím textu budeme pokračovat dalšími (složitějšími) případy.

1.6 Stanovení matice zobrazovací transformace – příklad B

Také tento příklad bude ještě věnován ukázce toho, jak lze složitější transformaci sestavit z transformací elementárních. Opět sestavíme matici středového promítání na základě velmi praktického zadání. Je dáno následující (obr. 1.7, 1.8).

$VRP = (vrp_x, vrp_y, vrp_z)$ (View Reference Point, souřadnice ve WC),

$VPN = (vpn_x, vpn_y, vpn_z)$ (View Plane Normal, souřadnice ve WC),

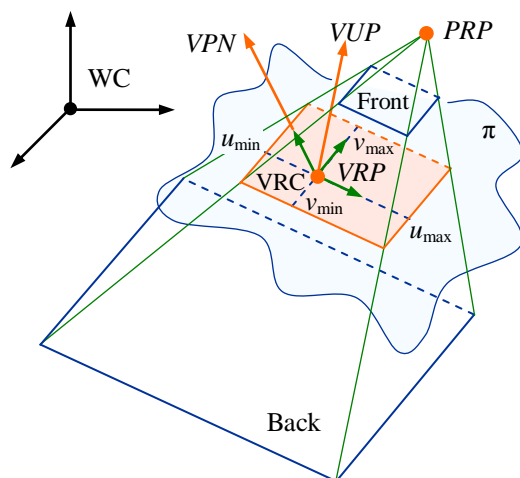
$VUP = (vup_x, vup_y, vup_z)$ (View Up Vector, souřadnice ve WC),

$PRP = (prp_x, prp_y, prp_z)$ (Projection Reference Point souřadnice ve VRC),

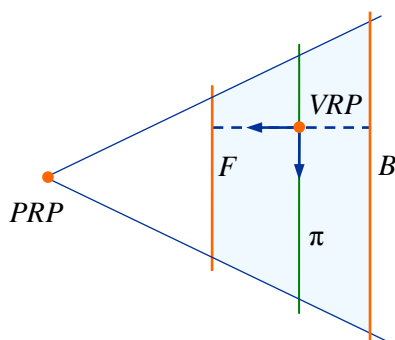
$Window = (u_{\min}, u_{\max}, v_{\min}, v_{\max})$ (Zorné pole souřadnice ve VRC),

F, B (Front, Back: přední a zadní ořezávací rovina; jedná se o dvě vzdálenosti vzhledem k VRC).

WC (world coordinate system) je souřadná soustava, která je zavedena v zobrazované scéně. Výše uvedeným zadáním je dále definována souřadná soustava obrazu VRC (view reference coordinate system), která umožňuje specifikaci dalších zadávacích prvků zobrazení. VRP je počátek souřadné soustavy VRC. Poloha počátku VRP je zadána ve WC. Bodem VRP také prochází průmětna, která splývá s rovinou xy soustavy VRC. Poloha průmětny (a tedy také roviny xy soustavy VRC) je dodefinována svojí normálou VPN (view plane normal). Souřadnice vektoru VPN



Obr. 1.7: Prvky zadávající uvažované středové promítání

Obr. 1.8: Prvky zadávající promítání - dvojrozměrné schéma; hodnoty F , B jsou včetně znaménka; zde je $F > 0$, $B < 0$

se zadávají v souřadné soustavě WC. Osa x souřadné soustavy VRC je kolmá jednak k vektoru VPN (který udává směr osy z soustavy VRC) a dále k vektoru VUP (view up vector). Vektor VUP je vektor, který se má na obrázku jevit svisle (kolmo k ose x obrazu). Souřadnice vektoru VUP se zadávají ve WC. PRP (projection reference point) je střed projekce. Poloha středu se zadává v souřadné soustavě VRC. Hodnoty u_{min} , u_{max} , v_{min} , v_{max} na osách x , y souřadné soustavy VRC definují v průmětně obdélník omezující obraz (jen obraz ležící uvnitř tohoto obdélníka bude získán). Uvedený obdélník spolu s bodem PRP definuje zorný jehlan. Zorný jehlan je dále omezen rovinami „front“ a „back“. Roviny jsou zadány svými vzdálenostmi F , B (opatřenými znaménkem) od průmětny (obr. 1.8). Zorný jehlan je tedy jehlanem komolým (obr. 1.7). Pouze objekty (nebo jejich části) ležící uvnitř zorného jehlanu budou zobrazeny.

Je možné, že se vám právě popsany způsob zadání bude zpočátku zdát dosti komplikovaný. Časem jej ale určitě shledáte velmi logickým. Je pravděpodobné, že pokud

byste sami měli vymyslet způsob zadání obecné projekce, vymysleli byste něco velmi podobného. Není také nezajímavé říci, že popsání způsobu zadání je využíváno grafickým standardem PHIGS.

Řešení. Hledanou transformaci sestavíme ze šesti elementárních transformací, jejichž transformační matice již umíme zapsat. Všechny dále podrobně popíšeme.

1. Posunutí počátku souřadné soustavy scény do *VRP*. Odpovídající matice má tvar

$$\mathbf{T}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -vrp_x & -vrp_y & -vrp_z & 1 \end{bmatrix}.$$

2. Natočení souřadné soustavy tak, aby osa z splynula s *VPN*. Vyjádříme bázevé vektory \mathbf{x}' , \mathbf{y}' , \mathbf{z}' soustavy po natočení vzhledem k soustavě před natočením. Dostaneme

$$\begin{aligned} \mathbf{z}' &= (z'_x, z'_y, z'_z) = \frac{VPN}{|VPN|}, \\ \mathbf{x}' &= (x'_x, x'_y, x'_z) = \frac{VUP \times VPN}{|VUP \times VPN|}, \\ \mathbf{y}' &= (y'_x, y'_y, y'_z) = \mathbf{z}' \times \mathbf{x}'. \end{aligned}$$

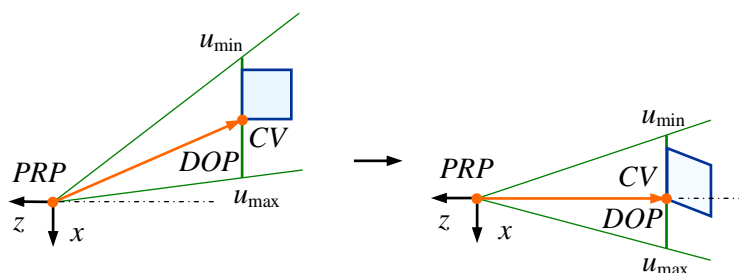
Matice rotace pak má tvar

$$\mathbf{T}_2 = \begin{bmatrix} x'_x & y'_x & z'_x & 0 \\ x'_y & y'_y & z'_y & 0 \\ x'_z & y'_z & z'_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

3. Posunutí počátku souřadné soustavy z *VRP* do *PRP*. Matice popisující tuto transformaci má tvar (povšimněte si, že se při konstrukci matice hodí, že souřadnice středu projekce *PRP* známe v soustavě *VRC*)

$$\mathbf{T}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -prp_x & -prp_y & -prp_z & 1 \end{bmatrix}.$$

4. Provedeme zkosení tak, aby zorný jehlan byl symetrický (obr. 1.9): Zavedme bod *CV* (center of view – střed obrazu) a vektor *DOP* (direction of projection) takto (souřadnice *CV* a *PRP* zde měříme v soustavě *VRC*)

Obr. 1.9: K činnosti transformace popsané maticí \mathbf{T}_4

$$CV = \left(\frac{u_{\max} + u_{\min}}{2}, \frac{v_{\max} + v_{\min}}{2}, 0 \right),$$

$$DOP = (dop_x, dop_y, dop_z) = CV - PRP.$$

Vektor DOP lze interpretovat také jako afinní souřadnice bodu CV v souřadné soustavě mající počátek v bodě PRP (do tohoto bodu jsme počátek soustavy předchozími třemi transformačními kroky přemístili). Matice realizující požadované zkosení má tvar

$$\mathbf{T}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{-dop_x}{dop_z} & \frac{-dop_y}{dop_z} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

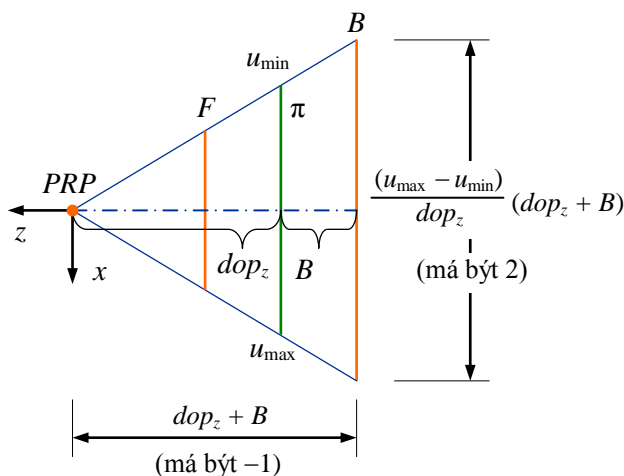
Správnost matice \mathbf{T}_4 můžeme snadno ověřit. Transformujme maticí \mathbf{T}_4 vektor $(dop_x, dop_y, dop_z, 1)$ (tedy bod CV). Dostaneme

$$(dop_x, dop_y, dop_z, 1)\mathbf{T}_4 = (0, 0, dop_z, 1).$$

Vidíme, že po transformaci bod CV (střed obrazu) leží na ose z . Zorný jehlan je tedy nyní symetrický vzhledem k souřadnicovým rovinám zx i yz .

5. Transformujeme velikost zorného jehlanu tak, aby měl jednotkové rozměry (obr. 1.10). Požadujeme, aby jeho větší podstava ležela v rovině $z = -1$ a aby délka její strany byla 2 (od -1 do $+1$ v obou směrech). Jedná se o změnu měřítka, a proto bude mít matice realizující uvedenou transformaci tvar

$$\mathbf{T}_5 = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Obr. 1.10: K činnosti transformace popsané maticí \mathbf{T}_5

Hodnoty s_x , s_y , s_z stanovíme jednoduše na základě následujících skutečností: Vzdálenost mezi středem projekce a větší podstavou je nyní $(dop_z + B)$, ale má být -1 (poznamenejme, že dop_z i B jsou obvykle záporné). Označme q_x , q_y rozměr větší podstavy jehlanu ve směru osy x , resp. y . Po transformaci mají být obě tyto šířky 2. Na základě podobnosti trojúhelníků máme (obr. 1.10)

$$q_x = \frac{u_{\max} - u_{\min}}{dop_z} (dop_z + B), \quad q_y = \frac{v_{\max} - v_{\min}}{dop_z} (dop_z + B).$$

Nyní již snadno určíme

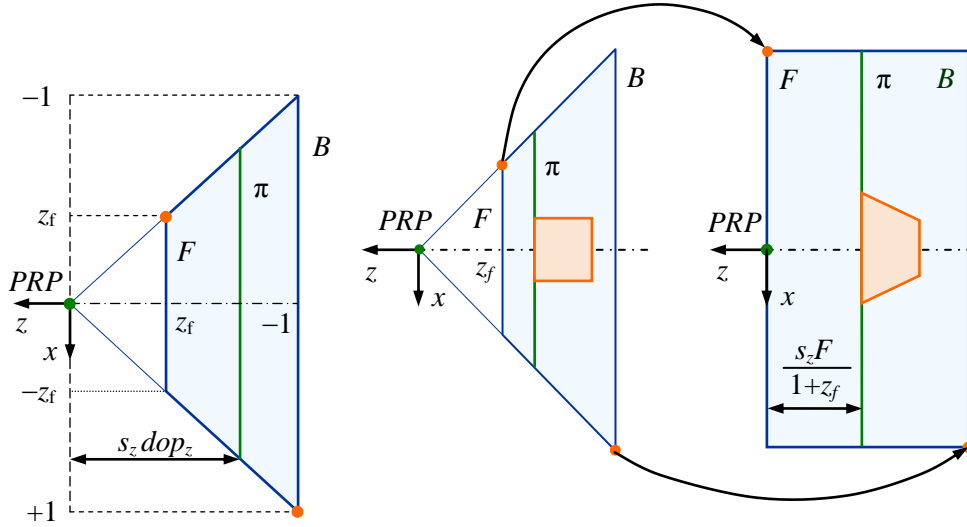
$$s_z = \frac{-1}{(dop_z + B)},$$

$$s_x = \frac{2dop_z}{(u_{\max} - u_{\min})(dop_z + B)},$$

$$s_y = \frac{2dop_z}{(v_{\max} - v_{\min})(dop_z + B)}.$$

6. Nakonec provedeme projekci. Abychom usnadnili následující výpočty (zejména výpočet viditelnosti metodou „z-buffer“), provedeme projekci tak, že scénu transformujeme (deformujeme) tak, aby rovnoběžné promítání dalo tentýž výsledek jako původně zamýšlené promítání středové (obr. 1.11). Místo středového promítání pak použijeme promítání rovnoběžného s promítacím pářskem rovnoběžným s osou z , které lze realizovat velmi snadno. Uvedená transformace transformuje zorný jehlan na hranol. Zavedme hodnotu

$$z_f = s_z(dop_z + F).$$



Obr. 1.11: K činnosti transformace popsané maticí \mathbf{T}_6 . Situace před transformací (obrázek vlevo). Transformace zorného jehlanu na kvádr (obrázek vpravo)

Dále zavedme matici

$$\mathbf{T}_6 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{1+z_f} & -1 \\ 0 & 0 & \frac{-z_f}{1+z_f} & 0 \end{bmatrix}.$$

Snadno ověříme, že matice \mathbf{T}_6 skutečně transformuje zorný jehlan na hranol. Můžeme to provést tak, že transformaci aplikujeme na některé význačné body zorného jehlanu. Např. (promyslete, které body jehlanu jsme transformovali):

$$(z_f, z_f, z_f, 1)\mathbf{T}_6 = \left(z_f, z_f, \frac{z_f}{1+z_f} - \frac{z_f}{1+z_f}, -z_f \right) \rightarrow (-1, -1, 0, 1),$$

$$(1, 1, -1, 1)\mathbf{T}_6 = \left(1, 1, \frac{-1}{1+z_f} - \frac{z_f}{1+z_f}, 1 \right) \rightarrow (1, 1, -1, 1).$$

Nechť $\mathbf{x}, \tilde{\mathbf{x}}$ značí vektory souřadnic bodu před a po provedení zobrazovací transformace zadané v tomto příkladě. V prvním případě se tedy jedná o souřadnice ve „světové“ souřadné soustavě scény, ve druhém případě o souřadnice v jednotkovém zobrazovacím hranolu, ke kterému jsme právě popsaným postupem dospěli. Platí

$$\tilde{\mathbf{x}} = \mathbf{x}\mathbf{T}_1\mathbf{T}_2\mathbf{T}_3\mathbf{T}_4\mathbf{T}_5\mathbf{T}_6 = \mathbf{x}\mathbf{T}, \quad (1.15)$$

kde je $\mathbf{T} = \mathbf{T}_1\mathbf{T}_2\mathbf{T}_3\mathbf{T}_4\mathbf{T}_5\mathbf{T}_6$. Poznamenejme ještě, že transformace a matice $\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3$ jsou ortonormální. Transformace $\mathbf{T}_4, \mathbf{T}_5, \mathbf{T}_6$ nikoli. ▲

Příklad 1.10. Spusťte si program „úkol1_10.exe“. Program předpokládá zadání zobrazení tak, jak bylo popsáno v této podkapitole, které čte ze souboru s názvem „úkol1_10.txt“. Program počítá a vypisuje dílčí matice \mathbf{T}_1 až \mathbf{T}_6 a také výslednou transformační matici, která je jejich součinem. Na konkrétních hodnotách si tak můžete ověřit, zda jste vše správně pochopili. V případě potřeby můžete zadání modifikovat (jedná se o textový soubor).



Shrnutí: Seznámili jste se s dalším příkladem, v němž byla matice zobrazovací transformace sestavena z dílčích matic transformací elementárních. Postup popsáný v tomto příkladě je velmi praktický a bude vám současně sloužit jako návod pro sestavení matice projekce v programu, který budete realizovat jako svoji samostatnou práci. Ujistěte se proto, zda všem krokům dokonale rozumíte.

1.7 Stanovení matice zobrazovací transformace – příklad C

V tomto příkladě ukážeme, že velmi podobně, jako jsme v předchozím případě B realizovali promítání středové, lze realizovat i promítání rovnoběžné. Promítání bude opět realizováno rovnicí (1.14). Drobné odchylky budou pouze ve stanovení matice \mathbf{T} . Provedeme následující: Prvky popsané v předchozím příkladě použijeme k zadání rovnoběžného promítání. Bod PRP spolu s bodem CV (viz. řešení příkladu B) v tomto případě určují směr promítacích paprsků. Význam ostatních prvků zůstává beze změny. Stanovíme matici realizující takto zadané promítání.

Řešení. Hledanou výslednou matici \mathbf{T} opět stanovíme jako součin matic realizujících jednotlivé dílčí transformace. I v tomto případě bude

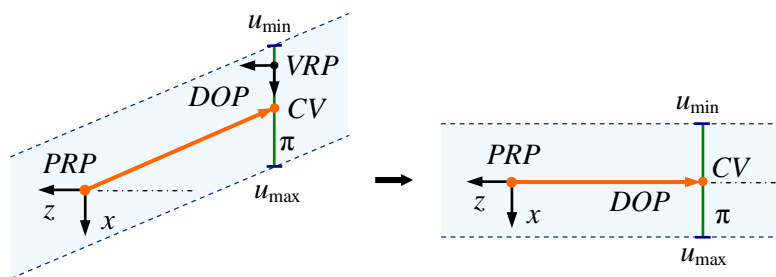
$$\mathbf{T} = \mathbf{T}_1 \mathbf{T}_2 \mathbf{T}_3 \mathbf{T}_4 \mathbf{T}_5 \mathbf{T}_6 .$$

Dílčí transformace 1) až 4) zůstávají zcela beze změny jako v příkladě B a beze změny zůstávají i jim odpovídající matice. Transformace 4) v tomto případě zajišťuje, aby promítací paprsky byly kolmé k průmětně (obr. 1.12). Transformace 5) převádí zorný kvádr na kvádr s jednotkovými rozměry. Jednotlivá měřítka v matici \mathbf{T}_5 se nyní stanoví takto (obr. 1.13)

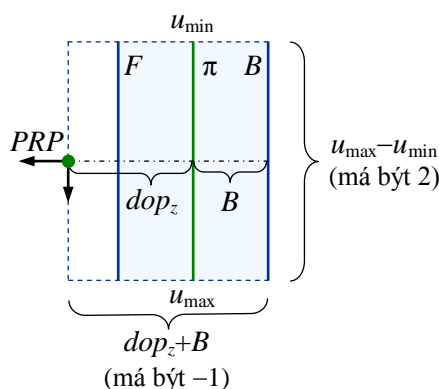
$$s_x = \frac{2}{u_{\max} - u_{\min}}, \quad s_y = \frac{2}{v_{\max} - v_{\min}}, \quad s_z = -\frac{1}{\text{dop}_z + B}. \quad (1.16)$$

Transformace 6) zajišťuje, aby přední rovina jednotkového zobrazovacího hranolu ležela v rovině $z = 0$ (až dosud ležela v rovině $z = z_f$) (obr. 1.14). Matice \mathbf{T}_6 má tvar

$$\mathbf{T}_6 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{1+z_f} & 0 \\ 0 & 0 & \frac{-z_f}{1+z_f} & 1 \end{bmatrix} .$$



Obr. 1.12: K transformaci reprezentované maticí \mathbf{T}_4



Obr. 1.13: K transformaci reprezentované maticí \mathbf{T}_5

Podobně jako v předchozím příkladě můžeme i nyní matici prověřit tak, že transformaci aplikujeme na některé významné body zorného kvádru. Snadno např. zjistíme, že je

$$(1, 1, z_f, 1) \mathbf{T}_6 = (1, 1, 0, 1) \text{ a také } (1, 1, -1, 1) \mathbf{T}_6 = (1, 1, -1, 1) .$$



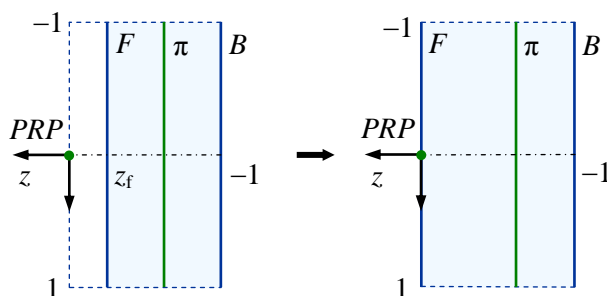
Příklad 1.11. Sestavte matici realizující rovnoběžné promítání, které je zadáno bodem P , kterým prochází průmětna, a normálou \mathbf{n} průmětny.



Řešení. Se znalostmi, které nyní již máte, by pro vás nalezení matice nemělo být problémem. Napovíme ale, že pomocí transformací popsanych maticemi $\mathbf{T}_1, \mathbf{T}_2$ z příkladu A v podkapitole 1.5 lze požadovanou projekci převést na projekci paprsky rovnoběžnými s osou z souřadné soustavy.



Shrnutí: Oba příklady (příklad B a C), které jste právě vyřešili, vám ukazují, že mezi realizací středového a rovnoběžného promítání není v grafických systémech zapotřebí činit žádných rozdílů. Oba případy lze realizovat vztahem (1.14) pro projektivní transformaci a liší se navzájem jen v konkrétních hodnotách matice \mathbf{T} .

Obr. 1.14: K transformaci reprezentované maticí \mathbf{T}_6

Příklady nám také ukázaly, že postup stanovení matice \mathbf{T} je si v obou případech velmi podobný.

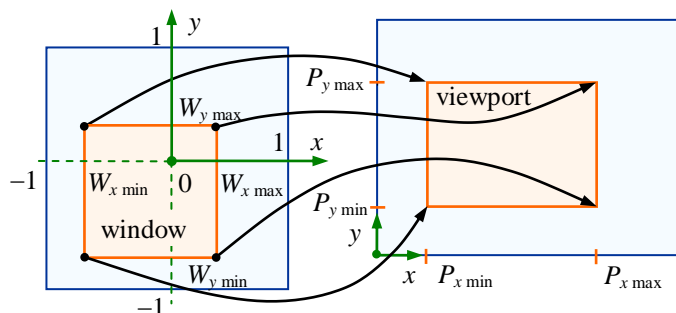
1.8 Transformace na výstupní zařízení

Při studiu příkladů B, C v předcházejících podkapitolách jste se mohli podívat, proč se zobrazovací transformace prováděla tak, aby souřadnice objektů, které mají být zobrazeny, padly do jednotkového zobrazovacího objemu. Mohli jste namítnout, že při vykreslování získaného obrazu na nějaké výstupní zařízení nebude pravděpodobně možné souřadnice z intervalu $\langle -1, 1 \rangle$ použít. Tyto pochybnosti byly na místě. Po zobrazovací transformaci do jednotkového zobrazovacího objemu totiž ještě následuje transformace na výstupní zařízení, kterou se v této podkapitole naučíte provádět.

Ačkoli je závěrečná transformace na výstupní zařízení z teoretického pohledu pouze jednoduchou afinní transformací, bývá často oddělena od předchozích transformačních kroků. Důvody jsou následující:

1. Předpokládá se, že může být více výstupních zařízení. Všechna pak mohou použít téhož výpočtu zobrazení, jen v závěrečné fázi, při transformaci na výstupní zařízení, se výpočet pro jednotlivá zařízení liší.
2. Předpokládá se, že závěrečnou část zobrazení obsluha mění častěji než části předchozí (např. prováděním výřezu atd.). V takovém případě se pak opakuje (pokud možno) pouze výpočet závěrečné části zobrazovací transformace, který je jednoduchý, a proto rychlý. Pro pořádek ještě dodejme, že z použití normalizovaného zorného objemu vyplývají i jisté výhody při provádění ořezávání.

Transformace na výstupní zařízení bývá zadána pomocí okna v normalizovaném zorném objemu. Toto okno říká, která část normalizovaného objemu se má zobrazit. (Poznamenejme, že v jednodušším případě lze zobrazit celý rozsah zorného objemu, okno pak není zapotřebí.) ViewPort naopak určuje oblast na výstupním zařízení, do které se má požadovaný obrázek umístit. Pro transformaci na výstupní zařízení tedy zpravidla známe následující hodnoty (obr. 1.15)



Obr. 1.15: K transformaci na výstupní zařízení

$$\begin{aligned} Window &= (W_{x \min}, W_{y \min}, W_{x \max}, W_{y \max}), \\ ViewPort &= (P_{x \min}, P_{y \min}, P_{x \max}, P_{y \max}). \end{aligned}$$

Nechť x', y', z' jsou souřadnice na výstupním zařízení a $\tilde{x}, \tilde{y}, \tilde{z}$ souřadnice v normalizovaném zorném objemu. Pro transformaci souřadnic na výstupní zařízení pak na základě úměry jednoduše máme

$$\begin{aligned} x' &= P_{x \min} + \frac{\tilde{x} - W_{x \min}}{W_{x \max} - W_{x \min}} (P_{x \max} - P_{x \min}), \\ y' &= P_{y \min} + \frac{\tilde{y} - W_{y \min}}{W_{y \max} - W_{y \min}} (P_{y \max} - P_{y \min}), \\ z' &= \tilde{z}. \end{aligned}$$

Poznamenejme, že souřadnici z' potřebujeme, má-li být řešena viditelnost. Předpis pro její výpočet může být stanoven s jistou volností. Musí však zůstat zachováno pořadí bodů a objektů, jak se jeví ve směru od pozorovatele. Výše uvedený jednoduchý postup tomuto kritériu vyhovuje. Dále poznamenejme, že někdy je požadováno nastavit okno tak, aby byly zobrazeny všechny objekty scény. V takovém případě je pak nutné nejprve vypočítat zobrazovací transformaci pro všechny objekty a pak nalézt extrémní hodnoty souřadnic x, y v normalizovaném zorném objemu, které pak budou použity pro specifikaci okna.

1.9 Inverze zobrazovací transformace – příklad D

Provádět inverzní transformaci k transformaci zobrazovací je zapotřebí např. při výpočtu Phongova stínování nebo při nanášení textury. V těchto případech je totiž nutné pro daný bod na výstupním zařízení (pixel obrazu) zjistit polohu jeho vzoru ve scéně (nebo jinak: je zapotřebí zjistit, který bod „skutečného světa“ se promítl do daného bodu obrazu). Detaily o použití této úlohy se lze dočíst [8]. Na tomto

místě si pouze připravíme inverzní transformaci k transformaci, která je složena ze zobrazovací transformace z podkapitoly 1.6 a z transformace na výstupní zařízení podle předchozí podkapitoly 1.8.

Řešení. Necht x', y', z' jsou souřadnice na výstupním zařízení a $\tilde{x}, \tilde{y}, \tilde{z}$ souřadnice v normalizovaném zobrazovacím hranolu zavedeném v podkapitole 1.6. Inverzí vztahů z podkapitoly 1.8 snadno získáme

$$\begin{aligned}\tilde{x} &= W_{x \min} + \frac{x' - P_{x \min}}{P_{x \max} - P_{x \min}}(W_{x \max} - W_{x \min}), \\ \tilde{y} &= W_{y \min} + \frac{y' - P_{y \min}}{P_{y \max} - P_{y \min}}(W_{y \max} - W_{y \min}), \\ \tilde{z} &= z'.\end{aligned}$$

Necht \mathbf{x} je vektor souřadnic v souřadné soustavě scény. Inverzí předpisu pro zobrazovací transformaci z příkladu v podkapitole 1.6 dostaneme

$$\mathbf{x} = \tilde{\mathbf{x}} \mathbf{T}_6^{-1} \mathbf{T}_5^{-1} \mathbf{T}_4^{-1} \mathbf{T}_3^{-1} \mathbf{T}_2^{-1} \mathbf{T}_1^{-1}.$$

Matice \mathbf{T}_1 až \mathbf{T}_6 byly již uvedeny v podkapitole 1.6. Inverzní matice mají tvar

$$\begin{aligned}\mathbf{T}_1^{-1} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ vrp_x & vrp_y & vrp_z & 1 \end{bmatrix}, \\ \mathbf{T}_2^{-1} &= \mathbf{T}_2^T, \\ \mathbf{T}_3^{-1} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ prp_x & prp_y & prp_z & 1 \end{bmatrix}, \\ \mathbf{T}_4^{-1} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{dop_x}{dop_z} & \frac{dop_y}{dop_z} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \\ \mathbf{T}_5^{-1} &= \begin{bmatrix} \frac{1}{s_x} & 0 & 0 & 0 \\ 0 & \frac{1}{s_y} & 0 & 0 \\ 0 & 0 & \frac{1}{s_z} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \\ \mathbf{T}_6^{-1} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1+z_f}{z_f} \\ 0 & 0 & -1 & -\frac{1}{z_f} \end{bmatrix}.\end{aligned}$$

▲

Poznamenejme, že ačkoli je uvedené matice možné získat „čistě mechanickou“ inverzí původních matic \mathbf{T}_1 až \mathbf{T}_6 , může být často pohodlnější zapsat je jako matice provádějící jednotlivé dílčí transformační kroky „v opačném směru“, než byl směr transformací odpovídajících maticím původním. Například: Jestliže matice \mathbf{T}_1 popisuje posunutí počátku souřadné soustavy scény do bodu VRP , bude matice \mathbf{T}_1^{-1} popisovat posunutí z bodu VRP zpět do původního počátku. Při praktické realizaci grafického systému byste samozřejmě také mohli invertovat až výslednou matici \mathbf{T} ze vztahu (1.15). Použili byste k tomu obvyklých algoritmů pro výpočet inverzní matice, které určitě znáte z matematiky.

Příklad 1.12. Stanovte inverzní transformaci k zobrazovací transformaci z příkladu C v podkapitole 1.7.



Řešení. Využijeme výsledků z předchozího příkladu D. Všechny vztahy zůstávají v platnosti kromě matice \mathbf{T}_6^{-1} . Zapište tuto matici. ▲

Shrnutí: Na tomto místě už byste měli vědět vše, co z teorie afinních a projektivních prostorů a transformací budete ke konstrukci grafických systémů potřebovat. Víte, že větší systémy pracují velmi často tak, že nejprve provedou zobrazovací transformaci do jednotkového zorného objemu tak, jak jsme to i my provedli v příkladech B nebo C. Pak následuje transformace na nějaké konkrétní výstupní zařízení, jak jsme ukázali v podkapitole 1.8. Víte, že grafické systémy zobrazování realizují jako projektivní transformaci podle vztahu (1.14). Umíte již také sestavit matici, která požadovanou projektivní transformaci realizuje. Víte, že je obvyklé sestavit požadovanou transformaci z transformací dílčích. Matice výsledné hledané transformace je pak vypočtena jako součin matic, které popisují použité dílčí transformace.

1.10 Dodatky ke kapitole o afinních prostorech a afinních transformacích

V této závěrečné podkapitole o afinních a projektivních prostorech a transformacích uvedeme nakonec ještě slíbené definice, kterým jsme se v předchozím textu záměrně pokusili vyhnout. Podkapitola je určena zejména pro teoreticky zaměřené čtenáře, kteří by snad mohli zjednodušení v předchozím textu pocítovat rušivě. Je také současně ale možné, že i pro ostatní by nyní mohlo být zajímavé vědět, jak jsou přesně definovány pojmy, s nimiž jsme v kapitole pracovali. Pokud ale takovou informaci nepostrádáte, můžete podkapitolu přeskočit.

Definice 1.13. *Afinní prostor* : Neprázdnou množinu \mathbf{A}_n nazveme afinním prostorem dimenze n (prvky množiny \mathbf{A}_n budeme nazývat body), jestliže je dán vektorový prostor \mathbf{V}_n dimenze n a zobrazení $\varphi: \mathbf{A}_n \times \mathbf{A}_n \rightarrow \mathbf{V}_n$, které má následující dvě vlastnosti:

1. Pro každý bod $A \in \mathbf{A}_n$ a pro každý vektor $\mathbf{v} \in \mathbf{V}_n$ existuje jediný bod $B \in \mathbf{A}_n$ tak, že $\varphi(A, B) = \mathbf{v}$.
2. Pro každé tři body $A, B, C \in \mathbf{A}_n$ platí, že $\varphi(A, C) = \varphi(A, B) + \varphi(B, C)$.

Přesně vzato je afinní prostor uspořádaná trojice $(\mathbf{A}_n, \mathbf{V}_n, \varphi)$. Často se však používá pouze stručného zápisu \mathbf{A}_n . Vektorový prostor \mathbf{V}_n bývá nazýván vektorovým zaměřením afinního prostoru \mathbf{A}_n .

Definice 1.14. *Afinní zobrazení* : Necht $\mathbf{A}_m, \mathbf{B}_n$ jsou dva afinní prostory dimenzí m, n a $\mathbf{V}_m, \mathbf{W}_n$ necht jsou jejich vektorová zaměření. Zobrazení $f: \mathbf{A}_m \rightarrow \mathbf{B}_n$ nazveme afinním zobrazením, jestliže existuje takový lineární operátor $\mathcal{A}: \mathbf{V}_m \rightarrow \mathbf{W}_n$, že pro každý bod $M \in \mathbf{A}_m$ a pro každý vektor $\mathbf{v} \in \mathbf{V}_m$ platí, že $f(M + \mathbf{v}) = f(M) + \mathcal{A}(\mathbf{v})$.

Definice 1.15. *Projektivní prostor* : Necht \mathbf{V}_{n+1} je vektorový prostor dimenze $n+1$. Množinu všech směrů ve \mathbf{V}_{n+1} nazveme projektivním prostorem dimenze n (a označíme $\langle \mathbf{V}_{n+1} \rangle$ nebo \mathbf{P}_n).

Definice 1.16. *Homogenní souřadnice* : Homogenními souřadnicemi v projektivním prostoru \mathbf{P}_n nazveme uspořádanou $(n+1)$ -tici, která určuje směr v prostoru \mathbf{V}_{n+1} . Každá taková $(n+1)$ -tice určuje jednoznačně nějaký bod v \mathbf{P}_n . Naopak každý bod může být reprezentován více (nekonečně mnoha) takovými $(n+1)$ -ticemi. Necht $\mathbf{x} \in \mathbf{V}_{n+1}$ je vektor reprezentující bod X , pak také vektor $\lambda \mathbf{x}$, $\lambda \neq 0 \in \mathbf{R}$, reprezentuje X .

Definice 1.17. *Projektivní zobrazení* : Necht $\mathbf{P}_n = \langle \mathbf{V}_{n+1} \rangle$, $\mathbf{P}'_n = \langle \mathbf{V}'_{n+1} \rangle$ jsou dva projektivní prostory a necht f je izomorfismus \mathbf{V}_{n+1} na \mathbf{V}'_{n+1} . Zobrazení $F: \mathbf{P}_n \rightarrow \mathbf{P}'_n$ je kolineací, jestliže pro každý bod $X \in \mathbf{P}_n$ a pro každého vektorového zástupce $\mathbf{x} \in \mathbf{V}_{n+1}$ tohoto bodu platí, že $F(X)$ je bod reprezentovaný vektorem $f(\mathbf{x})$.

Nejsou-li vám definice úplně proti mysli, můžete se nyní např. zamyslet nad tím, proč je afinní prostor definován právě tak, jak jsme uvedli. Co by se stalo, kdyby nebyly splněny ony dvě vlastnosti zobrazení φ , které definice požaduje?

Kapitola 2

Křivky a plochy v počítačové grafice

Průvodce studiem



Tato kapitola se bude zabývat popisem a vývojem křivek a ploch používaných v počítačové grafice. V první části se zaměříme na definování základních pojmů, rozdělení a popis obecných vlastností křivek a následně i ploch. V druhé části se zaměříme na jednotlivé konkrétní křivky a plochy, popíšeme si jejich princip a srovnáme jejich výhody a nevýhody.

V dnešní době se setkáváme s potřebou křivek a ploch napříč různými obory. Neobejdeme se bez nich snad v žádném vědním oboru (matematika, fyzika, statistika apod.) ani v technické praxi (strojírenství, stavebnictví apod.). Podle potřeby různých oborů se také odlišují požadavky na chování a vlastnosti použitých křivek. S tím jsou také spojeny různé představy o křivkách, kde třeba v případě fyziky se křivka popisuje jako dráha hmotného bodu (částice) a vlastnosti se pak přirovnávají k rychlosti, zrychlení apod.

Naším cílem je přiblížit čtenáři základní informace týkající se vývoje křivek a ploch používaných v počítačové grafice. Problém křivek a ploch je samozřejmě velmi rozsáhlý a stále se vyvíjející. Naším cílem bude shrnout nejdůležitější informace a demonstrovat vlastnosti vhodně vybraných zástupců křivek a ploch. Zajisté je možné si informace, které by byly nad rámec tohoto dokumentu, v případě zájmu kdykoli rozšířit, například dohledáním v některých z citovaných materiálů.

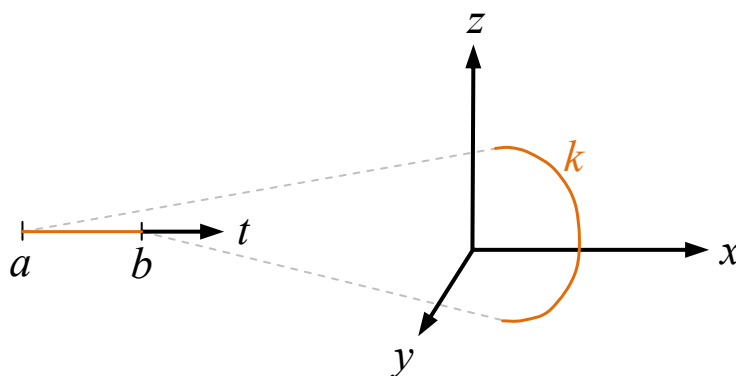
Po základním úvodu a seznámení se s vlastnostmi křivek a ploch se postupně seznámíte s křivkami polynomiálními, mezi které patří interpolace obecným algebraickým polynomem, Hermitovou interpolací a metodou nejmenších čtverců. Následně se budeme věnovat parametrickým křivkám, jako je Fergusonova křivka, Bézierova křivka, B-spline a NURBS křivky. Seznámíte se také se zástupci ploch, mezi které patří Bézierova plocha, B-spline plocha a NURBS plocha.

2.1 Křivka a její vlastnosti

Ačkoli se může zdát, že pojem „křivka“ je intuitivně jasný, ukázalo se, že najít definici matematicky přesnou a současně odpovídající tomu, co intuitivně chápeme pod tímto pojmem, není vůbec snadné. Počátek křivek a ploch sahá hluboko do naší minulosti spojené s každodenní činností člověka (stavba přístřeší, dráha pohybujících se těles apod.). Jak již také bylo zmíněno v úvodu, různé obory přistupují k pojmu křivka různými způsoby (geometrická topologie, diferenciální geometrie ve fyzice, počítačová grafika apod.).

V našem případě vyjdeme z následující definice parametrické křivky.

Definice 2.1. Mějme interval $I = \langle a, b \rangle \subseteq \mathbb{R}$. Parametricky vyjádřenou (parametrizovanou) křivku k v \mathbb{R}^n nazýváme diferencovatelné zobrazení $\phi : I \rightarrow \mathbb{R}^n$.



Obr. 2.1: Zobrazení $\phi : I \rightarrow \mathbb{R}^3$

Jedná se o parametrické vyjádření křivky, kde zobrazení ϕ je n -tice spojitých reálných funkcí jedné proměnné definovaných na intervalu $\langle a, b \rangle$. Derivace ϕ v krajních bodech a, b intervalu I chápeme jako jednostranné. V případě trojrozměrného euklidovského prostoru bude každému číslu t z intervalu I odpovídat na křivce příslušný bod $P(t) = [x(t), y(t), z(t)]$. Množina bodů $\{P(t), t \in I\}$ se nazývá obrazem křivky a číslo t nazýváme parametr.

Polohu bodu P lze v dané souřadné soustavě vyjádřit pomocí jeho polohového vektoru \mathbf{P} (vektor daný počátkem souřadné soustavy a souřadnicemi příslušného bodu P). To, zdali máme na mysli souřadnice bodu nebo jeho polohový vektor, bude zřejmé ze zápisu.

Parametricky vyjádřenou křivku můžeme tedy zapisovat její vektorovou rovnicí

$$\mathbf{P}(t) = (x(t), y(t), z(t)), \text{ kde } t \in I. \quad (2.1)$$

Poznámka 2.2. V našem případě, pokud nebude uvedeno jinak, budeme křivku popisovat její vektorovou rovnicí $\mathbf{P}(t)$.

Řekneme, že pro křivku zadanou svou vektorovou rovnicí $\mathbf{P}(t)$ na intervalu $I = \langle a, b \rangle$, kde $t \in \langle a, b \rangle$, je bod daný polohovým vektorem $\mathbf{P}(a)$ *počátečním* bodem křivky. Bod daný polohovým vektorem $\mathbf{P}(b)$ je *koncovým* bodem křivky. Parametrickou křivku nazýváme *uzavřenou*, pokud je počáteční bod roven koncovému bodu

$$\mathbf{P}(a) = \mathbf{P}(b). \quad (2.2)$$

Definice 2.3. Parametrickou křivku zadanou vektorovou rovnicí $\mathbf{P}(t)$ nazýváme *regulární*, jestliže má navíc nenulové první derivace pro všechna $t \in I$. V trojrozměrném prostoru tedy musí platit

$$\mathbf{P}'(t) \neq (0, 0, 0) \text{ pro všechna } t \in I. \quad (2.3)$$

Jinak řečeno: křivka je regulární, pokud pro všechna $t \in I$ platí

$$|\mathbf{P}'(t)| \neq 0. \quad (2.4)$$

Poznámka 2.4. Vektor $\mathbf{P}'(t)$ je tečným vektorem křivky v bodě $\mathbf{P}(t)$. Blíže se o tečném vektoru dozvíte v samostatné části popisující tečnu a tečný vektor.

Příklad 2.5. Zjistěte, zda se v případě křivky zadané vektorovou rovnicí $\mathbf{P}(t) = (t^2, t^2, 0)$, kde $t \in (-\infty, \infty)$, jedná o regulární křivku.



Řešení. Jak je vidět ze zadání, jedná se o přímku v rovině $z = 0$. O regulární křivku by se jednalo, pokud by platil vztah (2.3). V našem případě zkoumejme $t = 0$. Máme

$$\mathbf{P}'(0) = (0, 0, 0).$$

Pro zkoumané $t = 0$ bude mít křivka nulové první derivace, a nejedná se tedy o křivku regulární. ▲

Křivky můžeme rozdělit na

- *rovinné křivky* - křivka je popsána spojitými funkcemi $x, y : \langle a, b \rangle \rightarrow \mathbb{R}^2$ jako množina bodů $P(t) = [x(t), y(t)]$, kde parametr $t \in \langle a, b \rangle$;
- *prostorové křivky* - křivka je popsána třemi spojitými funkcemi $x, y, z : \langle a, b \rangle \rightarrow \mathbb{R}^3$ jako množina bodů $P(t) = [x(t), y(t), z(t)]$, kde $t \in \langle a, b \rangle$.

Implicitní a explicitní popis křivky

Mimo parametricky vyjádřené křivky existují také křivky zadané *implicitně* nebo *explicitně*. Explicitní vyjádření rovinné křivky je ve tvaru

$$y = f(x), \quad (2.5)$$

která je definována na intervalu I , na kterém má spojité své první derivace. Množinu všech bodů, které leží v prostoru \mathbb{R}^3 a které mají souřadnice $[x, f(x), g(x)]$, nazýváme křivkou definovanou explicitně.

Příklad 2.6. Je možné přejít od explicitního vyjádření k parametrickému v případě regulární křivky a opačně?



Řešení. Můžeme jednoduše ukázat, že v případě regulární křivky lze explicitní rovnici zapsat jednoduše v parametrickém tvaru jako

$$x = t, \quad y = f(t), \quad z = g(t), \quad \text{kde } t \in I.$$

Uvedené rovnice mají tvar parametrických rovnic křivky. Ukázali jsme tak, že regulární křivku zadanou explicitně lze popsat rovnicí v parametrickém tvaru. ▲

Poznámka 2.7. Regulární křivku definovanou parametricky již ale nelze vždy převést na rovnici v explicitním tvaru (příkladem takovéto křivky může být kružnice).

Pro úplnost dodejme popis křivky i implicitně vyjádřené. Mějme v E^3 množinu bodů $\{(x, y, z)\}$, určenou funkcemi

$$f(x, y, z) = 0, \quad g(x, y, z) = 0. \quad (2.6)$$

Křivka zadaná touto rovnicí je vlastně průnikem dvou ploch. Pokud je množina bodů neprázdná a následující matice (Jacobiho matice)

$$\begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \end{vmatrix} \quad (2.7)$$

má hodnotu rovnu dvěma, nazýváme tuto množinu *regulární křivkou definovanou implicitně*.

Poznámka 2.8. Pokud máme křivku zadanou v explicitním tvaru, převedením všech proměnných a konstant na jednu stranu rovnice dostaneme implicitní vyjádření. Obdobně lze implicitní vyjádření získat z parametrické rovnice pomocí eliminací parametru.

Křivky lze tedy podle způsobu popisu rozdělit na:

- 1) *Explicitně popsané křivky* $y = f(x)$, kde $x \in I$.
Příklad: $y = x^2 + x + 1$, jedná se o parabolu.
- 2) *Implicitně popsané křivky* $F(x, y) = 0$.
Příklad: $(x + 9)^2 + (y - 2)^2 - 4 = 0$, jedná se o kružnici se středem v bodě $[-9, 2]$ a s poloměrem 2.
- 3) *Parametricky popsané křivky* $x = x(t), y = y(t)$, kde parametr $t \in \langle a, b \rangle$.
Příklad: $x = t, y = t^2$, kde $t \in \mathbb{R}$ je parabola s vrcholem v počátku.



Příklad 2.9. Zapište rovnici kružnice k se středem v počátku souřadné soustavy a poloměrem r ve všech typech výše popsaných rovnic.

Řešení. Kružnici lze popsat jako

- množinu bodů $(x, y) \in \mathbb{R}^2 : x^2 + y^2 - r^2 = 0$,
- graf funkce $y = \sqrt{r^2 - x^2} \cup y = -\sqrt{r^2 - x^2}$,
- trajektorii bodu $x(t) = r \cos(t)$, $y(t) = r \sin(t)$, kde $t \in \langle 0, 2\pi \rangle$.



Jak lze již vidět z příkladu, mají různé typy vyjádření křivek různý význam a s tím spojené své výhody i nevýhody. Obecně lze říci, že v případě počítačové grafiky je nejvhodnějším tvarem parametrická rovnice, se kterou se budeme v další části setkávat nejčastěji. To ale neznamená, že by se ostatní tvary nepoužívaly. Implicitní tvar má svůj význam například v případě výpočtu průsečíků křivek a ploch. Explicitní křivky jsou zase nejjednodušší z hlediska konstrukce. Výhodou parametrických křivek je, že mají přirozeným způsobem určen směr jejich procházení. Například je snadné je nahradit jejich po částech lineární aproximací generováním posloupnosti jednotlivých bodů aproximačního polygonu.

Další způsob dělení křivek může být podle zadání:

- *křivky dané analytickým předpisem* (svým matematickým zápisem). Jedná se o nejvhodnější zadání, v praxi však obvykle analytický předpis nemáme přímo zadán a musíme ho nejprve nalézt.
- *křivky dané jednotlivými body*. Podle požadavků zadání pak v tomto případě obvykle hledáme interpolační nebo aproximační křivkou (obecný algebraický polynom, metoda nejmenších čtverců apod.).
- *křivky určené pro modelování*. V případě těchto křivek máme zadány řídicí body (popřípadě další vlastnosti jako např. tečné vektory). Hlavním cílem těchto křivek je, aby měly vhodné vlastnosti pro modelování a bylo je možné v rozumném čase spočítat. Mezi tyto křivky patří dále popsané Bézierovy křivky, B-spline křivky nebo NURBS křivky.

2.1.1 Parametrická a geometrická spojitost

Jednou z významných vlastností křivek je jejich spojitost. Výsledná křivka má parametrickou spojitost nulté třídy (označujeme C^0), je-li spojitá ve všech svých bodech. Křivka má parametrickou spojitost prvního řádu (křivka je hladká), pokud má ve všech bodech spojitě také své první derivace.

Definice 2.10. Křivka má parametrickou spojitost třídy n (značíme C^n), pokud pro každý její bod existují spojitě derivace až do řádu n .

Velmi často budeme řešit spojitost křivek, které vzniknou napojením jednotlivých částí (segmentů). Pro spojení třídy C^1 u křivky vzniklé napojením dvou částí $\mathbf{P}(t)$ a $\mathbf{Q}(t)$, musí pro tečné vektory $\mathbf{P}'(1)$ a $\mathbf{Q}'(0)$ ve společném bodě $\mathbf{P}(1) = \mathbf{Q}(0)$ platit, že jsou totožné, tj. $\mathbf{P}'(1) = \mathbf{Q}'(0)$.

Definice 2.11. Mějme křivku složenou ze dvou částí $\mathbf{P}(t)$ a $\mathbf{Q}(t)$. Křivka má ve společném bodě $\mathbf{P}(1) = \mathbf{Q}(0)$ spojení třídy C^n , pokud má v tomto bodě spojitě derivace až do řádu n , tj.

$$\frac{d^n}{dt^n} \mathbf{P}(1) = \frac{d^n}{dt^n} \mathbf{Q}(0). \quad (2.8)$$

Mimo popsanou parametrickou spojitost se v praxi setkáváme ještě se spojitostí geometrickou, kterou označujeme G^n , přičemž její dodržení bývá snazší. Oproti parametrické spojitosti, např. třídy C^1 , kdy byla zaručena totožnost tečných vektorů, postačí pro geometrickou spojitost G^1 , aby tečné vektory v navazovaném bodě byly pouze souhlasně kolineární. Obecně musí platit

$$\frac{d^n}{dt^n} \mathbf{P}(1) = k \frac{d^n}{dt^n} \mathbf{Q}(0), \text{ kde } k > 0. \quad (2.9)$$

2.1.2 Vlastnosti křivek v počítačové grafice

Při popisu konkrétních křivek budeme zkoumat některé další vlastnosti, jako jsou:

1. *lokalita změn* - změnou polohy řídicího bodu se změní pouze omezená část křivky, nikoli celá křivka.
2. *počáteční a koncový bod* - křivka může nebo nemusí procházet krajními (počátečním a koncovým) body svého řídicího polygonu.
3. *invariance k afinním transformacím* - vlastnost zaručuje, že afinní transformace řídicích bodů a následné generování křivky bude mít stejný výsledek, jako transformace jednotlivých bodů vygenerované křivky.
4. *vlastnost konvexní obálky* (ang. convex hull property) - celá křivka (nebo část křivky) leží v konvexní obálce řídicího polygonu.

2.1.3 Tečný vektor, tečna

V každém bodě křivky můžeme určit tečný vektor udávající aktuální směr pohybuujícího se hmotného bodu po křivce z počátečního do koncového bodu. *Tečný vektor*

(označme ho \mathbf{t}) parametricky zadané křivky $\mathbf{P}(t)$ v daném bodě $t = t_0$ získáme jako vektor

$$\mathbf{t} = \mathbf{P}'(t_0) = \frac{d\mathbf{P}}{dt}(t_0) = (x'(t_0), y'(t_0), z'(t_0)), \quad (2.10)$$

což znamená, že tečný vektor zadané křivky $\mathbf{P}(t)$ je určen derivacemi složek parametrické rovnice. Velikost vektoru $|\mathbf{P}'(t)|$ vypočteme jako

$$|\mathbf{P}'(t)| = \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2}. \quad (2.11)$$

Vektor

$$\hat{\mathbf{t}} = \frac{\mathbf{P}'(t)}{|\mathbf{P}'(t)|} \quad (2.12)$$

nazýváme *jednotkovým tečným vektorem* parametricky zadané křivky $\mathbf{P}(t)$.

Tečný vektor v bodě $\mathbf{P}(t_0)$ lze zapsat i jako

$$\mathbf{P}'(t_0) = \lim_{h \rightarrow 0} \frac{\mathbf{P}(t_0 + h) - \mathbf{P}(t_0)}{h}, \quad (2.13)$$

kde je hledaný tečný vektor určen polohovým vektorem pevného bodu $\mathbf{P}(t_0)$ a polohovým vektorem bodu $\mathbf{P}(t_0 + h)$, který se přibližuje (hodnota h se blíží k nule) k tomuto pevnému bodu. *Tečna křivky* je tedy limitní polohou sečny (přímka určená dvěma průsečíky s křivkou, pro kterou platí, že její dva průsečíky s křivkou splynou v jediný bod dotyku). Tečnu křivky v libovolném bodě $\mathbf{P}(t_0)$ určíme za pomoci tohoto bodu $\mathbf{P}(t_0)$ a tečného vektoru $\mathbf{P}'(t_0)$ jako

$$\mathbf{Q}(u) = \mathbf{P}(t_0) + u\mathbf{P}'(t_0), \quad \text{kde } u \in \mathbb{R}. \quad (2.14)$$

Bod $\mathbf{P}(t_0)$ se nazývá dotykovým bodem tečny. Hledaná tečna je tedy přímkou, která je určena dotykovým bodem $\mathbf{P}(t_0)$ a tečným vektorem $\mathbf{P}'(t_0)$, který je směrovým vektorem hledané tečny. Tečna křivky je nezávislá na zvolené parametrizaci křivky.

Regulární křivka má v každém bodě jedinou tečnu měnící se v závislosti na změně polohy bodu na křivce. Pokud v daném bodě $\mathbf{P}(t_0)$ existuje jediná tečna, nazveme tento bod *regulárním bodem*. V případě neregulární křivky mohou existovat body mající v jednom bodě více tečen, těmto bodům se říká *singulární body*.

Poznámka 2.12. Jak je patrné, lze z parametrické rovnice lehce vyjádřit tečný vektor a tečnu. Toho budeme později využívat při řešení křivek, které se vytváří napojováním jednotlivých segmentů (částí).

2.1.4 Délka křivky

Délku parametricky vyjádřené křivky $\mathbf{P}(t)$, kde $t \in I = \langle t_1, t_2 \rangle$ vypočteme pomocí integrálu jako

$$l = \int_{t_1}^{t_2} \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} dt = \int_{t_1}^{t_2} |\mathbf{P}'(t)| dt, \quad (2.15)$$

kde t_1 a t_2 jsou parametry krajních bodů křivky.



Příklad 2.13. Vypočtěte délku oblouku křivky $\mathbf{P}(t) = (r \cos t, r \sin t, vt)$, kde $r, v > 0$ pro $t \in \langle 0, 2\pi \rangle$.

Řešení. Derivace vektorové rovnice $\mathbf{P}(t)$ je rovna

$$\mathbf{P}'(t) = (-r \sin t, r \cos t, v).$$

Využijeme vztah (2.15), kde

$$\begin{aligned} l &= \int_0^{2\pi} \sqrt{(-r)^2 \sin^2 t + r^2 \cos^2 t + v^2} dt \\ &= \int_0^{2\pi} \sqrt{r^2(\sin^2 t + \cos^2 t) + v^2} dt \\ &= \int_0^{2\pi} \sqrt{r^2 + v^2} dt \\ &= 2\pi \sqrt{r^2 + v^2}. \end{aligned}$$

Výsledná délka křivky je rovna $2\pi\sqrt{r^2 + v^2}$. ▲

2.1.5 Parametrizace křivky

Někdy může být vhodnější pro regulární křivku, danou vektorovou rovnicí, přejít k jiné vektorové rovnici, která ale vyjadřuje tutéž křivku. Přejít mezi vektorovými rovnicemi se nazývá *transformace parametru*.

Definice 2.14. Necht $\mathbf{P}(t)$, $t \in I$ je křivka k třídy C^n . Na intervalu J necht je definována funkce $f(s)$, která má na J nenulové první derivace a zobrazuje J na interval I . Křivka $\mathbf{P}(f(s))$, $s \in J$ je křivkou, v níž byla provedena transformace parametru.



Příklad 2.15. Pro danou křivku $\mathbf{P}(t) = (2t^2, \sin t)$, $t \in (1, 4)$ transformujte vektorovou rovnici na novou $\mathbf{Q}(s)$, kde $t = s^2$, $s \in (1, 2)$.

Řešení. Hledaná rovnice bude mít po dosazení tvar

$$\mathbf{Q}(s) = (2s^4, \sin s^2), \quad s \in (1, 2).$$

▲

2.1.6 Parametrizace obloukem

Mezi významné parametrizace patří parametrizace obloukem, kdy je transformace parametru provedena tak, že parametr měří délku dané křivky. Vyjděme ze vztahu (2.15) a definujme funkci

$$s(t) = \int_a^t |\mathbf{P}'(u)| du, \quad (2.16)$$

která vyjadřuje délku křivky $\mathbf{P}(u)$ na intervalu $I = \langle a, t \rangle$ a nazýváme ji obloukem křivky. Oblouk křivky $s(t)$ je rostoucí funkce a existuje k ní inverzní funkce $t(s)$, kde $s : I \rightarrow J$ a $t : J \rightarrow I$. Pak o křivce

$$\mathbf{Q}(s) = \mathbf{P}(t(s)), s \in J \quad (2.17)$$

říkáme, že se jedná o křivku parametrizovanou obloukem.

Poznámka 2.16. Pro křivku parametrizovanou obloukem platí, že tečný vektor $\mathbf{Q}'(s)$ bude v každém jejím bodě jednotkový.

2.1.7 Oskulační rovina

Každá rovina procházející tečnou regulární křivky v dotykovém bodě $\mathbf{P}(t_0)$ se nazývá *tečná rovina*. Je zřejmé, že daným bodem křivky prochází svazek tečných rovin. Obdobně jako v případě tečny, je *oskulační rovina* limitní polohou tečných rovin. Mějme tedy v bodě $\mathbf{P}(t_0)$ tečnu \mathbf{Q} . Touto tečnou a dalším bodem $\mathbf{P}(t_0+h)$ proložíme tečnou rovinu α . Oskulační rovina (značíme τ) křivky $\mathbf{P}(t)$ v bodě t_0 je limitní případ tečné roviny, procházející dalším bodem $\mathbf{P}(t_0+h)$, kde $h \rightarrow 0$.

Oskulační rovinu snadno určíme pomocí pevného bodu $\mathbf{P}(t_0)$ a dvou vektorů $\mathbf{P}'(t_0)$ a $\mathbf{P}''(t_0)$. Platí, že jestliže jsou oba vektory lineárně nezávislé, existuje právě jedna oskulační rovina v daném bodě $\mathbf{P}(t_0)$. V opačném případě je oskulační rovinou každá tečná rovina.

Pro úplnost uvedme vztah pro druhou derivaci křivky $\mathbf{P}(t)$, který je dán jako

$$\mathbf{P}''(t_0) = \frac{d\mathbf{P}'}{dt}(t_0) = \lim_{h \rightarrow 0} \frac{\mathbf{P}'(t_0+h) - \mathbf{P}'(t_0)}{h}. \quad (2.18)$$

Definice 2.17. Mějme regulární křivku $\mathbf{P} = \mathbf{P}(t), t \in I$ a libovolný bod křivky $\mathbf{P}(t_0)$. Rovinu danou bodem $\mathbf{P}(t_0)$ a lineárně nezávislými vektory $\mathbf{P}'(t_0)$ a $\mathbf{P}''(t_0)$ nazýváme oskulační rovinou τ v bodě $\mathbf{P}(t_0)$. Platí tedy

$$\tau(u, v) = \mathbf{P}(t_0) + u\mathbf{P}'(t_0) + v\mathbf{P}''(t_0), \text{ kde } u, v \in \mathbb{R}. \quad (2.19)$$

Bod křivky, v němž jsou $\mathbf{P}'(t_0)$ a $\mathbf{P}''(t_0)$ kolineární (lineárně závislé), nazýváme *inflexní bod*.

Přímku procházející bodem $\mathbf{P}(t_0)$ a kolmou na oskulační rovinu nazýváme *binormálou* a značíme ji \mathbf{b} . Binormála je tedy kolmá na oba vektory $\mathbf{P}'(t_0)$ a $\mathbf{P}''(t_0)$, které definují oskulační rovinu τ . Jednotkový vektor binormály $\hat{\mathbf{b}}$ vypočteme na základě vektorového součinu obou vektorů jako

$$\hat{\mathbf{b}} = \frac{\mathbf{P}'(t_0) \times \mathbf{P}''(t_0)}{|\mathbf{P}'(t_0) \times \mathbf{P}''(t_0)|}. \quad (2.20)$$

Pro úplnost dodejme, že vektorový součin dvou vektorů $\mathbf{u} = (u_1, u_2, u_3)$ a $\mathbf{v} = (v_1, v_2, v_3)$ je vektor kolmý na oba vektory \mathbf{u}, \mathbf{v} . Definován je jako

$$\mathbf{u} \times \mathbf{v} = ((u_2v_3 - u_3v_2), (u_3v_1 - u_1v_3), (u_1v_2 - u_2v_1)). \quad (2.21)$$

Normálou křivky v daném bodě $\mathbf{P}(t_0)$ rozumíme každou přímku kolmou na tečnu. Existuje však jedna normála ležící navíc v oskulační rovině τ , kterou nazýváme *hlavní normálou* křivky a značíme ji \mathbf{n} . *Jednotkový vektor hlavní normály* $\hat{\mathbf{n}}$ je kolmý na tečný vektor a binormálu v daném bodě $\mathbf{P}(t_0)$. Platí tedy

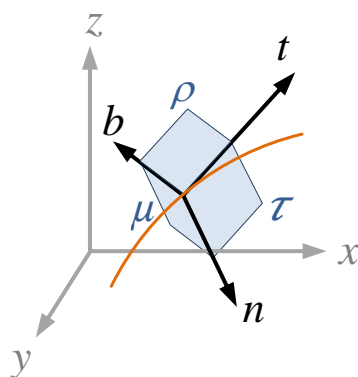
$$\hat{\mathbf{n}} = \frac{\mathbf{b} \times \mathbf{t}}{|\mathbf{b} \times \mathbf{t}|}. \quad (2.22)$$

Pro jednotkové vektory $\hat{\mathbf{t}}$ a $\hat{\mathbf{b}}$, stačí psát

$$\hat{\mathbf{n}} = \hat{\mathbf{b}} \times \hat{\mathbf{t}}. \quad (2.23)$$

Jak je již asi patrné, lze libovolný vektor vypočítat jako vektorový součin zbývajících dvou. Platí tedy

$$\mathbf{n} = \mathbf{b} \times \mathbf{t}, \quad \mathbf{b} = \mathbf{t} \times \mathbf{n}, \quad \mathbf{t} = \mathbf{n} \times \mathbf{b}. \quad (2.24)$$



Obr. 2.2: Zobrazení Frenetova průvodního (doprovodného) trojhranu

Rovina tvořená hlavní normálou \mathbf{n} a binormálou \mathbf{b} se nazývá *normálovou rovinou* (značíme ji μ). Rovina tvořená tečnou \mathbf{t} a binormálou \mathbf{b} nazýváme *rektifikační rovinou* a značíme ρ .

Oskulační, normálová a rektifikační rovina jsou navzájem kolmé roviny a tvoří Frenetův průvodní (doprovodný) trojhran křivky v bodě $\mathbf{P}(t_0)$ (obrázek 2.2).

Poznámka 2.18. Křivka, která má v každém svém bodě nekonečný počet oskulačních rovin (vektory $\mathbf{P}'(t_0)$ a $\mathbf{P}''(t_0)$ jsou kolineární), je přímkou nebo částí přímky. V takovém případě se názvy hlavní normály a binormály nezavádějí.

2.1.8 Křivost a torze křivky

Křivost křivky je jedna ze základních vlastností, které charakterizují jednotlivé křivky. U křivky rozlišujeme dva typy křivosti. *První křivost* (flexe), obvykle se označuje jen pojmem „křivost“ (ang. curvature), udává velikost odchýlení křivky od přímky v daném bodě $\mathbf{P}(t_0)$. Čím bude mít křivka v daném bodě větší křivost, tím více se v okolí zkoumaného bodu odchyluje od tečny.

Výpočet křivosti křivky zadané vektorovou rovnicí $\mathbf{P}(t), t \in I$ podle obecného parametru získáme jako

$$k_1(t) = \frac{|\mathbf{P}'(t) \times \mathbf{P}''(t)|}{|\mathbf{P}'(t)|^3}. \quad (2.25)$$

Poznámka 2.19. V inflexních bodech je křivost nulová. Lze tedy na základě první křivosti určovat inflexní body.

Poznámka 2.20. Pokud má regulární křivka v každém svém bodě křivost rovnou nule, bude tato křivka přímkou nebo částí přímky.

Příklad 2.21. Určete křivost pro křivku $\mathbf{P}(t) = (t, t^2, t^3)$.



Řešení. Využijeme vztahu (2.25), kde je

$$\begin{aligned} \mathbf{P}(t) &= (t, t^2, t^3), \\ \mathbf{P}'(t) &= (1, 2t, 3t^2), \\ |\mathbf{P}'(t)| &= \sqrt{(1 + 4t^2 + 9t^4)}, \\ \mathbf{P}''(t) &= (0, 2, 6t), \\ \mathbf{P}'(t) \times \mathbf{P}''(t) &= (6t^2, -6t, 2), \\ |\mathbf{P}'(t) \times \mathbf{P}''(t)| &= \sqrt{(4 + 36t^2 + 36t^4)}, \\ k_1(t) &= \frac{\sqrt{(4 + 36t^2 + 36t^4)}}{\left(\sqrt{(1 + 4t^2 + 9t^4)}\right)^3}. \end{aligned}$$

Výsledná křivost $k_1(t) = \frac{\sqrt{(4+36t^2+36t^4)}}{\left(\sqrt{(1+4t^2+9t^4)}\right)^3}$. ▲

Poznámka 2.22. Bude-li křivka parametrizovaná obloukem, bude její křivost v bodě $\mathbf{P}(s)$ rovna $|\mathbf{P}''(s)|$.

Druhá křivost nazývaná *torze* (ang. torsion) je mírou pro odchýlení křivky $\mathbf{P}(t)$ v daném bodě $\mathbf{P}(t_0)$ z její oskulační roviny do prostoru. Vypočteme ji jako

$$k_2(t) = \frac{(\mathbf{P}'(t) \times \mathbf{P}''(t)) \cdot \mathbf{P}'''(t)}{|\mathbf{P}'(t) \times \mathbf{P}''(t)|^2}. \quad (2.26)$$

Poznámka 2.23. Pokud má regulární křivka v každém svém bodě druhou křivost rovnou nule, bude tato křivka rovinnou křivkou.



Příklad 2.24. Určete druhou křivost pro křivku $\mathbf{P}(t) = (t, t^2, t^3)$ (křivka z předcházejícího příkladu 2.21).

Řešení. Využijeme vztahu (2.26), kde je

$$\begin{aligned} \mathbf{P}(t) &= (t, t^2, t^3), \\ \mathbf{P}'(t) &= (1, 2t, 3t^2), \\ \mathbf{P}''(t) &= (0, 2, 6t), \\ \mathbf{P}'''(t) &= (0, 0, 6), \\ \mathbf{P}'(t) \times \mathbf{P}''(t) &= (6t^2, -6t, 2), \\ (\mathbf{P}'(t) \times \mathbf{P}''(t)) \cdot \mathbf{P}'''(t) &= 12, \\ |\mathbf{P}'(t) \times \mathbf{P}''(t)| &= \sqrt{(4 + 36t^2 + 36t^4)}, \\ k_2(t) &= \frac{12}{4 + 36t^2 + 36t^4}. \end{aligned}$$

Výslednou druhou křivost (torzi) určíme jako $k_2(t) = \frac{12}{4+36t^2+36t^4}$. ▲

Poznámka 2.25. V případě zájmu lze odvození vztahů křivosti (2.25) a (2.26) nalézt v [1, 2].

Pro případ rovinné křivky zadané explicitně $y = f(x)$, lze vzorec pro první křivost v bodě $[x, y]$ zjednodušit na tvar

$$k_1 = \frac{|y''|}{(1 + (y')^2)^{\frac{3}{2}}}. \quad (2.27)$$



Příklad 2.26. Určete první křivost funkce $y = \sqrt{4 - x^2}$ (horní část kružnice $x^2 + y^2 = 4$) v bodě $x_1 = 0$ a $x_2 = \sqrt{2}$.

Řešení. Využijeme vzorce (2.27) pro výpočet první křivosti. Budeme potřebovat první y' a druhou derivaci y'' , kde první derivace bude mít tvar

$$y' = \frac{1}{2}(4 - x)^{-\frac{1}{2}}(-2x) = \frac{-x}{\sqrt{4 - x^2}}.$$

Druhá derivace bude rovna

$$y'' = -\frac{4}{(4-x^2)^{\frac{3}{2}}}.$$

Výslednou první křivost pro $x = 0$ vypočteme jako

$$k_1(0) = \frac{|y''|}{(1+(y')^2)^{\frac{3}{2}}} = \frac{\left|-\frac{4}{(4-x^2)^{\frac{3}{2}}}\right|}{\left(1+\left(\frac{-x}{\sqrt{4-x^2}}\right)^2\right)^{\frac{3}{2}}} = \frac{\left|-\frac{4}{(4-0^2)^{\frac{3}{2}}}\right|}{\left(1+\left(\frac{-0}{\sqrt{4-0^2}}\right)^2\right)^{\frac{3}{2}}} = \frac{1}{2}.$$

Výslednou první křivost pro $x = \sqrt{2}$ vypočteme jako

$$k_1(\sqrt{2}) = \frac{|y''|}{(1+(y')^2)^{\frac{3}{2}}} = \frac{\left|-\frac{4}{(4-x^2)^{\frac{3}{2}}}\right|}{\left(1+\left(\frac{-x}{\sqrt{4-x^2}}\right)^2\right)^{\frac{3}{2}}} = \frac{\left|-\frac{4}{(4-\sqrt{2}^2)^{\frac{3}{2}}}\right|}{\left(1+\left(\frac{-\sqrt{2}}{\sqrt{4-\sqrt{2}^2}}\right)^2\right)^{\frac{3}{2}}} = \frac{1}{2}.$$

Jelikož se jedná o kružnici, bude první křivost pro všechny x_i rovna $\frac{1}{2}$. ▲

2.1.9 Oskulační kružnice

Mějme zadanou parametrickou křivku $\mathbf{P}(t)$. Kružnice ležící v oskulační rovině bodu $\mathbf{P}(t_0)$ mající ve zkoumaném bodě $\mathbf{P}(t_0)$ stejnou tečnu s tečnou křivky v daném bodě, a která má střed na hlavní normále \mathbf{n} a poloměr r daný vztahem

$$r = \left| \frac{1}{k_1} \right|,$$

se nazývá oskulační kružnice. Jedná se o limitní případ kružnice, která se dotýká v daném bodě křivky a má stejnou křivost jako křivka v tomto bodě.

Příklad 2.27. Určete oskulační kružnici pro křivku v předcházejícím příkladě. 

Řešení. Protože je křivost konstantní a je rovna $\frac{1}{2}$, bude mít oskulační kružnice poloměr

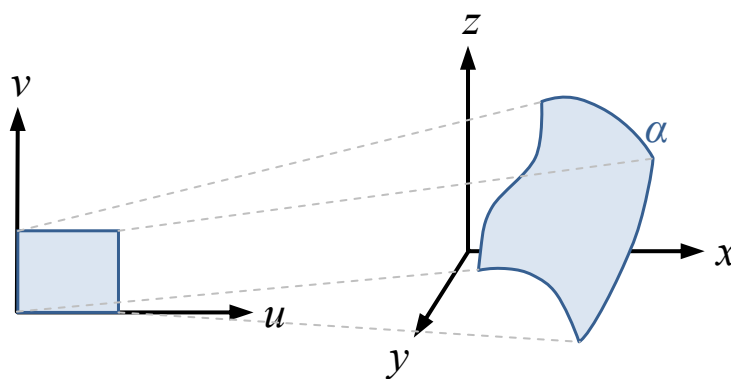
$$r = \frac{1}{k_1} = 2.$$
▲

Poznámka 2.28. Jelikož se jedná o část kružnice, bude oskulační kružnicí sama sobě.

2.2 Plochy v počítačové grafice

Rozšířením křivek se dostaneme k plochám, které však mají s křivkami hodně společného, zejména u konkrétních ploch, které vznikly rozšířením křivek (Bézierova plocha, NURBS plocha apod.). Podobně jako u křivek si před popisem konkrétních ploch přiblížíme jejich základní definice a vlastnosti.

Definice 2.29. *Parametrickou plochou* v \mathbb{R}^3 nazýváme diferencovatelné zobrazení $\theta : U \rightarrow \mathbb{R}^3$ na podmnožině $U \subseteq \mathbb{R}^2$.



Obr. 2.3: Zobrazení parametrické plochy

Plochu zadanou svou vektorovou rovnicí $\mathbf{Q}(u, v)$ můžeme zapsat jako

$$\mathbf{Q}(u, v) = (x(u, v), y(u, v), z(u, v)), \text{ kde } u, v \in I. \quad (2.28)$$

Analogicky jako křivky, můžeme rozdělit plochy podle vyjádření rovnice na:

- 1) **Explicitně popsané plochy** $z = f(x, y)$, kde $x, y \in J$.
Příklad: $z = 2x^3 - 2y^2 + x - 2y + 10$.
- 2) **Implicitně popsané plochy** $F(x, y, z) = 0$.
Příklad: $x^2 + y^2 + z^2 - r^2 = 0$ je kulová plocha se středem v počátku souřadné soustavy a o poloměru r .
- 3) **Parametricky popsané plochy** $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$ kde parametr $u, v \in \langle a, b \rangle$.
Příklad: $x = r \cos(u) \cos(v)$, $y = r \sin(u) \cos(v)$, $z = r \sin(v)$ kde $0 \leq u < 2\pi$, $-\frac{\pi}{2} < v < \frac{\pi}{2}$ je kulová plocha se středem v počátku a poloměrem r .

Obdobně jako u křivek jsou jednotlivá vyjádření rozdílná a mají rozdílné výhody a nevýhody. Z parametrického vyjádření je snadné získat jednotlivé body, z implicitního vyjádření zase můžeme jednoduše testovat, zda vybraný bod patří do zadané

plochy nebo ne. V našem případě se zaměříme na plochy parametrické, které jsou v počítačové grafice nejvyužívanější.

Dále lze plochy rozdělit podle toho, jak jsou zadané:

- plochy zadané analytickým předpisem (výše popsané typy vyjádření);
- plochy zadané hraničními křivkami;
- plochy zadané sítí bodů (Bézierova plocha, NURBS, apod.);
- plochy vytvářené kinematicky (rotační plochy, plochy vzniklé skládáním pohybu).

2.2.1 Vlastnosti ploch

Tečný vektor $\mathbf{t}_u(u, v)$ (resp. $\mathbf{t}_v(u, v)$) ve směru parametru u (resp. ve směru v) k parametricky zadané ploše $\mathbf{Q}(u, v)$ určíme jako

$$\mathbf{t}_u(u, v) = \frac{\partial \mathbf{Q}(u, v)}{\partial u} = \left(\frac{\partial x(u, v)}{\partial u}, \frac{\partial y(u, v)}{\partial u}, \frac{\partial z(u, v)}{\partial u} \right), \quad (2.29)$$

resp.

$$\mathbf{t}_v(u, v) = \frac{\partial \mathbf{Q}(u, v)}{\partial v} = \left(\frac{\partial x(u, v)}{\partial v}, \frac{\partial y(u, v)}{\partial v}, \frac{\partial z(u, v)}{\partial v} \right). \quad (2.30)$$

Pokud jsou vektory $\mathbf{t}_u(u, v)$ a $\mathbf{t}_v(u, v)$ nekolineární, nazýváme tento bod $\mathbf{Q}(u, v)$ regulárním bodem zadané plochy.

Rovina daná bodem $\mathbf{Q}(u, v)$ a tečnými vektory $\mathbf{t}_u(u, v)$, $\mathbf{t}_v(u, v)$ se nazývá *tečnou rovinou plochy* v bodě $\mathbf{Q}(u, v)$. Tečnou rovinu můžeme zapsat jako

$$\nu(r, s) = \mathbf{Q}(u, v) + r\mathbf{t}_u(u, v) + s\mathbf{t}_v(u, v), \quad \text{kde } r, s \in \mathbb{R}. \quad (2.31)$$

Normálu plochy v bodě $\mathbf{Q}(u, v)$ určíme jako normálu tečné roviny. K jejímu určení využijeme vektorového součinu tečných vektorů $\mathbf{t}_u(u, v)$ a $\mathbf{t}_v(u, v)$. Jednotkový vektor normály plochy v bodě $\mathbf{Q}(t)$ získáme jako

$$\hat{\mathbf{n}}(u, v) = \frac{\mathbf{t}_u(u, v) \times \mathbf{t}_v(u, v)}{|\mathbf{t}_u(u, v) \times \mathbf{t}_v(u, v)|}. \quad (2.32)$$

2.2.2 u-křivky, v-křivky

Nastavíme-li parametr u vektorové rovnice plochy $\mathbf{Q}(u, v)$ jako konstantu $u = u_0 = \text{konst.}$, dostaneme tzv. *u-křivku* plochy

$$\mathbf{Q}(u_0, v) = (x(u_0, v), y(u_0, v), z(u_0, v)). \quad (2.33)$$

Analogicky pro parametr $v = v_0 = \text{konst.}$, kde dostaneme tzv. *v-křivku*.

$$\mathbf{Q}(u, v_0) = (x(u, v_0), y(u, v_0), z(u, v_0)). \quad (2.34)$$

Mějme plochu zadanou vektorovou rovnicí $\mathbf{Q}(u, v) = (x(u, v), y(u, v), z(u, v))$, kde $u, v \in \langle 0, 1 \rangle$. Rohové body plochy jsou $\mathbf{Q}(0, 0)$, $\mathbf{Q}(1, 0)$, $\mathbf{Q}(0, 1)$ a $\mathbf{Q}(1, 1)$. Okrajové křivky plochy jsou u, v -křivky pro hodnoty parametru $u = \{0, 1\}$ a parametru $v = \{0, 1\}$. Tyto vlastnosti budou velmi důležité v případě navazování částí ploch na sebe a dodržení určité třídy parametrické nebo geometrické spojitosti (C^n a G^n).

2.2.3 Křivost plochy

V porovnání s křivkami budeme u ploch rozlišovat více druhů křivosti. Je to dáno tím, že vybraným bodem regulární plochy může procházet libovolný počet křivek. Jednotlivé křivosti vycházejí z normálové křivosti.

Definice 2.30. V regulárním bodě plochy můžeme pro konkrétní tečnu a křivku procházející tímto bodem a dotýkající se příslušné tečny určit tzv. normálovou křivost křivky v konkrétním bodě jako

$$k_n = k_1(\mathbf{n}_k \cdot \mathbf{n}) = k_1 \cos \gamma, \quad (2.35)$$

tedy jako první křivost křivky vynásobená kosinem úhlu mezi normálou plochy a hlavní normálou příslušné křivky. Všechny zkoumané křivky procházející tímto bodem dotýkající se společné tečny budou mít normálovou křivost stejnou.

Jelikož je však v tomto bodě celá tečná rovina (a tím i nekonečně mnoho tečen), budeme rozlišovat limitní případy normálové křivosti, kterými bude minimální normálová křivost ($k_{n,\min}$) a maximální normálová křivost ($k_{n,\max}$).

Gaussovou křivostí plochy v daném bodě pak rozumíme číslo

$$k_G = k_{n,\min} \cdot k_{n,\max}. \quad (2.36)$$

Střední křivost plochy je dána jako

$$k_H = \frac{k_{n,\min} + k_{n,\max}}{2}. \quad (2.37)$$

Absolutní křivost plochy je dána jako

$$k_{\text{abs}} = |k_{n,\min}| + |k_{n,\max}|. \quad (2.38)$$

Na základě křivosti plochy v daném bodě můžeme určit její přibližný tvar.

Poznámka 2.31. Dvě plochy se dotýkají ve zkoumaném bodě tehdy, platí-li, že obě plochy mají v daném bodě společnou tečnou rovinu. Bod se pak nazývá bodem dotyku. Pro společnou křivku platí, že jednotlivé body křivky jsou body dotyku obou ploch.

2.3 Polynomiální křivky

Základním prvkem teorie křivek v počítačové grafice jsou křivky polynomiální, které jsou popsány svým polynomem

$$\mathbf{P}_n(t) = a_0 + a_1t + \dots + a_nt^n. \quad (2.39)$$

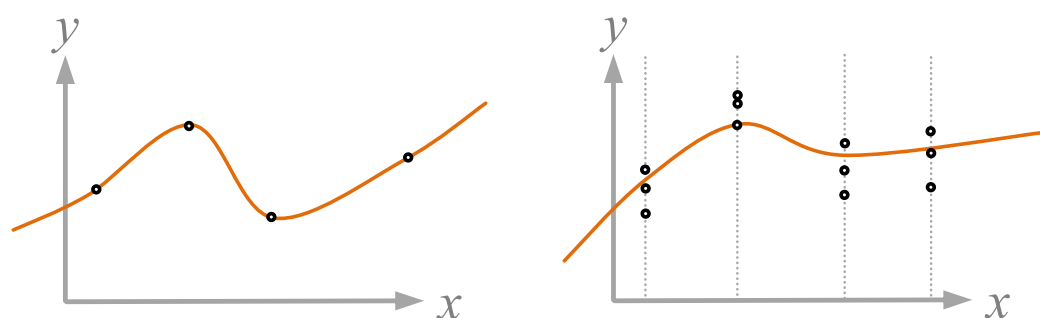
Stupeň polynomu je určen nejvyšší mocninou n . Nejčastěji používaným stupněm jsou křivky třetího stupně - kubické polynomy

$$\mathbf{P}_3(t) = a_0 + a_1t + a_2t^2 + a_3t^3, \quad (2.40)$$

kteří poskytují dostatečné možnosti při volbě tvarů a nemají nevýhody křivek vyššího stupně, jakými jsou nežádoucí vlnění a oscilace.

Obecně lze křivky rozdělit podle toho, zda procházejí zadanými body nebo prochází pouze v blízkosti těchto řídicích bodů. Podle toho je dělíme na:

- 1) interpolační křivky - prochází zadanými body přesně (obr. 2.4 vlevo);
- 2) aproximační křivky - prochází pouze v blízkosti zadaných bodů (obr. 2.4 vpravo).



Obr. 2.4: Ukázka a) interpolační křivky, b) aproximační křivky

I když bychom mohli říci, že obvykle chceme, aby výsledná křivka byla interpolační, existují případy, kdy je již ze zadání patrné, že interpolační křivka neexistuje a nezbývá než použít křivku aproximační. Blíže se dočteme v části o aproximační metodě nejmenších čtverců.

2.3.1 Interpolace algebraickým polynomem

Mezi základní interpolační křivky patří interpolace algebraickým polynomem. Cílem je nalézt analytický předpis křivky, která prochází všemi zadanými opěrnými body $[x_0, y_0], [x_1, y_1], \dots, [x_n, y_n]$.

Mějme zadáno $n + 1$ obecných bodů $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$, pro které platí $x_0 < x_1 < x_2 < \dots < x_n$, pak existuje výsledný interpolační polynom $\mathbf{P}_n(x)$ stupně n . To znamená, že platí

$$\mathbf{P}_n(x_i) = y_i, \text{ kde } i = 0, \dots, n. \quad (2.41)$$



Příklad 2.32. Určete výsledný interpolační polynom pro body $[0, 0], [1, 2], [3, 3]$.

Řešení. Námi hledaný polynom bude stupně dva, ve tvaru $\mathbf{P}_2(x) = a_0x^2 + a_1x + a_2$. Po dosazení dostaneme soustavu rovnic

$$\begin{aligned} a_2 &= 0 \\ a_0 + a_1 + a_2 &= 2 \\ 9a_0 + 3a_1 + a_2 &= 3. \end{aligned}$$

Výpočet vede na soustavu lineárních rovnic, jejímž řešením je hledaný interpolační polynom ve tvaru $\mathbf{P}(x) = -0.5x^2 + 2.5x$. ▲

Stupeň křivky tedy závisí na počtu obecně zadaných bodů

- pro dva body bude výsledkem lineární interpolace (přímka);
- pro tři body bude výsledkem kvadratická interpolace (parabola);
- pro čtyři body bude výsledkem kubická interpolace (kubika);
- pro $n+1$ obecných bodů bude výsledný interpolační polynomem n -tého stupně.

Jak je to s polynomy vyšších stupňů si lze lehce ukázat na jednoduchém zadání dvou bodů. Víme, že pro dva obecné body existuje jeden polynom prvního stupně, tedy přímka, která prochází oběma body. V případě polynomu druhého stupně (paraboly), bychom existenci jediného řešení dostali přidáním třetího obecného bodu. Je tedy patrné, že polynomů vyššího stupně existuje nekonečně mnoho a byly by závislé na volbě dalšího (v případě vyšších polynomů, dalších) bodů. Všimněte si, že snahou je získat polynom co nejmenšího stupně.

Velmi často se pro výpočet interpolačního polynomu používá maticový zápis, kde pro výpočet interpolačního polynomu n -tého stupně, využijeme následující zápis

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}. \quad (2.42)$$

Předcházející případ by tedy bylo možné zapsat pomocí maticového zápisu jako

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 3 & 9 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}.$$

Vyřešením soustavy dostaneme opět tvar $\mathbf{P}(x) = -0.5x^2 + 2.5x$.

Poznámka 2.33. Ve speciálních případech může výsledný stupeň polynomu vyjít menšího stupně. Mějme tři body, splňující podmínku $x_0 < x_1 < x_2$. Víme, že hledaný interpolační polynom bude obecně druhého stupně. Pokud však tyto body budou ležet na přímce, bude výsledná interpolační křivka přímo tato přímka, tedy polynom stupně prvního (hodnost matice (2.42) by v tomto případě byla rovna dvěma).

Zobecněním algebraického polynomu můžeme odvodit tzv. Lagrangeův tvar interpolačního polynomu, který umožňuje získat funkční hodnotu v libovolném bodě dosazením do následující rovnice

$$\mathbf{P}_n(x) = \sum_{i=0}^n y_i \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}. \quad (2.43)$$

Příklad 2.34. Určete výsledný Lagrangeův interpolační polynom pro body $P_0 = [0, 0]$, $P_1 = [1, 2]$, $P_2 = [3, 3]$.



Řešení. Podle vztahu (2.43) bude pro $n = 2$ platit

$$\mathbf{P}_2(x) = y_0 \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + y_1 \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + y_2 \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}.$$

Po dosazení jednotlivých bodů dostaneme

$$2 \frac{x(x - 3)}{-2} + 3 \frac{x(x - 1)}{6} = -x^2 + 3x + \frac{1}{2}x^2 - \frac{1}{2}x = -0.5x^2 + 2.5x.$$

Výsledný polynom bude mít tvar $\mathbf{P}(x) = -0.5x^2 + 2.5x$. ▲

Mezi základní nevýhody polynomiální interpolace patří změna celého výsledného polynomu (celé křivky) při změně kteréhokoli ze zadaných opěrných bodů. Další nevýhoda spočívá v nepřírozeném vlnění již u relativně nízkého stupně polynomu. Křivka se rozkmitá (osciluje). Proto se obvykle používá polynomiální interpolace pro křivky stupně maximálně pět a nejčastěji stupně tři (kubické křivky).

2.3.2 Interpolace po částech

V případě potřeby interpolace křivky s vyšším počtem bodů se mnohem častěji používá technika interpolace křivky po jejích částech (segmentech). Ta spočívá v tom, že celou křivku nahradíme několika částmi, které interpolují jednotlivé úseky. Celou křivku tak můžeme například aproximovat vždy pro každé čtyři body polynomem třetího stupně. Problém, který zde však vzniká, spočívá v navazování křivek v jednotlivých bodech, kde končí jedna část křivky a začíná druhá. Snahou je samozřejmě zaručit alespoň určitou třídu spojitosti C^n , popřípadě G^n . Příkladem, jak bude dále popsáno, může být interpolace Hermitovskými polynomy nebo Fergusonovými kubikami.

2.3.3 Hermitův interpolační polynom

Jednou z nevýhod interpolace obecným algebraickým polynomem je, že v případě polynomů vyšších stupňů, se může křivka rozkmitat. Jednou z možností je zadat do každého uzlového bodu ještě derivaci. Tím je křivka a její tvar ovlivněn nejen řídicími body, ale i derivacemi v těchto bodech.

Pro stupeň křivky zadané $n+1$ opěrnými body $[x_0, y_0], [x_1, y_1], \dots, [x_n, y_n]$ a hodnotou derivace v jednotlivých bodech y'_0, y'_1, \dots, y'_n , pak existuje interpolační polynom $\mathbf{P}_{2n+1}(x)$ stupně $2n+1$, pro který platí

$$\mathbf{P}_{2n+1}(x_i) = y_i \text{ a } \mathbf{P}'_{2n+1}(x_i) = y'_i, \text{ kde } i = 0, \dots, n. \quad (2.44)$$

Získáme tak soustavu rovnic, kde polovinu rovnic dostaneme dosazením řídicích bodů do

$$\mathbf{P}_{2n+1}(x_i) = y_i.$$

Druhou polovinu rovnic dostaneme dosazením do

$$\mathbf{P}'_{2n+1}(x_i) = y'_i.$$



Příklad 2.35. Určete výsledný interpolační polynom pro body $P_0 = [0, 1]$, $P_1 = [1, 0]$ a $P_2 = [3, 16]$ a jejich derivace $y'_0 = -1$, $y'_1 = -4$ a $y'_2 = 92$.

Řešení. Námi hledaný polynom bude podle výše uvedeného tvrzení stupně pátého. Výsledný polynom tedy bude tvaru $\mathbf{P}_5(x) = a_0x^5 + a_1x^4 + a_2x^3 + a_3x^2 + a_4x + a_5$. Po dosazení zadaných bodů dostaneme

$$\begin{aligned} 1 &= a_00^5 + a_10^4 + a_20^3 + a_30^2 + a_40 + a_5, \\ 0 &= a_01^5 + a_11^4 + a_21^3 + a_31^2 + a_41 + a_5, \\ 16 &= a_03^5 + a_13^4 + a_23^3 + a_33^2 + a_43 + a_5. \end{aligned}$$

Zbývající tři rovnice dostaneme po derivování polynomu, kde

$$\mathbf{P}'_5(x) = 5a_0x^4 + 4a_1x^3 + 3a_2x^2 + 2a_3x + a_4.$$

Po dosazení získáme další tři rovnice

$$\begin{aligned} -1 &= 5a_00^4 + 4a_10^3 + 3a_20^2 + 2a_30 + a_4, \\ -4 &= 5a_01^4 + 4a_11^3 + 3a_21^2 + 2a_31 + a_4, \\ 92 &= 5a_03^4 + 4a_13^3 + 3a_23^2 + 2a_33 + a_4. \end{aligned}$$

Výpočet tedy vede na řešení soustavy lineárních rovnic, jejímž řešením je Hermitův interpolační polynom ve tvaru $\mathbf{P}(x) = x^5 - 3x^4 + 2x^2 - x + 1$. ▲

Jak jsme již dříve naznačili, v praxi se v některých případech nahrazuje výpočet jednoho polynomu vyššího stupně aproximací více polynomy menších stupňů. Na předcházejícím příkladě si ukážeme rozdělení křivky na dvě části. Budeme tedy mít místo jednoho výsledného polynomu pátého stupně, dva polynomy stupně třetího.

Příklad 2.36. Určete výsledné po částech interpolační Hermitovy polynomy třetího stupně pro body $P_0 = [0, 1]$, $P_1 = [1, 0]$ a $P_2 = [3, 16]$ a jejich derivace $y'_0 = -1$, $y'_1 = -4$ a $y'_2 = 92$.



Řešení. V tomto případě dojde k rozdělení křivky na dvě části. První polynom bude zadán body P_0, P_1 a derivacemi v těchto bodech. Druhá část bude tvořena body P_1, P_2 a jejich derivacemi. Oba polynomy mají dva řídicí body a budou tedy stupně třetího. Hledané polynomy budou ve tvaru

$$\mathbf{P}_3(x) = a_0x^3 + a_1x^2 + a_2x + a_3, \text{ kde } x \in \langle 0, 1 \rangle$$

a derivovaná část

$$\mathbf{P}'_3(x) = 3a_0x^2 + 2a_1x + a_2, \text{ kde } x \in \langle 0, 1 \rangle. \quad (2.45)$$

Po dosazení a vyřešení soustavy lineárních rovnic dostaneme pro příslušné intervaly dva polynomy. Polynom $\mathbf{P}_a(x) = -3x^3 + 3x^2 - x + 1$ pro $x \in \langle 0, 1 \rangle$ a polynom $\mathbf{P}_b(x) = 18x^3 - 84x^2 - 10x + 16$ pro $x \in \langle 1, 2 \rangle$. ▲

Je třeba poznamenat, že aproximováním našeho zadání dvěma polynomy třetího stupně jsme získali odlišné řešení, než v předchozím případě, kdy byl použit pouze jeden polynom pátého stupně. Pokud použijeme více polynomů nižšího stupně, bude mít výsledná křivka složená z jednotlivých částí v napojovaných bodech z hlediska spojitosti zaručené spojitě derivace prvního řádu. Mezi hlavní výhody tohoto použití patří lokální změna křivky při změně zadání, neboť se změna provede pouze u konkrétní části křivky (oblouku). V případě bodu, který je společný pro dva oblouky, se změna projeví u obou částí, které mají tento bod společný.

2.4 Metoda nejmenších čtverců

Mimo interpolační polynomiální křivky, které prochází všemi zadanými body, existují i polynomiální křivky aproximační. Mezi jednu ze základních aproximačních metod patří metoda nejmenších čtverců. Jedná se o matematicko-statistickou metodu sloužící k minimalizaci chyb. Poprvé tuto metodu použil matematik a fyzik Carl Friedrich Gauss v roce 1795.

V praxi se setkáváme s tím, že naměřené hodnoty jsou obvykle zatíženy určitou chybou měření. Cílem metody nejmenších čtverců je naměřenými hodnotami proložit vhodnou (jednoduchou) funkcí, která bude respektovat charakter naměřených veličin tak, aby chyba aproximace naměřených hodnot byla co nejmenší.

Geometrický význam metody spočívá v tom, že se snažíme nalézt takové řešení, aby součet druhých mocnin chyb byl minimální. Příkladem můžou být zadané jednotlivé body P_0, P_1, \dots, P_n a naším cílem je proložit tyto body polynomem prvního stupně (přímkou). Jak je asi patrné, proložit přímkou třemi a více body nelze (mimo speciální případ, kdy všechny body leží na přímce).

Mějme polynom k -tého stupně

$$y(x) = a_0x^k + a_1x^{k-1} + \dots + a_k, \quad (2.46)$$

kterým chceme aproximovat zadanou množinu bodů $\{[x_0, y_0], [x_1, y_1], \dots, [x_n, y_n]\}$. Pokud bychom dosadili zadané body do této rovnice, získáme soustavu $n + 1$ rovnic o $k + 1$ koeficientech a_0, a_1, \dots, a_k . V případě, kdy $k < n$ získáme předeterminovanou soustavu rovnic, která nemá jednoznačné řešení. Využijeme tedy metody nejmenších čtverců a budeme se snažit nalézt přímkou (obecně polynom), která bude procházet co nejbližše těchto bodů.

Kritérium minimalizace bude vyjádřeno reziduální funkcí

$$\delta = \sum_{i=0}^n (a_0x_i^k + a_1x_i^{k-1} + \dots + a_k - y_i)^2. \quad (2.47)$$

Parciální derivací funkce δ podle proměnných a_0 až a_k získáme soustavu rovnic, jejíž řešením jsou hledané koeficienty a_i . Po dosazení získaných koeficientů do vztahu (2.46) získáme výsledný aproximační polynom.



Příklad 2.37. Určete výsledný aproximační polynom prvního stupně (přímkou) pro zadané body $P_0 = [0, 0]$, $P_1 = [1, 2]$, $P_2 = [3, 3]$.

Řešení. Hledaný polynom bude mít tvar $y(x) = a_0x + a_1$, po dosazení jednotlivých bodů získáme soustavu rovnic

$$\begin{aligned} a_1 &= 0 \\ a_0 + a_1 &= 2 \\ 3a_0 + a_1 &= 3. \end{aligned}$$

Dosazením koeficientů do vztahu (2.47) získáme funkci

$$\delta(a_0, a_1) = a_1^2 + (a_0 + a_1 - 2)^2 + (3a_0 + a_1 - 3)^2.$$

Minimum této funkce nalezneme pomocí následujících parciálních derivací

$$\begin{aligned} \frac{\partial}{\partial a_0} \delta &= 2(a_0 + a_1 - 2) + 6(3a_0 + a_1 - 3), \\ \frac{\partial}{\partial a_1} \delta &= 2a_1 + 2(a_0 + a_1 - 2) + 2(3a_0 + a_1 - 3). \end{aligned}$$

Hledané koeficienty a_i tedy získáme řešením soustavy rovnic


$$\begin{aligned} 10a_0 + 4a_1 &= 11 \\ 4a_0 + 3a_1 &= 5. \end{aligned}$$

Výsledná lineární funkce má tvar $y(x) = \frac{13}{14}x + \frac{6}{14}$. ▲

Pro výpočet neznámých koeficientů a_i hledaného aproximačního polynomu můžeme využít maticového tvaru soustavy lineárních rovnic

$$\begin{bmatrix} \sum_{i=1}^n x_i^{2k} & \dots & \sum_{i=1}^n x_i^{k+1} & \sum_{i=1}^n x_i^k \\ \vdots & \ddots & \vdots & \vdots \\ \sum_{i=1}^n x_i^{k+1} & \dots & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i^k & \dots & \sum_{i=1}^n x_i & n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i x_i^k \\ \vdots \\ \sum_{i=1}^n y_i x_i \\ \sum_{i=1}^n y_i \end{bmatrix}, \quad (2.48)$$

kde n je počet zadaných bodů, x_i a y_i jsou souřadnice jednotlivých bodů a k je stupeň výsledného aproximačního polynomu.

Příklad 2.38. Určete aproximační křivku prvního stupně (přímku) pro body $P_0 = [0, 0]$, $P_1 = [1, 2]$ a $P_2 = [3, 3]$. 

Řešení. Vyjdeme ze vztahu (2.48), kde počet zadaných bodů $n = 3$ a stupeň výsledného polynomu $k = 1$. Po dosazení dostaneme

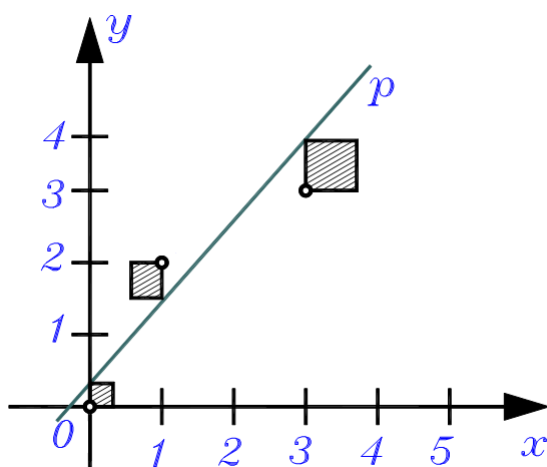
$$\begin{bmatrix} \sum_{i=1}^3 x_i^2 & \sum_{i=1}^3 x_i \\ \sum_{i=1}^3 x_i & 3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^3 y_i x_i \\ \sum_{i=1}^3 y_i \end{bmatrix}.$$

Po dosazení jednotlivých bodů $P_i = [x_i, y_i]$, pro $i = \{1, 2, 3\}$ dostaneme

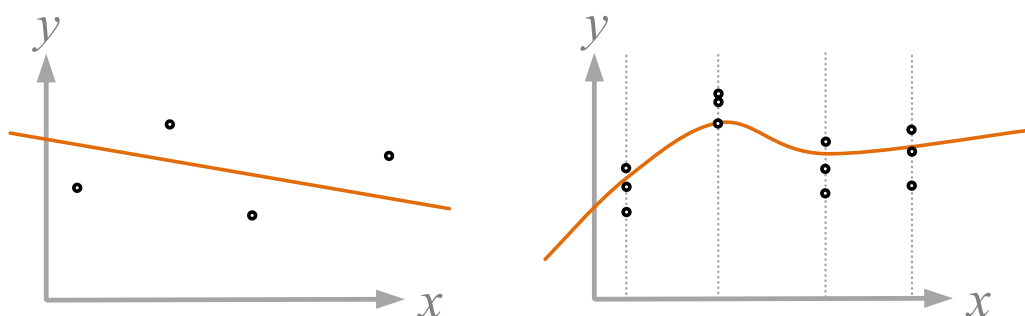
$$\begin{bmatrix} 10 & 4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 11 \\ 5 \end{bmatrix}.$$

Vyřešením této soustavy získáme výslednou aproximační přímku $f(x) = \frac{13}{14}x + \frac{6}{14}$ (výsledek je samozřejmě totožný jako v předchozím řešení). ▲

V případě polynomiálních křivek existují dva případy, kdy lze již ze zadání říct, že musíme použít polynom aproximační. V prvním případě (viz. obr. 2.6 vlevo) chceme aproximovat vstupní data nižším stupněm polynomu, než je potřebné pro interpolaci zadaných vstupních bodů. V druhém případě (viz. obr. 2.6 vpravo), kdy



Obr. 2.5: Výsledná aproximační přímka



Obr. 2.6: Použití aproximačních křivek

máme pro jedno x více funkčních hodnot $f(x)$, nelze použít interpolační polynom, což vyplývá z definice funkce.

Vzhledem k vlastnostem polynomiálních křivek je zřejmé, že v případě, kdy potřebujeme modelovat obecný tvar (profil křídla, tvar karosérie apod.), bude vhodnější využít křivky např. v parametrickém tvaru.

2.5 Fergusonova křivka

James C. Ferguson, jakožto analytik u amerického výrobce letadel Boeing, navrhl v roce 1964 použití parametrických křivek třetího stupně v leteckém průmyslu. Křivka je řízena dvěma body a tečnými vektory v těchto bodech.

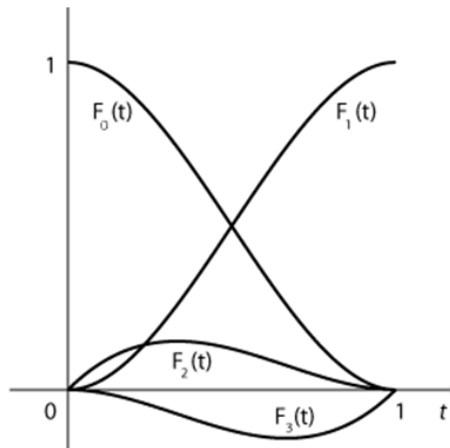
Definice 2.39. Mějme zadané dva body V_0 a V_1 a jejich polohové vektory \mathbf{V}_0 a \mathbf{V}_1 , dále mějme v těchto bodech tečné vektory \mathbf{v}_0 a \mathbf{v}_1 určující směr křivky v bodech V_0 a V_1 . *Fergusonova kubika* je daná jako

$$\mathbf{P}(t) = F_0(t)\mathbf{V}_0 + F_1(t)\mathbf{V}_1 + F_2(t)\mathbf{v}_0 + F_3(t)\mathbf{v}_1, \quad \text{kde } t \in \langle 0, 1 \rangle. \quad (2.49)$$

Jednotlivé bázové funkce F_0 , F_1 , F_2 a F_3 vypočteme jako

$$\begin{aligned} F_0(t) &= 2t^3 - 3t^2 + 1, \\ F_1(t) &= -2t^3 + 3t^2, \\ F_2(t) &= t^3 - 2t^2 + t, \\ F_3(t) &= t^3 - t^2. \end{aligned} \quad (2.50)$$

Průběh jednotlivých polynomů je na obr. 2.7.



Obr. 2.7: Průběh polynomů F_0 , F_1 , F_2 a F_3 Fergusonovy kubiky

Výsledný tvar Fergusonovy kubiky lze ovlivnit třemi způsoby:

- polohou řídicích bodů V_0 a V_1 ,
- směrem tečných vektorů \mathbf{v}_0 a \mathbf{v}_1 ,
- velikostí tečných vektorů \mathbf{v}_0 a \mathbf{v}_1 .

Poznámka 2.40. Velikost vektorů \mathbf{v}_0 a \mathbf{v}_1 významně ovlivňuje výsledný tvar křivky. Čím je délka tečných vektorů větší, tím více se křivka přimyká k příslušnému tečnému vektoru.

Výslednou rovnici Fergusonovy kubiky lze zapsat také v maticovém zápisu, který se skládá z vektoru obsahujícího mocniny parametru t , z bázové matice a z vektorů

geometrických podmínek. Nastavením vektorů podmínek (řídící body a tečné vektory) řídíme výsledný tvar křivky. Fergusonovu kubiku v maticovém tvaru zapíšeme jako

$$\mathbf{P}(t) = [t^3, t^2, t, 1] \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_0 \\ \mathbf{V}_1 \\ \mathbf{v}_0 \\ \mathbf{v}_1 \end{bmatrix}. \quad (2.51)$$

Mezi základní vlastnosti této křivky patří to, že začíná v bodě V_0 a končí v bodě V_1 . Vyjdeme-li ze vztahu (2.49) a dosadíme příslušné bázové funkce (2.50), dostaneme vektorovou rovnici

$$\mathbf{P}(t) = (2t^3 - 3t^2 + 1)\mathbf{V}_0 + (-2t^3 + 3t^2)\mathbf{V}_1 + (t^3 - 2t^2 + t)\mathbf{v}_0 + (t^3 - t^2)\mathbf{v}_1.$$

Položíme-li hodnotu parametru $t = 0$

$$\mathbf{P}(0) = (1)\mathbf{V}_0 + (0)\mathbf{V}_1 + (0)\mathbf{v}_0 + (0)\mathbf{v}_1 = \mathbf{V}_0,$$

bude výsledkem Fergusonovy kubiky počáteční bod \mathbf{V}_0 . V případě parametru $t = 1$, vyjde analogicky koncový bod \mathbf{V}_1 .

$$\mathbf{P}(1) = (0)\mathbf{V}_0 + (1)\mathbf{V}_1 + (0)\mathbf{v}_0 + (0)\mathbf{v}_1 = \mathbf{V}_1.$$

Ukázali jsme, že křivka začíná v počátečním řídicím bodě \mathbf{V}_0 a končí v koncovém řídicím bodě \mathbf{V}_1 (interpolace krajních bodů).

V případě derivace rovnice Fergusonovy křivky dostaneme

$$\mathbf{P}'(t) = (6t^2 - 6t)\mathbf{V}_0 + (-6t^2 + 6t)\mathbf{V}_1 + (3t^2 - 4t + 1)\mathbf{v}_0 + (3t^2 - 2t)\mathbf{v}_1. \quad (2.52)$$

Po dosazení parametru $t = 0$, resp. $t = 1$ do vztahu (2.52) dostaneme tečný vektor \mathbf{v}_0 v počátečním bodě, resp. tečný vektor \mathbf{v}_1 v koncovém bodě. Platí tedy

$$\begin{aligned} \mathbf{P}'(0) &= (0)\mathbf{V}_0 + (0)\mathbf{V}_1 + (1)\mathbf{v}_0 + (0)\mathbf{v}_1 = \mathbf{v}_0, \\ \mathbf{P}'(1) &= (0)\mathbf{V}_0 + (0)\mathbf{V}_1 + (0)\mathbf{v}_0 + (1)\mathbf{v}_1 = \mathbf{v}_1. \end{aligned}$$



Příklad 2.41. Vypočtěte rovnice Fergusonovy kubiky pro zadané řídicí body $V_0 = [1, 4]$, $V_1 = [6, 4]$ a tečné vektory $\mathbf{v}'_0 = (3, 1)$ a $\mathbf{v}_1 = (-3, -3)$.

Řešení. Výslednou křivku vypočteme dosazením do vztahu (2.49). Dostaneme

$$\begin{aligned} x(t) &= (2t^3 - 3t^2 + 1)1 + (-2t^3 + 3t^2)6 + (t^3 - 2t^2 + t)3 + (t^3 - t^2)(-3) \\ &= -10t^3 + 12t^2 + 3t + 1, \\ y(t) &= (2t^3 - 3t^2 + 1)4 + (-2t^3 + 3t^2)4 + (t^3 - 2t^2 + t)1 + (t^3 - t^2)(-3) \\ &= -2t^3 + t^2 + t + 4. \end{aligned}$$

V případě maticového zápisu dostaneme

$$x(t) = [t^3, t^2, t, 1] \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 6 \\ 3 \\ -3 \end{bmatrix} = -10t^3 + 12t^2 + 3t + 1,$$

$$y(t) = [t^3, t^2, t, 1] \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \\ 1 \\ -3 \end{bmatrix} = -2t^3 + t^2 + t + 4.$$

Výsledná parametrická rovnice Fergusonovy kubiky má tvar $\mathbf{P}(t) = (-10t^3 + 12t^2 + 3t + 1, -2t^3 + t^2 + t + 4)$, kde $t \in \langle 0, 1 \rangle$. ▲

Poznámka 2.42. V případě nulových tečných vektorů \mathbf{v}_0 a \mathbf{v}_1 , bude výslednou křivkou úsečka spojující body \mathbf{V}_0 a \mathbf{V}_1 .

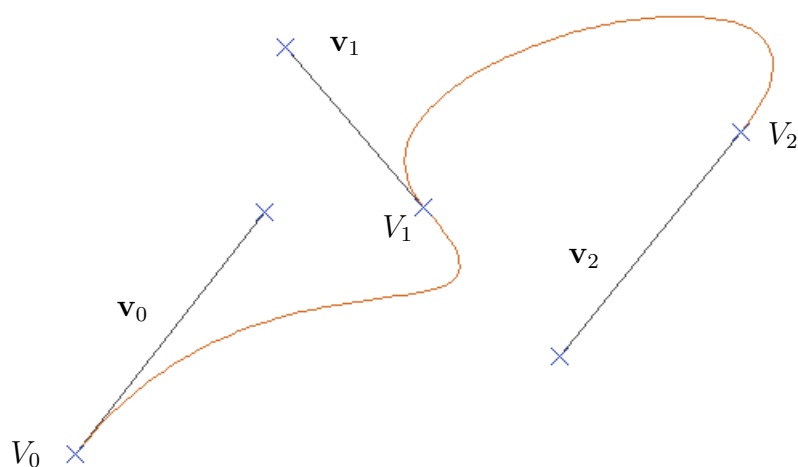
2.5.1 Navazování Fergusonových kubik

Navazováním jednotlivých Fergusonových kubik můžeme získat *kubickou spline křivku*, která má parametrickou spojitost třídy C^1 . Spojitost třídy C^0 v navazovaném bodě dvou segmentů zajistíme totožností posledního bodu prvního segmentu (druhý řídicí bod) a počátečního bodu následujícího segmentu (první řídicí bod následujícího oblouku). Spojitost třídy C^1 (křivka je hladká) v navazovaném bodě zaručíme identitou tečných vektorů sousedních oblouků. V případě, kdy požadujeme pouze geometrická spojitost třídy G^1 , stačí, aby oba vektory byly pouze souhlasně kolineární. Detailnějšímu popisu spline křivek se bude věnovat následující podkapitola.

2.6 Spline křivky

Slovo „spline“ vzniklo z označení pružného dřevěného nebo kovového laťkového křívítka, které užívali konstruktéři trupu lodí, když potřebovali při návrhu trupu lodí proložit hladkou křivku zadanými body. Označení spline poprvé použil Isaac Jacob Schoenberg v roce 1946. Intuitivně jsou spline křivky chápány jako po částech polynomiální funkce stupně k , které mají zajištěnou spojitost až do $(k - 1)$ derivací. Nejjednodušším příkladem je lineární spline křivka, která je složena z úseček spojujících řídicí body. Jedná se tedy o lomenou čáru. Nejčastěji používanou křivkou je *kubická spline křivka*, ale nejprve nadefinujeme kubickou spline funkci.

Definice 2.43. Mějme zadanou posloupnost $x_0 < x_1 < \dots < x_n$ a funkční hodnoty y_0, y_1, \dots, y_n .
Kubickou spline funkcí $f(x)$ budeme nazývat interpolační funkcí (tj. $f(x_i) = y_i$), pro kterou platí, že nad každým intervalem $\langle x_i, x_{i+1} \rangle$ je tvořena polynomem třetího stupně a dále má spojitě první a druhé derivace na intervalu (x_0, x_n) .



Obr. 2.8: Kubická spline křivka vzniklá napojením dvou Fergusonových kubik

Pro jednoznačné zadání kubické spline funkce je potřeba doplnit ještě další dvě podmínky. Máme totiž pro n kubických polynomů celkem $4n$ neznámých koeficientů.

Ze zadaných bodů, které funkce interpoluje, získáme $n + 1$ podmínek. Ze spojitosti C^0 ve vnitřních bodech získáme dalších $n - 1$ podmínek a ze spojitosti první a druhé derivace ve vnitřních $n - 1$ bodech získáme posledních $2(n - 1)$ podmínek. Celkem tedy máme $4n - 2$ podmínek pro $4n$ neznámých koeficientů. Doplněnými podmínkami bývají obvykle hodnoty prvních derivací v počátečním a koncovém bodě nebo hodnoty druhých derivací v těchto bodech.

Poznámka 2.44. Přidání podmínek $f''(x_0) = f''(x_n) = 0$ dostaneme tzv. přirozený kubický spline (podmínka volného konce).

Definice 2.45. Kubickou spline křivkou pro zadané hodnoty parametru $t_0 < t_1 < \dots < t_n$ a řídicí body $\mathbf{P}_0, \mathbf{P}_1, \dots, \mathbf{P}_n$ nazýváme křivku $\mathbf{P} = \mathbf{P}(t)$, $t \in \langle t_0, t_n \rangle$, kde je každá ze složek vektorové funkce \mathbf{P}_t kubickou spline funkcí.

Podle charakteru hodnot parametru t můžeme parametrizaci rozdělit na

- uniformní,
- neuniformní.

V případě uniformní parametrizace je krok parametru konstantní a platí

$$t_{i+1} - t_i = \text{konst.}, \text{ kde } i = 0, 1, \dots, n - 1.$$

2.7 Bézierova křivka

Bézierova křivka byla pojmenována po francouzském inženýru Pierru Bézierovi, který pracoval jako konstruktér od roku 1933 u automobilky Renault. Jedná se o parametricky zadanou křivku, která se využívá zejména díky svým vlastnostem v oblasti modelování. Mějme zadáno $n + 1$ řídicích bodů P_0, P_1, \dots, P_n , kde $n \geq 1$. Bézierova křivka je zadána jako

$$\mathbf{P}(t) = \sum_{i=0}^n \mathbf{P}_i B_i^n(t), \quad \text{kde } t \in \langle 0, 1 \rangle, \quad (2.53)$$

přičemž pro výpočet bázových funkcí $B_i^n(t)$ se využívá *Bernsteinových polynomů*. Výpočet i -tého Bernsteinova polynomu n -tého stupně je definován jako

$$B_i^n(t) = \binom{n}{i} t^i (1-t)^{n-i} = \frac{n!}{i!(n-i)!} t^i (1-t)^{n-i}, \quad \text{kde } \binom{n}{0} = 1 \text{ a } 0^0 = 1. \quad (2.54)$$

Mezi základní vlastnosti Bernsteinových polynomů patří, že součet všech Bernsteinových polynomů pro libovolné $t \in \langle 0, 1 \rangle$ je roven jedné.

Zvyšováním počtu řídicích bodů obecné Bézierovy křivky s sebou přináší některé nevýhody, jako je změna tvaru celé křivky při změně jednoho řídicího bodu. S rostoucím počtem řídicích bodů, a tedy i stupněm polynomů, se křivka stále více vzdaluje od řídicích bodů. To může být v případě modelování někdy nepříjemné. V praxi se nejčastěji využívá Bézierova křivka třetího stupně, tedy *Bézierova kubika*, která je zadána jako

$$\mathbf{P}(t) = \sum_{i=0}^3 \mathbf{P}_i B_i^3(t), \quad \text{kde } t \in \langle 0, 1 \rangle. \quad (2.55)$$

Bázové funkce Bézierovy kubiky mají tvar

$$\begin{aligned} B_0^3(t) &= (1-t)^3, \\ B_1^3(t) &= 3t(1-t)^2, \\ B_2^3(t) &= 3t^2(1-t), \\ B_3^3(t) &= t^3. \end{aligned}$$

Obdobně jako v předcházejících případech lze Bézierovu křivku zapsat v maticovém tvaru. Omezíme se v tomto případě na Bézierovu kubiku, pro kterou máme

$$\mathbf{P}(t) = [t^3, t^2, t, 1] \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{P}_0 \\ \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_3 \end{bmatrix}. \quad (2.56)$$

Příklad 2.46. Určete Bézierovu křivku zadanou čtyřmi body $P_0 = [0, 0]$, $P_1 = [10, 15]$, $P_2 = [20, 15]$ a $P_3 = [30, 0]$.



Řešení. Využijeme vztahu (2.55) pro výpočet Bézierovy kubiky, neboť je křivka zadána čtyřmi řídicími body. Máme

$$\begin{aligned}x(t) &= 0(1-t)^3 + 30t(1-t)^2 + 60(1-t) + 30t^3 = 30t, \\y(t) &= 0(1-t)^3 + 45t(1-t)^2 + 45(1-t) + 0t^3 = 45(t-t^2).\end{aligned}$$

Výsledná křivka je dána parametrickou rovnicí $\mathbf{P}(t) = (30t, 45(t-t^2))$.



Příklad 2.47. Určete Bézierovu křivku zadanou následujícími body řídicího polygonu $P_0 = [0, 0, 0]$, $P_1 = [1, 1, 0]$, $P_2 = [0, 1, 1]$ a $P_3 = [0, 0, 2]$.

Řešení. Využijeme vztahu (2.55) pro výpočet Bézierovy kubiky a získáme

$$\begin{aligned}x(t) &= 0B_0^3(t) + 1B_1^3(t) + 0B_2^3(t) + 0B_3^3(t) = 3t^3 - 6t^2 + 3t, \\y(t) &= 0B_0^3(t) + 1B_1^3(t) + 1B_2^3(t) + 0B_3^3(t) = 3t^3 - 3t^2, \\z(t) &= 0B_0^3(t) + 0B_1^3(t) + 1B_2^3(t) + 2B_3^3(t) = 5t^3 - 3t^2.\end{aligned}$$

Výsledná křivka má parametrickou rovnicí $\mathbf{P}(t) = (3t^3 - 6t^2 + 3t, 3t^3 - 3t^2, 5t^3 - 3t^2)$.



2.7.1 Základní vlastnosti Bézierovy křivky

Křivka prochází prvním a posledním řídicím bodem. Vyjdeme ze vztahu (2.53), kde lze ukázat, že pro $t = 0$ platí

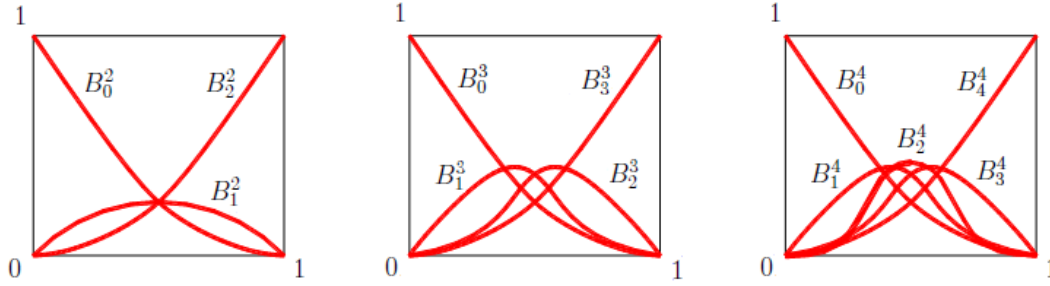
$$\mathbf{P}(0) = \sum_{i=0}^n \mathbf{P}_i B_i^n(0) = \mathbf{P}_0$$

a analogicky pro $t = 1$ platí

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

Ostatní řídicí body mezi počátečním a koncovým jsou určeny k modelování tvaru výsledné křivky a křivka jimi obecně neprochází.

Také z obr. 2.9 s grafickým znázorněním Bernsteinových polynomů je patrné, že křivka prochází prvním a koncovým bodem. Pro $t = 0$ má pouze první polynom hodnotu jedna a ostatní mají hodnotu nula. Pro $t = 1$ je analogicky pouze hodnota posledního polynomu rovna jedné, ostatní polynomy mají hodnotu nula.

Obr. 2.9: Grafické zobrazení Bernsteinových polynomů pro vybrané n ($n = 3, 4$ a 5)

Pro určení tečného vektoru v počátečním a koncovém bodě křivky vyjdeme z derivace Bernsteinových polynomů

$$\frac{d}{dt} B_i^n(t) = n(B_{i-1}^{n-1}(t) - B_i^{n-1}(t)). \quad (2.57)$$

Pro výslednou derivaci Bézierovy křivky platí

$$\frac{d}{dt} \mathbf{P}(t) = \mathbf{P}'(t) = \sum_{i=0}^{n-1} B_i^{n-1}(t) (n(\mathbf{P}_{i+1} - \mathbf{P}_i)).$$

V případě Bézierovy kubiky je tečna v počátečním (resp. koncovém) bodě dána rovnicí

$$\mathbf{P}'(0) = 3(\mathbf{P}_1 - \mathbf{P}_0), \quad (2.58)$$

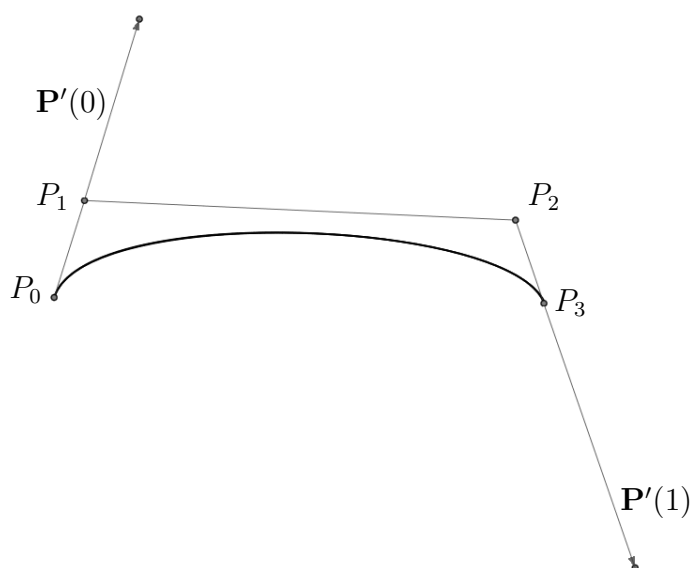
resp.

$$\mathbf{P}'(1) = 3(\mathbf{P}_3 - \mathbf{P}_2). \quad (2.59)$$

Na obr. 2.10 je zobrazena Bézierova kubika a její tečné vektory v počátečním a koncovém bodě. Ty budou velmi důležité při napojování křivek po obloucích pro dodržení spojitosti C^1 (viz. Bézierův spline).

Mezi další základní vlastnosti Bézierových křivek patří to, že křivka leží v konvexním obalu bodů, které tvoří řídicí polygon. Tvar výsledné křivky je nezávislý na orientaci řídicího polygonu a křivka tvořena řídicím polygonem P_0, P_1, \dots, P_n bude mít stejný tvar jako křivka určena řídicím polygonem P_n, P_{n-1}, \dots, P_0 .

Výsledná křivka je invariantní k afinním transformacím, což znamená, že je jedno, jestli nejdříve pomocí afinních transformací transformujeme řídicí polygon a pak vypočteme výslednou křivku nebo nejdříve vypočteme Bézierovu křivku a pak ji pomocí afinních transformací transformujeme. Výsledek bude v obou případech tožný.



Obr. 2.10: Tečné vektory v počátečním a koncovém bodě Bézierovy kubiky

2.7.2 Konverze Fergusonovy kubiky na Bézierovu kubiku

Každou Fergusonovu kubiku je možné reprezentovat jako kubickou Bézierovu křivku. Jelikož je Fergusonova kubika zadaná dvěma body V_0, V_1 a dvěma tečnými vektory $\mathbf{v}_0, \mathbf{v}_1$, stačí nastavit Bézierovu kubiku tak, aby se ztotožnily počáteční body, koncové body a tečné vektory v těchto bodech podle následujících vztahů

$$\mathbf{P}_0 = \mathbf{V}_0, \mathbf{P}_1 = \mathbf{V}_0 + \frac{1}{3}\mathbf{v}_0, \mathbf{P}_2 = \mathbf{V}_1 - \frac{1}{3}\mathbf{v}_1, \mathbf{P}_3 = \mathbf{V}_1. \quad (2.60)$$

Po dosazení do vztahu (2.55) pro výpočet Bézierovy kubiky dostaneme

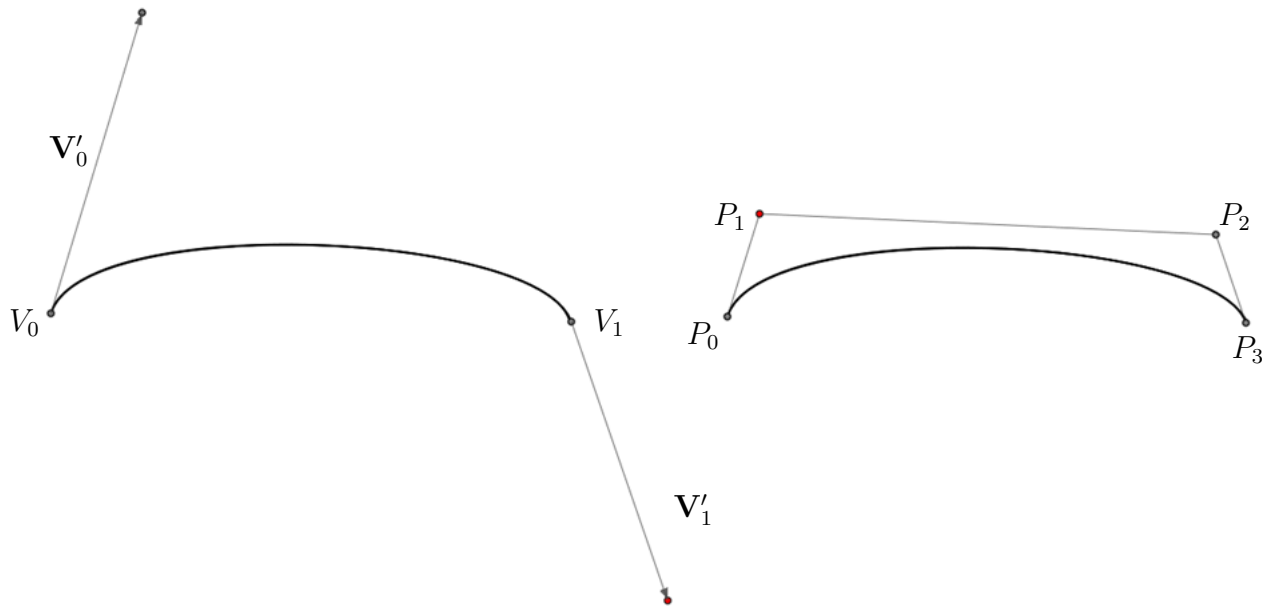
$$\begin{aligned} \mathbf{P}(t) &= \mathbf{V}_0(1-t)^3 + (\mathbf{V}_0 + \frac{1}{3}\mathbf{v}_0)3t(1-t)^2 + (\mathbf{V}_1 - \frac{1}{3}\mathbf{v}_1)3t^2(1-t) + \mathbf{V}_1t^3 \\ &= (6t^2 - 6t)\mathbf{V}_0 + (-6t^2 + 6t)\mathbf{V}_1 + (3t^2 - 4t + 1)\mathbf{v}_0 + (3t^2 - 2t)\mathbf{v}_1, \end{aligned}$$

kde $t \in \langle 0, 1 \rangle$. Výsledkem je vztah 2.49 a jedná se tedy o Fergusonovu kubiku.



Příklad 2.48. Určete Bézierovu kubiku pro nahrazení Fergusonovy kubiky $\mathbf{V}_0 = (1, 4)$, $\mathbf{V}_1 = (6, 4)$, $\mathbf{v}_0 = (3, 1)$ a $\mathbf{v}_1 = (-3, -3)$.

Řešení. Převědeme zadání pro Fergusonovu kubiku na Bézierovu kubiku podle vztahu (2.60).



Obr. 2.11: Převod Fergusonovy kubiky na Bézierovu kubiku

$$\begin{aligned}\mathbf{P}_0 &= (1, 4), \\ \mathbf{P}_1 &= (1, 4) + \frac{1}{3}(3, 1) = \left(2, \frac{13}{3}\right), \\ \mathbf{P}_2 &= (6, 4) - \frac{1}{3}(-3, -3) = (7, 5), \\ \mathbf{P}_3 &= (6, 4).\end{aligned}$$

Pro takto získané body \mathbf{P}_0 , \mathbf{P}_1 , \mathbf{P}_2 a \mathbf{P}_3 vypočteme Bézierovu kubiku

$$\mathbf{P}(t) = [t^3, t^2, t, 1] \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} (1, 4) \\ (2, \frac{13}{3}) \\ (7, 5) \\ (6, 4) \end{bmatrix}.$$

Výsledná parametrická rovnice Bézierovy kubiky má tvar $\mathbf{P}(t) = (-10t^3 + 126t^2 + 3t + 1, -2t^3 + t^2 + t + 4)$.

▲

2.7.3 Bézierův spline

Beziérův spline je spline křivkou, u které se pro výpočet jejich jednotlivých částí používá Bézierova křivka. Podle použitého stupně polynomu rozlišujeme lineární

Bézierovu spline křivku, kvadratickou Bézierovu spline křivku, kubickou Bézierovu spline křivku apod.

V případě kubické Bézierovy spline křivky je potřeba zaručit spojitost třídy C^2 , to znamená, že pro každé dva sousední segmenty, definované řídicími body P_0, P_1, P_2, P_3 a Q_0, Q_1, Q_2, Q_3 , musíme zaručit spojitost v napojovaném bodě až do druhých derivací.

Spojitost třídy C^0 je zajištěna napojením koncového bodu prvního segmentu s prvním bodem následujícího segmentu. Musí platit

$$\mathbf{P}_3 = \mathbf{Q}_0. \quad (2.61)$$

Spojitost třídy C^1 zaručíme identitou tečných vektorů (jako v případě Fergusonovy křivky), a to tak, že společný bod pro dva oblouky se bude nacházet v polovině úsečky tvořené předcházejícím a následujícím bodem. Musí tedy platit

$$\mathbf{P}_3 - \mathbf{P}_2 = \mathbf{Q}_1 - \mathbf{Q}_0. \quad (2.62)$$

Pro spojitost druhých derivací (třída C^2) vyjdeme z obecného vztahu pro druhou derivaci

$$\mathbf{P}''(t) = (n-1)n \sum_{i=0}^{n-2} B_i^{n-2}(t)(\mathbf{P}_{i+2} - 2\mathbf{P}_{i+1} + \mathbf{P}_i), \quad (2.63)$$

kde pro rovnost druhých derivací v místě styku oblouků $\mathbf{P}''(1) = \mathbf{Q}''(0)$ musí v případě kubických křivek platit

$$\begin{aligned} \mathbf{P}_3 - 2\mathbf{P}_2 + \mathbf{P}_1 &= \mathbf{Q}_2 - 2\mathbf{Q}_1 + \mathbf{Q}_0, \\ (\mathbf{P}_3 - \mathbf{P}_2) - (\mathbf{P}_2 - \mathbf{P}_1) &= (\mathbf{Q}_2 - \mathbf{Q}_1) - (\mathbf{Q}_1 - \mathbf{Q}_0), \\ (\mathbf{P}_3 - \mathbf{P}_2) + (\mathbf{Q}_1 - \mathbf{Q}_0) &= (\mathbf{P}_2 - \mathbf{P}_1) + (\mathbf{Q}_2 - \mathbf{Q}_1). \end{aligned}$$

Protože pro spojitou první derivaci musí platit $(\mathbf{P}_3 - \mathbf{P}_2) = (\mathbf{Q}_1 - \mathbf{Q}_0)$, je pro spojitost třídy C^2 požadována podmínka

$$2(\mathbf{P}_3 - \mathbf{P}_2) = (\mathbf{P}_2 - \mathbf{P}_1) + (\mathbf{Q}_2 - \mathbf{Q}_1). \quad (2.64)$$

Pokud má křivka v navazovaných bodech spojitost do druhé derivace (křivka je ve třídě C^2) můžeme křivku vzniklou navazováním jednotlivých Bézierových kubik nazvat Bézierova kubická spline křivka. Podmínka pro spojitost třídy C^2 již ale velmi omezuje možnosti změny polohy řídicích bodů, neboť délka vektoru $(\mathbf{Q}_1 - \mathbf{P}_2)$ musí být stejná s délkou vektoru $(\mathbf{P}_2 - \mathbf{P}_1) + (\mathbf{Q}_2 - \mathbf{Q}_1)$. Proto se v praxi využívá v případě Bézierových křivek obvykle maximálně spojitost prvních derivací, tedy C^1 .

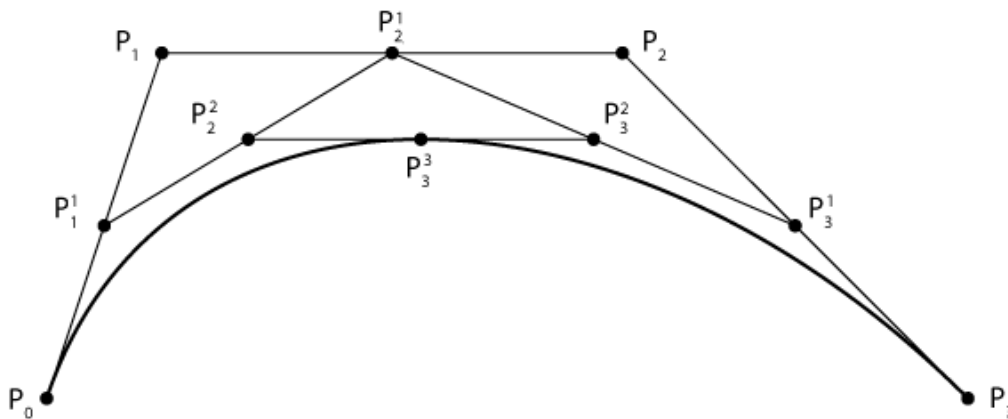
Poznámka 2.49. V případě potřeby křivek se spojitostí C^2 se nabízí Coonsova B-spline křivka, která bude popsána následně.

2.7.4 Geometrická konstrukce bodů Bézierových křivek

Paul de Casteljaou (narozen 1930 v Besançon, Francie), matematik a fyzik v Citroënu, využíval pro Bézierovy křivky geometrickou konstrukci nezávisle na Bézierovi. Tento algoritmus dává stejný výsledek jako dříve popsany analytický postup. Pokud je $\mathbf{P}_n(t)$ bod na křivce n -tého stupně pro daný parametr t , pak platí rekurentní vztah

$$\mathbf{P}_i^j(t) = \begin{cases} (1-t)\mathbf{P}_{i-1}^{j-1}(t) + t\mathbf{P}_i^{j-1}(t) & j > 0 \\ \mathbf{P}_i & \text{jinak} \end{cases} \quad (2.65)$$

Pro geometrickou konstrukci tedy platí, že zvolíme poměr dělení (odpovídá parametru t) a následně rozdělíme jednotlivé úsečky řídicího polygonu daným poměrem. Opětovně dělíme vzniklé úsečky až do chvíle, kdy získáme hledaný bod Bézierovy křivky $\mathbf{P}_n(t)$, což je bod na křivce n -tého stupně pro parametr t .



Obr. 2.12: Bod na Bézierově kubice pro parametr $t = 0.5$

Výhodu tohoto algoritmu oceníme při vykreslování v diskretním rastru (obrazovka monitoru). Jak je vidět na obrázku 2.12. Bod \mathbf{P}_3^3 je nejen bod Bézierovy křivky pro parametr $t = 0.5$, ale také rozděluje danou křivku na dvě části. Vygenerované body tvoří řídicí body nových polygonů těchto dvou částí (body $\mathbf{P}_0, \mathbf{P}_1^1, \mathbf{P}_2^2$ a \mathbf{P}_3^3 tvoří řídicí polygon levé části, body $\mathbf{P}_3^3, \mathbf{P}_3^2, \mathbf{P}_3^1$ a \mathbf{P}_3 tvoří řídicí polygon pravé části). Následně můžeme rekurzivně aplikovat algoritmus na obě části, čímž získáme z každé další dvě části. V případě potřeby můžeme rekurzivní algoritmus ukončit s ohledem na vlastnosti křivky. V místech většího zakřivení můžeme generovat více bodů křivky a v místech menšího zakřivení zase méně. Pokud bychom použili pro výpočet přímo analytický postup, nastavujeme obvykle krok a následně počítáme jednotlivé body křivky se stejným krokem bez závislosti na aktuálním zakřivení.

2.7.5 Racionální Bézierova křivka

Zobecněním Bézierových křivek jsou *Racionální Bézierovy křivky*. Motivací je odstranit nevýhodu spočívající v nutnosti změny polohy některého řídicího bodu, kdykoli chceme změnit tvar výsledné křivky. V případě prostorové křivky je situace ještě komplikovanější obtížnějším nastavováním 3D souřadnic (např. pomocí kurzoru myši) na 2D monitoru. Hlavní idea spočívá v tom, že každému bodu řídicího polygonu přiřadíme reálné číslo (váhu), které bude ovlivňovat výsledný tvar křivky. Můžeme tedy měnit výsledný tvar křivky bez nutnosti změny polohy bodů řídicího polynomu.

Mějme zadanou posloupnost bodů $\mathbf{P}_0, \mathbf{P}_1, \dots, \mathbf{P}_n$, kde $n \geq 1$ a posloupnost reálných čísel (váhy v jednotlivých bodech) w_0, w_1, \dots, w_n . Bézierova racionální křivka řádu n je dána vztahem

$$\mathbf{P}(t) = \frac{\sum_{i=0}^n w_i \mathbf{P}_i B_i^n(t)}{\sum_{i=0}^n w_i B_i^n(t)} = \sum_{i=0}^n \mathbf{P}_i R_i^n(t), \quad (2.66)$$

kde $i = 0, 1, \dots, n \in \mathbb{N}$, $w_i \geq 0$, $w_i \in \mathbb{R}$. Polynomy R_i^n jsou *racionální Bernsteinovy polynomy*.

$$R_i^n(t) = \frac{w_i B_i^n(t)}{\sum_{j=0}^n w_j B_j^n(t)}. \quad (2.67)$$

Bézierovy křivky jsou speciálním případem racionálních Bézierových křivek, kdy koeficienty w_i budou pro všechna i rovny jedné. Platí tedy

$$w_i = 1, \quad i = 0, 1, \dots, n. \quad (2.68)$$

Mezi základní vlastnosti racionální Bézierovy křivky patří, obdobně jako v případě Bézierových křivek, že začíná a končí v počátečním a koncovém bodě. Tvar výsledné křivky můžeme ovlivnit nejen změnou polohy bodů řídicího polygonu, ale také pomocí změny váhy nastavené pro jednotlivé řídicí body.

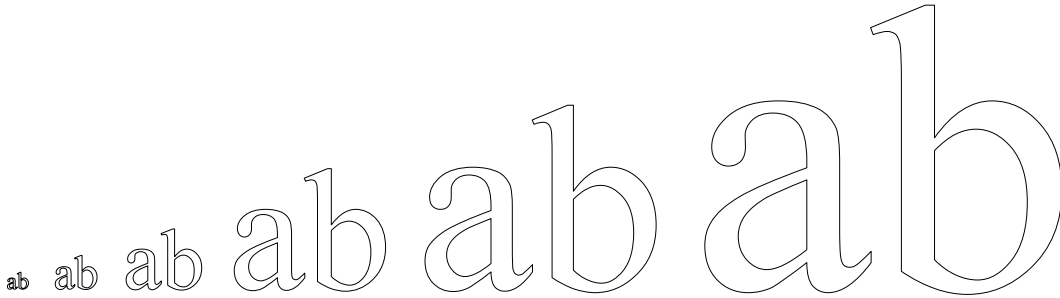
Poznámka 2.50. Blíže se racionálním tvarům budeme věnovat u NURBS křivek.

2.7.6 Použití Bézierových křivek

Bézierovy křivky se například používají pro popis TrueType fontů (kvadratické křivky) a postscriptových fontů (kubické křivky). Výhoda vektorového popisu fontu pomocí křivek je zřejmá a vychází z výhod vektorové grafiky.

Každý znak (písmeno) je popsán množinou křivek a každá z těchto křivek je popsána svými řídicími body. V případě nastavení velikosti nastavíme měřítko pro

řídící body a podle nich vypočteme příslušné křivky. Znaky se tak přizpůsobí konkrétní velikosti. Proto můžeme v různých editorech libovolně měnit velikost písma a to klidně i na desetinná místa. Ukázka možnosti spojitě změny fontu je na obrázku 2.13.



Obr. 2.13: Změna velikosti fontu

2.8 Bézierova plocha

Rozšířením Bézierovy křivky získáme Bézierovu plochu. Ta patří mezi parametrické plochy určené sítí bodů. Vznik byl, stejně jako u Bézierových křivek, motivován potřebou vhodných ploch pro modelování složitějších objektů (automobily, letadla, apod.). Prvotní využití také našly, obdobně jako křivky, u Renaultu (Bézier) a Citroenu (de Casteljaou).

Mějme zadanou obecně obdélníkovou síť řídicích bodů, zapsanou v matici \mathbf{M} o velikosti $m \times n$

$$\mathbf{M} = \begin{bmatrix} \mathbf{P}_{0,0} & \mathbf{P}_{0,1} & \cdots & \mathbf{P}_{0,m} \\ \mathbf{P}_{1,0} & \mathbf{P}_{1,1} & \cdots & \mathbf{P}_{1,m} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{n,0} & \mathbf{P}_{n,1} & \cdots & \mathbf{P}_{n,m} \end{bmatrix}. \quad (2.69)$$

Výslednou Bézierovu plochu vypočteme jako

$$\mathbf{P}(u, v) = \sum_{j=0}^m \sum_{i=0}^n \mathbf{P}_{i,j} B_i^n(u) B_j^m(v), \text{ kde } u, v \in \langle 0, 1 \rangle. \quad (2.70)$$

Obdobně jako v případě Bézierovy křivky, jsou B_i^n Bernsteinovy polynomy. Bézierovu plochu lze, stejně jako Bézierovu křivku, zapsat v maticovém tvaru

$$\mathbf{P}(u, v) = [B_0^n(u), B_1^n(u), \dots, B_n^n(u)] \begin{bmatrix} \mathbf{P}_{0,0} & \mathbf{P}_{0,1} & \dots & \mathbf{P}_{0,m} \\ \mathbf{P}_{1,0} & \mathbf{P}_{1,1} & \dots & \mathbf{P}_{1,m} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{n,0} & \mathbf{P}_{n,1} & \dots & \mathbf{P}_{n,m} \end{bmatrix} \begin{bmatrix} B_0^m(v) \\ B_1^m(v) \\ \vdots \\ B_m^m(v) \end{bmatrix}. \quad (2.71)$$

Mezi základní vlastností patří

- Bézierova plocha prochází čtyřmi rohovými body sítě. Tyto body jsou součástí hledané Bézierovy plochy. Ostatními řídicími body sítě plocha obecně neprochází.
- Okrajové řídicí body jednotlivých stran zadané sítě bodů tvoří také zadání pro Bézierovy křivky. Plocha je tedy ohraničena čtyřmi Bézierovými křivkami.
- Všechny body výsledné Bézierovy plochy leží v konvexním obalu řídicích bodů.



Příklad 2.51. Určete Bézierovu plochu zadanou čtvercovou sítí bodů 3x3. Jednotlivé polohové vektory bodů jsou zadány jako $\mathbf{P}_{0,0} = (0, 0, 0)$, $\mathbf{P}_{0,1} = (1, -1, 0)$, $\mathbf{P}_{0,2} = (3, 0, 1)$, $\mathbf{P}_{1,0} = (0, 1, 0)$, $\mathbf{P}_{1,1} = (1, 0, 2)$, $\mathbf{P}_{1,2} = (2, 2, 4)$, $\mathbf{P}_{2,0} = (-1, 2, -1)$, $\mathbf{P}_{2,1} = (1, 1, 3)$ a $\mathbf{P}_{2,2} = (2, 3, 5)$.

Řešení. Využijeme vztahu (2.71) pro výpočet Bézierovy plochy. Dostaneme

$$\mathbf{P}(u, v) = [B_0^2(u), B_1^2(u), B_2^2(u)] \begin{bmatrix} \mathbf{P}_{0,0} & \mathbf{P}_{0,1} & \mathbf{P}_{0,2} \\ \mathbf{P}_{1,0} & \mathbf{P}_{1,1} & \mathbf{P}_{1,2} \\ \mathbf{P}_{2,0} & \mathbf{P}_{2,1} & \mathbf{P}_{2,2} \end{bmatrix} \begin{bmatrix} B_0^2(v) \\ B_1^2(v) \\ B_2^2(v) \end{bmatrix}.$$

Pro Bernsteinovy polynomy druhého stupně podle vztahu (2.54) platí

$$\begin{aligned} B_0^2(t) &= (1-t)^2, \\ B_1^2(t) &= 2t(1-t), \\ B_2^2(t) &= t^2. \end{aligned}$$

Po dosazení a vynásobení dostaneme hledané parametrické rovnice ve tvaru

$$\begin{aligned} x(u, v) &= -v^2(1-u)^2 + 2u(1-u)(1-v)^2 + 4uv(1-u)(1-v) \\ &\quad + 2u(1-u)v^2 + 3u^2(1-v)^2 + 12u^2v(1-v) + 6u^2v^2, \\ y(u, v) &= 2v(1-v)(1-u)^2 + 2v^2(1-u)^2 - 2u(1-u)(1-v) \\ &\quad + 2u(1-u)v^2 + 4vu^2(1-v) + 3v^2u^2, \\ z(u, v) &= -v^2(1-u)^2 + 2u(1-u)(1-v) + 3uv^2(1-u) \\ &\quad + u^2(1-v)^2 + 8vu^2(1-v) + 5v^2u^2, \end{aligned}$$

kde $u, v \in \langle 0, 1 \rangle$.

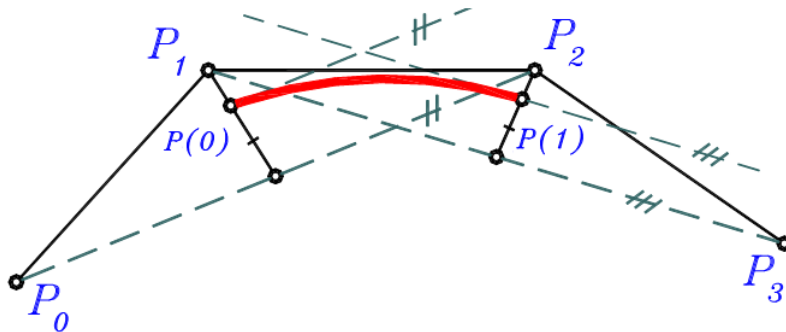


Poznámka 2.52. Z vlastností Bézierovy plochy víme, že pro hodnoty parametru $u, v = \{0, 1\}$, budou výsledkem rohové řídicí body, což může sloužit pro kontrolu výpočetní správnosti našeho řešení. Ověřte na příkladě.

2.9 Coonsova B-spline křivka

Tato křivka patří mezi obecné B-spline křivky, se kterými se seznámíte v další části. V tomto případě se zaměříme na Coonsovou kubiku a následně Coonsovou B-spline křivku. Coonsova kubika je parametrickou křivkou, pojmenovanou po Steven Anson Coonsovi, která je určena čtyřmi řídicími body $\mathbf{P}_0, \mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$, a definována vztahem

$$\mathbf{P}(t) = \frac{1}{6} \sum_{i=0}^3 \mathbf{P}_i C_i(t), \quad \text{kde } t \in \langle 0, 1 \rangle, \quad (2.72)$$



Obr. 2.14: Coonsova kubika zadána čtyřmi řídicími body $\mathbf{P}_0, \mathbf{P}_1, \mathbf{P}_2$ a \mathbf{P}_3

přičemž pro výpočet bázových funkcí platí

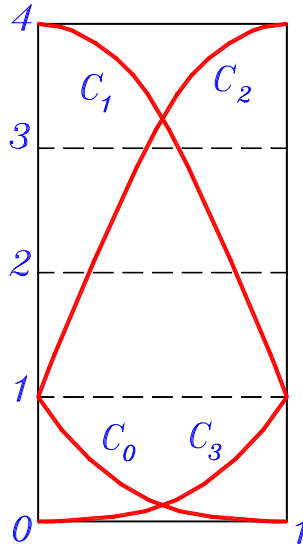
$$\begin{aligned} C_0(t) &= -t^3 + 3t^2 - 3t + 1, \\ C_1(t) &= 3t^3 - 6t^2 + 4, \\ C_2(t) &= -3t^3 + 3t^2 + 3t + 1, \\ C_3(t) &= t^3. \end{aligned}$$

Na obrázku 2.15 jsou zobrazeny jednotlivé průběhy bázových funkcí. Opět můžeme křivku zapsat v maticovém vyjádření, a to jako

$$\mathbf{P}(t) = [t^3, t^2, t, 1] \frac{1}{6} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{P}_0 \\ \mathbf{P}_1 \\ \mathbf{P}_2 \\ \mathbf{P}_3 \end{bmatrix}. \quad (2.73)$$

Příklad 2.53. Najděte Coonsovou kubickou B-spline křivku určenou řídicím polygonem $\mathbf{P}_0 = (0, 0)$, $\mathbf{P}_1 = (2, 5)$, $\mathbf{P}_2 = (4, 4)$ a $\mathbf{P}_3 = (6, 3)$.





Obr. 2.15: Průběh bázových funkcí C_0 , C_1 , C_2 a C_3 Coonsovy kubiky

Řešení. Vyjdeme z rovnice (2.72)

$$\begin{aligned}
 x(t) &= \frac{1}{6} (0 + 2(3t^3 - 6t^2 + 4) + 4(-3t^3 + 3t^2 + 3t + 1) + 6(t^3)) \\
 &= \frac{1}{6} (6t^3 - 12t^2 + 8 - 12t^3 + 12t^2 + 12t + 4 + 6t^3) \\
 &= \frac{1}{6} (12t + 12) = 2t + 2, \\
 y(t) &= \frac{1}{6} ((-t^3 + 3t^2 - 3t + 1) + 5(3t^3 - 6t^2 + 4) + 4(-3t^3 + 3t^2 + 3t + 1) + 0) \\
 &= \frac{1}{6} (15t^3 - 30t^2 + 20 - 12t^3 + 12t^2 + 12t + 4) \\
 &= \frac{1}{2} t^3 - 3t^2 + 2t + 4.
 \end{aligned}$$

Výsledná parametrická rovnice má tvar $\mathbf{P}(t) = (2t + 2, \frac{1}{2}t^3 - 3t^2 + 2t + 4)$. ▲

Pro určení počátku křivky vypočteme bázové funkce pro parametr $t = 0$. Dostaneme $C_0(0) = 1$, $C_1(0) = 4$, $C_2(0) = 1$ a $C_3(0) = 0$. Pro počáteční bod křivky tedy musí platit

$$\mathbf{P}(0) = \frac{1}{6}(\mathbf{P}_0 + 4\mathbf{P}_1 + \mathbf{P}_2).$$

Počáteční bod leží v tzv. antitěžišti trojúhelníka zadaného prvními třemi body \mathbf{P}_0 , \mathbf{P}_1 a \mathbf{P}_2 , to znamená, že leží na těžnici trojúhelníka procházející vrcholem \mathbf{P}_1 a vzdálenost vrcholu \mathbf{P}_1 od počátečního bodu $\mathbf{P}(0)$ je rovna jedné třetině délky těžnice.

Obdobně lze ukázat, že pro koncový bod $\mathbf{P}(1)$ platí, že leží v antitěžišti trojúhelníku daného posledními třemi body \mathbf{P}_1 , \mathbf{P}_2 a \mathbf{P}_3 . Máme totiž

$$\mathbf{P}(1) = \frac{1}{6}(\mathbf{P}_1 + 4\mathbf{P}_2 + \mathbf{P}_3).$$

Příklad 2.54. Ověřte počáteční a koncový bod u předcházejícího příkladu.



Řešení. Dosazením za $t = 0$ resp. $t = 1$ do

$$\mathbf{P}(t) = (2t + 2, \frac{1}{2}t^3 - 3t^2 + 2t + 4),$$

získáme počáteční bod $\mathbf{P}(0) = (2, 4)$, resp. koncový bod $\mathbf{P}(1) = (4, 4)$. ▲

Pro výpočet tečných vektorů v počátečním a koncovém bodě vyjdeme z derivací bázových funkcí

$$\begin{aligned} C'_0(t) &= -3(1-t)^2, \\ C'_1(t) &= 9t^2 - 12t, \\ C'_2(t) &= -9t^2 + 6t + 3, \\ C'_3(t) &= 3t^2. \end{aligned}$$

Tečný vektor v počátečním bodě $\mathbf{P}(0)$ určíme jako

$$\mathbf{P}'(0) = \frac{1}{6}(-3\mathbf{P}_0 + 3\mathbf{P}_2) = \frac{1}{2}(\mathbf{P}_2 - \mathbf{P}_0).$$

Směr tečny v počátečním bodě $\mathbf{P}(0)$ je tedy totožný se směrem přímky dané řídicími body \mathbf{P}_0 , \mathbf{P}_2 . Tečna v počátečním bodě je tedy rovnoběžná s přímkou danou prvním a třetím bodem.

Obdobně pro tečný vektor v koncovém bodě platí

$$\mathbf{P}'(1) = \frac{1}{6}(-3\mathbf{P}_0 + 3\mathbf{P}_2) = \frac{1}{2}(\mathbf{P}_2 - \mathbf{P}_0).$$

Tečna v koncovém bodě je tedy rovnoběžná s přímkou danou druhým a čtvrtým bodem.

Mezi další vlastnosti Coonsovy kubiky patří, že křivka leží v konvexním obalu řídicích bodů \mathbf{P}_0 , \mathbf{P}_1 , \mathbf{P}_2 a \mathbf{P}_3 .

2.9.1 Coonsův kubický B-spline

Coonsův kubický B-spline vznikne tak, že budeme na sebe napojovat Coonsovy kubiky a přitom zajistíme spojitě napojování do druhé derivace (spojitost druhého řádu). Mějme dvě Coonsovy kubiky

$$\mathbf{P}(t) = \frac{1}{6} \sum_{i=0}^3 \mathbf{P}_i C_i(t), \quad \mathbf{R}(s) = \frac{1}{6} \sum_{i=0}^3 \mathbf{R}_i C_i(s), \quad \text{kde } t, s \in \langle 0, 1 \rangle. \quad (2.74)$$

Pro jednotlivé spojitosti musí platit následující podmínky.

- Pro spojitost C^0 musí platit $\mathbf{P}(1) = \mathbf{R}(0)$, to znamená, že

$$\mathbf{P}_1 + 4\mathbf{P}_2 + \mathbf{P}_3 = \mathbf{R}_0 + 4\mathbf{R}_1 + \mathbf{R}_2.$$

- Pro spojitost prvních derivací (C^1) musí platit $\mathbf{P}'(1) = \mathbf{R}'(0)$, to znamená, že

$$\mathbf{P}_3 - \mathbf{P}_1 = \mathbf{R}_2 - \mathbf{R}_0.$$

- Pro spojitost druhých derivací (C^2) musí platit $\mathbf{P}''(1) = \mathbf{R}''(0)$, to znamená, že

$$\mathbf{P}_1 - 2\mathbf{P}_2 + \mathbf{P}_3 = \mathbf{R}_0 - 2\mathbf{R}_1 + \mathbf{R}_2.$$

Abychom tedy dodrželi všechny spojitosti až do druhého řádu, musí platit $\mathbf{P}_1 = \mathbf{R}_0$, $\mathbf{P}_2 = \mathbf{R}_1$, $\mathbf{P}_3 = \mathbf{R}_2$, jinak řečeno poslední tři řídicí body prvního segmentu musí být totožné s prvními třemi body druhého segmentu.

Poznámka 2.55. Coonsův kubický B-spline je speciálním případem B-spline křivky pro stupeň $p = 3$, uzlový vektor $\mathbf{t} = (-3, -2, -1, 0, 1, 2, 3, 4)$ a $m = 7$. Jedná se o neracionální uniformní B-spline křivku. Více se dozvíte v kapitole o obecných B-spline křivkách.

2.10 Coonsova B-spline plocha

Rozšířením Coonsovy B-spline křivky je Coonsova B-spline plocha, která má podobné chování jako křivka. Jedná se o plochu zadanou sítí $m \times n$ bodů a vychází z Coonsovy kubické B-spline křivky. Jedná se tedy o plochu řešenou po částech, kde výsledný stupeň, na rozdíl od Bézierovy plochy, neroste s počtem řídicích bodů. Plochu lze popsat vztahem

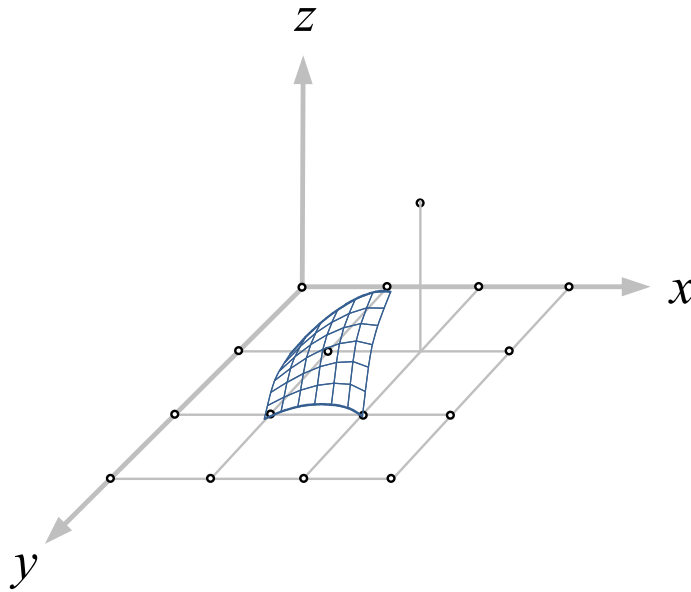
$$\mathbf{P}(u, v) = \frac{1}{36} \mathbf{c}_u \mathbf{M} \mathbf{c}_v, \quad (2.75)$$

kde u, v jsou parametry, \mathbf{M} je matice řídicích bodů a \mathbf{c}_u a \mathbf{c}_v jsou vektorové funkce parametru u a v . Funkce \mathbf{c}_u a \mathbf{c}_v lze zapsat ve tvaru

$$\mathbf{c}_t = (C_0(t), C_1(t), C_2(t), C_3(t)), \quad \text{kde } t \in \langle 0, 1 \rangle. \quad (2.76)$$

Jednotlivé složky jsou dány předpisem

$$\begin{aligned}
 C_0(t) &= -t^3 + 3t^2 - 3t + 1, \\
 C_1(t) &= 3t^3 - 6t^2 + 4, \\
 C_2(t) &= -3t^3 + 3t^2 + 3t + 1, \\
 C_3(t) &= t^3.
 \end{aligned}$$



Obr. 2.16: Ukázka Coonsovy bikubické B-spline plochy

Hranice bikubické Coonsovy B-spline plochy jsou kubické B-spline křivky. Při určování okrajové křivky pro $u = 0$ platí

$$C_0(0) = 1, \quad C_1(0) = 4, \quad C_2(0) = 1, \quad C_3(0) = 0.$$

Po dosazení do rovnice (2.75) dostaneme

$$\mathbf{P}(0, v) = \frac{1}{36} [1, 4, 1, 0] \mathbf{M} \mathbf{c}_v \quad (2.77)$$

a po úpravě získáme tvar

$$\mathbf{P}(0, v) = \frac{1}{36} \sum_{i=0}^3 (\mathbf{P}_{0,i} + 4\mathbf{P}_{1,i} + \mathbf{P}_{2,i}) \mathbf{c}_v. \quad (2.78)$$

Jedná se tedy o Coonsovu kubickou B-spline křivku, která je určena vrcholy \mathbf{Q}_i , kde

$$\mathbf{Q}_i = \frac{1}{6}(\mathbf{P}_{0,i} + 4\mathbf{P}_{1,i} + \mathbf{P}_{2,i}), \text{ kde } i = 0, 1, 2, 3. \quad (2.79)$$

Analogicky by totéž platilo pro ostatní tři okrajové křivky.

Výhoda Coonsovy B-spline plochy spočívá v napojování jednotlivých částí na sebe. Výsledná plocha se v podstatě skládá z jednotlivých částí o rozměru 4×4 bodů $(\mathbf{P}_{0,0}, \dots, \mathbf{P}_{0,3}) \times (\mathbf{P}_{0,0}, \dots, \mathbf{P}_{3,0})$. Pro další část nemusíme jako v případě Bézierovy plochy přidávat další část plochy o rozměru 4×4 , ale zvětšíme rozsah v jednom ze směrů o jedna (místo 4×4 bude mít např. 5×4). Výsledná plocha se pak bude skládat ze dvou samostatných částí, kde první bude dána řídicími body $(\mathbf{P}_{0,0}, \dots, \mathbf{P}_{0,3}) \times (\mathbf{P}_{0,0}, \dots, \mathbf{P}_{3,0})$ a druhá $(\mathbf{P}_{1,0}, \dots, \mathbf{P}_{0,4}) \times (\mathbf{P}_{1,0}, \dots, \mathbf{P}_{4,0})$.

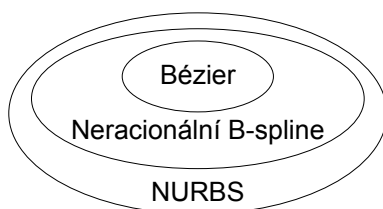
Výslednou obecnou plochu $m \times n$ bodů budeme analogicky jako Coonsovy B-spline křivky počítat a vykreslovat postupně, vždy po částech 4×4 bodů. Znamená to tedy, že například pro síť bodů 5×7 řídicích bodů budeme počítat 2×4 samostatných částí, tedy celkem 8 bikubických Coonsových ploch. Výsledná bikubická B-spline plocha bude mít parametrickou spojitost třídy C^2 a z hlediska modelování, bude změna jednoho řídicího bodu ovlivňovat pouze segmenty s ním spojené.

Poznámka 2.56. Coonsova B-spline plocha je speciálním případem obecných B-spline ploch, kterými se budeme zabývat v následující kapitole.

2.11 NURBS křivky a plochy

NURBS křivky a plochy jsou v současnosti zřejmě nejpoužívanější reprezentací obecných křivek a ploch v řadě aplikací zahrnujících oblasti designu a navrhování. Jsou podporovány prakticky všemi významnými softwarovými CAD/CAM nástroji. Tato forma matematické reprezentace křivek a ploch skýtá řadu výhod, a to jak v rámci reprezentace běžných analytických tvarů jako jsou např. kuželosečky, tak pro popis složitých obecných tvarů např. v podobě karoserie automobilů. U složitých modelů se uvádí, že použití NURBS ploch přináší až 75 % úsporu v počtu ploch nebo plátů, které je nutno vytvořit. V extrémních případech může být toto zjednodušení procesu modelování ještě mnohem výraznější. Výraznému rozšíření napomohly rovněž vynikající editační možnosti. Tvorba křivek i ploch je intuitivní a manipulace s řídicími body má pouze omezený lokální charakter. Neméně důležitá je také podpora NURBS entit v řadě nejpoužívanějších výměnných formátů, jakými jsou standardy STEP a IGES. Následující kapitoly se budou snažit podat základní přehled užitého matematického aparátu, který je ukryt za poněkud kostrbatou zkratkou NURBS, tedy Neuniformní Racionální B-Spline křivky a plochy.

Tato kapitola navazuje na předešlý výklad Bézierových křivek a ploch, které jsou speciálním případem neracionálních B-spline křivek a ploch. Neracionální B-spline jsou podmnožinou obecnějších racionálních B-spline. Vzájemný vztah těchto tříd je zachycen na obr. 2.11.



Obr. 2.17: Vzájemný vztah mezi Bézierovými, neracionálními B-spline a NURBS křivkami a plochami

2.11.1 B-spline bázové funkce

Význam bázových funkcí spočívá v reprezentaci interpolačního nebo aproximačního schématu vyjadřujícího vztah mezi průběhem křivky a řídicím polygonem. Vlastnosti tohoto schématu jsou dány volbou bázových funkcí. Např. pro Bézierovy křivky jsou využity Bernsteinovy báze. Ty mají ovšem dvě zásadní omezení. Počet řídicích bodů určuje řád křivky. To znamená, že např. kubická křivka musí být definována pomocí právě čtyř kontrolních bodů. Chceme-li tedy změnit řád křivky, musíme změnit počet kontrolních bodů. Dalším faktorem komplikujícím praktické využití je globální povaha těchto bází. Hodnota funkce je nenulová nad otevřeným intervalem $(0, 1)$ a výsledná podoba parametrické křivky nad tímto intervalem je závislá na poloze všech řídicích bodů. To znemožňuje provádět snadno lokální změny průběhu křivky nebo plochy.

Před samotným popisem NURBS křivek a ploch zavedeme v úvodní podkapitole pojem bázových spline funkcí (ang. Basis Spline). Častěji jsou označovány zkráceně jako *B-spline* a zahrnují již zmíněné Bernsteinovy báze jako svůj speciální případ. Obecně je spline funkcí nekonečně mnoho. Jako B-spline bývá v praxi také často nazývána parametrická křivka, která je zobecněním Bézierovy křivky. V této části textu budeme mít při použití výrazů B-spline na mysli výhradně systém bázových spline funkcí.

Křivky, které jsou reprezentovány pomocí pouze jednoho polynomiálního nebo racionálního segmentu, jsou limitovány v rozsahu možného použití. A to zejména z důvodů potřeby vysokého počtu řídicích bodů a s tím souvisejícím rovněž vysokým stupněm křivky. To přináší komplikace s numerickou stabilitou a efektivností výpočtu. Také komplexního tvaru je možno dosáhnout pouze za cenu vysokého stupně křivky. Komplikovaná je také interaktivní manipulace s takovými křivkami v softwarových nástrojích, jelikož jejich modifikace nemá lokální charakter. Řešení uvedených problémů spočívá v zavedení křivek, které jsou po částech polynomiální nebo po částech racionální. Parametrický prostor, nad kterým je křivka definována, je rozdělen pomocí zlomových bodů (ang. breakpoints) ve kterých je zaručen určitý stupeň návaznosti jednotlivých segmentů. Tímto rozčleněním zajistíme možnost lokální manipulace s jedním segmentem při zachování požadované návaznosti se sousedními segmenty, což činí práci s takovouto křivkou mnohem intuitivnější a snazší.

Pro možnost měnit vliv jednotlivých řídicích bodů na průběh křivky je důle-

žitý také tzv. uzlový vektor. Uzlový vektor nebo také vektor parametrizací $\mathbf{u} = (u_0, u_1, \dots, u_m)$ je definován jako neklesající posloupnost $m + 1$ reálných čísel, přičemž platí $u_0 \leq u_1 \leq \dots \leq u_m$.

Nyní můžeme definovat i -tou normalizovanou bázovou B-spline funkci stupně p označenou jako $N_i^p(u)$ pomocí rekurentního Cox-de Boorova vztahu (2.80). Existují i jiné formulace, ale z hlediska následné počítačové implementace se rekurentní vztah jeví jako nejvýhodnější.

$$\begin{aligned} N_i^0(u) &= \begin{cases} 1 & u_i \leq u < u_{i+1} \\ 0 & \text{jinak} \end{cases} \\ N_i^p(u) &= \frac{u - u_i}{u_{i+p} - u_i} N_i^{p-1}(u) + \frac{u_{i+p+1} - u}{u_{i+p+1} - u_{i+1}} N_{i+1}^{p-1}(u) \end{aligned} \quad (2.80)$$

Někdy se místo stupně (ang. degree) hovoří o řádu křivky (ang. order). Mezi řádem k a stupněm p platí vztah $k = p + 1$. Úsečka je tedy (lineární) křivka prvního stupně a druhého řádu. Nejčastěji se používají křivky druhého (kvadrík, $p = 2$ a $k = 3$) a třetího (kubík, $p = 3$ a $k = 4$) stupně. Křivky a plochy vyšších řádů však mohou vzniknout při manipulacích s původně jednoduchými entitami v modelovacích nástrojích. Obecně platí, že čím je řád vyšší, tím je do výpočtu zahrnuto více řídicích bodů a výpočet trvá déle.

Nyní uvedme některá přímá pozorování plynoucí z výše uvedených definic.

- Pro výpočet i -té bázové funkce potřebujeme znát uzlový vektor \mathbf{u} a stupeň bázové funkce p .
- Funkce $N_i^p(u)$ je nezáporná pro libovolné i, p a u .
- S výjimkou $p = 0$ nabývá funkce $N_i^p(u)$ právě jednoho maxima.
- Z uvedené definice je zřejmé, že $N_i^0(u)$ je tzv. skoková funkce, která nabývá jednotkových hodnot v intervalu $\langle u_i, u_{i+1} \rangle$. Uvedený interval se často označuje jako rozpětí uzlů (ang. knot span).
- Předchozí pozorování lze zobecnit pro funkci $N_i^p(u)$ libovolného stupně p . Tato funkce nabývá nenulových hodnot pouze pro $u \in \langle u_i, u_{i+p+1} \rangle$ (ang. local support).
- Pro libovolné rozpětí uzlů $\langle u_i, u_{i+1} \rangle$ je nejvýše $p + 1$ bázových funkcí stupně p nenulových (viz. 2.1). Konkrétně se jedná o funkce

$$N_{i-p}^p(u), N_{i-p+1}^p(u), \dots, N_i^p(u).$$

Díky tomu je B-spline křivka závislá pouze na k nejbližších kontrolních bodech a vliv kontrolních bodů na výsledný tvar křivky, potažmo plochy, má pouze lokální charakter.

- Součet všech báзовých funkcí stupně p na intervalu $\langle u_i, u_{i+1} \rangle$ je roven 1

$$\sum_{j=i-p}^i N_j^p(u) = 1, \quad u \in \langle u_i, u_{i+1} \rangle.$$

- Označíme-li počet uzlů jako $m + 1$ a počet báзовých funkcí stupně p pomocí $n + 1$, pak platí $m = n + p + 1$.
- Pozorný čtenář si jistě povšimnul, že definice uzlového vektoru nijak nebrání v možném dělení nulou v (2.80). Pro tento případ zavedeme pravidlo $0/0 \stackrel{\text{def}}{=} 0$.
- Funkce $N_i^p(u)$ jsou po částech polynomiální funkce definované nad celým oborem reálných čísel \mathbb{R} . Podstatný je však jejich průběh jen na intervalu $\langle u_0, u_m \rangle$.
- V uzlu s násobností k je báзовая funkce $N_i^p(u)$ C^{p-k} spojitá. Připomeňme, že křivka je C^n spojitá, pokud má spojitě derivate podle parametru u až do n -tého řádu.

Průběh báзовой funkce $N_i^p(u)$ je zcela určen relativním rozmístěním uzlů na pomyslné reálné ose. Posunutí těchto bodů ani změna měřítka se nijak neprojeví na průběhu báзовой funkce a ani na tvaru křivky nebo plochy, která je nad těmito báзовými funkcemi definována. Neformálně řečeno, uzlový vektor přiřazuje jednotlivým polynomům začátky a konce jejich vlivu v rámci celého intervalu vymezeného prvním a posledním uzlem.

Rozlišujeme mezi dvěma typy uzlových vektorů, a to periodickým a otevřeným (neperiodickým). Oba tyto typy se dále dělí na uniformní a neuniformní.

1. Uniformní uzlový vektor (ang. uniform) splňuje podmínku rovnoměrného rozložení uzlů. Platí tedy, že vzdálenosti mezi jednotlivými uzly jsou konstantní

$$u_{i+1} - u_i = \text{konst.}, \quad 0 \leq i < m.$$

Příklad: (0.0, 0.25, 0.5, 0.75, 1.0)

2. Otevřený uniformní vektor (ang. open uniform) je uniformní uzlový vektor s výjimkou k totožných počátečních a koncových uzlů. Připomeňme, že k značí řád báзовой funkce.

$$\begin{aligned} u_0 &= u_i, & 0 < i < k \\ u_{i+1} - u_i &= \text{konst.}, & k - 1 \leq i \leq m - k \\ u_{m-k+1} &= u_i, & m - k + 2 \leq i \leq m \end{aligned}$$

Příklad: (1, 1, 1, 2, 3, 4, 5, 6, 6, 6), (0, 0, 0, 1, 1, 1) pro $k = 3$

3. Neperiodický uzlový vektor (ang. non-periodic)

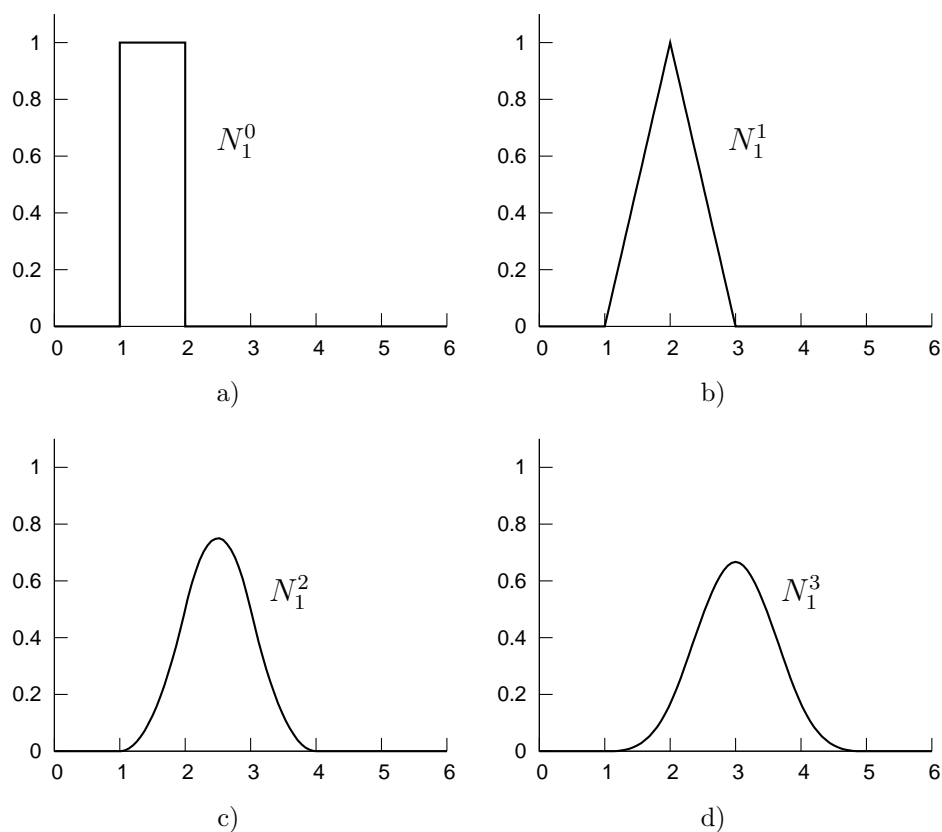
$$\underbrace{(a, \dots, a)}_{p+1}, u_{p+1}, \dots, u_{m-p-1}, \underbrace{(b, \dots, b)}_{p+1}. \quad (2.81)$$

Příklad: (1, 1, 2, 3, 4, 4, 6, 6)

4. Všechny ostatní uzlové vektory se označují jako periodické neuniformní (ang. non-uniform).

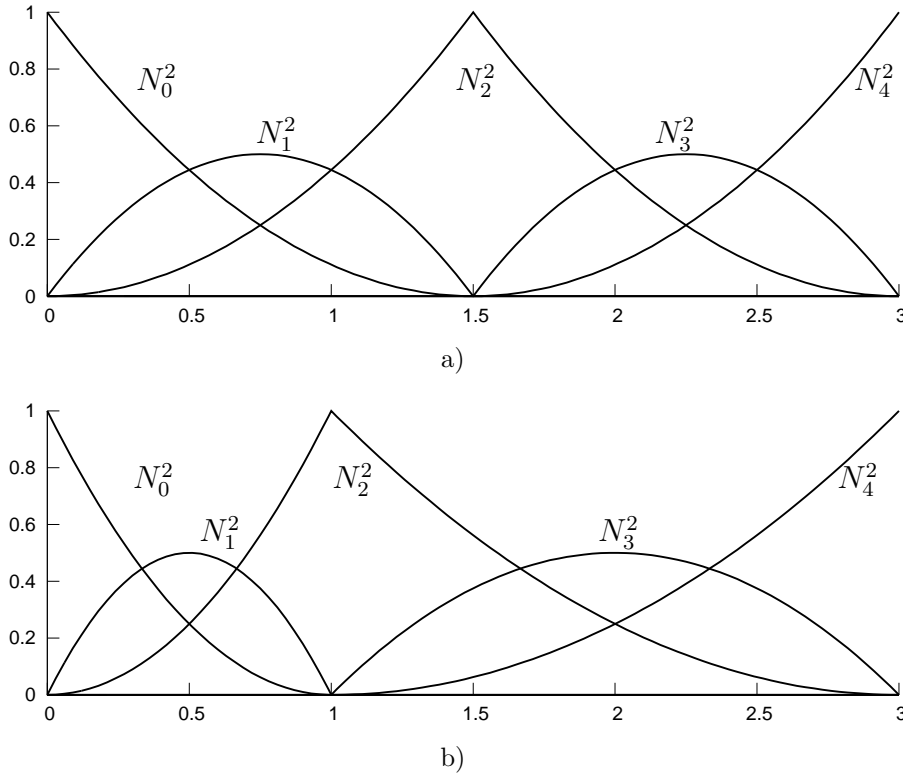
Příklad: (1, 1, 1, 2, 2, 3, 5, 6, 6, 7, 7)

V dalším výkladu se z praktických důvodů omezíme pouze na neperiodické (obecně neuniformní) uzlové vektory. Odůvodnění této volby ponecháme na čtenáři.



Obr. 2.18: Ukázka čtyř bázových funkcí pro uniformní uzlový vektor $(0, 1, \dots, 6)$.

a) N_i^0 je skoková funkce nabývající nenulové hodnoty pouze nad intervalem (u_i, u_{i+1}) . b) N_i^1 je po částech lineární funkce ($p = 1$) rozpínající se přes dvě uzlová rozpětí. c) N_i^2 je po částech kvadratická funkce nad třemi uzlovými rozpětími. d) N_i^3 je po částech kubická funkce nabývající nenulových hodnot nad čtyřmi uzlovými rozpětími. A takto bychom mohli pokračovat dále pro vyšší stupně bázových funkcí



Obr. 2.19: Bázové funkce neuniformního uzlového vektoru a) $(0, 0, 0, 1.5, 1.5, 3, 3, 3)$
 a) b) $(0, 0, 0, 1, 1, 3, 3, 3)$. Pro oba případy platí $m = 7$ a $p = 2$

2.11.2 Derivace B-spline bázových funkcí

Při praktickém použití neracionálních i racionálních B-spline křivek a ploch se neobejdeme bez znalosti výpočtu jejich derivací. Často nastane potřeba vypočítat směr tečny křivky nebo normálový vektor plochy v zadaném bodě. Derivace bázové funkce je dána

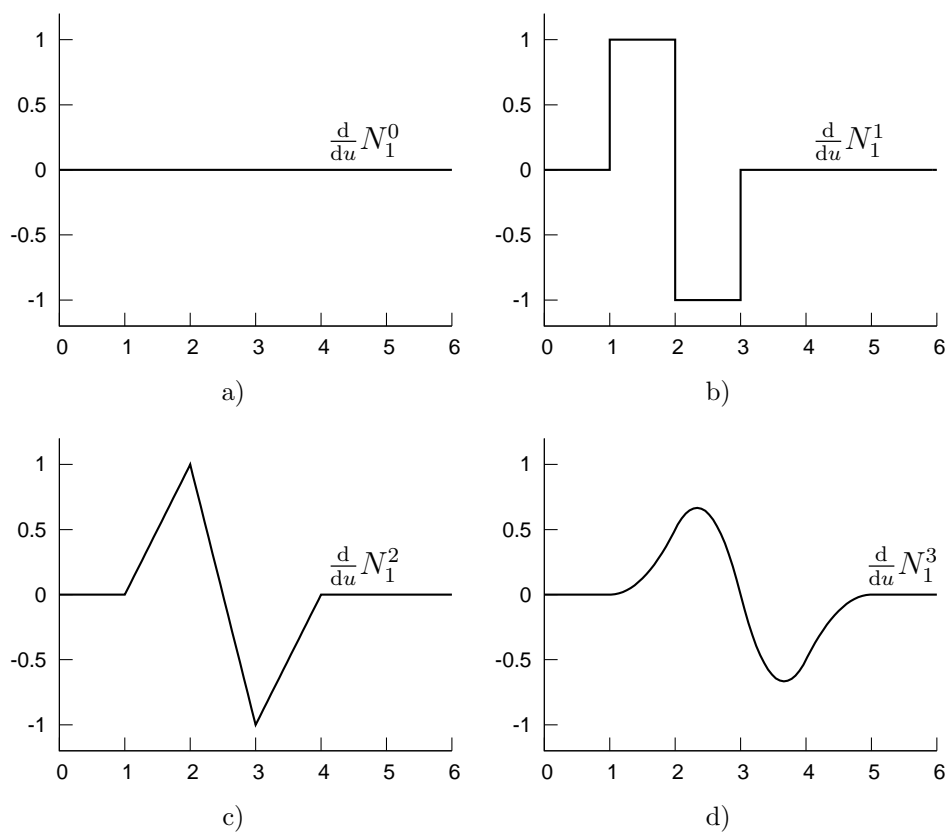
$$\frac{d}{du} N_i^p(u) = \frac{p}{u_{i+p} - u_i} N_i^{p-1}(u) - \frac{p}{u_{i+p+1} - u_{i+1}} N_{i+1}^{p-1}(u). \quad (2.82)$$

Pro výpočet derivací vyšších řádů vyjdeme z rovnice (2.82) a dalším derivováním obdržíme

$$\frac{d^k}{du^k} N_i^p(u) = \frac{p}{u_{i+p} - u_i} \frac{d^{k-1}}{du^{k-1}} N_i^{p-1}(u) - \frac{p}{u_{i+p+1} - u_{i+1}} \frac{d^{k-1}}{du^{k-1}} N_{i+1}^{p-1}(u), \quad (2.83)$$

přičemž platí, že derivace vyšších řádů než p jsou rovny nule

$$\frac{d^k}{du^k} N_i^p(u) = 0 \text{ pro } k > p. \quad (2.84)$$



Obr. 2.20: Ukázka derivací bázových funkcí z obr. 2.19

$$\begin{array}{ccccccc}
| & N_0^2 & \cdots & \boxed{N_i^2} & \cdots & N_{m-3}^2 & | \\
| & N_0^1 & \cdots & N_i^1 & N_{i+1}^1 & \cdots & N_{m-2}^1 & | \\
| & N_0^0 & \cdots & N_i^0 & N_{i+1}^0 & N_{i+2}^0 & \cdots & N_{m-1}^0 & | \\
\hline
& u_0 & \cdots & u_i & u_{i+1} & u_{i+2} & u_{i+3} & \cdots & u_m
\end{array}$$

Tab. 2.1: Bázová funkce $N_i^2(u) \neq 0$ jen pro $u \in \langle u_i, u_{i+2+1} \rangle$. Důvodem je rekurentní definice B-spline bázových funkcí, kdy bázová funkce stupně p závisí na funkcích nižšího stupně až do $p = 0$. Výsledkem je trojúhelníkový vzor zachycený v uvedené tabulce

$$\begin{array}{ccccccc}
| & N_0^2 & \cdots & \boxed{N_{i-2}^2} & \boxed{N_{i-1}^2} & \boxed{N_i^2} & \cdots & N_{m-3}^2 & | \\
| & N_0^1 & \cdots & & N_{i-1}^1 & N_i^1 & \cdots & & N_{m-2}^1 & | \\
| & N_0^0 & \cdots & & & N_i^0 & \cdots & & & N_{m-1}^0 & | \\
\hline
& u_0 & \cdots & & & u_i & u_{i+1} & \cdots & & u_m
\end{array}$$

Tab. 2.2: Pro libovolné $u \in \langle u_i, u_{i+1} \rangle$ je nenulových nejvýše $p + 1$ bázových funkcí, konkrétně $N_{i-p}^p(u), N_{i-p+1}^p(u), \dots, N_i^p(u)$

2.11.3 Neracionální B-spline křivky

V této kapitole se seznámíme s *neracionálními B-spline křivkami*. Prvním, kdo se zabýval významem B-spline funkcí byl v 19. století N. Lobachevsky. B-spline byly konstruovány pomocí konvoluce určitých hustot rozdělení pravděpodobnosti definovaných nad speciálním uzlovým vektorem. V roce 1946 I. J. Schoenberg použil B-spline pro vyhlazení statistických dat. Z hlediska dalšího aplikačního rozšíření byl významný objev rekurentní formule C. de Boorem, M. Coxem a L. Mansfieldem. B-spline křivky představují zobecnění Bézierových křivek.

Parametrická B-spline křivka stupně p v d -dimenzionálním prostoru je zobrazení $\mathbf{C}(u) : \langle u_0, u_m \rangle \rightarrow \mathbb{R}^d$ a definujeme ji jako lineární kombinaci $m-p$ bázových B-spline funkcí N_i^p

$$\mathbf{C}(u) = \sum_{i=0}^n \mathbf{P}_i N_i^p(u). \quad (2.85)$$

Body $\mathbf{P}_i \in \mathbb{R}^d$ tvoří množinu $n+1$ řídicích bodů řídicího polygonu křivky \mathbf{C} . Křivka je definována nad intervalem $\langle a, b \rangle$, kde a a b jsou krajní uzly uzlového vektoru \mathbf{u} . V další části textu se omezíme na speciální případ neperiodického neuniformního uzlového vektoru, který je v praxi používán nejčastěji

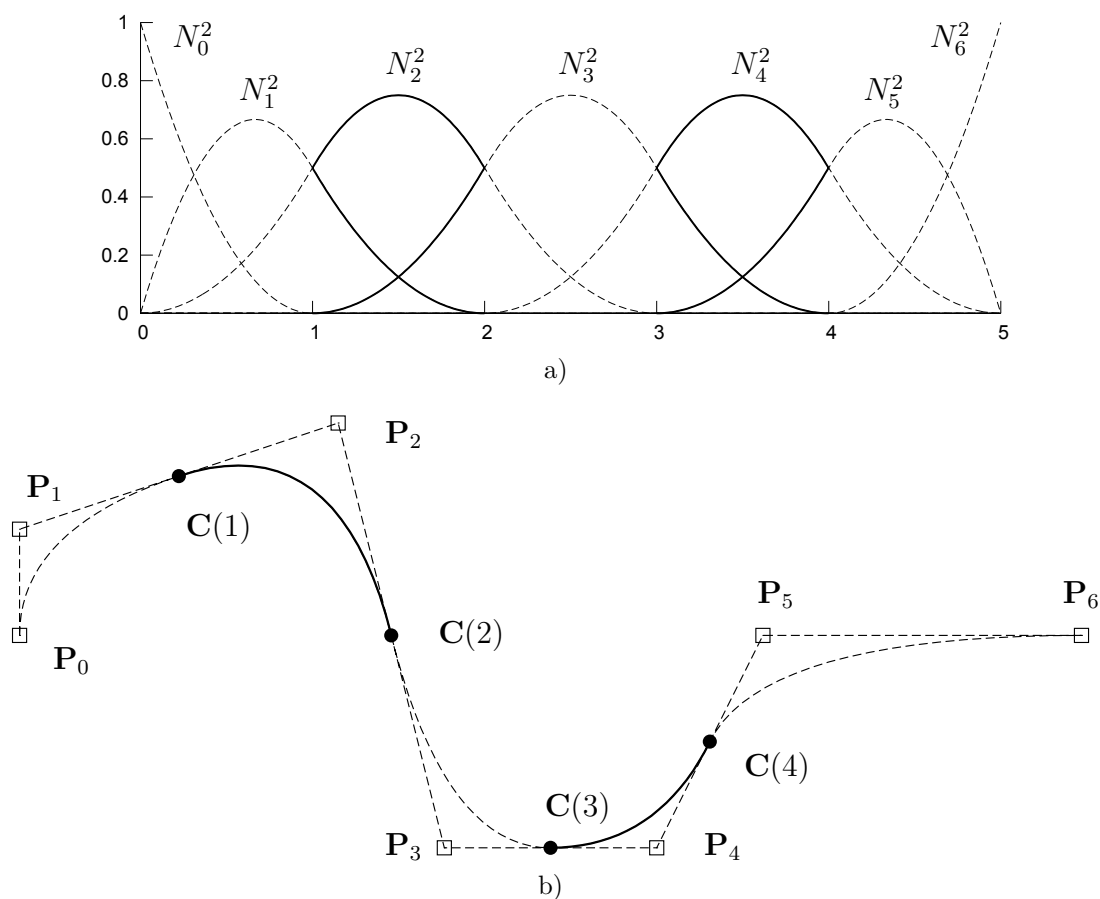
$$\mathbf{u} = \underbrace{(a, \dots, a)}_{p+1}, u_{p+1}, \dots, u_{m-p-1}, \underbrace{(b, \dots, b)}_{p+1}.$$

Násobnost krajních uzlů zajistí, že výsledná křivka vždy prochází počátečním \mathbf{P}_0 a koncovým \mathbf{P}_n řídicím bodem.

Opět uvedme některá důležitá pozorování.

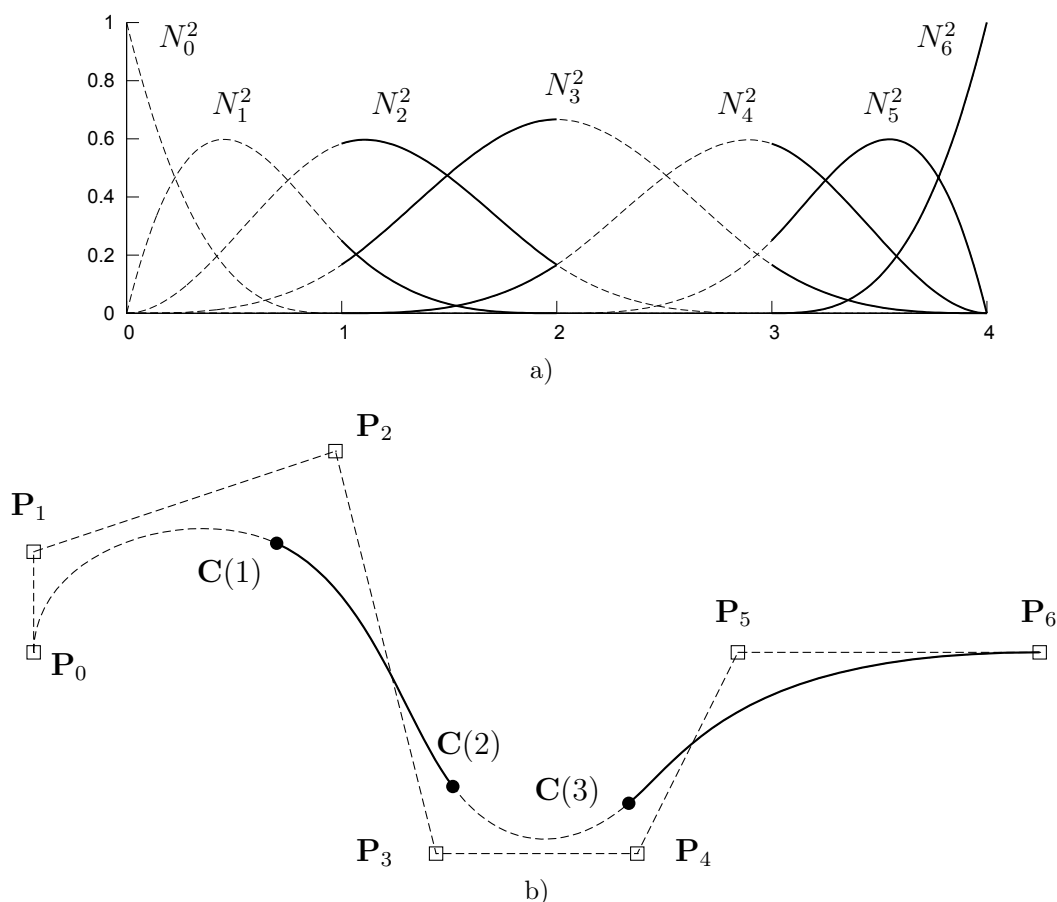
- \mathbf{C} je po částech polynomiální křivka, jelikož i její bázové funkce N_i^p jsou polynomiální funkce.
- Stupeň křivky p , počet řídicích bodů $n + 1$ a délka uzlového vektoru $m + 1$ jsou vzájemně svázány vztahem

$$n + 1 = m - p.$$



Obr. 2.21: a) Kvadratické bázové funkce pro uzlový vektor $(0, 0, 0, 1, 2, 3, 4, 5, 5, 5)$.
 b) Kvadratická B-spline křivka nad uvedeným uzlovým vektorem s vyznačeným vlivem bázových funkcí. Porovnejte vliv jednotlivých bází s tab. 2.2

Vyjmenujme alespoň některé základní vlastnosti B-spline křivek.



Obr. 2.22: a) Kubické bázové funkce pro uzlový vektor $(0, 0, 0, 0, 1, 2, 3, 4, 4, 4, 4)$.
 b) Kubická B-spline křivka nad uvedeným uzlovým vektorem s vyznačeným vlivem bázových funkcí. Opět porovnejte vliv jednotlivých bází s tab. 2.2

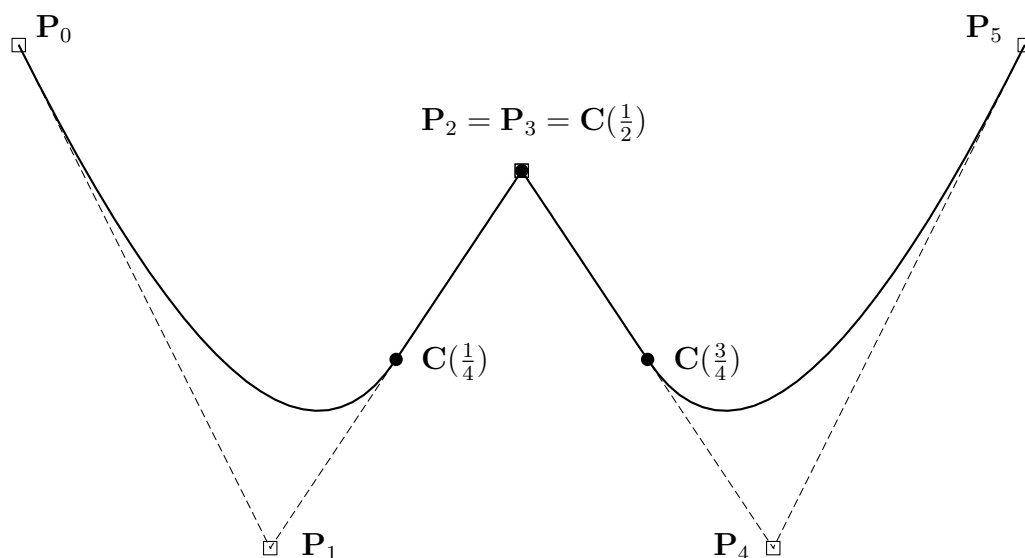
- Suma všech bázových funkcí je pro jakýkoli parametr $u \in \langle u_0, u_m \rangle$ rovna 1.

$$\sum_{i=0}^{m-p-1} N_i^p(u) = 1 \quad (2.86)$$

- S výjimkou bázové funkce druhého stupně, má každá bázová funkce právě jeden extrém.
- Nejvyšší stupeň křivky je roven počtu kontrolních bodů mínus 1.

$$p \leq m - p - 1 \quad (2.87)$$

- Libovolná afinní transformace kontrolních bodů má stejný efekt, jako aplikace téže transformace na všechny body křivky.
- Celá křivka se nachází v interiéru konvexního obalu jejich řídicích bodů.



Obr. 2.23: Kvadratická křivka s dvojnásobným řídicím bodem $P_2 = P_3$ a uzlovým vektorem $(0, 0, 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, 1, 1)$

Derivace

Vztah pro výpočet k -té derivace B-spline křivky získáme snadno pomocí derivace báze funkce (2.83)

$$\frac{d^k}{du^k} \mathbf{C}(u) = \sum_{i=0}^n \mathbf{P}_i \frac{d^k}{du^k} N_i^p(u). \quad (2.88)$$

Ukázku derivací B-spline křivky si lze prohlédnout na obr. 2.26.

2.11.4 Konvexní obal B-spline křivek

Tvrzení o vlastnostech konvexního obalu je v případě B-spline křivek silnější, než je tomu u Bézierových křivek. Pro B-spline křivku stupně p platí, že bod křivky leží uvnitř konvexního obalu $p + 1$ okolních bodů. Tudíž všechny body B-spline křivky musí ležet uvnitř sjednocení všech konvexních obalů vytvořených kolem $p + 1$ následujících vrcholů řídicího polygonu. Na obr. 2.27 lze vidět ukázky konvexních obalů křivek různých stupňů se shodnými řídicími body.

2.11.5 Plocha jako tenzorový součin dvou křivek

Před následující kapitolou si ujasněme klíčový pojem nutný k přechodu od křivek k plochám, a tím je tenzorový součin. *Tenzorový součin* dvou množin, např. $\{1, 2, 3\}$ a $\{x, y\}$, je množina všech dvojic, v našem případě $\{(1, x), (1, y), (2, x), (2, y), (3, x), (3, y)\}$. Pro vektory $\mathbf{a} = (a_1, a_2, a_3)^T$ a $\mathbf{v} = (b_1, b_2, b_3, b_4)^T$ by tenzorový součin nabýval podoby

$$\mathbf{a} \otimes \mathbf{b} = \mathbf{a}\mathbf{b}^T = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 & a_1b_4 \\ a_2b_1 & a_2b_2 & a_2b_3 & a_2b_4 \\ a_3b_1 & a_3b_2 & a_3b_3 & a_3b_4 \end{bmatrix}.$$

Podobně je tomu v případě následujících ploch, jejichž nové báze funkce dvou parametrů u a v jsou výsledkem součinu páru B-spline báze funkcí.

2.11.6 Neracionální B-spline plochy

Neracionální B-spline plochy jsou často využívány v oblastech, které vyžadují reprezentaci hladkých a plynule navazujících ploch. Takovéto plochy jsou typické pro oblasti designu, automobilového, leteckého či lodářského průmyslu. B-spline plocha je definována jako následující tenzorový součin

$$\mathbf{S}(u, v) = \sum_{i=0}^n \sum_{j=0}^m \mathbf{P}_{i,j} N_i^p(u) N_j^q(v), \quad (2.89)$$

kde uzlové vektory \mathbf{u} a \mathbf{v} tvaru

$$\begin{aligned} \mathbf{u} &= \underbrace{(a, \dots, a)}_{p+1}, u_{p+1}, \dots, u_{r-p-1}, \underbrace{(b, \dots, b)}_{p+1} \\ \mathbf{v} &= \underbrace{(c, \dots, c)}_{q+1}, u_{q+1}, \dots, u_{s-q-1}, \underbrace{(d, \dots, d)}_{q+1} \end{aligned}$$

čítají celkem $r + 1$, resp. $s + 1$ uzlů. Síť řídicích bodů $\mathbf{P}_{i,j}$ tvoří pravidelnou mřížku o rozměrech $(n + 1) \times (m + 1)$. Vzájemné vztahy mezi počty kontrolních bodů a délkami uzlových vektorů jsou dány následujícími předpisy

$$r = n + p + 1 \quad \text{a} \quad s = m + q + 1.$$

Poznamenejme, že při praktickém použití se často setkáme s normovaným tvarem uzlového vektoru, např. $(0, 0, 1, 2, 3, 4, 4) = (0, 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, 1)$. Pak pro krajní uzly platí, že $u_0 = 0$ a $u_m = 1$, přičemž relativní rozložení jednotlivých vnitřních uzlů zůstává zachováno.

Nyní uvedme některé základní vlastnosti tenzorového součinu báze funkcí, které vychází z již uvedených vlastností jednotlivých báze funkcí jedné proměnné.

- Obě báze funkce $N_i^p(u)$ a $N_j^q(v)$ jsou nezáporné, a tedy i jejich součin je nezáporný

$$N_i^p(u) N_j^q(v) \geq 0. \quad (2.90)$$

- Součet všech bázových funkcí je v daném bodě parametrického prostoru roven 1

$$\sum_{i=0}^n \sum_{j=0}^m N_i^p(u) N_j^q(v) = 1 \quad \text{pro všechny } (u, v) \in \langle a, b \rangle \times \langle c, d \rangle. \quad (2.91)$$

- Součin bázových funkcí mimo meze parametrického prostoru je nulový

$$\sum_{i=0}^n \sum_{j=0}^m N_i^p(u) N_j^q(v) = 0 \quad \text{pro všechny } (u, v) \notin \langle a, b \rangle \times \langle c, d \rangle. \quad (2.92)$$

- Pro všechny množiny $\langle u_{i_0}, u_{i_0+1} \rangle \times \langle v_{j_0}, v_{j_0+1} \rangle$ platí, že nejvýše $(p+1)(q+1)$ bázových funkcí je nenulových, konkrétně se jedná o funkce

$$N_i^p(u) N_j^q(v) \geq 0 \quad \text{pro } i_0 - p \leq i \leq i_0 \quad \text{a} \quad j_0 - q \leq j \leq j_0. \quad (2.93)$$

- Je-li $p > 0$ a $q > 0$, pak $N_i^p(u) N_j^q(v)$ nabývá právě jedno maximum.

Nyní se můžeme podívat alespoň na některé vlastnosti neracionálních B-spline ploch.

- Je-li $n = p$, $m = q$, $\mathbf{u} = (0, \dots, 0, 1, \dots, 1)$ a $\mathbf{v} = (0, \dots, 0, 1, \dots, 1)$, pak $\mathbf{S}(u, v)$ je Bézierova plocha.
- Plocha prochází (interpoluje) čtyři rohové body mřížky řídicích vrcholů: $\mathbf{S}(a, c) = \mathbf{P}_{0,0}$, $\mathbf{S}(b, c) = \mathbf{P}_{n,0}$, $\mathbf{S}(a, d) = \mathbf{P}_{0,m}$ a $\mathbf{S}(b, d) = \mathbf{P}_{n,m}$, což plyne z identity

$$N_0^p(a) N_0^q(c) = N_n^p(b) N_0^q(c) = N_0^p(a) N_m^q(d) = N_n^p(b) N_m^q(d) = 1. \quad (2.94)$$

- Chceme-li aplikovat afinní transformaci na plochu $\mathbf{S}(u, v)$, stačí transformovat pouze řídicí vrcholy.
- Je-li $(u, v) \in \langle u_{i_0}, u_{i_0+1} \rangle \times \langle v_{j_0}, v_{j_0+1} \rangle$, pak bod $\mathbf{S}(u, v)$ leží uvnitř konvexního obalu řídicích bodů $\mathbf{P}_{i,j}$, kde $i_0 - p \leq i \leq i_0$ a $j_0 - q \leq j \leq j_0$.
- Triangulovaná síť řídicích bodů tvoří po částech planární aproximaci \mathbf{S} a stejně jako v případě křivek (viz. obr. 2.27) platí, že čím je nižší stupeň plochy \mathbf{S} , tím je aproximace lepší.
- Díky lokalitě bázových funkcí se změna polohy řídicího bodu $\mathbf{P}_{i,j}$ projeví pouze nad podoblastí $\langle u_i, u_{i+p+1} \rangle \times \langle v_j, v_{j+q+1} \rangle$ parametrického prostoru.

Derivace

Parciální derivace neracionálních B-spline ploch až do řádu d včetně je možno získat podle vztahu

$$\frac{\partial^{k+l}}{\partial u^k \partial v^l} \mathbf{S}(u, v) = \sum_{i=0}^n \sum_{j=0}^m \mathbf{P}_{i,j} \frac{d^k}{du^k} N_i^p(u) \frac{d^l}{dv^l} N_j^q(v), \quad (2.95)$$

kde $0 \leq k + l \leq d$.

2.11.7 Racionální B-spline křivky (NURBS)

V této podkapitole definujeme racionální zobecnění předchozích B-spline forem a jedná se o B-spline v projektivním rozšíření euklidovského prostoru. Zkratka NURBS znamená *neuniformní racionální B-spline* (ang. NonUniform Rational B-Spline) křivky a plochy. To, že je křivka či plocha neuniformní znamená, stejně jako v předchozích případech, že jednotlivé uzly nemusí být rozloženy v rámci uzlového vektoru rovnoměrně. Racionalita přináší možnost specifikovat váhu každého řídicího bodu, což dále rozšiřuje možnosti manipulace s podobou výsledné křivky nebo plochy. Racionální B-spline plocha je jednotná matematická reprezentace schopná popsat běžné analytické plochy jakými jsou roviny, kónické plochy zahrnující kulové plochy, všechny regulární kuželosečky, kvadriky a volné plochy. Např. Bézierovy křivky umožňují popis parabol, ale další kuželosečky je jimi možné nahradit pouze přibližně. V současnosti se jedná o průmyslový standard v oblasti návrhu a reprezentace ploch tzv. A třídy. Mezi zásadní výhody patří:

- Jednotná a flexibilní forma pro reprezentaci široké škály tvarů.
- Numericky robustní a poměrně rychlé algoritmy pro výpočet bodů a derivací.
- Invariantní vůči afinní i perspektivní transformaci.
- Jsou zobecněním neracionálních B-spline a neracionálních i racionálních Bézierových křivek a ploch.
- Neracionální B-spline nebo Bézierovy křivky mohou aproximovat kružnice, ale nelze je jimi reprezentovat přesně. Pomocí racionálních B-spline lze reprezentovat jakoukoli kuželosečku, ačkoli tato reprezentace není jedinečná.

NURBS křivka řádu p je na intervalu $u \in \langle a, b \rangle$ definována následovně

$$\mathbf{C}(u) = \frac{\sum_{i=0}^n \mathbf{P}_i w_i N_i^p(u)}{\sum_{i=0}^n w_i N_i^p(u)}, \quad (2.96)$$

kde \mathbf{P}_i je $n + 1$ vrcholů řídicího polygonu, $w_i \geq 0$ jsou váhy přidružené k jednotlivým bodům a $N_i^p(u)$ jsou bázové funkce stupně p definované nad parametrizací neperiodickým a neuniformním uzlovým vektorem tvaru

$$\mathbf{u} = (\underbrace{a, \dots, a}_{p+1}, u_{p+1}, \dots, u_{m-p-1}, \underbrace{b, \dots, b}_{p+1}),$$

čítajícího $m + 1$ uzlů. Pomocí racionálních bázových funkcí

$$R_i^p(u) = \frac{w_i N_i^p(u)}{\sum_{k=0}^n w_k N_k^p(u)} \quad (2.97)$$

můžeme rovnici (2.96) přepsat do následujícího tvaru

$$\mathbf{C}(u) = \sum_{i=0}^n \mathbf{P}_i R_i^p(u). \quad (2.98)$$

Význam váhy w_i spočívá v možnosti řídit míru vlivu bodu \mathbf{P}_i na výsledný průběh křivky. Neformálně řečeno, čím je váha větší, tím se křivka více přimyká k příslušnému řídicímu bodu. Čím je hodnota váhy nižší, tím má příslušný řídicí bod menší vliv na výslednou podobu křivky. V případě, že je váha w_i nulová, nemají souřadnice bodu \mathbf{P}_i vliv na výsledný tvar křivky. Váha může být také záporná, mnohé nástroje však tuto možnost vůbec nepodporují. Dochází k porušení vlastností konvexního obalu, bod křivky může ležet mimo něj. Jelikož lze váhy w_i bodů \mathbf{P}_i reprezentovat pomocí homogenních souřadnic $\mathbf{P}_i^w = (w_i x_i, w_i y_i, w_i z_i, w_i)$, lze na racionální NURBS křivku nahlížet jako na neracionální B-spline křivku ve 4-dimenzionálním prostoru. Ke každému bodu \mathbf{P}_i^w v homogenních souřadnicích lze nalézt korespondující bod v 3D euklidovském prostoru pomocí perspektivního mapování $H : (wx, wy, wz, w) \rightarrow (wx/w, wy/w, wz/w, w/w) = (x, y, z, 1) \equiv (x, y, z) \in \mathbb{R}^3$. Vztah (2.96) můžeme upravit do podoby

$$\mathbf{C}^w(u) = \sum_{i=0}^n \mathbf{P}_i^w N_i^p(u) \quad (2.99)$$

a obdobně můžeme získat z (2.88)

$$\frac{d^k}{du^k} \mathbf{C}^w(u) = \sum_{i=0}^n \mathbf{P}_i^w \frac{d^k}{du^k} N_i^p(u). \quad (2.100)$$

Racionální křivku obdržíme zpět pomocí H

$$\mathbf{C}(u) = H \{ \mathbf{C}^w(u) \} = H \left\{ \sum_{i=0}^n \mathbf{P}_i^w N_i^p(u) \right\}. \quad (2.101)$$

Derivace

Derivace racionálních funkcí jsou komplikovanější. Zavedme následující značení

$$\mathbf{C}(u) = \frac{w(u)\mathbf{C}(u)}{w(u)} = \frac{\mathbf{A}(u)}{w(u)}, \quad (2.102)$$

kde vektorová funkce $\mathbf{A}(u) = \langle x^w(u), y^w(u), z^w(u) \rangle$ je tvořena prvními třemi složkami $\mathbf{C}^w(u)$, tedy

$$\mathbf{A}(u) = \sum_{i=0}^n N_i^p(u) w_i \mathbf{P}_i \quad (2.103)$$

a funkce $w(u)$ reprezentuje homogenní souřadnici

$$w(u) = \sum_{i=0}^n N_i^p(u) w_i. \quad (2.104)$$

První derivaci vypočteme z (2.102) pomocí vzorce pro derivaci podílu a obdržíme

$$\begin{aligned} \frac{d}{du} \mathbf{C}(u) &= \frac{w(u) \frac{d}{du} \mathbf{A}(u) - \frac{d}{du} w(u) \mathbf{A}(u)}{w(u)^2} \\ &= \frac{w(u) \left(\frac{d}{du} \mathbf{A}(u) - \frac{d}{du} w(u) \mathbf{C}(u) \right)}{w(u)^2} \\ &= \frac{\frac{d}{du} \mathbf{A}(u) - \frac{d}{du} w(u) \mathbf{C}(u)}{w(u)}, \end{aligned} \quad (2.105)$$

kde obecně k -tou derivaci funkce $\mathbf{A}(u)$ vypočteme snadno s pomocí (2.83)

$$\frac{d^k}{du^k} \mathbf{A}(u) = \sum_{i=0}^n \frac{d^k}{du^k} N_i^p(u) w_i \mathbf{P}_i \quad (2.106)$$

a obdobně získáme také derivace funkce $w(u)$

$$\frac{d^k}{du^k} w(u) = \sum_{i=0}^n \frac{d^k}{du^k} N_i^p(u) w_i. \quad (2.107)$$

Pro vyjádření vztahu pro derivace k -tého řádu křivky $\mathbf{C}(u)$ využijeme rovnost $\mathbf{A}(u) = w(u)\mathbf{C}(u)$ z (2.102) a Leibnizův vzorec (derivace součinu dvou funkcí)

$$\begin{aligned}
\frac{d^k}{du^k} \mathbf{A}(u) &= \frac{d^k}{du^k} (w(u) \mathbf{C}(u)) = \sum_{i=0}^k \binom{k}{i} \frac{d^i}{du^i} w(u) \frac{d^{k-i}}{du^{k-i}} \mathbf{C}(u) \\
&= w(u) \frac{d^k}{du^k} \mathbf{C}(u) + \sum_{i=1}^k \binom{k}{i} \frac{d^i}{du^i} w(u) \frac{d^{k-i}}{du^{k-i}} \mathbf{C}(u),
\end{aligned} \tag{2.108}$$

z čehož přímo plyne hledaná derivace

$$\frac{d^k}{du^k} \mathbf{C}(u) = \frac{\frac{d^k}{du^k} \mathbf{A}(u) - \sum_{i=1}^k \binom{k}{i} \frac{d^i}{du^i} w(u) \frac{d^{k-i}}{du^{k-i}} \mathbf{C}(u)}{w(u)}. \tag{2.109}$$

V praxi se nejčastěji zajímáme o první a druhou derivaci. První derivaci máme již vyjádřenou rovnicí (2.105). Druhou derivaci získáme položením $k = 2$ v obecném vztahu (2.109) a obdržíme

$$\frac{d^2}{du^2} \mathbf{C}(u) = \frac{\frac{d^2}{du^2} \mathbf{A}(u) - 2 \frac{d}{du} w(u) \frac{d}{du} \mathbf{C}(u) - \frac{d^2}{du^2} w(u) \mathbf{C}(u)}{w(u)}. \tag{2.110}$$

2.11.8 Racionální B-spline plochy (NURBS)

V této kapitole se konečně dostáváme k *NURBS plochám*, které definujeme obdobným způsobem jako tomu bylo u neracionálních ploch na základě tenzorového součinu racionálních B-spline křivek z předchozí kapitoly. NURBS plocha stupně p ve směru u a stupně q ve směru v je na intervalu $(u, v) \in \langle a, b \rangle \times \langle c, d \rangle$ po částech racionální vektorová funkce dvou proměnných

$$\mathbf{S}(u, v) = \frac{\sum_{i=0}^n \sum_{j=0}^m \mathbf{P}_{i,j} w_{i,j} N_i^p(u) N_j^q(v)}{\sum_{i=0}^n \sum_{j=0}^m w_{i,j} N_i^p(u) N_j^q(v)}, \tag{2.111}$$

kde $\mathbf{P}_{i,j}$ je síť $(n+1) \times (m+1)$ vrcholů řídicí mřížky, $w_{i,j}$ jsou váhy přidružené k jednotlivým bodům (jejich význam je obdobný jako v případě NURBS křivek) a $N_i^p(u)$, resp. $N_j^q(v)$ jsou B-spline bázové funkce stupně p , resp. q definované nad parametrizací neperiodickými a neuniformními uzlovými vektory tvaru

$$\begin{aligned}
\mathbf{u} &= \left(\underbrace{a, \dots, a}_{p+1}, u_{p+1}, \dots, u_{r-p-1}, \underbrace{b, \dots, b}_{p+1} \right) \\
\mathbf{v} &= \left(\underbrace{c, \dots, c}_{q+1}, v_{q+1}, \dots, v_{s-q-1}, \underbrace{d, \dots, d}_{q+1} \right)
\end{aligned}$$

kde $r = n + p + 1$ a $s = m + q + 1$. Zavedením po částech racionální bázové funkce

$$R_{i,j}^{p,q}(u,v) = \frac{w_{i,j} N_i^p(u) N_j^q(v)}{\sum_{k=0}^n \sum_{l=0}^m w_{k,l} N_k^p(u) N_l^q(v)} \quad (2.112)$$

lze rovnici (2.111) přepsat do tvaru

$$\mathbf{S}(u,v) = \sum_{i=0}^n \sum_{j=0}^m \mathbf{P}_{i,j} R_{i,j}^{p,q}(u,v). \quad (2.113)$$

Vlastnosti racionálních bázových funkcí $R_{i,j}^{p,q}(u,v)$ plynou z již zmíněných vlastností neracionálních bázových funkcí $N_i^p(u)$ a $N_j^q(v)$, a proto se jimi nemusíme na tomto místě znovu detailněji zabývat, jen připomeňme, že se jedná o nezápornost, normalizovaný tvar, lokalitu, podmínku existence právě jednoho extrému a interpolaci rohových bodů mřížky. Zmínit však můžeme jednu novou vlastnost.

- Jsou-li všechny váhy $w_i = a$ a $a \neq 0$, pak $R_{i,j}^{p,q}(u,v) = N_i^p(u) N_j^q(v)$ pro všechny i, j .

Nyní přímo navážeme některými vlastnostmi NURBS ploch.

- Předpokládejme, že $w_{i,j} \geq 0$ pro všechna i, j a $(u,v) \in \langle u_{i_0}, u_{i_0+1} \rangle \times \langle v_{j_0}, v_{j_0+1} \rangle$, pak se pro $i_0 - p \leq i \leq i_0$ a $j_0 - q \leq j \leq j_0$ bod $\mathbf{S}(u,v)$ nachází uvnitř konvexní obálky řídicích bodů $\mathbf{P}_{i,j}$.
- NURBS jsou invariantní vůči afinní i perspektivní transformaci.
- Změní-li se poloha bodu $\mathbf{P}_{i,j}$ nebo váha $w_{i,j}$, projeví se změna tvaru plochy pouze v rámci parametrické oblasti $\langle u_i, u_{i+p+1} \rangle \times \langle v_j, v_{j+q+1} \rangle$.
- Neracionální B-spline a neracionální i racionální Bézierova plocha jsou speciální případy NURBS ploch.
- NURBS plocha $\mathbf{S}(u,v)$ je $(p-k)$ -krát, resp. $(q-k)$ -krát spojitě diferencovatelná v uzlovém bodě u , resp. v o násobnosti k .

Jelikož lze váhy bodů reprezentovat pomocí homogenních souřadnic tak, že $\mathbf{P}_{i,j}^w = (w_{i,j} x_{i,j}, w_{i,j} y_{i,j}, w_{i,j} z_{i,j}, w_{i,j})$, lze na NURBS plochu nahlížet jako na neracionální plochu ve 4-dimenzionálním prostoru a vztah (2.111) můžeme upravit do podoby

$$\mathbf{S}^w(u,v) = \sum_{i=0}^n \sum_{j=0}^m \mathbf{P}_{i,j}^w N_i^p(u) N_j^q(v). \quad (2.114)$$

kde $\mathbf{S}^w(u,v)$ je bod na ploše v homogenních souřadnicích a platí $\mathbf{S}(u,v) = H \{ \mathbf{S}^w(u,v) \}$. Poznamenejme, že $\mathbf{S}^w(u,v)$ je po částech polynomiální (neracionální) plocha v 4-rozměrném prostoru.

Derivace

Derivace racionálních funkcí jsou komplikovanější. V předchozí kapitole jsme odvodili derivace k -tého řádu pro racionální B-spline křivky (2.109). Obdobným postupem můžeme dojít také k předpisu pro výpočet parciálních derivací NURBS ploch. Vzhledem k tomu, že se jedná o poměrně zdlouhavé odvození, omezíme se zde pouze na uvedení výsledného vztahu

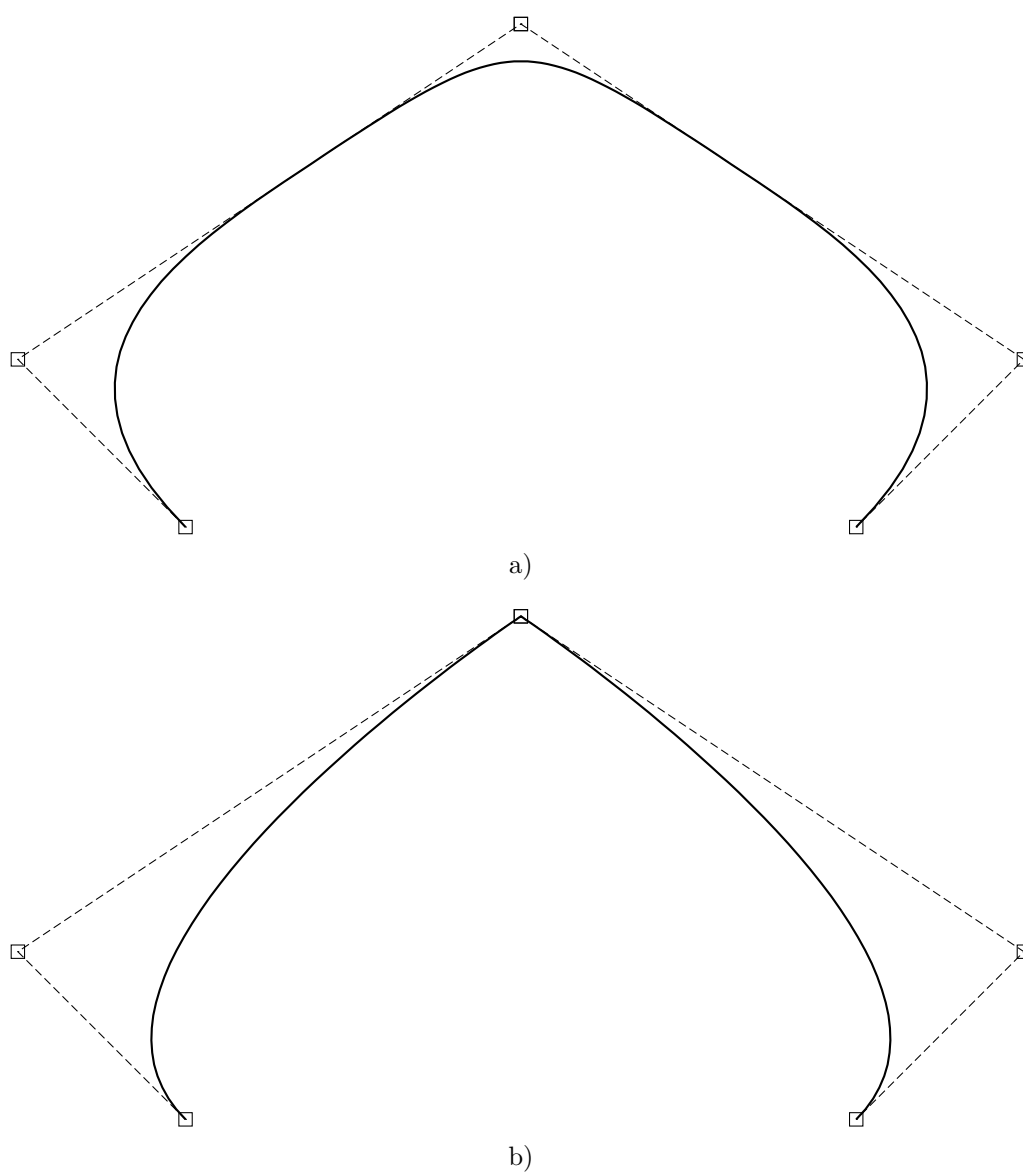
$$\begin{aligned} \frac{\partial^{k+l}}{\partial u^k \partial v^l} \mathbf{S}(u, v) &= \frac{1}{w(u, v)} \left(\frac{\partial^{k+l}}{\partial u^k \partial v^l} \mathbf{A}(u, v) \right. \\ &\quad - \sum_{i=1}^k \binom{k}{i} \frac{\partial^i}{\partial u^i} w(u, v) \frac{\partial^{k-i+l}}{\partial u^{k-i} \partial v^l} \mathbf{S}(u, v) \\ &\quad - \sum_{j=1}^l \binom{l}{j} \frac{\partial^j}{\partial v^j} w(u, v) \frac{\partial^{k+l-j}}{\partial u^k \partial v^{l-j}} \mathbf{S}(u, v) \\ &\quad \left. - \sum_{i=1}^k \binom{k}{i} \sum_{j=1}^l \binom{l}{j} \frac{\partial^{i+j}}{\partial u^i \partial v^j} w(u, v) \frac{\partial^{k-i+l-j}}{\partial u^{k-i} \partial v^{l-j}} \mathbf{S}(u, v) \right). \end{aligned} \quad (2.115)$$

Závěrem se ještě pokusme vyjádřit z rovnice (2.115) následující často používané parciální derivace

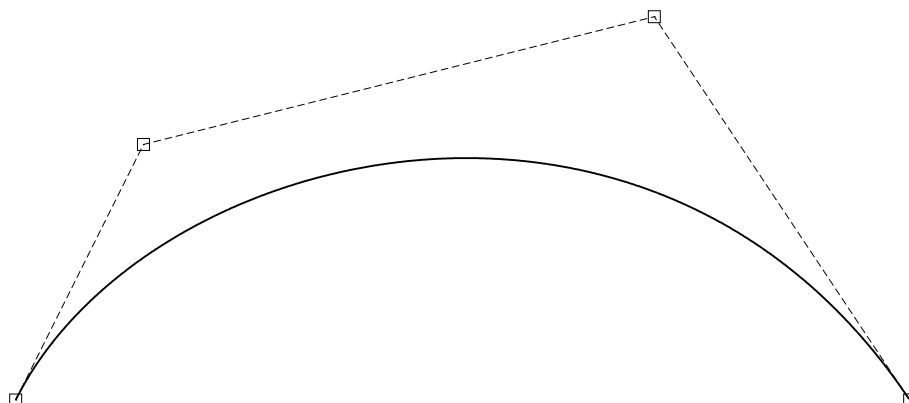
$$\begin{aligned} \frac{\partial^2}{\partial u \partial v} \mathbf{S}(u, v) &= \frac{1}{w(u, v)} \left(\frac{\partial^2}{\partial u \partial v} \mathbf{A}(u, v) - \frac{\partial^2}{\partial u \partial v} w(u, v) \mathbf{S}(u, v) \right. \\ &\quad \left. - \frac{\partial}{\partial u} w(u, v) \frac{\partial}{\partial v} \mathbf{S}(u, v) - \frac{\partial}{\partial v} w(u, v) \frac{\partial}{\partial u} \mathbf{S}(u, v) \right), \end{aligned} \quad (2.116)$$

$$\begin{aligned} \frac{\partial^2}{\partial u^2} \mathbf{S}(u, v) &= \frac{1}{w(u, v)} \left(\frac{\partial^2}{\partial u^2} \mathbf{A}(u, v) - 2 \frac{\partial}{\partial u} w(u, v) \frac{\partial}{\partial u} \mathbf{S}(u, v) \right. \\ &\quad \left. - \frac{\partial^2}{\partial u^2} w(u, v) \mathbf{S}(u, v) \right), \end{aligned} \quad (2.117)$$

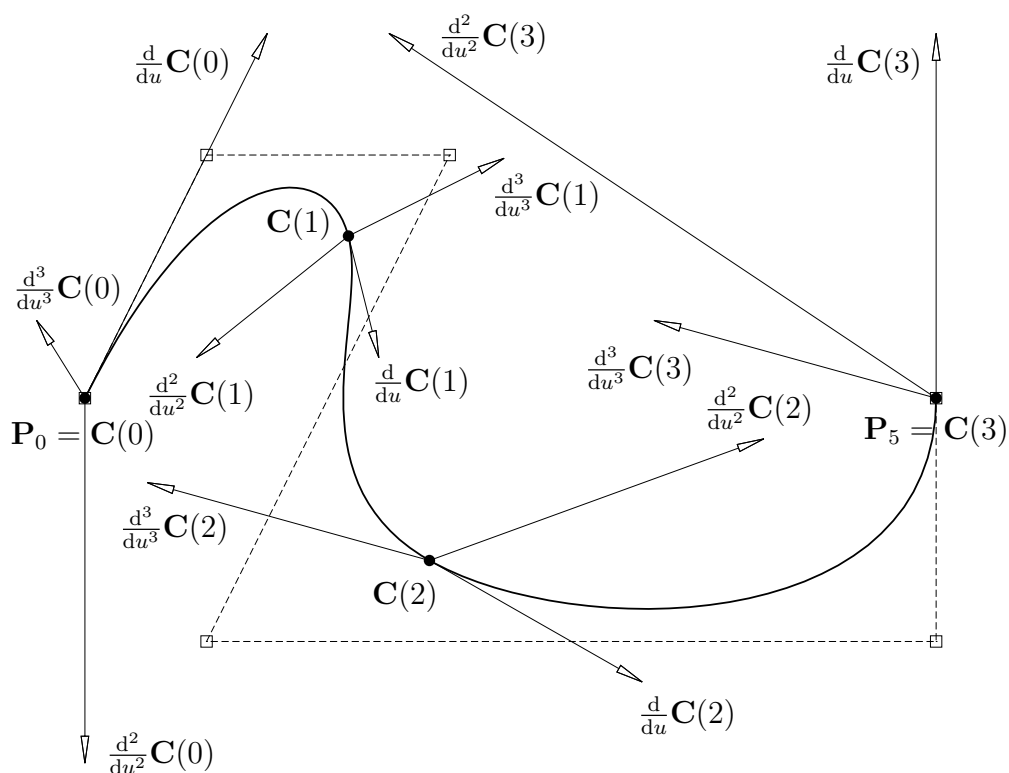
$$\begin{aligned} \frac{\partial^2}{\partial v^2} \mathbf{S}(u, v) &= \frac{1}{w(u, v)} \left(\frac{\partial^2}{\partial v^2} \mathbf{A}(u, v) - 2 \frac{\partial}{\partial v} w(u, v) \frac{\partial}{\partial v} \mathbf{S}(u, v) \right. \\ &\quad \left. - \frac{\partial^2}{\partial v^2} w(u, v) \mathbf{S}(u, v) \right). \end{aligned} \quad (2.118)$$



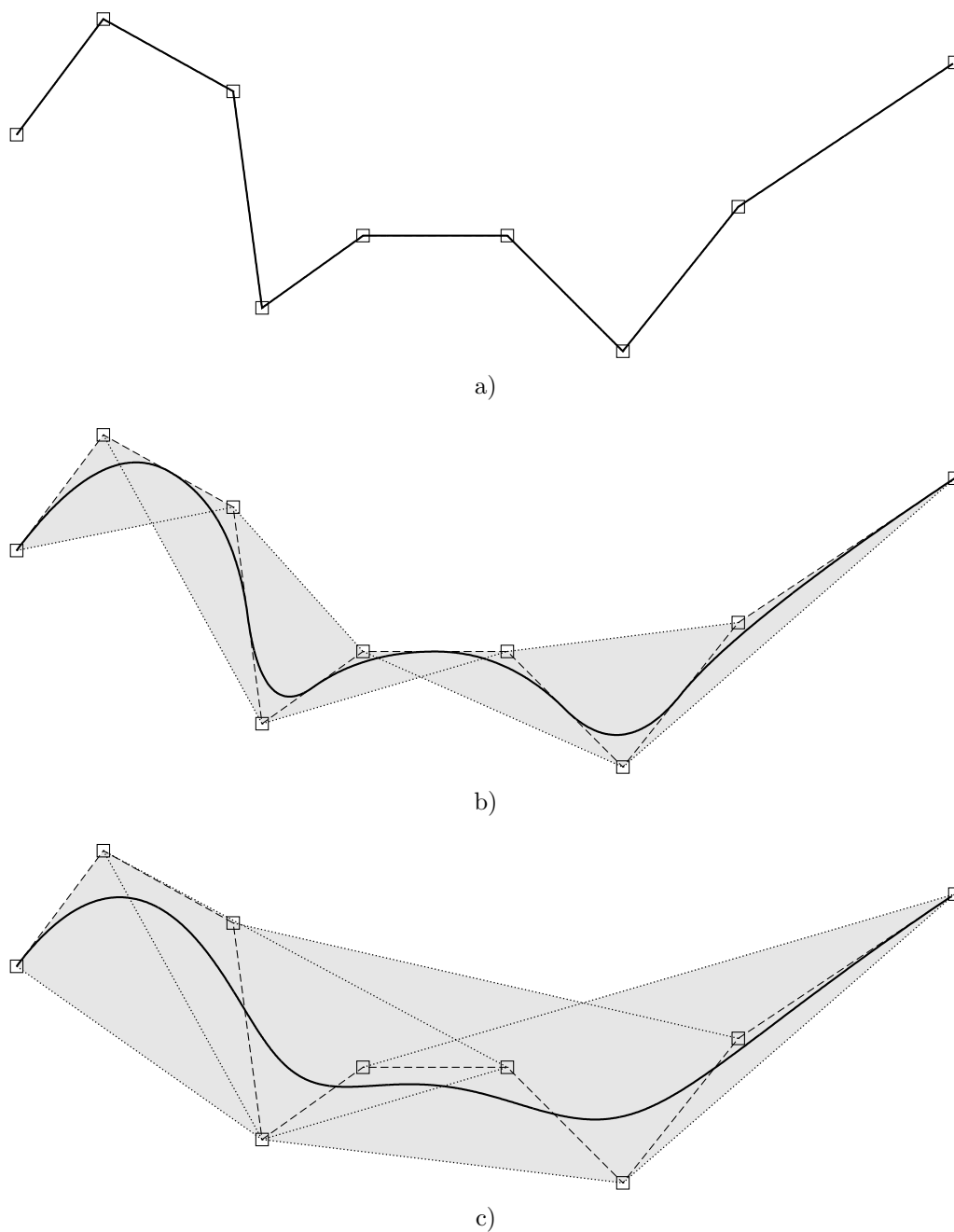
Obr. 2.24: Příklady kubických křivek s dvojnásobným řídicím bodem $\mathbf{P}_2 = \mathbf{P}_3$



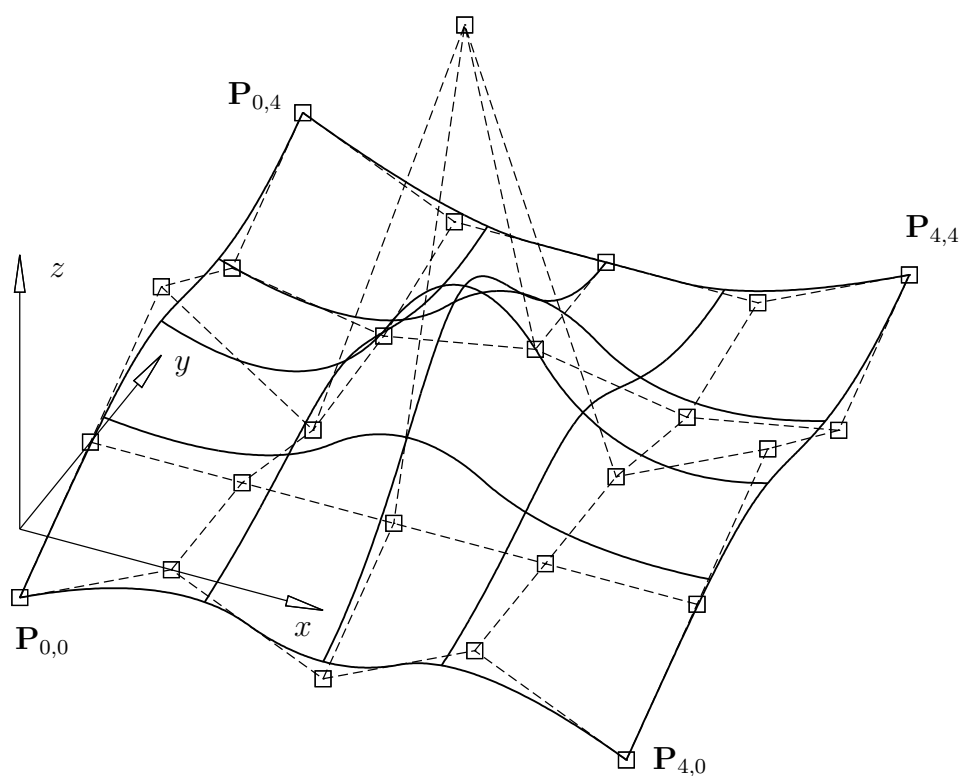
Obr. 2.25: Kubická B-spline křivka s uzlovým vektorem $(0, 0, 0, 0, 1, 1, 1, 1)$ nebo také kubická Bézierova křivka jakožto speciálním případ



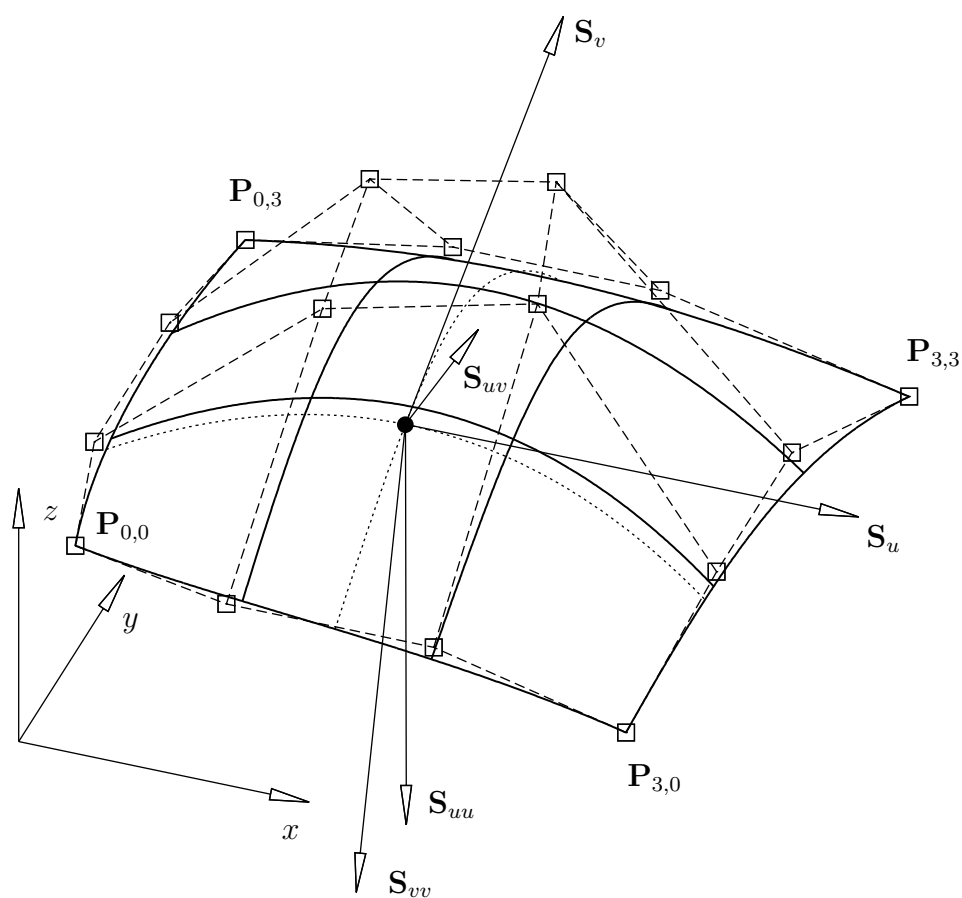
Obr. 2.26: Kubická B-spline křivka s uzlovým vektorem $(0, 0, 0, 0, 1, 2, 3, 3, 3, 3)$. První, druhá a třetí derivace jsou vypočteny pro parametry $u \in \{0, 1, 2, 3\}$. Délky vektorů jednotlivých stupňů derivací jsou upraveny koeficienty 0.5, 0.25 a 0.08



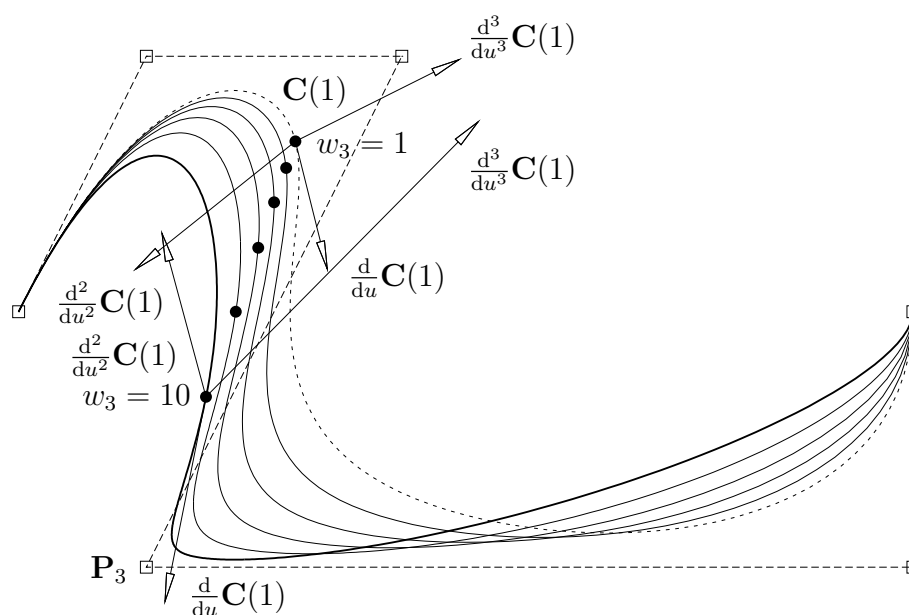
Obr. 2.27: Konvexní obal B-spline křivky stupně a) $p = 1$, b) $p = 2$ a c) $p = 5$. Všimněte si, že v případě křivky stupně 1 je konvexní obálka totožná s řídicím polygonem



Obr. 2.28: Kvadratický \times kubický B-spline povrch. Příslušné bázové funkce jsou parametrizovány uzlovými vektory $\mathbf{u} = (0, 0, 0, 1, 2, 3, 3, 3)$ a $\mathbf{v} = (0, 0, 0, 0, 1, 2, 2, 2, 2)$. Řídicí síť obsahuje 5×5 bodů



Obr. 2.29: Bikubická B-spline plocha nad uzlovými vektory $\mathbf{u} = \mathbf{v} = (0, 0, 0, 0, 1, 1, 1, 1)$. Řídící síť obsahuje 4×4 bodů. Zobrazené parciální derivace jsou vypočteny v bodě $\mathbf{S}(0.5, 0.3)$



Obr. 2.30: Racionální kubická B-spline křivka s uzlovým vektorem $(0, 0, 0, 0, 1, 2, 3, 3, 3, 3)$. Váha řídicího bodu \mathbf{P}_3 nabývá pro každou křivku postupně hodnoty $\{1, 1.4, 2, 3, 5, 10\}$, ostatní váhy jsou pevně nastaveny na hodnotu 1. Čárkovaná křivka je ekvivalentní s neracionální B-spline křivkou z obr. 2.26. Délky vektorů jednotlivých stupňů derivací vypočtených podle (2.109) jsou upraveny koeficienty 0.5, 0.25 a 0.08

Kapitola 3

Modelování těles

Modelování těles je v počítačové grafice úlohou zásadní důležitosti. Práce s tělesy (byť různého technického významu) je také společnou vlastností mnoha CAD systémů pro strojírenství, stavebnictví atd. Existence trojrozměrného modelu umožňuje získat jakoukoli geometrickou informaci o vytvářeném objektu. Lze např. získat informace pro libovolné zobrazení objektu včetně řezů. Lze vypočítat takové charakteristiky jako jsou objem, hmotnost, poloha těžiště, momenty a poloměry setrvačnosti atd. Dále může trojrozměrný model sloužit k automatizované přípravě zadání pro výpočty metodou konečných prvků. S využitím trojrozměrného modelu lze též uvažovat o automatizovaném rozpoznávání tvarů vzniklých v průběhu modelování - geometrický model se tak může stát východiskem pro systémy CAD s jistými prvky inteligence.

Modelování těles zabezpečují systémy resp. podsystémy, které bývají nazývány objemovými modeláři (solid modelery). Problematika modelování těles (reprezentace těles v počítači) a realizace odpovídajících modelářů tvoří významnou oblast počítačové grafiky, která se vyvinula v samostatné odvětví. Počátky vývoje této oblasti spadají přibližně do sedmdesátých let dvacátého století. Od té doby bylo dosaženo pozoruhodných výsledků, které do značné míry podmínily rozmach systémů CAD. Přesto však nelze vývoj v této oblasti zatím považovat za zcela uzavřený.

V této kapitole se budeme problematikou objemového modelování zabývat podrobněji. Ze známých metod modelování se zaměříme zejména na dvě z nich, a to na reprezentaci těles pomocí povrchu (boundary representation) a na tzv. konstruktivní geometrii těles (constructive solid geometry). Vždy se též budeme zabývat řešením problémů vyvolaných potřebou provádět výpočty (např. zobrazení, booleovské operace) nad modely uvedených typů.

Ještě dřív, než se pustíme do řešení uvedených praktických problémů, musíme se nejprve věnovat otázce, co považujeme za těleso. Namítnete, že to je jasné. Víte to však pravděpodobně jen intuitivně. To pro počítačové zpracování nestačí, protože počítač vaši intuici nemá. Teoretický úvod se může čtenáři zdát poněkud nepříjemný. Pro řádné pochopení problematiky je však alespoň nějaká znalost v tomto ohledu nezbytná. Věříme, že cílevědomý čtenář jistě nalezne dost síly k tomu, aby úvodní

podkapitoly „přežil“. Závěry z teoretického úvodu využijeme později, a to zejména při popisu těles pomocí hraniční reprezentace.

3.1 Topologické prostory a topologická ekvivalence

Chceme-li se seriózněji zabývat problémem, jak je možné definovat a popisovat tělesa, sotva se můžeme obejít bez pojmů jako je topologický prostor, topologické zobrazení a n -manifold v E^m . Uvedené pojmy zavedeme v této podkapitole. Poznáváme, že se jedná o vcelku rozsáhlou a náročnou partii matematiky. My se zde proto musíme spokojit jen se základními informacemi a terminologií, které budeme později využívat.

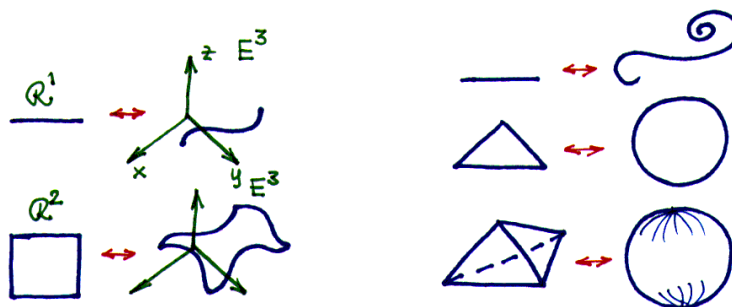
Topologický prostor je množina X spolu se systémem svých podmnožin τ . Systém τ musí splňovat následující axiomy: 1) $\emptyset \in \tau$, $X \in \tau$; 2) průnik konečného počtu množin z τ musí být opět v τ ; 3) sjednocení libovolného počtu množin (i nespočetného) z τ je opět v τ . Systém τ množin tedy musí být uzavřený vůči konečnému průniku a libovolnému součtu (a to i s nespočetným počtem prvků). Souboru τ říkáme topologie na X . Množiny v τ pak nazveme otevřené množiny a jejich doplňky v X uzavřené množiny. Topologický prostor budeme značit (X, τ) .

Podmnožinu $U \subset X$ topologického prostoru (X, τ) nazveme *okolím bodu* p , pokud existuje prvek $O \in \tau$ takový, že $p \in O$ a platí $O \subset U$. Okolí bodu p budeme značit $U(p)$. Zobrazení f topologického prostoru (X, τ) do topologického prostoru (X', τ') je *spojité*, jestliže obraz každého okolí bodu p v (X, τ) je současně také okolím bodu $f(p)$ v (X', τ') .

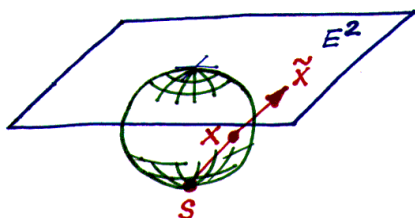
Jestliže f je bijekcí a f i f^{-1} jsou spojitá zobrazení, pak se f nazývá *homeomorfizmem* (topologickým zobrazením). Dva topologické prostory jsou topologicky ekvivalentní, jestliže mezi nimi existuje *homeomorfismus*. Homeomorfismus je vzájemně jednoznačné zobrazení mezi topologickými prostory, které zachovává topologické vlastnosti. Z pohledu topologie mají tedy oba prostory stejné vlastnosti (jsou stejné). Homeomorfismus má pro definici pojmu těleso zásadní význam (má ale také význam při definici pojmů jako je křivka či plocha). Obr. 3.1 ukazuje několik prakticky významných příkladů homeomorfismu.

Při popisu těles, a to zejména jejich hranice, využijeme pojmu n -manifold v E^m (konkrétně zejména 2-manifold v E^3). Seznámíme se proto s tímto pojmem podrobněji. N -manifold v E^m , kde $m \geq n$, je taková podmnožina E^m , která je homeomorfní s E^n . V tomto případě bychom mohli ještě doplnit, že se jedná o tzv. manifold bez hranice. N -manifold v E^m s hranicí je pak podmnožina E^m , která je homeomorfní s pozitivním poloprostorem $E^{n+} = \{(x_1, \dots, x_n) \in E^n | x_1 \geq 0\}$. Praktický dopad ukazuje obr. 3.2. Jako těleso uvažujeme kouli. Hranicí tělesa je kulová plocha. Kulová plocha je 2-manifoldem v E^3 . Obr. 3.2 přímo ukazuje homeomorfismus prokazující uvedené tvrzení.

Právě uvedenou definici by bylo dále možné poněkud modifikovat, a to takto: N -manifold v E^m , je taková podmnožina E^m , v níž okolí každého bodu je homeomorfní



Obr. 3.1: Některé prakticky zajímavé příklady homeomorfismu: Křivka v E^3 je homeomorfní s intervalem v R ; plocha je homeomorfní s oblastí v R^2 ; úsečka je topologicky ekvivalentní křivce; trojúhelník je ekvivalentní kruhu; čtyřstěn je ekvivalentní kouli.



Obr. 3.2: Kulová plocha jako 2-manifold v E^3 ; homeomorfismem je zde projekce bodů kulové plochy z bodu S na rovinu E^2 .

s E^n případně s E^{n+} . Pokuste se sami promyslet důsledky této modifikace; kterou později využijeme.

Příklad 3.1. Protože kruh je homeomorfní s E^2 , můžeme také říci, že 2-manifold bez hranice v E^3 je taková množina bodů v E^3 , v níž je okolí každého bodu homeomorfní s kruhem. Jak uvidíme později, je toto tvrzení názorné praktické.



3.2 Soubor simplexů

Vysvětleme dále, jak lze pojem těleso a jeho hranici zavést kombinatoricky. Je to zajímavé i praktické zároveň. Při troše trpělivosti čtenář jistě shledá, že všechny pojmy, které jsou zde zavedeny, plně odpovídají intuitivní představě, že povrch tělesa lze pokrýt trojúhelníkovou sítí nebo že celé těleso lze vyplnit čtyřstěny (trojúhelník i čtyřstěn zde ale chápeme z pohledu topologické ekvivalence; hrany ani stěny nemusí být rovinné). Jen snaha o určitou formální preciznost a systematickosti se snad může zdát při prvním čtení poněkud nepřátelská. Jestliže si ale čtenář během čtení výkladu bude jeho obsah současně také zakreslovat, jistě tuto potíž snadno překoná.

Nechť p_0 a p_1 jsou dva různé body. Jejich konvexní kombinací je množina $\langle p_0, p_1 \rangle = \{\lambda p_0 + (1 - \lambda)p_1 \mid 0 \leq \lambda \leq 1\}$. Z geometrického pohledu je $\langle p_0, p_1 \rangle$ úsečka. Podobně můžeme definovat konvexní kombinaci $d + 1$ bodů vztahem $\langle p_0, \dots, p_d \rangle = \left\{ \sum_{k=1}^d \lambda_k p_k \mid \lambda_k \geq 0, \sum \lambda_k = 1 \right\}$. Koeficienty λ_k v předchozím výrazu se nazývají barycentrické souřadnice. Řekneme, že body jsou lineárně nezávislé, jestliže žádný z nich není obsažen v konvexní kombinaci zbývajících. Jestliže p_k je množina lineárně nezávislých bodů, pak jsou barycentrické souřadnice každého bodu v $\langle p_0, \dots, p_d \rangle$ jedinečné.

Pojmem d -simplex rozumíme konvexní kombinaci $d + 1$ lineárně nezávislých bodů, d je pak dimenzí simplexu. Je zřejmé, že 0-simplex je bod, 1-simplex je úsečka, 2-simplex je trojúhelník, 3-simplex je čtyřstěn atd. Hranice d -simplexu (nazvěme ji S) sestává ze všech $(d - k)$ -simplexů obsažených v S , kde $0 < k \leq d$. Každý simplex hranice budeme dále nazývat stěnou, k -simplex, který je součástí hranice, pak také k -stěnou. Není obtížné ověřit následující tvrzení: 1) každý d -simplex má právě $\binom{d+1}{k+1}$ k -simplexů, které jsou jeho k -stěnami; 2) každé dva d -simplexy jsou homeomorfní. Soubor simplexů (označme jej C) je konečná množina simplexů vyhovující následujícím podmínkám: 1) Nechť S je simplex a S' jeho libovolná stěna. Je-li $S \in C$, pak musí také být $S' \in C$. 2) Nechť S_1 a S_2 jsou dva simplexy z C . Pak průnik S_1, S_2 je buď prázdný nebo musí být opět simplexem z C . Dimenze souboru simplexů je definována jako maximum dimenze jednotlivých simplexů, které jsou v souboru obsaženy. Dimenze simplexu je invariantní vůči spojitému zobrazení. Jestliže S je d -simplex a $d > 0$, pak hranice S je soubor simplexů dimenze $d - 1$.

Podmnožina E^n je triangulovatelná, jestliže je homeomorfní se souborem simplexů. Triangulovatelná množina je také nazývána topologickým mnohostěnem. Poznamenejme, že tohoto termínu zde není užito v geometrickém smyslu. Nemusíme si tedy představovat skutečný mnohostěn, na jaký jsme byli doposud zvyklí, protože homeomorfismus může zobrazovat lineární útvary do zakřivených. Tak např. koule v E^3 je také topologický mnohostěn, protože je homeomorfní s 3-simplexem.

Nechť S je d -simplex. Vlastní podmnožina $\{q_0, \dots, q_r\}$ množiny vrcholů simplexu S definuje r -simplex S_1 , který je stěnou S . Nechť $\{q_{r+1}, \dots, q_d\}$ jsou zbývající vrcholy S . Tyto vrcholy definují další stěnu S_2 . Stěny S_1 a S_2 nazýváme protilehlými. Nechť S a S' jsou simplexy v souboru simplexů. S, S' nazvěme přilehlými, jestliže oba obsahují tutéž stěnu S'' . Řekneme také, že S a S' incidují s S'' . Soubor simplexů C nazvěme souvislým, jestliže pro každou dvojici S, S' z C lze v C nalézt posloupnost simplexů S_1, \dots, S_r tak, že $S = S_1, S' = S_r$ a pro $1 \leq k < r$ S_k inciduje s S_{k+1} nebo S_k je stěnou S_{k+1} . Dále nechť S je simplex v souboru simplexů C . Předpokládejme, že s S inciduje s konečnou množinou simplexů $\{S_1, \dots, S_r\}$ v C (pro každý incidující simplex je S jeho stěnou). Nechť T_i je stěna simplexu S_i protilehlá stěně S . Množinu všech protilehlých stěn T_i ($1 \leq i \leq r$) nazvěme link S .

Nyní již můžeme definovat 2 a 3-manifolds v E^3 s využitím souboru simplexů: 2-manifold v E^3 bez hranice je homeomorfní se souborem simplexů C dimenze 2 splňujícím následující podmínky: 1) s každým 1-simplexem v C incidují právě dva

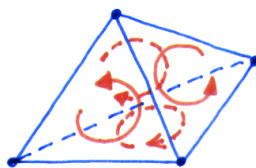
2-simplexy; 2) link každého 0-simplexu v C je kružnicí (triangulací kružnice). 3-manifold v E^3 bez hranice je homeomorfní se souborem simplexů C dimenze 3 splňujícím následující podmínky: 1) s každým 2-simplexem v C incidují právě dva 3-simplexy; 2) link každého 0-simplexu v C je triangulací kulové plochy. (Poznámenejme, že z hlediska modelování těles je 3-manifold v E^3 útvarem poněkud nepraktickým. Uvádíme jej ale pro úplnost.) Při popisu 3-manifolds E^3 s hranicí je třeba rozlišovat simplex ležící na hranici a simplex ležící uvnitř. Nechtě je např. S 2-simplex souboru simplexů C . S je vnitřní stěnou, jestliže inciduje právě se dvěma 3-simplexy z C , nebo je stěnou na hranici, jestliže inciduje právě s jedním 3-simplexem z C . Podobně 0-simplex je vnitřní, jestliže je jeho link triangulací kulové plochy, nebo leží na hranici, jestliže je jeho link triangulací kružnice. Můžeme tedy shrnout, že 3-manifold v E^3 s hranicí je homeomorfní se souborem simplexů C splňujícím následující podmínky: 1) s každým 2-simplexem z C incidují jeden nebo dva 3-simplexy z C ; 2) link každého 0-simplexu je triangulací kružnice nebo kulové plochy.

3.3 Tělesa s hranicí 2-manifold v E^3

Intuitivně uvažujeme tělesa jako celek - tj. trojdimenzionální objekt s vnitřkem a s hranicí. V hraniční reprezentaci těles, jak později uvidíme, máme však k dispozici pouze hranici. Je proto nutné studovat vztahy mezi vlastnostmi tělesa a vlastnostmi jeho hranice. Přesněji lze říci, že předpokládáme, že pracujeme s tělesy 3-manifold v E^3 s hranicí. Hranice takového tělesa je 2-manifold v E^3 . Chceme zjistit, jaké vlastnosti, kterých lze prakticky využít při implementaci na počítači, má 2-manifold (v E^3), který tvoří hranici tělesa. Důležité vlastnosti jsou následující.

Orientovatelnost: Orientovatelnost (což je schopnost „být orientován“) nejnázorněji přiblížíme prostřednictvím triangulace manifoldu. Pojem orientovatelnosti 2-manifolds E^3 lze tedy vysvětlit pomocí orientace souboru simplexů, který je triangulací manifoldu. 2-simplex orientujeme cyklickým uspořádáním vrcholů. Např. simplex $\langle p_0, p_1, p_2 \rangle$ může být orientován dvěma způsoby, a to buď p_0, p_1, p_2 nebo p_0, p_2, p_1 . Orientace 2-simplexu indukují také orientaci každé z jeho stěn, v našem případě 1-simplexů. Nechtě S a S' jsou dva přilehlé 2-simplexy v souboru simplexů C . Předpokládáme, že oba přiléhají k 1-simplexu S'' . Řekneme, že S a S' jsou orientovány konzistentně, jestliže orientace S'' indukované orientacemi S a S' jsou opačné (obr. 3.3). Soubor simplexů dimenze 2 považujeme za orientovatelný, jestliže všechny navzájem přiléhající 2-simplexy mohou být orientovány konzistentně.

Zjištění, že soubor simplexů a 2-manifold v E^3 tvořící hranici tělesa je orientovatelný, je prakticky významné. Pokud by tomu tak nebylo, nebyla by možná tzv. hraniční reprezentace tělesa v podobě, jak se s ní později seznámíme. Na základě dosud uvedeného by také bylo možné zkonstruovat algoritmus pro testování orientovatelnosti souboru simplexů. Algoritmus začne volbou orientace libovolného 2-simplexu. Dále jsou postupně orientovány sousedící 2-simplexy tak, aby jejich ori-



Obr. 3.3: Příklad orientace čtyřstěnu: Všechny trojúhelníky (2-simplexy) jsou orientovány tak, že na hranách (1-simplexech) běží orientace proti sobě a je tedy konzistentní. Jsou možné dvě konzistentní orientace. Jedna je vyznačena; druhá by vznikla otočením směru šipek.

entace byla konzistentní. Jestliže se v procesu „šíření orientace“ narazí na 2-simplex, který je již orientovaný a u kterého by orientaci bylo zapotřebí obrátit, pak soubor simplexů není orientovatelný. My ale nyní již víme, že tato situace by u „slušných“ těles neměla nastat. Lze snadno ukázat, že 2-manifold v E^3 tvořící hranici tělesa je orientovatelný právě dvěma způsoby. I toho při praktické implementaci využijeme. Orientaci můžeme zvolit např. tak, abychom při pohledu zevnitř tělesa na jeho stěnu (2-simplex) viděli její vrcholy v pořadí proti směru hodinových ručiček.

Dále se věnujme obrácenému problému, kdy bychom chtěli formulovat požadavky, které musí být splněny, aby souvislý a orientovaný 2-manifold v E^3 bez hranice bylo možné považovat za hranici tělesa (přesněji řečeno souvislého 3-manifolds v E^3 s hranicí). K tomu uvedeme několik dalších pojmů a pozorování.

Omezená variace: Necht' M je 3-manifold s hranicí v E^3 . Řekneme, že M vykazuje omezenou variaci, jestliže každá přímka v E^3 protíná M jen v konečně mnoha segmentech a každá rovina v E^3 jen v konečně mnoha oblastech. Analogicky můžeme říci, že kompaktní 2-manifold v E^3 vykazuje omezenou variaci, jestliže jej každá přímka protíná jen v konečně mnoha bodech a každá rovina jen v konečně mnoha křivkách. Je zřejmé, že takovou vlastnost můžeme v běžných případech předpokládat. Pokud bychom měli za úkol představit si nějaké těleso, jistě si představíme takové, které uvedenou podmínku splňuje. Potíže by spíše naopak mohly nastat s představením si opačného případu, kdy je útvarů vzniklých protínáním nekonečně mnoho.

Vnitřek tělesa: Necht' M je orientovaný, souvislý 2-manifold v E^3 . Bod p nazveme vnitřním bodem, jestliže splňuje následující podmínky: 1) Existuje takový 2-simplex S souboru C , že se jeho vrcholy p_0, p_1, p_2 při pozorování z p jeví v pořadí proti směru hodinových ručiček; 2) na S existuje bod q tak, že úsečka pq neprotíná M jinde než v q . 2-manifold tak rozděluje E^3 na dvě otevřené množiny. Jedna obsahuje všechny vnitřní body tělesa (vnitřek). Druhá obsahuje body, které nenáleží ani vnitřku ani nejsou součástí M ; jedná se o vnějšek tělesa. Za těleso typu manifold s jedním povrchem považujeme souvislý 3-manifold v E^3 s hranicí. Jeho hranice je kompaktní souvislý 2-manifold bez hranice v E^3 s omezenou variací. Těleso sestává

z bodů uvnitř a z bodů na hranici. Hranice je orientována tak, aby orientace správně určovala, které body leží uvnitř tělesa a které leží vně.

Konzistentní orientace obálek: K tomu, abychom mohli pracovat s tělesy s více povrchy (můžeme mít např. těleso, uvnitř kterého je dutina), musíme zavést ještě pojem konzistentní orientace obálek (v našem příkladě je jednou obálkou vnější hranice tělesa, druhou obálkou je hranice dutiny). Necht M_1 a M_2 jsou dva souvislé, orientované a navzájem disjunktní 2-manifolds v E^3 . Řekneme, že M_1 a M_2 jsou orientovány konzistentně, jestliže jsou orientovány tak, aby obě orientace ukazovaly dovnitř tělesa, jak je ukázáno na obr. 3.4 (vnitřek tělesa je vyznačen šrafováním).



Obr. 3.4: Konzistentní (vlevo) a nekonzistentní (vpravo) orientace obálek.

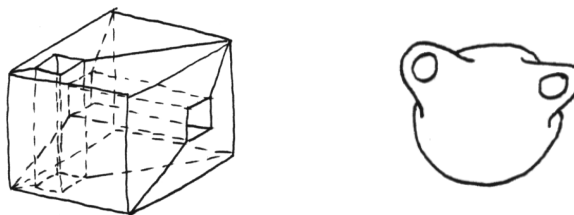
3.4 Eulerova formule

Při praktické implementaci modelerů bývá zapotřebí v průběhu výpočtů kontrolovat topologickou správnost těles. Ke kontrole lze použít Eulerovu formuli, která je nutnou (bohužel však ale nikoli postačující) podmínkou správnosti. Tato formule říká, že počet stěn, hran a vrcholů tělesa je vázán jistým vztahem. Uvažujme nejprve tělesa, která mají jediný povrch, neprochází jimi žádné díry, neobsahují žádné vnitřní dutiny a každá jejich stěna je homeomorfní s kruhem. Taková tělesa jsou z topologického pohledu nejjednodušší - jsou všechna homeomorfní s koulí. Jestliže označíme počet vrcholů V , počet hran E , počet stěn F , pak lze Eulerovu formuli pro tato tělesa zapsat ve tvaru

$$V + F = E + 2. \quad (3.1)$$

Uvažujme dále možnost, že tělesa mají díry (nikoli ale vnitřní dutiny), ale že jsou stále ohraničena jediným souvislým povrchem. Stále také předpokládáme, že každá stěna tělesa je homeomorfní s kruhem. Příklad tělesa se dvěma otvory je na obr. 3.5. Tělesa tohoto typu jsou homeomorfní s koulí s jistým počtem „držadel“ (na obr. 3.5 je počet držadel 2). Počet držadel se nazývá genus. Pro tělesa, jejichž genus je G , má Eulerova formule tvar

$$V + F = E + 2(1 - G). \quad (3.2)$$

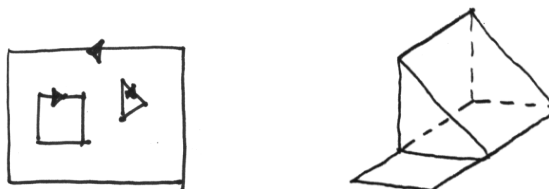


Obr. 3.5: Těleso ve tvaru krychle se dvěma průběžnými dírami (*vlevo*) je homeomorfní s koulí se dvěma „držadly“ (*vpravo*); počet držadel udává hodnota nazývaná genus, která je v tomto případě 2. (Stěny krychle jsou zde záměrně rozděleny diagonálami, aby byly ohraničeny vždy jen jedinou smyčkou. Jinak bychom museli použít složitější tvar Eulerovy formule, o němž je pojednáno dále v textu.)

Dále ještě navíc uvažujme tělesa s vnitřními dutinami. Taková tělesa jsou ohraničena několika oddělenými uzavřenými povrchy. Počet povrchů označme S . Konečně též upuštěme od požadavku, že každá stěna tělesa je homeomorfní s kruhem (tj., že každá stěna je ohraničena jedinou smyčkou hran; na obr. 3.6 je znázorněna stěna ohraničená třemi smyčkami). Označme L celkový počet smyček na plochách tělesa a G součet genů jednotlivých povrchů. Eulerova formule má pak tvar

$$V + F = E + (L - F) + 2(S - G). \quad (3.3)$$

Ačkoliv každý povrch typu 2-manifold v E^3 musí splňovat Eulerovu formuli, opak neplatí. Ne každý povrch, který Eulerovu formuli splňuje, je povrchem typu 2-manifold v E^3 . Obr. 3.6 ukazuje takový případ.

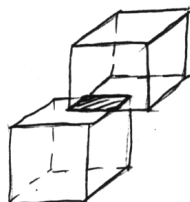


Obr. 3.6: Stěna ohraničená více (třemi) smyčkami (*vlevo*) a těleso (*vpravo*), které sice splňuje Eulerovu formuli ($F = 6$, $V = 8$, $E = 12$), ale ve smyslu předchozích definic tělesem není kvůli „trčící“ stěně (také intuitivně určitě máme pochybnosti, zda je těleso v pořádku).

3.5 Regularizované booleovské operace

Ať se pro modelování těles používá jakéhokoli přístupu, bývá často požadována možnost vytvářet nová tělesa aplikací booleovských operací na tělesa již existující.

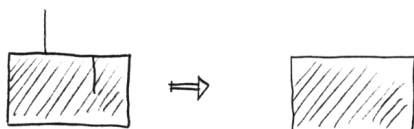
Zpravidla bývá v objemových modelářích k dispozici operace součtu, součinu a rozdílu (\cup, \cap, \setminus). Předpokládejme nejprve, že tělesa chápeme jako množiny bodů a že zmíněné operace mají význam součtu, součinu nebo rozdílu množin. Pak např. při provádění operace součinu nad dvěma navzájem se dotýkajícími krychlemi, je výsledkem operace „těleso“ tvořené pouze společnou částí stěn obou krychlí (obr. 3.7; předpokládáme, že taková společná „vrstva“ bodů existuje). Podobně by bylo možné snadno ukázat případy, kdy je výsledkem operace úsečka nebo dokonce jediný bod. Taková tělesa však chápeme jako nepatřičná, a to jednak ve smyslu obsahu předchozí podkapitoly, ale také intuitivně. Při implementacích objemových modelářů se tato degenerovaná tělesa často vylučují. Požaduje se, aby tělesa vždy měla nenulový objem.



Obr. 3.7: Vznik útvaru, které není tělesem; průnikem dvou těles vznikla plocha.

Zavádí se proto pojem *regularizovaná množina* a *regularizovaná booleovská operace*. Regularizace množiny (označme ji A) je definována jako uzávěr vnitřku množiny podle předpisu $\text{closure}(\text{interior}(A))$. Množina, která se rovná své vlastní regularizaci, se nazývá regulární. Účinek regularizace je ilustrován na obr. 3.8. Zavedením regularizace booleovských operací se sleduje, aby výsledkem operací byly pouze regulární množiny. Regularizované operace zde značíme $\cup^*, \cap^*, \setminus^*$. Regularizovaná operace op^* (op je některou z operací $\{\cup, \cap, \setminus\}$) je definována předpisem

$$A op^* B = \text{closure}(\text{interior}(A op B)). \quad (3.4)$$

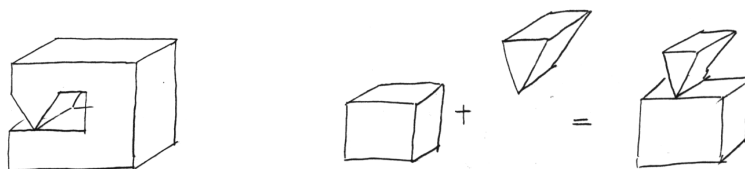


Obr. 3.8: Regularizace tělesa (množiny): nepatřičné stěny trčící vně nebo dovnitř tělesa jsou regularizací odstraněny.

3.6 Repräsentace tělesa pomocí hranice

Hraniční reprezentace tělesa (boundary representation, B-rept) používá k popisu tělesa jeho hranice (povrchu). Jestliže je hranice např. 2-manifoldem bez hranice v E^3 , pak rozděluje prostor E^3 na dvě části: vnitřek a vnějšek tělesa. Hranice je orientovaná, což umožňuje rozlišit, na které straně je vnitřek a na které vnějšek. Popis tělesa má dvě části: topologickou a geometrickou. Zatímco v topologické části jsou popisovány vzájemné vztahy (sousednost, incidence) mezi stěnami, hranami a vrcholy, geometrický popis popisuje vestavění uvedených entit do prostoru (např. souřadnice vrcholů, rovnice ploch). Podrobnosti uvedeme později.

Historicky se metody hraniční reprezentace vyvíjely z popisu mnohostěnu. Mnoho nejstarších objemových modelářů pracujících s hraniční reprezentací požadovalo hranici typu 2-manifold v E^3 . Nevýhodou bylo, že tak byla vyloučena některá tělesa, která by jinak konstruktér možná rád použil, např. tělesa, která se sama sebe dotýkají (obr. 3.9 vlevo). Na druhou stranu byly ale pro uvedenou třídu těles velmi dobře prozkoumány jejich topologické vlastnosti, takže bylo možné čerpat z bohaté matematické teorie. Později byl koncept téměř výhradního používání hranice typu 2-manifold revidován. Mimo jiné nepochybně také proto, že regularizovaná booleovská operace nad 2-manifolds může vést k výsledku, který 2-manifoldem není (obr. 3.9 vpravo).



Obr. 3.9: Prakticky snad přijatelné těleso (vlevo), jehož hranice ale není 2-manifoldem v E^3 ; těleso, jehož hranice není 2-manifoldem, může bohužel vzniknout množinovou operací s tělesy, která tuto podmínku splňují (vpravo).

V historii konstrukce objemových modelářů využívajících hraniční reprezentace lze tak vysledovat tři přístupy: 1) Tělesa musí mít hranici typu 2-manifold v E^3 . Pokud také výsledek booleovské operace nesplňuje tuto podmínku, dojde k chybě. 2) Tělesa jsou sice topologicky popsána tak, jako by se jednalo o objekty s hranicí typu 2-manifold v E^3 , avšak jejich vestavění do prostoru dovoluje geometrickou koincidence entit (vrcholy, hrany, plochy) hranice, která ale není zachycena v topologickém popisu. 3) Jsou plně povoleny objekty, jejichž hranice není 2-manifold v E^3 (jako např. na obr. 3.9 vlevo). Je očividné, že přístup 2) není zcela korektní a že může vést k některým potížím s robustností algoritmů. Ošetření geometrických koincidenčí je pracné a problematické.



Příklad 3.2. Vysvětlete, proč hranice tělesa na obr. 3.9 vlevo nebo hranice výsledného tělesa na témže obrázku vpravo není 2-manifoldem v E^3 . Návod: Soustředte se na body na hraně ve „špičaté“ části tělesa a uvažujte jejich okolí. Je toto okolí homeomorfní s kruhem? Porovnejte s body a jejich okolím na jiných hranách.

3.6.1 Implementace hraniční reprezentace

Dále popíšeme, jak lze popis tělesa pomocí hranice realizovat prakticky. Pro jednoduchost se omezíme na mnohostěny (tělesa s rovinnými stěnami). Víme však již ale, že toto omezení není zásadní. Mezi tělesy s rovinnými stěnami a tělesy se zakřivenými stěnami existuje topologická ekvivalence. O tělesech s rovinnými stěnami se nám ale bude o něco lépe uvažovat.

Principiálně je popis tělesa tvořen seznamem povrchů (shell), které těleso omezuji. Každý povrch je popsán seznamem jeho stěn (face). Každá stěna je popsána pomocí smyček (loop), které ji obklopují. Každá smyčka je popsána seznamem hran (edge), které ji vytvářejí. Poloha každé hrany je určena dvěma vrcholy.

Při úvahách o konkrétní podobě datové struktury tělesa je samozřejmě nutné vzít v úvahu také požadavky operací, které budou s tělesem následně prováděny. Jedná se zejména o booleovské operace. Jak uvidíme později, budou od datové struktury vyžadovány zejména odpovědi na následující otázky: 1) Pro každý vrchol určit hrany a stěny, které s vrcholem incidují. 2) Pro každou hranu určit, se kterými vrcholy inciduje. Určit také, které stěny incidují s danou hranou (v tělesech s hranicí 2-manifold v E^3 přiléhají k hraně pouze dvě stěny; obecně se ale může jednat o více stěn). 3) Pro každou stěnu určit hrany, které ji omezují. Hrany jsou organizovány do smyček. Pro stěny bez otvorů je jedna smyčka; pro stěny s otvory je více smyček. Směr, ve kterém jsou hrany a vrcholy ve smyčce procházeny, určuje normálu stěny, a tedy vnitřek a vnějšek tělesa. Ne všechny uvedené informace musí být ovšem v datové struktuře zaznamenány explicitně. Je však žádoucí, aby je bylo možné poměrně rychle dopočítat.

Dalším hlediskem při návrhu datové struktury je hledisko robustnosti operací, které jsou nad strukturou prováděny. Při implementaci booleovských operací se např. ukázalo, že splnit požadavek robustnosti odpovídajících algoritmů je dosti obtížné. Z hlediska robustnosti je např. žádoucí, aby v datové struktuře nebyly informace uchovávány duplicitně, jinak hrozí nebezpečí vzniku nekonzistencí, které mohou vzniknout zejména vlivem numerických nepřesností při výpočtech. Jednoduše řečeno: Může se stát, že duplicitní informace popisují dvě různá tělesa místo toho, aby popisovaly těleso jediné. Potíže by pak nastaly tehdy, jestliže by některý z algoritmů současně použil informace, které si vzájemně odporují. Jako příklad uvedme, že lze např. upustit od explicitního uchovávání souřadnic vrcholů tělesa a rovnic hran, jsou-li známy rovnice stěn. Souřadnice vrcholů a rovnice hran lze vždy podle potřeby a v odpovídající přesnosti dopočítat (tuto variantu bychom mohli použít např. u těles ohraničených obecnými plochami). Podobně naopak u mnohostěnů postačí uchovávat souřadnice vrcholů. Rovnice hran a stěn mohou být vždy podle

potřeby a v odpovídající přesnosti dopočítány.

Jednu z možných variant konkrétní implementace hraniční reprezentace navrhl již na počátku sedmdesátých let Baumgart. V jím navržené datové struktuře je nejvíce topologických informací nesenou hranou (obr. 3.10). Pro datovou strukturu se vžil název okřídlená hrana (winged edge; k hraně přiléhají dvě plochy, což připomíná křídla, viz. též obr. 3.10). Ani po mnoha letech nepozbyla tato datová struktura na aktuálnosti. Konkrétní příklad definice datové struktury pro popis tělesa uvádíme v následujícím fragmentu zdrojového textu. Text je maximálně zjednodušen. Dovolena jsou zde pouze tělesa s jediným povrchem a plochy s jedinou smyčkou. (Slovní popis prezentovaný v úvodu této podkapitoly byl o něco obecnější.)

```
//*****
//*          Popis telesa metodou hranicni reprezentace          *
//*****

typedef struct{float x, y, z;} Coord3; // Souradnice v E3

struct Vertex {          // ***   Vrchol telesa   ***
    Vertex *next;        // Dalsi vrchol v telese
    Coord3 coord;        // Souradnice vrcholu
};

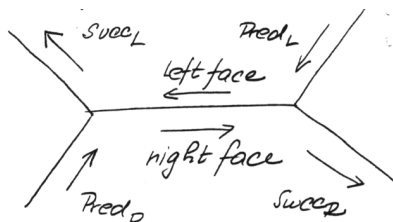
struct Edge {           // ***   Hrana telesa   ***
    Edge *pred1, *succ1; // Predch. a nasl. hrana plochy 1
    Edge *pred2, *succ2; // Predch. a nasl. hrana plochy 2
    Face *face1, *face2; // Prvni a druha prilehajici face
    Vertex *vert1, *vert2; // Vrcholy hrany
};

struct Face {           //***   Stena telesa   ***
    Face *next;         // Dalsi stena tehoz telesa
    Edge *edge;         // Prvni hrana smycky
};

struct Solid {          // ***   Teleso B rept   ***
    Face *face;         // Prvni stena povrchu
    Vertex *vertex;     // Prvni vrchol telesa
};
//*****
```

Za povšimnutí stojí, že struktura je použitelná pouze pro tělesa, jejichž povrch je 2-manifold v E^3 . Stojí ale na druhé straně rovněž za zmínku, že popisovaná struktura nachází uplatnění nejen při modelování těles, ale i jinde, např. při popisu planárních map. Pozorný čtenář jistě dokáže obojí zdůvodnit. V průběhu času se také objevily

různé modifikace, které sledovaly určitá zlepšení při provádění některých operací s modelem tělesa. Objevila se též rozšíření pro tělesa, jejichž povrchy nejsou 2-manifold v E^3 . Systematický popis těchto modifikací bohužel ale přesahuje rámec tohoto textu.



Obr. 3.10: Okřídlená hrana (vodorovná hrana) a k ní přiléhající dvě plochy („křídla“). Hrana je součástí dvou smyček hran, které jsou na obrázku vyznačeny šipkami. Pospojování hran do smyček je zajištěno pointery na předchozí a následující hranu podél každé z obou ploch, jak plyne ze zdrojového kódu.

Příklad 3.3. Rozšiřte dříve uvedenou ukázkou zdrojového kódu tak, aby těleso mohlo mít více navzájem nesouvisejících povrchů (např. těleso s dutinou uvnitř) a dále ještě tak, aby každá stěna mohla být ohraničena více než jednou smyčkou (to je zapotřebí, pokud mají být ve stěnách otvory, jako je tomu např. na obr. 3.6 vlevo).



Příklad 3.4. Navrhněte, jak by bylo možné dříve uvedenou datovou strukturu rozšířit tak, aby mohla popisovat také tělesa, jejichž hranice není 2-manifoldem v E^3 , jako je tomu např. na obr. 3.9. Návod: Musí být možné, aby k jedné hraně přiléhaly více než dvě plochy. Obecně není počet přiléhajících ploch omezen. Za jistých okolností (jakých?) bychom mohli usuzovat, že by měl být sudý.



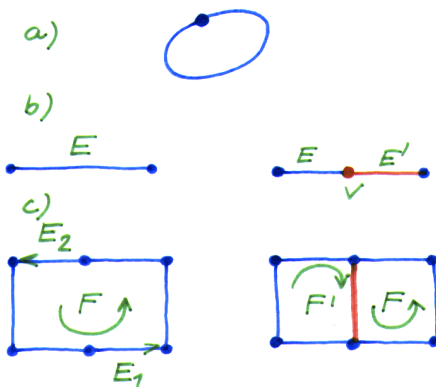
3.6.2 Eulerovy operátory

K obsluze datových struktur hraniční reprezentace se často používají tzv. Eulerovy operátory. Bylo tomu tak už i v minulosti, kdy např. ještě nebyla známa technika objektového programování. To proto, že přímá a bezchybná obsluha datové struktury i v její nejjednodušší zde prezentované variantě by byla bez operátorů téměř nemožná. Použití Eulerových operátorů je výhodné, protože do značné míry odstraňuje nutnost zabývat se konkrétní realizací datových struktur při implementaci operací na vyšší úrovni. Eulerovy operátory modifikují datovou strukturu tak, aby vždy popisovala nějaké těleso, pro které je splněna Eulerova formule (třebaže už víme, že ne všechna tělesa splňující Eulerovu formuli jsou tělesa, ke kterým bychom neměli výhrady).

Uvedme několik konkrétních příkladů možných operátorů. Za operátory nejnižší úrovně lze považovat např. následující operátory: Operátor *createSolid* vytvoří minimální (ve smyslu Eulerovy formule) neprázdné těleso, které obsahuje jeden vrchol, jednu hranu a dvě plochy (obě plochy přiléhají k jediné hraně a leží tedy geometricky zatím „na sobě“) (obr. 3.11a); operátor *splitEdge* „rozštípne“ zadanou hranu vložením nového vrcholu (obr. 3.11b); operátor *splitFace* „rozštípne“ zadanou stěnu vložením nové hrany (obr. 3.11c). (Poznamenejme, že jsme se zde při pojmenovávání operátorů nedrželi tradice a nepoužili jsme jména ve tvaru *mkxy*). Deklarace uvedených operátorů by mohla být následující:

```
Solid *createSolid(Vertex *v);
Edge *splitEdge(Edge *e, Vertex *v);
Face *splitFace(Face *f, Edge *e1, Edge *e2);
```

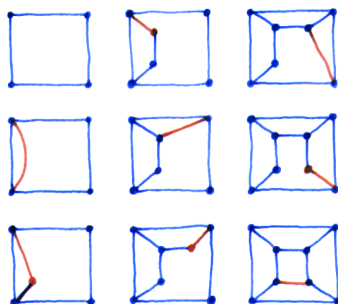
Všechny uvedené parametry operátorů jsou vstupní: **v* je pointer na místo v paměti, v němž buď už jsou nebo někdy později budou souřadnice vrcholu, **e* je pointer na hranu, která má být rozštípnuta, **e₁*, **e₂* jsou pointery na hrany, které svými počátky určují polohu, kam má být vložena nová hrana při rozštípnutí plochy (obr. 3.11c). Každý operátor vrací pointer na nový útvar, který vytvořil.



Obr. 3.11: Eulerovy operátory *createSolid*, *splitEdge* a *splitFace* (viz text)

Jako příklad aplikace uvedených operátorů ukážeme postup, při němž je vytvořen kvádr. Proces začíná použitím operátoru *createSolid* a třikrát opakovaným použitím operátoru *splitEdge*. Další postup je již znázorněn a vysvětlen na obr. 3.12.

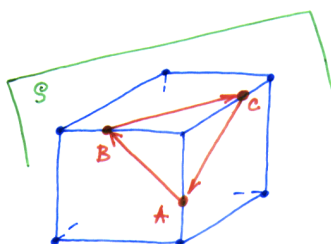
Jiným příkladem aplikace uvedených operátorů je vepsání řezového polygonu do tělesa. Mějme těleso a rovinu. Pro jednoduchost předpokládejme, že těleso má jeden povrch a že každá stěna tělesa je ohraničena jedinou konvexní smyčkou hran. (Prosíme čtenáře, aby po přečtení odstavce promyslel, jaké potíže by nastaly, kdyby smyčka nebyla konvexní.) Chceme najít řezový polygon a vepsat jej do modelu (obr. 3.13). Postup řešení je následující: Probíráme všechny stěny tělesa tak dlouho, až nalezneme libovolnou stěnu, která protíná zadanou řezovou rovinu. Některé dvě



Obr. 3.12: Vytvoření kvádrů: Nejprve je aplikován operátor *createSolid*, pak třikrát operátor *splitEdge*. Výsledek této akce je výchozím obrázkem vlevo nahoře. Dále se pak střídá použití operátorů *splitFace* a *splitEdge*, dokud není vytvořena taková struktura vrcholů, hran a stěn, jak vyžaduje kvádr (obrázek si prohlížejte po sloupcích). K tomu, abychom skutečně obdrželi kvádr, je nutné ještě zadat odpovídající souřadnice vrcholů, ale to je již jednoduché (stejným postupem bychom vytvořili také obecný šestistěn - rozdíl je jen v souřadnicích, které zadáme).

hrany této stěny protínají řezovou rovinu v bodech, které se v modelu stanou vrcholy (označme tyto body *A*, *B*). Rozštípíme hrany, na nichž body *A*, *B* leží. Nalezenou stěnu také rozštípíme, a to vložením nové hrany spojující vrcholy *A*, *B*. Ve vytváření řezové křivky pokračujeme přechodem do sousední stěny, která inciduje s vrcholem *B* (v diskutované datové struktuře je takový přechod snadný). Postupným prohledáváním hran ležících na obvodu této stěny nalezneme vrchol *C*. Rozštípíme hranu, na níž bod *C* leží a plochu rozštípíme hranou spojující vrcholy *B*, *C*. Podobně pokračujeme tak dlouho, dokud se řezová křivka neuzavře ve vrcholu *A*. Deklarace popsání operátoru může být následující:

```
EdgeList *inscribeCuttingPolygon(Solid *s, PlaneEq *plane);
```

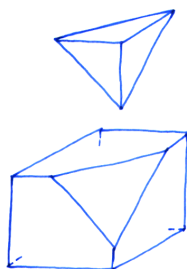


Obr. 3.13: Vepsání řezového polygonu (viz text)

Po provedení operátoru *inscribeCuttingPolygon* je řezový polygon vepsán. Funce navíc vrací seznam hran, které byly vepsány (řezový polygon). Tohoto seznamu lze

využít zejména při provádění řezu tělesa pomocí operátoru *splitSolid*, který těleso podél řezové křivky rozdělí (obr. 3.14). Operátor *splitSolid* může být deklarován takto:

```
Solid *splitSolid(Solid *s, EdgeList *cuttingPolygon);
```



Obr. 3.14: Operátor *splitSolid*



Příklad 3.5. Implementujte operátory *createSolid*, *splitEdge*, *splitFace* a *inscribeCuttingPolygon* pro datovou strukturu uvedenou v této podkapitole.



Příklad 3.6. Pro datovou strukturu uvedenou v této podkapitole implementujte funkci, která zjistí, zda bod o zadaných souřadnicích leží uvnitř či vně tělesa, které datová struktura popisuje. Návod: Ze zadaného bodu vedte libovolnou polopřímku a spočítejte, kolikrát protíná povrch tělesa. Třebaže je princip řešení jednoduchý, je detailní implementace s ošetřením všech speciálních případů, ke kterým může při výpočtu dojít, poměrně pracná.

3.6.3 Provádění booleovských operací

Provádění booleovských operací s tělesy je při hraniční reprezentaci těles obávaným implementačním problémem. Pokud by snad čtenář byl někdy postaven před problém jej řešit, doporučujeme velmi dobře tento krok uvážit. Třebaže se může zdát situace z teoretického pohledu jasná, je praktická a přiměřeně bezchybná implementace velmi obtížná. I renomovaným firmám trvalo poměrně dlouho, než se dopracovaly ke stabilním implementacím.

Při nástupu řešení se zde omezíme na operace mezi takovými tělesy, jejichž povrch je tvořen vždy jen jedinou obálkou. To je možné, protože problém operací s tělesy s více obávkami lze na tento problém, který lze považovat za základní, vždy převést. Další zjednodušení bude spočívat v redukci repertoáru operací, které budeme uvažovat. Třebaže v systémech pro modelování těles bývají obvykle k dispozici operace \cap^* , \cup^* a \setminus^* , není zapotřebí implementovat všechny. Uvážíme-li, že je možné snadno

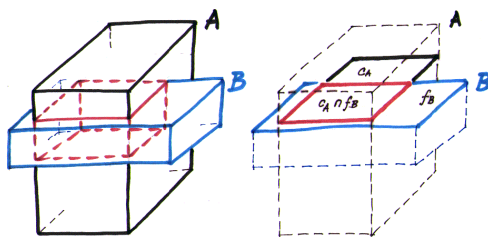
realizovat doplněk tělesa, pak je zřejmé, že postačí implementovat jednu z uvedených operací, např. operaci \cap^* . Ostatní operace pak snadno získáme pomocí operace \cap^* a doplňku těles (je označen pomocí \neg)

$$A \cup^* B = \neg(\neg A \cap^* \neg B), \quad (3.5)$$

$$A \setminus^* B = A \cap^* \neg B. \quad (3.6)$$

Vidíme tedy, že se můžeme omezit pouze na operaci regularizovaného součinu. Vytvoření doplňku je jednoduché. Jak jsme vysvětlili již dříve, je zapotřebí obrátit orientaci normál všech stěn tělesa. To se provede otočením orientace (pořadí, v němž jsou procházeny vrcholy) všech smyček na všech stěnách. Poznamenejme, že zavedením doplňku získáváme možnost pracovat s nekonečnými tělesy, jejichž povrch však musí být konečný.

Princip řešení úlohy regularizovaného průniku dvou těles A , B , každé s jedním povrchem, lze popsat takto: 1) Urči, zda se některá dvojice stěn $f \in A$ a $g \in B$ protíná. Jestliže se neprotíná žádná dvojice stěn, pak proved pouze test, zda povrchy těles leží jeden uvnitř druhého a zbývající kroky přeskoč (výsledkem operace je pak buď „menší“ těleso nebo těleso prázdné). 2) Pro každou stěnu f tělesa A , která protíná těleso B , proved řez tělesa B plochou stěny f . Nalezenou řezovou oblast omez na tu část, která současně leží uvnitř stěny f . 3) Pro každou stěnu g tělesa B , která protíná těleso A , proved řez tělesa A plochou stěny g . Nalezenou řezovou oblast omez na tu část, která současně leží uvnitř stěny g (obr. 3.15). 4) Sestav hranici výsledného tělesa $A \cap^* B$. Hledaná hranice bude obsahovat ty části hranice tělesa A , které leží uvnitř tělesa B , ty části hranice tělesa B , které leží uvnitř tělesa A a ty části společné hranice těles A a B , kde je normála hranic obou těles souhlasně orientována. Tyto části byly určeny v krocích 2), 3).



Obr. 3.15: K realizaci operace průniku (popis je uveden v textu)

Přestože popsany postup nevypadá obzvláště komplikovaně, není jeho implementace vůbec jednoduchou záležitostí. Závažným problémem je dosažení robustnosti algoritmu. Jaké potíže zde v tomto ohledu vznikají, osvětlíme na příkladě. Předpokládejme, že chceme najít průsečíky přímky s krychlí dle obr. 3.16. Při tom se může např. stát, že průsečík přímky s rovinou stěny f_1 je vyhodnocen jako bod na hraně stěny f_1 , ale naproti tomu průsečík s rovinou stěny f_2 je vyhodnocen jako bod uvnitř stěny f_2 (obr. 3.16). Tato nekonzistence je způsobena numerickými nepřesnostmi,

kteřé nikdy nemohou být zcela odstraněny, a bez dalších opatření způsobí algoritmu, v němž by k takové situaci došlo, vážné potíže. Robustní algoritmus se musí vyhnout vzniku nekonzistencí. Jednou zjištěná informace musí být „rozeslána“ do všech míst, kde je jí zapotřebí. V našem příkladě to znamená, že jestliže je zjištěno, že průsečík přímky s rovinou stěny f_1 leží na hraně stěny f_1 , pak také průsečík s f_2 nutně musí ležet na hraně stěny f_2 (hrana je společná), a proto se průsečík s f_2 již nepočítá; polohu na společné hraně je zapotřebí vynutit. Je velmi obtížné všechny možné podobné situace při implementaci odhalit a ošetřit. Obtížnost situace je ilustrována skutečností, že bod může vzhledem k tělesu zaujímat následující polohy: uvnitř tělesa, na ploše stěny, na hraně a ve vrcholu. Uvedené čtyři možnosti pro jedno těleso dávají vzniknout šestnácti případům, jsou-li tělesa dvě.



Obr. 3.16: K problémům vznikajícím při implementaci množinových operací



Příklad 3.7. Pro datovou strukturu uvedenou v této podkapitole implementujte operaci doplněk tělesa.

3.6.4 Zobrazení tělesa popsaného hraniční reprezentací

Zobrazení tělesa popsaného hraniční reprezentací nečiní potíže. V každém okamžiku je totiž k dispozici popis hranice, kterou je třeba zobrazit. Stačí z něho pouze přečíst potřebné údaje. Tak např. při zobrazení pomocí tzv. drátěného modelu stačí při průchodu datovou strukturou získat všechny hrany (tj. úsečky) a ty vykreslit. Je-li požadováno zobrazení s řešením viditelnosti, pak se při průchodu datovou strukturou zjistí všechny plochy tělesa. Po vyřešení viditelnosti standardními metodami se viditelné plochy zobrazí.



Příklad 3.8. Pro datovou strukturu uvedenou v této podkapitole implementujte zobrazení tělesa. (Tento úkol je ale poněkud pracnější.)

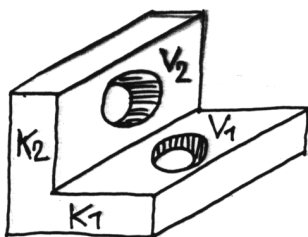
3.7 CSG reprezentace tělesa

Computer Solid Geometry (CSG) je další často používanou metodou reprezentace těles. Současně je tímto názvem také charakterizována jistá metodologie vytváření

těles. Tělesa jsou vytvářena za pomoci standardních primitivních těles, regularizovaných booleovských operací a geometrických transformací. Standardními primitivními tělesy bývají zpravidla kvádr, hranol, koule, válec, kužel, toroid atd. Uživatel vytváří instance primitivních těles zadáním požadovaných rozměrů a jejich umístěním v prostoru. Regularizovanými booleovskými operacemi bývají i zde regularizovaný součin, součet a rozdíl (\cap^* , \cup^* , \setminus^*). Zdá se logické připustit, aby se booleovských operací mohly kromě primitivních těles zúčastňovat také tělesa uživatelem již dříve vytvořená. Jako příklad CSG popisu uveďme těleso dle obr. 3.17. Popis tohoto tělesa by bylo možné zapsat ve tvaru

$$[\text{translate}_y(\text{block}(6, 2, 1), 1) \setminus^* \text{translate}_x(\text{translate}_y(\text{cylinder}_z(1, 1), 2), 3)] \\ \cup^* [\text{block}(6, 1, 3) \setminus^* \text{translate}_x(\text{translate}_z(\text{rotate}_x(\text{cylinder}_z(1, 1), -90), 2), 3)],$$

kde translate_x znamená translaci tělesa ve směru osy x o danou vzdálenost, rotate_x znamená rotaci kolem osy x o daný úhel, cylinder_z značí válec s osou totožnou s osou z s daným poloměrem a výškou, block je kvádr s vrcholem v počátku a s danými rozměry ve směru jednotlivých os. Na obr. 3.17 jde osa x dopředu a osa y zleva do prava.



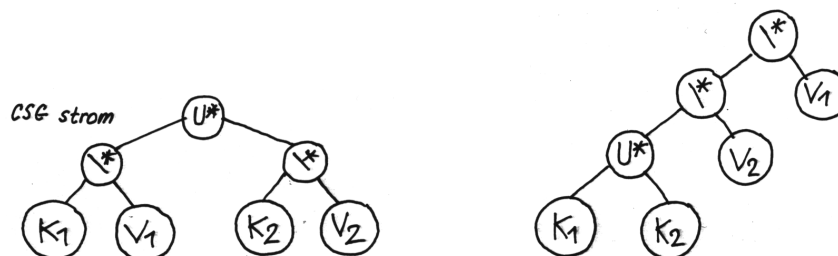
Obr. 3.17: K modelování tělesa technikou CSG

V tomto případě jsme těleso vytvořili spojením dvou kvádrů, do nichž jsme předem „vyvrtali“ otvory odečtením válců umístěných do správné polohy. Operace, které bylo pro vytvoření tělesa nutné provést, lze znázornit pomocí CSG stromu (obr. 3.18 vlevo). Poznamenejme, že k témuž tělesu lze dospět různými postupy. Např. naše těleso je možné vytvořit také postupem, kdy nejprve spojíme oba kvádry a pak teprve „vrtáme“ díry, což lze zapsat takto

$$[\text{translate}_y(\text{block}(6, 2, 1), 1) \cup^* \text{block}(6, 1, 3)] \\ \setminus^* \text{translate}_x(\text{translate}_z(\text{rotate}_x(\text{cylinder}_z(1, 1), -90), 2), 3) \\ \setminus^* \text{translate}_x(\text{translate}_y(\text{cylinder}_z(1, 1), 2), 3).$$

Tomuto postupu odpovídá CSG strom dle obr. 3.18 vpravo. Teoreticky jsou oba postupy rovnocenné a bylo by možné vymyslet ještě postupy další. Při spojování obou kvádrů nebylo dále např. zapotřebí, aby se právě dotýkaly. Mohly se i překrývat. Operace sjednocení by překryv odstranila. V praxi je při vytváření možné

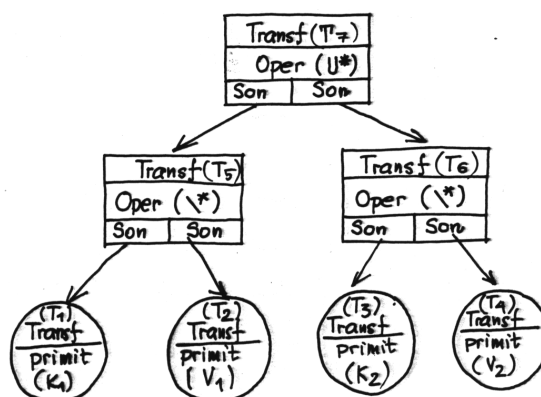
dát přednost např. tomu postupu, který více odpovídá předpokládané technologii výroby tělesa.



Obr. 3.18: Stromy popisující vytvoření tělesa různými postupy popsány v textu

3.7.1 Realizace CSG reprezentace

Nyní již můžeme přistoupit k příkladu datové struktury pro modelování tělesa metodou CSG. Ve shodě s obr. 3.18 je datovou strukturou strom (obr. 3.19). Vnitřní uzly stromu odpovídají booleovským operacím a transformacím. Listy odpovídají transformacím a primitivním tělesům. Chápeme-li booleovské operace jako binární, pak je datová struktura binárním stromem. Zobecníme-li booleovskou operaci pro n argumentů, pak je datová struktura stromem n -árním. Stojí za povšimnutí, že CSG strom a jemu odpovídající datová struktura uchovávají také informaci o historii modelování (vytváření) tělesa. To je významná vlastnost, ze které se těží např. při pozdějších modifikacích tělesa (již vytvořená tělesa lze snadno editovat). V hraniční reprezentaci tato informace k dispozici není.



Obr. 3.19: Počítačová realizace CSG modelu: V listech jsou primitivní tělesa a jejich transformace; v uzlech operace a transformace; transformace v uzlu se využije pokud je s vytvořeným tělesem nějak manipulováno.

3.7.2 Provádění booleovských operací v reprezentaci CSG

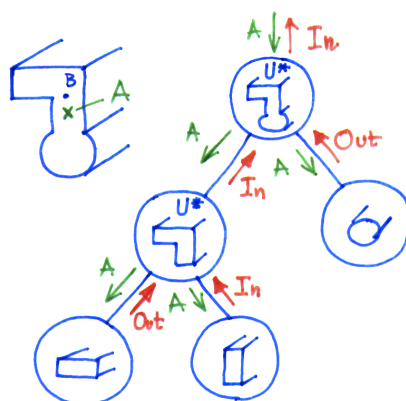
Jak vyplynulo z předchozího výkladu, je provedení booleovských operací v CSG reprezentaci velmi jednoduché. Stačí vytvořit CSG strom, v němž kořen odpovídá požadované booleovské operaci. CSG stromy, které odpovídají tělesům vstupujícím do booleovské operace, se stanou podstromy tohoto nově vytvořeného stromu. Jednoduchost provádění booleovských operací v reprezentaci CSG je způsobena tím, že CSG reprezentace explicitně neuchovává informaci o hranici tělesa. Jak víme z odstavce o hraniční reprezentaci, byl to právě výpočet hranice tělesa vzniklého booleovskou operací, který vyžadoval značného úsilí. Bohužel, ať už používá objemový modelář jakoukoli reprezentaci, je hranice tělesa důležitou a často požadovanou informací. Jestliže se zjištění hranice požaduje, pak je nutné všechny booleovské operace předepsané CSG stromem skutečně provést. To se děje procházením CSG stromu a realizací všech předepsaných operací podobně, jak to bylo naznačeno pro reprezentaci hraniční. Problémy s komplikovaným výpočtem hranice se tak pouze přesunou do jiného času a do jiných míst systému. Při řešení některých problémů je však možné se výpočtu hranice vyhnout. Ilustraci takového postupu lze nalézt v následujících odstavcích.

3.7.3 Určení vzájemné polohy bodu a CSG tělesa

Určení vzájemné polohy bodu a tělesa patří k základním úlohám při využití objemového modelu. Metoda, která se používá pro hraniční reprezentaci, spočívá v tom, že se zvolí libovolná polopřímka začínající/končící v klasifikovaném bodě a vyšetřují se její průsečíky s hranicí tělesa. Je-li počet průsečíků lichý, pak leží klasifikovaný bod uvnitř tělesa. Použití této metody v CSG modelu by vyžadovalo výpočet hranice, ale tomu se chceme pokud možno vyhnout. Ukážeme proto jiný postup, který je současně ilustrací možného řešení celé třídy podobných problémů.

Předpokládejme, že konstruujeme funkci, jejíž argumenty jsou CSG popis tělesa a souřadnice klasifikovaného bodu. Funkce vrací hodnoty In, Out, On (bod leží uvnitř, vně nebo na hranici tělesa). Předpokládejme, že by funkce mohla pracovat takto (obr. 3.20): 1) Je-li testované těleso primitivem (je reprezentováno listem CSG stromu), pak lze vzájemnou polohu tělesa a bodu jednoduše určit. 2) Není-li testované těleso primitivem (je reprezentováno vnitřním uzlem CSG stromu), pak funkce zjistí vzájemnou polohu bodu a těles reprezentovaných synovskými uzly vyšetřovaného uzlu (to může spustit rekurzivní provádění uvažované funkce). Vzájemná poloha bodu a tělesa reprezentovaného vyšetřovaným uzlem se určí podle pravidel plynoucích z operace, kterou uzel reprezentuje. Je-li např. bod uvnitř obou těles reprezentovaných synovskými uzly a je-li operací v uzlu např. průnik, bude bod ležet také uvnitř tělesa reprezentovaného vyšetřovaným uzlem. Myšlenka postupu je elegantní. Stanovení pravidel pro výslednou polohu vzhledem k uzlu však není tak jednoduché. Obtížné je zejména promýšlení všech speciálních případů, které mohou nastat. Řekněme, že vyšetřujeme např. operaci sjednocení a že vzhledem k objek-

tům korespondujícím se synovskými uzly leží uvažovaný bod na povrchu, a to vždy ve vrcholech. Vzhledem k výslednému tělesu může pak uvažovaný bod ležet v různých polohách: uvnitř, na povrchu uvnitř stěny, na povrchu na hraně a konečně také na povrchu ve vrcholu. (Prosíme laskavého čtenáře, aby si všechny uvedené varianty pokusil představit.) Pro správnou funkci algoritmu bude nezbytné rozpoznat, která z variant nastala. Podobně bude zapotřebí postupovat i pro jiné polohy bodu vzhledem k tělesům reprezentovaným synovskými uzly.

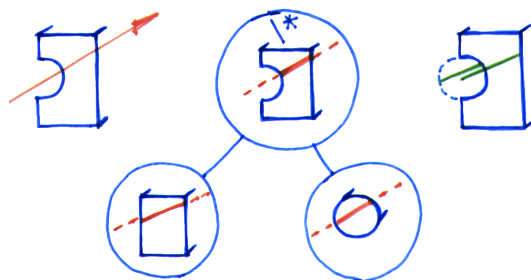


Obr. 3.20: Určování vzájemné polohy bodu a tělesa reprezentovaného CSG modelem: Bod A je uvnitř jednoho z primitivních těles. Je také uvnitř celého tělesa, protože při operaci sjednocení, která je zde k vytvoření tělesa použita, k tomu stačí, aby byl bod uvnitř alespoň jednoho z primitivních těles, která do operace vstupují.

3.7.4 Zobrazení tělesa reprezentovaného pomocí CSG

Zde jsou problémy do značné míry podobné těm, které jsme naznačili v předchozím odstavci. Nejpřímočařejším postupem je požadovat pro zobrazení tělesa výpočet jeho hranice. Tento postup se volí dosti často, a to i přesto, že implementace příslušných algoritmů je, jak již víme, dosti nepříjemná. Jako ukázkou postupu, kdy se hranice nepočítá, uvedeme příklad zobrazení CSG tělesa metodou sledování paprsku. Naznačíme princip funkce, která vypočítá průnik paprsku s daným tělesem, tedy tu část paprsku, která leží uvnitř tělesa (z ní bod nejbližší pozorovateli je zapotřebí zobrazit). Vstupními parametry funkce jsou paprsek (zadaný např. počátečním bodem a směrem) a CSG popis tělesa. Výstupem je seznam přímkových segmentů paprsku, které leží uvnitř zadaného tělesa. Funkce může pracovat takto: 1) Je-li testované těleso primitivem (listem CSG stromu), pak lze průnik paprsku s tělesem jednoduše určit. 2) Není-li testované těleso primitivem, ale vnitřním uzlem CSG stromu, pak se zjistí průnik paprsku s tělesy reprezentovanými synovskými uzly vyšetřovaného uzlu (to může spustit rekurzivní provádění uvažované funkce). Nad takto zjištěnými

průniky pro synovské se pak provede booleovská operace předepsaná ve vyšetřovaném uzlu. Protože se jedná o booleovskou operaci nad jednorozměrnou oblastí, lze ji snadno implementovat. Realizaci a činnost popsané funkce ukazuje následující fragment zdrojového textu a obr. 3.21.



Obr. 3.21: Přímé zobrazování CSG modelu (bez explicitního výpočtu hranice) metodou sledování paprsku. Vlevo je vyznačeno těleso a paprsek. Uprostřed je CSG strom pro těleso; vyznačeno je také řešení průsečíků paprsku s jednotlivými primitivami (kvádr a válec) i s výsledným tělesem. Vpravo je naznačeno, že úsečku, v níž paprsek protíná těleso, lze získat jako rozdíl úsečky, v níž paprsek protíná kvádr, a úsečky, v níž paprsek protíná válec. Rozdíl proto, že těleso bylo získáno odečtením válce od hranolu.

```
SgmtList *rayIntersec(Coord3 *orig, Coord3 *dir, CSGnode *node){
    if (isList(node)) return intersecPrimitive(orig, dir, node);
    sl:= rayIntersec(orig, dir, node >lSon);
    sr:= rayIntersec(orig, dir, node >rSon);
    return boolOperOnSgmtLists(sl, sr, node >operation);
}
```

Příklad 3.9. Pokuste se o implementaci zobrazení tělesa dle této podkapitoly. (Úkol je samozřejmě ale poněkud pracnější.)





Literatura

- [1] Budinský, B. – Kepr, B. *Základy diferenciální geometrie s technickými aplikacemi*. 2. vydání. Praha: Státní nakladatelství technické literatury, 1970. 344 s.
- [2] Budinský, B. *Analytická a diferenciální geometrie*. 1. vydání. Praha: Státní nakladatelství technické literatury, 1983. 296 s.
- [3] Ježek, F. *Diferenciální geometrie*. Skripta. Plzeň: ZČU, 2008. 58 s.
- [4] Martišek, D. *Matematické principy grafických systémů*. 1. vydání. Brno: Nakladatelství Littera. 2002. 296 s. ISBN 80-85763-19-2
- [5] Piegl, L. – Tiller, W. *The NURBS book*. 2. vydání. New York: Springer-Verlag New York, Inc. 1997. 660 s. ISBN 3-540-61545-8
- [6] Rogers, David F. *An introduction to NURBS: with historical perspective*. San Francisco: Morgan Kaufmann Publishers, Inc. 2001. 344 s. ISBN 1-55860-669-6
- [7] Saxena, A. – Sahay, B. *Computer Aided Engineering Design*. 2. vydání. The Netherlands: Springer. 2005. 387 s. ISBN 978-1-4020-2555-6
- [8] Žára, J. – Beneš, B. – Sochor, J. – Felkel, P. *Moderní počítačová grafika*. 2. vydání. Brno: Computer Press. 2004. 579 s. ISBN 80-251-0454-0

Rejstřík



- afinní prostor, 30
- afinní zobrazení, 30
- binormála, 40
- bod
 - inflexní, 39
 - regulární, 37
 - singulární, 37
- body
 - nevlastní, 7
- funkce
 - B-spline, 75
 - kubická spline, 57
- hlavní normála, 40
- homeomorfizmus, 100
- homogenní souřadnice, 30
- konzistentní orientace obálek, 105
- křivka
 - B-spline
 - neracionální, 81
 - racionální, 87
 - Bézierova, 59
 - kubika, 59
 - racionální, 66
 - explicitní, 33
 - Fergusonova
 - kubika, 55
 - implicitní, 33, 34
 - kubický spline, 57
 - křivost, 41
 - prostorová, 33
 - regulární, 33
 - rovinná, 33
 - spline
 - kubická, 57
 - tečna, 37
 - uzavřená, 33
 - křivost
 - druhá, 42
 - první, 41
 - matice
 - inverzní, 5
 - jednotková, 5
 - množina
 - regularizovaná, 107
 - omezená variace, 104
 - orientovatelnost, 103
 - parametr
 - transformace, 38
 - plocha, 44
 - B-spline
 - neracionální, 85
 - racionální, 90
 - tečná rovina, 45
 - polynomy
 - Bernsteinovy, 59
 - racionální, 66
 - projektivní prostor, 30
 - projektivní zobrazení, 30
 - prostor
 - projektivní, 7
 - topologický, 100
 - rovina
 - normálová, 40
 - oskulační, 39

- rektifikační, 40
- tečná, 39
- rovnice
 - explicitní, 34
 - implicitní, 34
 - parametrická, 34
- skalární součin, 5
- součin
 - tenzorový, 84
- torze, 42
- transformace
 - ortonormální, 4
 - projektivní, 7
- vektor
 - hlavní normály, 40
 - tečný, 36
 - jednotkový, 37
- vektory
 - bázové, 3
- vnitřek tělesa, 104