

Kapitola 6

Linearita v aplikacích aneb lineární algebra do třetice

Potřetí, a v tomto dílu naposledy, se vracíme k lineární algebře. Spojíme-li následující poznatky se znalostmi z předchozích dvou kapitol o lineární algebře, budeme moci říci, že jsme se seznámili se základy této matematické disciplíny. Naše setkávání s ní tím však zdaleka nekončí. Každý, kdo se bude zabývat fyzikou, ať již teoretickou nebo experimentální, a zejména pak jejími praktickými aplikacemi v inženýrství, přístrojové technice, medicínských oborech vyžadujících znalost fyziky, apod., bude potřebovat lineární algebru neustále. V této kapitole obohatíme náš dosavadní vektorový prostor o další výbavu — skalární součin, se kterou se ještě jednou vrátíme k problému projekcí, tentokrát v geometricky názornější podobě. A nakonec si všimneme lineárních operátorů, které vůči skalárnímu součinu splňují jisté „zákony zachování“. Tyto speciální vlastnosti, kterými většinou oplývají fyzikální „operátorové“ veličiny, zajistí existenci báze tvořené vlastními vektory každého takového operátoru.

6.1 Skalární součin — nadstandardní výbava vektorového prostoru

Se skalárním součinem vektorů jsme se již setkali, vzpomínáte? V každém případě se o tom můžete přesvědčit nalistováním odstavce 1.4.3. Situace, již jsme se tehdy zabývali, byla však velmi speciální. Pracovali jsme s vázanými vektory definovanými jako orientované úsečky v trojrozměrném euklidovském prostoru, které jsme sčítali a násobili skalárem pomocí jednoduchých geometrických definic — grafického sčítání orientovaných úseček a jejich násobení číslem, také v podstatě grafického. Volné vektory pak byly množinami ekvivalentních (tj. stejně dlouhých a stejně orientovaných) úseček, operace sčítání a násobení skalárem byly přirozeným způsobem zobecněny i na volné vektory. Prostor takových vektorů samozřejmě splňoval všechny požadavky obecné algebraické definice vektorového prostoru. Obsahoval však i něco navíc — možnost měření délek vektorů a úhlů mezi nimi, vyplývající z použití euklidovského prostoru pro vytvoření konkrétní geometrické představy o vektorech. Pomocí délek vektorů a úhlu mezi nimi jsme pak definovali skalární součin vektorů a zjistili jsme, jaké má tato operace vlastnosti.

V tomto odstavci definici skalárního součinu vektorů formulujeme algebraicky a zobecníme ji na libovolný vektorový prostor.

6.1.1 Skalární součin a jeho reprezentace v bázích

Zopakujme si definici (1.32) skalárního součinu vektorů z odstavce 1.4.3, v němž jsme jako s vektory pracovali s množinami orientovaných úseček. Podstatné bylo, že skalární součin přiřazoval uspořádané dvojici vektorů číslo. V daném konkrétním případě bylo toto číslo určeno jako součin velikostí obou vektorů a kosinu úhlu, který svíraly, tj.

$$V_3 \times V_3 \ni [\vec{u}, \vec{v}] \longrightarrow \vec{u}\vec{v} = |\vec{u}||\vec{v}| \cos \varphi, \quad \varphi = \sphericalangle(\vec{u}, \vec{v}) \in \mathbf{R}.$$

Základními vlastnostmi takto definovaného skalárního součinu vektorů byly komutativita, linearity a pozitivní definitnost:

$$\begin{aligned} \vec{u}\vec{v} &= \vec{v}\vec{u}, \\ (k\vec{u})\vec{v} &= \vec{u}(k\vec{v}) = k(\vec{u}\vec{v}), \\ \vec{u}(\vec{v} + \vec{w}) &= \vec{u}\vec{v} + \vec{u}\vec{w}, \\ \vec{u}\vec{u} &\geq 0, \quad \text{rovnost} \iff \vec{u} = \vec{o}. \end{aligned} \tag{6.1}$$

Pokud jsme zapsali vektory \vec{u} a \vec{v} ve složkách v bázi $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ navzájem kolmých jednotkových vektorů, tj. v *ortonormální bázi*, zjistili jsme, že skalární součin vektorů $\vec{u} = u_1 \vec{e}_1 + u_2 \vec{e}_2 + u_3 \vec{e}_3$ a $\vec{v} = v_1 \vec{e}_1 + v_2 \vec{e}_2 + v_3 \vec{e}_3$ (v prvním dílu ještě s dolními indexy u složek), můžeme vyjádřit ve tvaru (1.34), tj. $\vec{u}\vec{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$. Pokusme se najít různá zobrazení z týmiž vlastnostmi, aniž bychom se omezovali nějakou zcela určitou představou o vektorech jako konkrétních objektech. Nebudeme tedy trvat na tom, že vektory jsou zrovna orientované úsečky.

Příklad 6.1: Skalární součin uspořádaných n -tic

V příkladu 4.20 jsme viděli, že uspořádané n -tice komplexních čísel vytvoří vektorový prostor, definujeme-li jejich sčítání a násobení číslem „po složkách“, tj. $(\alpha^1, \alpha^2, \dots, \alpha^n) + (\beta^1, \beta^2, \dots, \beta^n) = (\alpha^1 + \beta^1, \alpha^2 + \beta^2, \dots, \alpha^n + \beta^n)$, $\gamma(\alpha^1, \alpha^2, \dots, \alpha^n) = (\gamma\alpha^1, \gamma\alpha^2, \dots, \gamma\alpha^n)$. Na této skutečnosti se nic nezmění, budeme-li pracovat pouze s reálnými čísly. Získáme tak vektorový prostor dimenze n nad polem reálných čísel. Přiřadíme-li nyní libovolné dvojici n -tic reálných čísel $(\alpha) = (\alpha^1, \alpha^2, \dots, \alpha^n)$ a $(\beta) = (\beta^1, \beta^2, \dots, \beta^n)$ číslo

$$[(\alpha), (\beta)] \longrightarrow ((\alpha), (\beta)) = \alpha^1 \beta^1 + \alpha^2 \beta^2 + \dots + \alpha^n \beta^n,$$

snadno zjistíme, že toto zobrazení má všechny čtyři vlastnosti skalárního součinu, které jsme objevili u orientovaných úseček.

Skalární součin jsme v příkladu 6.1 označili symbolem $(\ , \)$, který budeme používat i dále.

Příklad 6.2: Skalární součin matic typu m/n

Také matice typu m/n s operacemi sčítání a násobení skalárem tvoří vektorový prostor (viz příklad 4.21). Jeho dimenze je mn . Uvažujme o tomto vektorovém prostoru v případě, že polem skalárů je množina reálných čísel. Přiřadíme-li dvojici matic $A = (\alpha_i^j)$ a $B = (\beta_i^j)$, $1 \leq i \leq m$, $1 \leq j \leq n$, číslo

$$\mathcal{M}(m/n) \times \mathcal{M}(m/n) \ni [A, B] \longrightarrow (A, B) = \sum_{i=1}^m \sum_{j=1}^n \alpha_i^j \beta_i^j \in \mathbf{R},$$

nebo třeba

$$\mathcal{M}(n/n) \times \mathcal{M}(n/n) \ni [A, B] \longrightarrow (A, B) = \sum_{i=1}^n \alpha_i^i \beta_i^i \in \mathbf{R},$$

zjistíme v obou případech, že takové zobrazení má opět čtyři vlastnosti skalárního součinu orientovaných úseček.

Příklad 6.3: Skalární součin polynomů

Také polynomy lze sčítat a násobit číslem. Uvažujme o množině $\mathbf{P}[n]$ polynomů jedné proměnné stupně nejvýše n s reálnými koeficienty. Každý z nich lze zapsat ve tvaru

$$P(x) = p_0 + p_1x + p_2x^2 + \cdots + p_nx^n.$$

S operací sčítání polynomů a násobení polynomu reálným číslem se tato množina stává vektorovým prostorem nad polem reálných čísel. Označíme-li

$$e_1(x) = 1, \quad e_2(x) = x, \quad \dots, \quad e_{i+1}(x) = x^i, \quad \dots, \quad e_{n+1}(x) = x^n,$$

můžeme každý polynom zapsat ve tvaru lineární kombinace $P(x) = p_0 e_1(x) + p_1 e_2(x) + \cdots + p_n e_{n+1}(x)$. Vektorový prostor polynomů je tedy $(n+1)$ -rozměrný. Každý polynom je určen souborem $(n+1)$ čísel — složek v bázi $(e_1(x), e_2(x), \dots, e_{n+1}(x))$. Přiřadíme-li dvojici polynomů $P(x) = p_0 + p_1x + \cdots + p_nx^n$ a $Q(x) = q_0 + q_1x + \cdots + q_nx^n$ číslo

$$\mathbf{P}[n] \times \mathbf{P}[n] \ni [P(x), Q(x)] \longrightarrow (P(x), Q(x)) = p_0q_0 + p_1q_1 + \cdots + p_nq_n \in \mathbf{R},$$

opět snadno prověříme vlastnosti (6.1).

Příklad 6.4: Skalární součin polynomů jinak

Pro polynomy z předchozího příkladu zavedme následující operaci: Dvojici polynomů $P(x)$ a $Q(x)$ přiřadíme číslo

$$\mathbf{P}[n] \times \mathbf{P}[n] \ni [P(x), Q(x)] \longrightarrow (P(x), Q(x)) = \int_0^1 P(x)Q(x) dx \in \mathbf{R}.$$

I tentokrát jsou splněny vlastnosti (6.1).

Pozn.: Nepozastavujte se nad tím, že indexy složek jsou nyní dole. U zápisu polynomů je to obvyklé. Navíc je indexování složek standardně posunuto vůči indexování prvků báze.

Příklad 6.5: Skalární součin polynomů ještě jinak

Uvažujme teď o polynomech trochu obecněji. Opět nechť jsou stupně nejvýše n , připuštěme však u nich komplexní koeficienty. Množina takových polynomů s obvyklými operacemi sčítání polynomů a násobení polynomu komplexním číslem je opět vektorovým prostorem dimenze $n+1$ (tentokrát však nad polem komplexních čísel). Zavedme pro ně operaci, která bude dvojici polynomů $P(x)$ a $Q(x)$ přiřazovat (komplexní) číslo

$$\mathbf{P}[n] \times \mathbf{P}[n] \ni [P(x), Q(x)] \longrightarrow (P(x), Q(x)) = \int_0^1 P(x)Q(x) dx,$$

stejně jako tomu bylo v příkladu 6.4 (při integraci zacházíme s imaginární jednotkou jako s konstantou). Zkusme nyní prověřit vlastnosti (6.1),

$$(P(x), Q(x)) = \int_0^1 P(x)Q(x) dx = (Q(x), P(x)),$$

$$(P_1(x) + P_2(x), Q(x)) = \int_0^1 [P_1(x) + P_2(x)]Q(x) dx =$$

$$= \int_0^1 P_1(x)Q(x) dx + \int_0^1 P_2(x)Q(x) dx = (P_1(x), Q(x)) + (P_2(x), Q(x)),$$

$$(\alpha P(x), Q(x)) = \int_0^1 \alpha P(x)Q(x) dx = (P(x), \alpha Q(x)) = \alpha(P(x), Q(x)),$$

$$(P(x), P(x)) = \int_0^1 [P(x)]^2 dx = \text{komplexní číslo.}$$

Vida, první tři ze vztahů (6.1) jsou opět splněny, čtvrtý však nikoliv, neboť i pro $P(x) = Q(x)$ je výsledkem operace obecně komplexní číslo a komplexní čísla nelze porovnávat. Můžeme se pokusit definici operace opravit tak, abychom platnosti čtvrtého vztahu docílili. Třeba takto:

$$\mathbf{P}[n] \times \mathbf{P}[n] \ni [P(x), Q(x)] \longrightarrow (P(x), Q(x)) = \int_0^1 P^*(x)Q(x) dx.$$

Teď však zase zjistíme, že neplatí první vztah. Skutečně,

$$(P(x), Q(x)) = \int_0^1 P^*(x)Q(x) dx = \int_0^1 [Q^*(x)P(x)]^* dx = (Q(x), P(x))^*.$$

Je vidět, že druhá definice operace sice porušuje vlastnost komutativity, ale „ne příliš“. Stačí provést operaci komplexního sdružení a nesrovnalost je napravena. Zato první definice se neshoduje se čtvrtým vztahem velmi výrazně. Neshoda se nedá napravit nějakou jeho snadnou modifikací.

A tak bychom mohli dál vymýšlet. Vraťme se však k příkladům vektorových prostorů nad reálnými čísly. Jsou v něčem odlišné a v něčem shodné. Shodují se v tom, že jejich objekty jsou vždy prvky n -rozměrného vektorového prostoru. Odlišují se konkrétními objekty (jednou jde o orientované úsečky, jindy o n -tice čísel, pak zase o matice či polynomy) a konkrétní definicí operace, která dvěma vektorům přiřazuje číslo. Co však je opět společné, jsou vlastnosti této operace, popsané vztahy (6.1). O tuto shodu se opřeme při formulaci obecné definice skalárního součinu. Nebude při ní důležité, jaké jsou konkrétní prvky vektorového prostoru, ani jak je operace konkrétně definována. Budeme ji charakterizovat pouze souborem vlastností typu (6.1), zobecněných však na případ vektorového prostoru nad polem komplexních čísel.

Skalárním součinem ve vektorovém prostoru V_n nad \mathbf{C} rozumíme zobrazení

$$V_n \times V_n \ni [a, b] \longrightarrow (a, b) \in \mathbf{C} \quad (6.2)$$

s vlastnostmi

- $(a, b) = (b, a)^*$ (kososymetrie, nebo jen symetrie),
- $(a_1 + a_2, b) = (a_1, b) + (a_2, b)$ (první vlastnost linearity),
- $(\alpha a, b) = \alpha(a, b)$ (druhá vlastnost linearity),
- $(a, a) \geq 0$, rovnost $\iff a = 0_{V_n}$ (pozitivní definitnost)

pro libovolné vektory $a, a_1, a_2, b \in V_n$ a libovolný skalár $\alpha \in \mathbf{C}$.

Pokud nastane speciální případ a vektorový prostor bude konstruován nad polem reálných čísel, odpadne operace komplexního sdružení v první z vlastností (6.3) a dostaneme opět vlastnosti (6.1). Nové obecné definici vyhovuje pro případ polynomů v příkladu 6.5 druhý způsob, jak přiřadit dvojici polynomů komplexní číslo. Vektorový prostor nad \mathbf{C} , v němž jsme definovali skalární součin, se nazývá *unitární* a značí se U_n . Nad \mathbf{R} jde o prostor *euklidovský* s obvyklým označením E_n .

Příklad 6.6: Znovu n -tice, tentokrát komplexní

S vektorovým prostorem uspořádaných n -tic komplexních čísel jsme se setkali v příkladu 4.20. Přiřadíme nyní dvojici takových n -tic $[(\alpha), (\beta)]$, $(\alpha) = (\alpha^1, \alpha^2, \dots, \alpha^n)$, $(\beta) = (\beta^1, \beta^2, \dots, \beta^n)$ komplexní číslo

$$\begin{aligned} [(\alpha), (\beta)] &\longrightarrow ((\alpha), (\beta)) = \alpha^1 \beta^{1*} + \alpha^2 \beta^{2*} + \dots + \alpha^n \beta^{n*} = \\ &= (\alpha)(\beta)^{T*} = \begin{pmatrix} \alpha^1 & \alpha^2 & \dots & \alpha^n \end{pmatrix} \begin{pmatrix} \beta^{1*} \\ \beta^{2*} \\ \vdots \\ \beta^{n*} \end{pmatrix}. \end{aligned}$$

Poslední zjednodušený zápis představuje násobení řádkové matice (α) sloupcovou maticí $(\beta)^{T*}$ vzniklou transpozicí a komplexním sdružením řádkové matice (β) . Ověřte sami, že toto zobrazení má vlastnosti 6.3 a je tedy

skalárním součinem.

Příklad 6.7: Je definice skalárního součinu úplná?

Nepostrádáte něco mezi vlastnostmi (6.3) definujícími skalární součin? Jak si poradíme s výrazy $(a, b_1 + b_2)$ a $(a, \beta b)$? Je třeba zahrnout pravidla pro jejich úpravu mezi definiční vlastnosti, nebo tato pravidla z definice již vyplývají? Pokusíme se to zjistit. Platí

$$(a, b_1 + b_2) = [(b_1 + b_2, a)]^* = [(b_1, a) + (b_2, a)]^* = (a, b_1) + (a, b_2),$$

$$(a, \beta b) = (\beta b, a)^* = [\beta(b, a)]^* = \beta^*(b, a)^* = \beta^*(a, b).$$

Vidíme, že tyto vlastnosti jsou již důsledkem definice (6.3). Můžeme je shrnout do jediného vzorce

$$(\alpha^1 a_1 + \alpha^2 a_2, \beta^1 b_1 + \beta^2 b_2) = \begin{pmatrix} \alpha^1 & \alpha^2 \end{pmatrix} \begin{pmatrix} (a_1, b_1) & (a_1, b_2) \\ (a_2, b_1) & (a_2, b_2) \end{pmatrix} \begin{pmatrix} \beta^1 \\ \beta^2 \end{pmatrix}^*,$$

který představuje *seskvilinearitu*, neboli linearitu v prvním a *antilinearitu* v druhém činiteli skalárního součinu.

Pochopitelně nás napadnou otázky: Kolik takových skalárních součinů — zobrazení určených vztahy (6.2) a (6.3) — můžeme ve vektorovém prostoru V_n zadat? A existuje vůbec nějaké takové zobrazení v každém vektorovém prostoru? Jak všechny takové skalární součiny popíšeme? A zadáme-li konkrétně skalární součin, tj. konkrétní zobrazení typu (6.2) a (6.3), jak postupujeme prakticky, chceme-li umět libovolné dvojici vektorů zadaných jejich složkami v nějaké bázi opravdu přiřadit odpovídající konkrétní komplexní číslo — hodnotu jejich skalárního součinu? Na tyto otázky postupně odpovíme. Odpovědi na první dvě otázky již určitě tušíte. Napovídají je příklady 6.1 až 6.6. V nich jsme vždy skalární součin dokázali sestrotit a u polynomů jsme našli dvojí konkrétní možnost volby skalárního součinu. Z odstavce 4.1.3 také víme, že vektorové prostory téže dimenze n jsou izomorfní a tedy „algebraicky stejné“. Zdá se tedy, že stačí na naše otázky odpovědět například pro případ prostoru n -tic. V tomto prostoru jsme skalární součin definovali v příkladu 6.6 a dokázali tak jeho existenci. Zkuste v tomto prostoru najít ještě jiný skalární součin. Dokážete tím nejednoznačnost volby, tj. existenci více skalárních součinů.

Nyní však již přistupme k obecnému vyjádření *všech* skalárních součinů ve vektorovém prostoru V_n nad \mathbf{C} . Budeme zase používat Einsteinovu sčítací symboliku, bude-li to pro výpočty vhodné. Zvolme ve V_n bázi (e_1, e_2, \dots, e_n) a vektory

$$a = \alpha^i e_i = \alpha^1 e_1 + \dots + \alpha^n e_n, \quad b = \beta^j e_j = \beta^1 e_1 + \dots + \beta^n e_n.$$

A počítejme s využitím druhé a třetí vlastnosti (6.3) a jejich důsledků z příkladu 6.7:

$$(a, b) = (\alpha^i e_i, \beta^j e_j) = \alpha^i \beta^{j*} (e_i, e_j).$$

(Uvědomme si, že získaný výraz je dán dvojnásobnou sumou obsahující n^2 sčítanců, v níž sčítací indexy i a j nabývají nezávisle hodnot 1 až n .) Výsledek napovídá, že k tomu, abychom

byli schopni dvojici libovolných vektorů a a b prostoru V_n , zadaných složkami, přiřadit konkrétní číslo, které bude jejich skalárním součinem, je třeba znát, resp. zadat hodnoty skalárních součinů všech dvojic vektorů báze $g_{ij} = (e_i, e_j)$, $i, j \in \{1, 2, \dots, n\}$. Tyto hodnoty tvoří matici G typu n/n . Nemohou být ovšem zadány zcela libovolně, nýbrž musí splňovat ještě zbývající požadavky skalárního součinu, konkrétně prvou (kososymetrii) a čtvrtou (pozitivní definitnost) z vlastností (6.3). Tyto vlastnosti vedou k omezení možností volby matice G :

$$g_{ij} = (e_i, e_j) = (e_j, e_i)^* = g_{ji}^* \implies G = G^{T*}.$$

Ze symetrie skalárního součinu je vidět, že matice G je shodná se svou maticí transponovanou komplexně sdruženou. Taková matice se nazývá *samoadjungovaná* (sama k sobě sdružená). V případě vektorového prostoru nad \mathbf{R} je $G = G^T$ a jedná se o matici *symetrickou*. Pozitivní definitnost skalárního součinu vede okamžitě k závěru, že $g_{ii} > 0$ (přemýšlejte, proč v tomto případě nemůžeme připustit rovnost), v diagonále matice G jsou tedy reálná kladná čísla. Navíc lze pomocí této vlastnosti, platné pro libovolný vektor, ukázat, že matice G je *pozitivně definitní*, tj. její levé horní subdeterminanty všech řádů jsou kladné.

Matice G reprezentující skalární součin v bázi unitárního, resp. euklidovského prostoru je samoadjungovaná a pozitivně definitní, resp. symetrická a pozitivně definitní.

Vidíme, že pro zadání skalárního součinu je nutné, ale také stačí, zadat libovolnou samoadjungovanou pozitivně definitní matici G řádu n a prohlásit její prvky g_{ij} za skalární součiny vektorů zvolené báze s odpovídajícími indexy. Skalární součin vektorů zadaných v této bázi složkami pak již vyjádříme snadno:

$$\begin{aligned} (a, b) &= (\alpha)G(\beta)^{T*} = \\ &= \begin{pmatrix} \alpha^1 & \alpha^2 & \cdots & \alpha^n \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{12}^* & g_{22} & \cdots & g_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ g_{1n}^* & g_{2n}^* & \cdots & g_{nn} \end{pmatrix} \begin{pmatrix} \beta^{1*} \\ \beta^{2*} \\ \vdots \\ \beta^{n*} \end{pmatrix}. \end{aligned} \quad (6.4)$$

V případě pole reálných skalárů lze zápis všech předchozích vztahů zjednodušit vypuštěním operace komplexního sdružení, tj. odstraněním hvězdiček.

Další otázka je nasnadě: Jestliže se skalární součin *reprezentuje* v jedné bázi, řekněme $\mathcal{B} = (e_1, e_2, \dots, e_n)$, maticí G , jaký tvar bude mít reprezentující matice tohoto skalárního součinu \bar{G} v bázi $\bar{\mathcal{B}} = (\bar{e}_1, \bar{e}_2, \dots, \bar{e}_n)$, ke které přejdeme pomocí matice přechodu T v souladu se vztahy (4.12) a (4.13)? Platí

$$\bar{g}_{ij} = (\bar{e}_i, \bar{e}_j) = (\tau_i^k e_k, \tau_j^\ell e_\ell) = \tau_i^k \tau_j^{\ell*} (e_k, e_\ell) = \tau_i^k g_{k\ell} \tau_j^{\ell*} = (TGT^{T*})_{ij},$$

nebo

$$(a, b) = (\alpha)G(\beta)^{T*} = (\bar{\alpha})TG((\bar{\beta})T)^{T*} = (\bar{\alpha})(TGT^{T*})(\bar{\beta})^{T*}.$$

Z obou předchozích výpočtů lze vztah mezi G a \bar{G} ihned přecíst:

$$\begin{aligned} \bar{G} &= TGT^{T*} = \\ &= \begin{pmatrix} \tau_1^1 & \tau_1^2 & \cdots & \tau_1^n \\ \tau_2^1 & \tau_2^2 & \cdots & \tau_2^n \\ \cdots & \cdots & \cdots & \cdots \\ \tau_n^1 & \tau_n^2 & \cdots & \tau_n^n \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix} \begin{pmatrix} \tau_1^{1*} & \tau_1^{2*} & \cdots & \tau_1^{n*} \\ \tau_2^{1*} & \tau_2^{2*} & \cdots & \tau_2^{n*} \\ \cdots & \cdots & \cdots & \cdots \\ \tau_n^{1*} & \tau_n^{2*} & \cdots & \tau_n^{n*} \end{pmatrix}. \end{aligned} \quad (6.5)$$

Z těchto výsledků lze vyvodit ještě další závěr: skalárních součinů v daném n -rozměrném vektorovém prostoru můžeme definovat nekonečně mnoho.

6.1.2 Ortonormální báze

Terminologie názvu tohoto odstavce nám není neznámá. Setkali jsme se s ní v odstavci 1.4. Nyní si ji nejprve připomeneme názorným příkladem a poté ji zobecníme.

Příklad 6.8: Kolmost, jednotkový vektor

Vraťme se ke vztahům (6.1) pro skalární součin vektorů jako orientovaných úseček v prostoru, kde je definováno měření úhlů a délek. Z definice skalárního součinu, kterou jsme pomocí úhlů a délek zavedli, je vidět, že dva nenulové vektory svírají pravý úhel, jsou *kolmé*, právě tehdy, je-li jejich skalární součin nulový ($\cos 90^\circ = 0$). Vektor má jednotkovou délku právě tehdy, je-li jeho skalární součin sama se sebou roven jedné. Tuto skutečnost použijeme pro zobecnění definice a zavedení pojmů ortogonální vektory a normovaný vektor.

Vektory $a \in U_n$ a $b \in U_n$ se nazývají *ortogonální vzhledem ke zvolenému skalárnímu součinu*, platí-li $(a, b) = 0$. Vektor a se nazývá *normovaný vzhledem ke skalárnímu součinu*, je-li $(a, a) = 1$. Báze (e_1, \dots, e_n) v U_n se nazývá *ortonormální vzhledem ke zvolenému skalárnímu součinu*, je-li $(e_i, e_j) = \delta_{ij}$, $i, j = 1, \dots, n$.

Slova „vzhledem ke skalárnímu součinu“ jsou v této definici významná. V tomtéž prostoru V_n by totiž mohl být zvolen jiný konkrétní skalární součin, vůči kterému by již vektory nemusely být ortogonální nebo normované. Pomocí definice skalárního součinu naopak můžeme zavést pojem délky vektoru a úhlu mezi vektory i ve vícerozměrných prostorech, kde si třeba úhel ani neumíme představit. *Délkou* vektoru nazveme $|a| = \sqrt{(a, a)}$, *úhel* mezi nenulovými vektory definujeme pomocí jeho kosinu vztahem

$$\cos \alpha = \frac{(a, b)}{|a||b|}.$$

Příklad 6.9: Znovu polynomy

V příkladu 6.3 jsme v $(n+1)$ -rozměrném vektorovém prostoru polynomů stupně nejvýše n nad \mathbf{R} zvolili skalární součin takto:

$$(P(x), Q(x)) = p_0q_0 + p_1q_1 + \cdots + p_nq_n$$

pro $P(x) = p_0 + p_1x + p_2x^2 + \cdots + p_nx^n$, $Q(x) = q_0 + q_1x + q_2x^2 + \cdots + q_nx^n$. Báze

$$(e_1(x), e_2(x), \dots, e_{n+1}(x)) = (1, x, x^2, \dots, x^n)$$

je tvořena vektory, které jsou vůči tomuto součinu normované a navzájem ortogonální. Tvoří *ortonormální bázi*. Zvolíme-li však skalární součin polynomů podle příkladu 6.4, tj.

$$(P(x), Q(x)) = \int_0^1 P(x)Q(x) dx,$$

pak tato báze již vzhledem k němu ortonormální nebude. Spočítejme skalární součiny vektorů báze, které definují matici G .

$$(e_i(x), e_j(x)) = \int_0^1 x^{i-1}x^{j-1} dx = \int_0^1 x^{i+j-2} dx = \left[\frac{x^{i+j-1}}{i+j-1} \right]_0^1 = \frac{1}{i+j-1},$$

$$G = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n-1} \end{pmatrix}.$$

Vraťme se k závěru předchozího odstavce, který nyní, jak se zdá, nabízí výbornou myšlenku: Dejme tomu, že jsme se rozhodli zadávat vektory v předem zvolené bázi (e_1, e_2, \dots, e_n) . Přitom chceme vektorový prostor opatřit skalárním součinem. Při zadané bázi máme ovšem nekonečně mnoho možností, jak jej zvolit, právě prostřednictvím volby matice G . Nejjednodušší matice, která může definovat skalární součin, je matice jednotková, tj. $G = E$. Pak $(e_i, e_j) = \delta_{ij}$, tj. $(e_i, e_i) = 1$ a $(e_i, e_j) = 0$ pro $i \neq j$. Zvolíme-li tedy takový skalární součin, který je v předem zadané bázi definován jednotkovou maticí, stanou se vůči němu vektory této báze automaticky ortogonálními a normovanými. Přejdeme-li k jiné ortonormální bázi $(\bar{e}_1, \bar{e}_2, \dots, \bar{e}_n)$ pomocí matice přechodu T , zjistíme díky (6.5), že matice přechodu má dodatečnou vlastnost:

$$E = \bar{G} = TGT^{T*} = TET^{T*} = TT^{T*} \implies T^{T*} = T^{-1}.$$

Matice transponovaná a komplexně sdružená k T tedy hraje roli matice inverzní. S touto vlastností ve speciální podobě (pro případ matic reálných čísel) jsme se již setkali ve vztahu (1.41), kdy se jednalo o přechod mezi bázemi tvořenými jednotkovými a navzájem kolmými vektory v trojrozměrném euklidovském prostoru. Čtvercová matice z komplexních čísel, pro kterou je $T^{T*} = T^{-1}$, se nazývá *unitární*, matice z reálných čísel, pro niž $T^T = T^{-1}$, se nazývá, stejně jako v odst. 1.4.4, *ortogonální*.

Matice přechodu mezi ortonormálními bázemi v unitárním prostoru je vždy unitární, v prostoru euklidovském ortogonální.

Je velmi výhodné vyjadřovat vektory unitárního, resp. euklidovského prostoru v ortonormálních bázích. Skalární součin je v nich reprezentován jednotkovou maticí a při přechodu mezi nimi si ušetříme nepříjemný výpočet inverzní matice přechodu, stačí jen transponovat a komplexně sdružit matici T . Skalární součin dvou vektorů je vyjádřen také velmi jednoduše — dosazením $G = E$ do (6.4) dostaneme následující závěr:

V ortonormální bázi je

$$(a, b) = (\alpha)(\beta)^{T*} = \begin{pmatrix} \alpha^1 & \alpha^2 & \cdots & \alpha^n \end{pmatrix} \begin{pmatrix} \beta^{1*} \\ \beta^{2*} \\ \vdots \\ \beta^{n*} \end{pmatrix} = \alpha^1\beta^{1*} + \alpha^2\beta^{2*} + \cdots + \alpha^n\beta^{n*}, \quad (6.6)$$

v případě euklidovského prostoru není třeba psát hvězdičku,

$$(a, b) = (\alpha)(\beta)^T = \alpha^1\beta^1 + \alpha^2\beta^2 + \cdots + \alpha^n\beta^n.$$

Co když však máme ve vektorovém prostoru skalární součin již zadán nějak nezávisle na bázi, například jako tomu bylo u polynomů v příkladu 6.5? Co když pak žádná ortonormální báze neexistuje? Není třeba se obávat. Lze ukázat, že ortonormální bázi lze zkonstruovat vždy. Existují procedury, jak takovou bázi získat z výchozí báze obecné. Nejznámější z nich je *Grammův–Schmidtův ortogonalizační proces*. Předpokládejme, že v unitárním prostoru U_n , tj. takovém, který již je vybaven skalárním součinem, je zadána obecná báze $\mathcal{A} = (a_1, a_2, \dots, a_n)$. Nejprve pomocí ní sestrojíme bázi $\mathcal{C} = (c_1, c_2, \dots, c_n)$ tvořenou ortogonálními vektory, které potom budeme normovat. Konstrukce se děje v postupných krocích. Konkrétní proceduru ukážeme nejprve na praktickém příkladu, který se vyskytuje velmi často. Dejme tomu, že máme složkami zadány (a třeba i v nějaké základní ortonormální bázi (e_1, e_2, \dots, e_n) , aby to bylo jednoduché) lineárně nezávislé vektory (a_1, a_2, \dots, a_k) , $1 \leq k \leq n$, které generují vektorový podprostor $L = [[a_1, a_2, \dots, a_k]]$. Chceme však zjistit, zda a jak v tomto podprostoru můžeme najít ortonormální bázi.

Příklad 6.10: Grammův–Schmidtův ortogonalizační proces v praxi

Nechť je trojrozměrný vektorový podprostor $L \subset U_4$ generován vektory zadanými složkami v bázi $\mathcal{B} = (e_1, e_2, e_3, e_4)$ takto:

$$L = [[a_1, a_2, a_3]] = [[(1, 1, 0, 0), (0, i, 1, 0), (0, 0, i, -i)]].$$

(Ověřte nezávislost vektorů a_1, a_2, a_3 .) Pro jednoduchost předpokládejme, že báze \mathcal{B} je vzhledem ke skalárnímu

součinu v U_4 bázi ortonormální, takže v ní skalární součin reprezentuje jednotková matice. Pro výpočet hodnot skalárních součinů dvojice vektorů pomocí složek tedy použijeme vztahu (6.6). Jak nyní vypadají jednotlivé kroky konstrukce:

1. krok: Položíme $c_1 = a_1 = (1, 1, 0, 0)$.

2. krok: Hledáme vektor c_2 ve tvaru součtu $c_2 = \gamma c_1 + a_2$ a požadujeme, aby byl ortogonální k vektoru c_1 :

$$0 = (c_2, c_1) = \gamma(c_1, c_1) + (a_2, c_1) \implies \gamma = -\frac{(a_2, c_1)}{(c_1, c_1)},$$

$$(c_1, c_1) = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = 2, \quad (a_2, c_1) = \begin{pmatrix} 0 & i & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = i,$$

$$\gamma = -\frac{i}{2}, \quad c_2 = -\frac{i}{2}(1, 1, 0, 0) + (0, i, 1, 0) = \left(-\frac{i}{2}, \frac{i}{2}, 1, 0\right).$$

3. krok: Vektor c_3 hledáme ve tvaru součtu lineární kombinace již získaných ortogonálních vektorů c_1 a c_2 a dalšího vektoru původního systému, tj. $c_3 = \gamma^1 c_1 + \gamma^2 c_2 + a_3$. Z požadavků $(c_3, c_1) = 0$ a $(c_3, c_2) = 0$ dostáváme neznámé koeficienty γ^1 a γ^2 :

$$\gamma^1 = -\frac{(a_3, c_1)}{(c_1, c_1)}, \quad \gamma^2 = -\frac{(a_3, c_2)}{(c_2, c_2)}, \quad \text{kde}$$

$$(a_3, c_1) = \begin{pmatrix} 0 & 0 & i & -i \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = 0, \quad (a_3, c_2) = \begin{pmatrix} 0 & 0 & i & -i \end{pmatrix} \begin{pmatrix} \frac{i}{2} \\ -\frac{i}{2} \\ 1 \\ 0 \end{pmatrix} = i,$$

$$(c_2, c_2) = \begin{pmatrix} -\frac{i}{2} & \frac{i}{2} & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{i}{2} \\ -\frac{i}{2} \\ 1 \\ 0 \end{pmatrix} = \frac{3}{2}, \quad \gamma^1 = 0, \quad \gamma^2 = -\frac{2i}{3},$$

$$c_3 = 0(1, 1, 0, 0) - \frac{2i}{3} \left(-\frac{i}{2}, \frac{i}{2}, 1, 0\right) + (0, 0, i, -i) = \left(-\frac{1}{3}, \frac{1}{3}, \frac{i}{3}, -i\right),$$

$$(c_3, c_3) = \begin{pmatrix} -\frac{1}{3} & \frac{1}{3} & \frac{i}{3} & -i \end{pmatrix} \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ -\frac{i}{3} \\ i \end{pmatrix} = \frac{4}{3}.$$

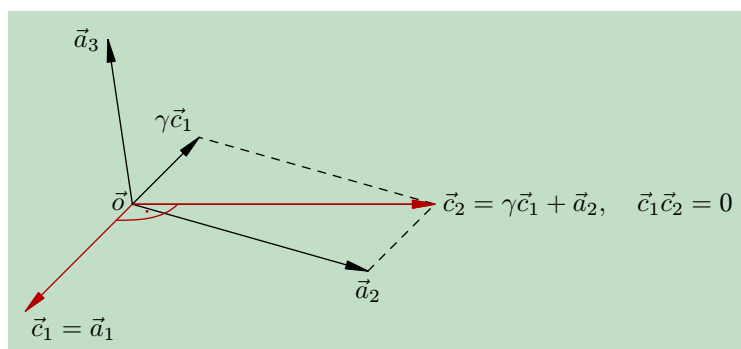
Získali jsme v podprostoru L bázi tvořenou navzájem ortogonálními vektory c_1 , c_2 a c_3 . Ty je ještě třeba normovat. Hledejme vektory $f_1 = \alpha c_1$, $f_2 = \beta c_2$, $f_3 = \gamma c_3$ tak, aby byly normované. Požadujeme-li například

$(f_1, f_1) = 1$, dostaneme $(\alpha c_1, \alpha c_1) = \alpha \alpha^* (c_1, c_1) = 1$. Tomuto požadavku vyhovují všechna komplexní čísla α , jejichž modul je $|\alpha| = \left(\sqrt{(c_1, c_1)}\right)^{-1}$. Obdobně postupujeme při nalezení β a γ . Pak

$$\begin{aligned} f_1 &= \frac{c_1}{\sqrt{(c_1, c_1)}} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right), \\ f_2 &= \frac{c_2}{\sqrt{(c_2, c_2)}} = \left(-\frac{i}{\sqrt{6}}, \frac{i}{\sqrt{6}}, \sqrt{\frac{2}{3}}, 0 \right), \\ f_3 &= \frac{c_3}{\sqrt{(c_3, c_3)}} = \left(-\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{i}{2\sqrt{3}}, -\frac{i\sqrt{3}}{2} \right). \end{aligned}$$

Ve vektorovém podprostoru L jsme nyní vytvořili jeho ortonormální bázi (f_1, f_2, f_3) namísto původní báze obecné. Ještě se nabízí poslední zajímavá otázka: V našem příkladu jsme vektorový podprostor dimenze k generovali opravdu k nezávislými vektory, tj. bází. Víme však, že v praxi existují i jiné možnosti — často generujeme vektorový podprostor systémem vektorů, jejichž počet je vyšší než dimenze podprostoru. Jak postupovat v takovémto případě? Možnosti jsou dvě. Při první z nich závislé vektory vyřadíme, popřípadě získáme bázi podprostoru L úpravou matice tvořené složkami generujících vektorů na schodovitý tvar. Poté uplatníme ortogonalizační proces. Druhá možnost je rovnou začít ortogonalizovat. Závislé vektory se při proceduře vyřadí samy tím, že vzniknou vektory nulové, které nemohou být prvky báze. Přesvědčíme se o tom za chvíli.

Již příklad naznačuje, že při libovolně zvoleném skalárním součinu ve vektorovém prostoru vždy půjde najít ortonormální báze. Postup jejich hledání nyní shrneme. Mějme v prostoru U_n obecnou bázi $\mathcal{A} = (a_1, a_2, \dots, a_n)$. Hledáme bázi $\mathcal{C} = (c_1, c_2, \dots, c_n)$, která bude nejprve



Obr. 6.1 Ortogonalizace v E_3 — první dva kroky.

ortogonální, její vektory však nebudou ještě normovány. Opět postupujeme v krocích.

Grammův–Schmidtův ortogonalizační proces:

1. krok: Položíme $c_1 = a_1$.

2. krok: Hledáme vektor c_2 ve tvaru součtu $c_2 = \gamma c_1 + a_2$ a požadujeme, aby byl ortogonální k vektoru c_1 , tj.

$$0 = (c_2, c_1) = \gamma(c_1, c_1) + (a_2, c_1) \implies \gamma = -\frac{(a_2, c_1)}{(c_1, c_1)}, \quad c_2 = -\frac{(a_2, c_1)}{(c_1, c_1)}c_1 + a_2.$$

Obrázek 6.1 znázorňuje první dva kroky v trojrozměrném prostoru.

3. krok: Vektor c_3 hledáme ve tvaru součtu lineární kombinace již získaných ortogonálních vektorů c_1 a c_2 a dalšího vektoru původního systému, tj.

$$c_3 = \gamma^1 c_1 + \gamma^2 c_2 + a_3.$$

Z požadavků $(c_3, c_1) = 0$ a $(c_3, c_2) = 0$ dostáváme

$$\gamma_1 = -\frac{(a_3, c_1)}{(c_1, c_1)}, \quad \gamma_2 = -\frac{(a_3, c_2)}{(c_2, c_2)}, \quad c_3 = -\frac{(a_3, c_1)}{(c_1, c_1)}c_1 - \frac{(a_3, c_2)}{(c_2, c_2)}c_2 + a_3.$$

k -tý krok:

$$c_k = -\frac{(a_k, c_1)}{(c_1, c_1)}c_1 - \frac{(a_k, c_2)}{(c_2, c_2)}c_2 - \dots - \frac{(a_k, c_{k-1})}{(c_{k-1}, c_{k-1})}c_{k-1} + a_k \quad (6.7)$$

pro $k = 2, 3, \dots, n$.

Pozn.: Uvědomte si, že výsledek procedury ortogonalizace není jednoznačný. Jistě jste si všimli, že ortogonální báze zkonstruovaná použitím Grammova–Schmidtova procesu závisí na konkrétní volbě výchozího systému vektorů generujícího daný podprostor L , včetně pořadí vektorů. Navíc lze zvolit i jiné ortogonalizační procedury. Co je však na zvoleném postupu nezávislé, je to, že získaná ortogonální, resp. ortonormální báze generuje opět vektorový podprostor L .

Zbývá ještě prověřit, že takto sestrojené vektory tvoří opravdu bázi. Počítejme:

$$\gamma^1 c_1 + \gamma^2 c_2 + \dots + \gamma^n c_n = 0_{U_n} \implies$$

$$\implies \gamma^1 (c_1, c_j) + \gamma^2 (c_2, c_j) + \dots + \gamma^n (c_n, c_j) = (0_{U_n}, c_j) = 0.$$

Protože jsou vektory c_1 až c_n navzájem ortogonální, platí $(c_i, c_j) = 0$ pro $i \neq j$ a $(c_j, c_j) \geq 0$. Pak $\gamma^j (c_j, c_j) = 0$. Číslo γ^j by tedy mohlo být nenulové jedině kdyby byl nulový vektor c_j . Ze

vztahu (6.7) pro j -tý krok však plyne, že v takovém případě by platilo

$$-\frac{(a_j, c_1)}{(c_1, c_1)} c_1 - \frac{(a_j, c_2)}{(c_2, c_2)} c_2 - \dots - \frac{(a_j, c_{j-1})}{(c_{j-1}, c_{j-1})} c_{j-1} + a_j = 0_{U_n}.$$

Vektor a_j by byl lineární kombinací vektorů c_1 až c_{j-1} . Tyto vektory však byly postupně sestrojeny jako lineární kombinace vektorů a_1, a_1 a a_2 , atd. až a_1, a_2, \dots, a_{j-1} , takže vektor a_j by byl nakonec lineární kombinací vektorů a_1, a_2, \dots, a_{j-1} . To však není možné, neboť původní systém vektorů jsme volili jako bázi, tj. systém lineárně nezávislý. Zdůrazněme ještě jeden praktický výsledek předchozí úvahy: Ukáže-li se, že je nějaký systém ortogonálních vektorů závislý, obsahuje s jistotou vektor nulový (úloha 13 ve cvičení 6.1.4). Můžeme tedy odpovědět na otázku, kterou jsme formulovali při řešení příkladu 6.10: Jak by dopadl ortogonalizační proces, kdyby vektory původního systému byly závislé? Odpověď je jednoduchá. Dejme tomu, že bychom na závislý vektor a_k narazili v k -tém kroku. Vektor c_k vypočtený podle vztahu (6.7) by tak byl lineárně závislý na vektorech c_1 až c_{k-1} získaných v předchozích krocích, současně by však k nim byl ortogonální. Nutně by tedy musel být vektorem nulovým. Zkusme to ověřit i prakticky.

Příklad 6.11: Ortogonalizace závislých vektorů

V trojrozměrném euklidovském prostoru orientovaných úseček (E_3 nad \mathbf{R}) se skalárním součinem zadaným vztahem $\vec{u}\vec{v} = |\vec{u}||\vec{v}|\cos \sphericalangle(\vec{u}, \vec{v})$ je v bázi jednotkových a navzájem kolmých vektorů $\vec{e}_1, \vec{e}_2, \vec{e}_3$ zadán systém vektorů $\vec{a}_1 = (1, -1, 0)$, $\vec{a}_2 = (1, 0, 1)$ a $\vec{a}_3 = (0, -1, -1)$. Hned vidíme, že $\vec{a}_3 = \vec{a}_1 - \vec{a}_2$, vektory jsou závislé. Můžeme to ověřit i výpočtem hodnoty matice tvořené jejich složkami:

$$h(A) = h \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & -1 \end{pmatrix} = h \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{pmatrix} = h \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} = 2.$$

(Matici jsme upravovali tak, že jsme první řádek odečetli od druhého.) Pokud je pravdivé předchozí tvrzení o anulování vektorů vznikajících ortogonalizačním procesem pro případ závislého systému, měli bychom ortogonalizačním procesem dostat pouze dva nenulové ortogonální vektory. V prvním kroku položíme $\vec{c}_1 = \vec{a}_1$, tj. $\vec{c}_1 = (1, -1, 0)$. Platí

$$\vec{c}_1\vec{c}_1 = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = 2, \quad \vec{a}_2\vec{c}_1 = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = 1.$$

V druhém kroku dostáváme

$$\vec{c}_2 = -\frac{\vec{a}_2\vec{c}_1}{\vec{c}_1\vec{c}_1} \vec{c}_1 + \vec{a}_2 = -\frac{1}{2}(1, -1, 0) + (1, 0, 1) = \left(\frac{1}{2}, \frac{1}{2}, 1\right).$$

Přemýšlejme, jak proces dopadne v třetím kroku. Hledaný vektor \vec{c}_3 je součtem vektoru \vec{a}_3 a lineární kombinace ortogonálních vektorů \vec{c}_1 a \vec{c}_2 , které jsme již vytvořili a které leží v podprostoru $L = [|\vec{a}_1, \vec{a}_2|]$. Vektor \vec{a}_3 však v tomto podprostoru leží také. Proto $\vec{c}_3 \in L$. Současně však, díky ortogonalizačnímu procesu, bude vektor \vec{c}_3 ke

všem vektorům podprostoru L ortogonální. To je možné splnit jen tak, že vektor \vec{c}_3 vyjde nulový. Zkusme to. Platí

$$\begin{aligned}\vec{a}_3\vec{c}_1 &= (0 \ -1 \ -1) \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = 1, \\ \vec{a}_3\vec{c}_2 &= (0 \ -1 \ -1) \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix} = -\frac{3}{2}, \quad \vec{c}_2\vec{c}_2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix} = \frac{3}{2}, \\ \vec{c}_3 &= -\frac{\vec{a}_3\vec{c}_1}{\vec{c}_1\vec{c}_1}\vec{c}_1 - \frac{\vec{a}_3\vec{c}_2}{\vec{c}_2\vec{c}_2}\vec{c}_2 + \vec{a}_3 = -\frac{1}{2}(1, -1, 0) + \left(\frac{1}{2}, \frac{1}{2}, 1\right) + (0, -1, -1) = (0, 0, 0).\end{aligned}$$

Očekávání se tedy splnilo.

Nakonec jsme po provedení všech kroků procedury dostali *ortogonální* bázi $\mathcal{C} = (c_1, c_2, \dots, c_n)$. Její prvky budeme nyní *normovat*: Označme

$$e_i = \frac{c_i}{\sqrt{(c_i, c_i)}} \implies (e_i, e_i) = \left(\frac{c_i}{\sqrt{(c_i, c_i)}}, \frac{c_i}{\sqrt{(c_i, c_i)}} \right) = 1. \quad (6.8)$$

Vektory e_1, e_2, \dots, e_n jsou tedy ortogonální i normované. Tvoří ortonormální bázi.

Příklad 6.12: Grammův–Schmidtův ortogonalizační proces v praxi ještě jednou

Vektorový prostor polynomů poskytuje z hlediska algebry opravdu mnoho poučení. Ještě jednou se k němu vraťme, tentokrát se skalárním součinem definovaným pomocí integrálu v příkladu 6.4. Postačí, budeme-li se pro jednoduchost pohybovat pouze nad \mathbf{R} a nejvyšším stupněm polynomů bude $n = 3$. Prostor je tedy čtyřrozměrný. Již jsme zjistili, že nejjednodušší báze

$$\mathcal{A} = (a_1(x), a_2(x), a_3(x), a_4(x)) = (1, x, x^2, x^3)$$

není, bohužel, vzhledem ke skalárnímu součinu danému integrálem z příkladu 6.4 ortonormální. Chceme-li pracovat v ortonormální bázi, musíme si ji vyrobit, opět po jednotlivých krocích.

1. krok:

$$c_1(x) = a_1(x) = 1, \quad (c_1(x), c_1(x)) = \int_0^1 dx = 1.$$

2. krok:

$$\begin{aligned}c_2(x) &= -\frac{(a_2(x), c_1(x))}{(c_1(x), c_1(x))}c_1(x) + a_2(x), \\ (a_2(x), c_1(x)) &= \int_0^1 x dx = \frac{1}{2}, \quad c_2(x) = -\frac{1}{2} + x, \\ (c_2(x), c_2(x)) &= \int_0^1 \left(-\frac{1}{2} + x\right)^2 dx = \frac{1}{12}.\end{aligned}$$

3. krok:

$$c_3(x) = -\frac{(a_3(x), c_1(x))}{(c_1(x), c_1(x))} c_1(x) - \frac{(a_3(x), c_2(x))}{(c_2(x), c_2(x))} c_2(x) + a_3(x),$$

$$(a_3(x), c_1(x)) = \int_0^1 x^2 dx = \frac{1}{3},$$

$$(a_3(x), c_2(x)) = \int_0^1 x^2 \left(-\frac{1}{2} + x\right) dx = \frac{1}{12},$$

$$c_3(x) = -\frac{1}{3} - \left(-\frac{1}{2} + x\right) + x^2 = \frac{1}{6} - x + x^2,$$

$$(c_3(x), c_3(x)) = \int_0^1 \left(\frac{1}{6} - x + x^2\right)^2 dx = \frac{1}{180}.$$

4. krok:

$$c_4(x) = -\frac{(a_4(x), c_1(x))}{(c_1(x), c_1(x))} c_1(x) - \frac{(a_4(x), c_2(x))}{(c_2(x), c_2(x))} c_2(x) - \frac{(a_4(x), c_3(x))}{(c_3(x), c_3(x))} c_3(x) + a_4(x),$$

$$(a_4(x), c_1(x)) = \int_0^1 x^3 dx = \frac{1}{4},$$

$$(a_4(x), c_2(x)) = \int_0^1 x^3 \left(-\frac{1}{2} + x\right) dx = \frac{3}{40},$$

$$(a_4(x), c_3(x)) = \int_0^1 x^3 \left(\frac{1}{6} - x + x^2\right) dx = \frac{1}{120},$$

$$c_4(x) = -\frac{1}{4} - \frac{9}{10} \left(-\frac{1}{2} + x\right) - \frac{3}{2} \left(\frac{1}{6} - x + x^2\right) + x^3 = -\frac{1}{20} + \frac{3}{5}x - \frac{3}{2}x^2 + x^3,$$

$$(c_4(x), c_4(x)) = \int_0^1 \left(-\frac{1}{20} + \frac{3}{5}x - \frac{3}{2}x^2 + x^3\right)^2 dx = \frac{1}{2800}.$$

Zbývá provést normování, pro něž jsou již všechny skalární součiny polynomů $(c_j(x), c_j(x))$ připraveny. Dostáváme

$$e_1(x) = 1,$$

$$e_2(x) = -\sqrt{3} + 2\sqrt{3}x,$$

$$e_3(x) = \sqrt{5} - 6\sqrt{5}x + 6\sqrt{5}x^2,$$

$$e_4(x) = -\sqrt{7} + 12\sqrt{7}x - 30\sqrt{7}x^2 + 20\sqrt{7}x^3.$$

Hned také můžeme přecíst matici přechodu od původní báze \mathcal{A} k nové ortonormální bázi \mathcal{B} ,

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\sqrt{3} & 2\sqrt{3} & 0 & 0 \\ \sqrt{5} & -6\sqrt{5} & 6\sqrt{5} & 0 \\ -\sqrt{7} & 12\sqrt{7} & -30\sqrt{7} & 20\sqrt{7} \end{pmatrix}.$$

Kdybychom vyjádřili složky obecného polynomu $P(x) = p_0 + p_1x + p_2x^2 + p_3x^3$ v této nové bázi, mohli bychom se přesvědčit, že pro výpočet skalárního součinu funguje vztah (6.6). K tomu bychom však potřebovali matici T^{-1} , kterou se v tuto chvíli nikomu nechce počítat. V dalším odstavci však uvidíme efektivní způsob, jak její prvky zjistit, aniž bychom museli přímo použít procedury inverze.

6.1.3 Ortogonální projekce

Projekcemi na podprostory jsme se obecně zabývali v odstavci 4.2. To jsme ale neměli k dispozici skalární součin. Proto při každém promítání na podprostor $L \subset V_n$ bylo nutné také zadávat jeho konkrétní doplněk, jehož volba nebyla a priori jednoznačná. Při různé volbě doplňků jsme pak dostávali různé průměty jednoho a téhož vektoru do podprostoru L . Podstatou této nejednoznačnosti byla skutečnost, že šlo o lineární zobrazení, které by se z geometrického hlediska dalo nazvat *rovnoběžným promítáním*, přičemž se do podprostoru L promítalo „rovnoběžně“ právě s jeho předem zvoleným konkrétním doplňkem. Při jiné volbě doplňku dopadlo „rovnoběžně“ promítání jinak. Situaci si znovu můžeme připomenout návratem k obrázku 4.21 nebo 4.25. V unitárním prostoru však mezi všemi doplňky vektorového podprostoru vždy existuje právě jeden, který má význačné postavení:

Podprostor L_\perp vektorového prostoru U_n definovaný jako soubor všech vektorů ortogonálních ke každému prvku podprostoru L , tj.

$$L_\perp = \{b \in U_n \mid (a, b) = 0 \text{ pro libovolné } a \in L\},$$

nazýváme *ortogonálním doplňkem vektorového podprostoru L* .

Konstrukce tohoto doplňku je jednoduchá. Zvolme ortonormální bázi (u_1, u_2, \dots, u_k) k -rozměrného vektorového podprostoru L , $1 \leq k \leq n$, a libovolně ji doplňte tak, aby vznikla báze $(u_1, u_2, \dots, u_k, a_{k+1}, \dots, a_n)$ celého prostoru. Ortogonalizačním procesem získáme ortonormální bázi $\mathcal{U} = (u_1, u_2, \dots, u_n)$ v prostoru U_n . Položíme $L_\perp = [u_{k+1}, u_{k+2}, \dots, u_n]$. Vznikne $(n - k)$ -rozměrný doplněk podprostoru L , splňující předchozí definici. Co kdybychom však v U_n zvolili ortonormální bázi $(u_1, \dots, u_k, v_{k+1}, \dots, v_n)$, která by se od báze $(u_1, \dots, u_k, u_{k+1}, \dots, u_n)$ lišila vektory na posledních $n - k$ pozicích? Nic by se nestalo. Vektory v_{k+1} až v_n by generovaly též podprostor jako vektory u_{k+1} až u_n . Skutečně, označíme-li $L'_\perp = [v_{k+1}, \dots, v_n]$, kde

$\dim L'_\perp = n - k$, hned vidíme, že každý z vektorů v_{k+1} až v_n je prvkem prostoru L_\perp a naopak, každý z vektorů u_{k+1} až u_n je prvkem prostoru L'_\perp . Proto je $L_\perp = L'_\perp$. Z těchto úvah vyplývá:

Ortogonální doplněk L_\perp k podprostoru L je určen jednoznačně.

Pro libovolný vektor $a \in U_n$ existuje jednoznačný rozklad

$$a = a_L + a_{L_\perp}, \quad a_L \in L, \quad a_{L_\perp} \in L_\perp.$$

Zobrazení

$$\begin{aligned} \pi_L : U_n \ni a &\longrightarrow \pi_L(a) = a_L \in L, \\ \pi_{L_\perp} : U_n \ni a &\longrightarrow \pi_{L_\perp}(a) = a_{L_\perp} \in L_\perp \end{aligned}$$

se nazývají *ortogonální projekce* na podprostor L resp. L_\perp .

Vektor a_L resp. a_{L_\perp} je *ortogonální průmět* vektoru a do podprostoru L resp. L_\perp . Fakticky opět promítáme do L rovnoběžně s doplňkem, který je však tentokrát ortogonální k L — odtud terminologie. Jako každá projekce, i tyto jsou lineárními zobrazeními prostoru U_n do sebe. V bázích musí být obě projekce reprezentovány maticemi, pomocí kterých dokážeme promítnout každý vektor do L a do L_\perp . Tyto matice nyní najdeme. Než se pustíme do počítání, odvoďme obecné vlastnosti projekcí π_L a π_{L_\perp} , které se samozřejmě přenesou i na reprezentující matice.

Obrazem projekce π_L je podprostor L a jádrem L_\perp , u projekce π_{L_\perp} je tomu naopak. Z vlastností projekcí, které jsme studovali obecně v odstavci 4.2.4, plyne

$$\begin{aligned} \pi_L + \pi_{L_\perp} &= \text{id}_{U_n}, \quad \pi_L^2 = \pi_L, \quad \pi_{L_\perp}^2 = \pi_{L_\perp}, \\ \pi_L \circ \pi_{L_\perp} &= \pi_{L_\perp} \circ \pi_L = 0_{L(U_n, U_n)}. \end{aligned} \tag{6.9}$$

Označíme-li odpovídající matice těchto projekcí (v kterékoli bázi) jako P_L resp. P_{L_\perp} , platí pro ně odpovídající vztahy, tj.

$$\begin{aligned} P_L + P_{L_\perp} &= E, \quad P_L^2 = P_L, \quad P_{L_\perp}^2 = P_{L_\perp}, \\ P_L P_{L_\perp} &= P_{L_\perp} P_L = O. \end{aligned} \tag{6.10}$$

Nyní sestrojíme matice ortogonálních projekcí. Budeme používat ortonormálních bází, neboť to, jak již víme, přispěje ke zjednodušení výpočtů. Připravíme si k tomu jeden užitečný zápis rozkladu vektoru do složek v ortonormální bázi.

Příklad 6.13: Užitečný zápis složek vektoru

Zvolme ortonormální bázi $\mathcal{B} = (e_1, e_2, \dots, e_n)$, v níž budeme vyjadřovat všechny vektory i zobrazení.

Vynásobíme-li libovolný vektor $a \in U_n$ zprava některým vektorem báze, například e_1 , dostaneme

$$\begin{aligned} a &= \alpha^1 e_1 + \alpha^2 e_2 + \cdots + \alpha^n e_n \implies \\ \implies (a, e_1) &= \alpha^1 (e_1, e_1) + \alpha^2 (e_2, e_1) + \cdots + \alpha^n (e_n, e_1) = \alpha^1, \\ a &= (a, e_1) e_1 + (a, e_2) e_2 + \cdots + (a, e_n) e_n = \sum_{i=1}^n (a, e_i) e_i. \end{aligned} \quad (6.11)$$

Uvědomme si, že toto vyjádření můžeme použít jenom v ortonormální bázi, v obecné nikoliv.

Zvolme nyní vektorový podprostor L dimenze k a v něm bázi $\mathcal{U} = (u_1, u_2, \dots, u_k)$, rovněž ortonormální. Vektory báze \mathcal{U} jsou v bázi \mathcal{B} také zadány svými složkami, konkrétně

$$u_s = \gamma_s^1 e_1 + \gamma_s^2 e_2 + \cdots + \gamma_s^n e_n = \gamma_s^j e_j, \quad 1 \leq s \leq k.$$

Zapišme tyto složky do matice typu k/n ,

$$C = \begin{pmatrix} \gamma_1^1 & \gamma_1^2 & \cdots & \gamma_1^n \\ \gamma_2^1 & \gamma_2^2 & \cdots & \gamma_2^n \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_k^1 & \gamma_k^2 & \cdots & \gamma_k^n \end{pmatrix}.$$

Potřebujeme najít složky ortogonálního průmětu a_L vektoru a do podprostoru L v základní bázi \mathcal{B} . Označme

$$\begin{aligned} a &= \alpha^1 e_1 + \alpha^2 e_2 + \cdots + \alpha^n e_n \sim (\alpha), \\ a_L &= \alpha_L^1 e_1 + \alpha_L^2 e_2 + \cdots + \alpha_L^n e_n \sim (\alpha_L). \end{aligned}$$

Současně však vektor a_L leží v L , proto jej můžeme pomocí vztahu (6.11) zapsat jako

$$a_L = (a_L, u_1) u_1 + (a_L, u_2) u_2 + \cdots + (a_L, u_k) u_k = \sum_{s=1}^k (a_L, u_s) u_s = \sum_{s=1}^k (a, u_s) u_s.$$

V posledním výrazu předchozího vztahu — to není chyba. Skutečně platí $(a, u_s) = (a_L + a_{L^\perp}, u_s) = (a_L, u_s)$, neboť $(a_{L^\perp}, u_s) = 0$. Tento zápis je pro další postup důležitý. Průmět vektoru a do podprostoru L se nám totiž podařilo vyjádřit pomocí vektorů, jejichž složky v bázi \mathcal{B} známe. Při sčítání přes hodnoty indexu s jsme úmyslně nepoužili Einsteinovu symboliku, neboť sčítání probíhá pouze do hodnoty $k \leq n$. U sčítacích indexů nabývajících hodnot od 1 do n se jí však budeme držet. Po rozepsání vektorů do složek dostáváme

$$a_L = \sum_{s=1}^k (\alpha^i e_i, \gamma_s^j e_j) \gamma_s^\ell e_\ell = \alpha^i \left(\sum_{s=1}^k \gamma_s^{j*} \gamma_s^\ell \right) (e_i, e_j) e_\ell = \alpha^i (C^{T*} C)_j^\ell \delta_{ij} e_\ell \implies$$

$$\implies \alpha_L^\ell = \alpha^i P_i^\ell, \quad P_i^\ell = (C^{T*}C)_i^\ell,$$

kde jsme označili $P_L = (P_i^\ell)$, $i, \ell = 1, \dots, n$. V maticovém zápisu je

$$\begin{pmatrix} \alpha_L^1 & \alpha_L^2 & \dots & \alpha_L^n \end{pmatrix} = (\alpha^1 \ \alpha^2 \ \dots \ \alpha^n) \begin{pmatrix} P_1^1 & P_1^2 & \dots & P_1^n \\ P_2^1 & P_2^2 & \dots & P_2^n \\ \dots & \dots & \dots & \dots \\ P_n^1 & P_n^2 & \dots & P_n^n \end{pmatrix}, \quad \text{neboli}$$

$$(\alpha_L) = (\alpha)P_L, \quad P_L = C^{T*}C. \quad (6.12)$$

P_L je matice ortogonální projekce na podprostor L .

Příklad 6.14: Polynomů se nezabýváte

V příkladu 6.12 jsme potřebovali vyjádřit prvky standardní báze $\mathcal{A} = (1, x, x^2, x^3)$ v prostoru polynomů nejvýše stupně 3 pomocí báze $\mathcal{B} = (e_1(x), e_2(x), e_3(x), e_4(x))$, získané ortogonalizačním procesem. Nechtělo se nám však invertovat matici přechodu T od \mathcal{A} k \mathcal{B} a nyní se pokusíme tomuto nepříjemnému výpočtu vyhnout. Použijeme k tomu vztah (6.11). Označme jako obvykle prvky inverzní matice k T takto: $T^{-1} = S = (\sigma_i^j)$, $1 \leq i, j \leq n$. Pro i -tý polynom $a_i(x)$ báze \mathcal{A} platí

$$a_i(x) = \sum_{j=1}^4 (a_i(x), e_j(x)) e_j(x) = \sigma_i^j e_j(x) \implies \sigma_i^j = (a_i(x), e_j(x)) = \int_0^1 a_i(x) e_j(x) dx.$$

Stačí tedy vypočítat všechny potřebné skalární součiny, které jsou ovšem jednoduchými integrály z polynomů:

$$\sigma_1^1 = \int_0^1 dx = 1,$$

$$\sigma_1^2 = \int_0^1 (-\sqrt{3} + 2\sqrt{3}x) dx = 0,$$

$$\sigma_1^3 = \int_0^1 (\sqrt{5} - 6\sqrt{5}x + 6\sqrt{5}x^2) dx = 0,$$

$$\sigma_1^4 = \int_0^1 (-\sqrt{7} + 12\sqrt{7}x - 30\sqrt{7}x^2 + 20\sqrt{7}x^3) dx = 0,$$

$$\sigma_2^1 = \int_0^1 x dx = \frac{1}{2},$$

$$\sigma_2^2 = \int_0^1 x (-\sqrt{3} + 2\sqrt{3}x) dx = \frac{1}{2\sqrt{3}},$$

$$\sigma_2^3 = \int_0^1 x (\sqrt{5} - 6\sqrt{5}x + 6\sqrt{5}x^2) dx = 0,$$

$$\sigma_2^4 = \int_0^1 x (-\sqrt{7} + 12\sqrt{7}x - 30\sqrt{7}x^2 + 20\sqrt{7}x^3) dx = 0,$$

$$\sigma_3^1 = \int_0^1 x^2 dx = \frac{1}{3},$$

$$\sigma_3^2 = \int_0^1 x^2 (-\sqrt{3} + 2\sqrt{3}x) dx = \frac{1}{2\sqrt{3}},$$

$$\sigma_3^3 = \int_0^1 x^2 (\sqrt{5} - 6\sqrt{5}x + 6\sqrt{5}x^2) dx = \frac{1}{6\sqrt{5}},$$

$$\sigma_3^4 = \int_0^1 x^2 (-\sqrt{7} + 12\sqrt{7}x - 30\sqrt{7}x^2 + 20\sqrt{7}x^3) dx = 0,$$

$$\sigma_4^1 = \int_0^1 x^3 dx = \frac{1}{4},$$

$$\sigma_4^2 = \int_0^1 x^3 (-\sqrt{3} + 2\sqrt{3}x) dx = \frac{3\sqrt{3}}{20},$$

$$\sigma_4^3 = \int_0^1 x^3 (\sqrt{5} - 6\sqrt{5}x + 6\sqrt{5}x^2) dx = \frac{1}{4\sqrt{5}},$$

$$\sigma_4^4 = \int_0^1 x^3 (-\sqrt{7} + 12\sqrt{7}x - 30\sqrt{7}x^2 + 20\sqrt{7}x^3) dx = \frac{\sqrt{7}}{140}.$$

Je tedy

$$T^{-1} = S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & 0 & 0 \\ \frac{1}{3} & \frac{1}{2\sqrt{3}} & \frac{1}{6\sqrt{5}} & 0 \\ \frac{1}{4} & \frac{3\sqrt{3}}{20} & \frac{1}{4\sqrt{5}} & \frac{\sqrt{7}}{140} \end{pmatrix}.$$

Příklad 6.15: Pomocí algebry lze opravdu promítat

V příkladu 6.10 jsme sestrojili ortonormální bázi ve vektorovém podprostoru $L \subset U_4$,

$$L = [a_1, a_2, a_3] = [(1, 1, 0, 0), (0, i, 1, 0), (0, 0, i, -i)],$$

$$f_1 = \frac{c_1}{\sqrt{(c_1, c_1)}} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right),$$

$$f_2 = \frac{c_2}{\sqrt{(c_2, c_2)}} = \left(-\frac{i}{\sqrt{6}}, \frac{i}{\sqrt{6}}, \sqrt{\frac{2}{3}}, 0 \right),$$

$$f_3 = \frac{c_3}{\sqrt{(c_3, c_3)}} = \left(-\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{i}{2\sqrt{3}}, -\frac{i\sqrt{3}}{2} \right).$$

Matice složek těchto vektorů v původní ortonormální bázi (e_1, e_2, e_3, e_4) prostoru U_4 je

$$C = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ -\frac{i}{\sqrt{6}} & \frac{i}{\sqrt{6}} & \sqrt{\frac{2}{3}} & 0 \\ -\frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{i}{2\sqrt{3}} & -\frac{i\sqrt{3}}{2} \end{pmatrix}.$$

Matice ortogonální projekce na podprostor L je

$$P_L = C^{T*}C = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{6}} & -\frac{1}{2\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{6}} & \frac{1}{2\sqrt{3}} \\ 0 & \sqrt{\frac{2}{3}} & -\frac{i}{2\sqrt{3}} \\ 0 & 0 & \frac{i\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ -\frac{i}{\sqrt{6}} & \frac{i}{\sqrt{6}} & \sqrt{\frac{2}{3}} & 0 \\ -\frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{i}{2\sqrt{3}} & -\frac{i\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & \frac{i}{4} & \frac{i}{4} \\ \frac{1}{4} & \frac{3}{4} & -\frac{i}{4} & -\frac{i}{4} \\ -\frac{i}{4} & \frac{i}{4} & \frac{3}{4} & -\frac{1}{4} \\ -\frac{i}{4} & \frac{i}{4} & -\frac{1}{4} & \frac{3}{4} \end{pmatrix}.$$

Všimněte si, že matice P_L vyšla samoadjungovaná. Je to náhoda? Ze vztahu (6.12) plyne, že tomu tak být musí. Můžeme se přesvědčit:

$$P_L^{T*} = (C^{T*}C)^{T*} = C^{T*}(C^{T*})^{T*} = C^{T*}C = P_L.$$

V případě vektorového prostoru nad \mathbf{C} je matice ortogonální projekce samoadjungovaná, v případě vektorového prostoru nad \mathbf{R} je matice ortogonální projekce symetrická.

Máme-li matici projekce, můžeme už promítat libovolný vektor. Stačí zadat jeho složky v té bázi, ve které jsme vyjádřili $P(L)$, tj. v (e_1, e_2, e_3, e_4) . Můžeme tedy provádět geometrické operace, v tomto případě promítání, čistě algebraickou, dalo by se říci rutinní, procedurou. To je výhodné ve vícerozměrných prostorech a v prostorech nad komplexními čísly, kde geometrická představivost selhává. Máme však nějakou možnost kontroly, že jsme v pracném číselném výpočtu neudělali chybu? Kontrola nebude dokonalá, může však odhalení chyby napomoci. Známe totiž základní vlastnosti projekce (6.9) a matice P_L musí mít odpovídající vlastnosti (6.10), je-li správně určena. Dále víme, že vektory, které již v podprostoru L leží, by se promítnutím neměly změnit. Můžeme se o tom přesvědčit u vektorů a_1, a_2, a_3 , kterými jsme prostor L definovali. Tak třeba pro a_1 dostaneme

$$\pi_L(a_1) = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & \frac{i}{4} & \frac{i}{4} \\ \frac{1}{4} & \frac{3}{4} & -\frac{i}{4} & -\frac{i}{4} \\ -\frac{i}{4} & \frac{i}{4} & \frac{3}{4} & -\frac{1}{4} \\ -\frac{i}{4} & \frac{i}{4} & -\frac{1}{4} & \frac{3}{4} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix}.$$

A ještě jedna zkuška. Vektor $b = (-1, 1, i, i)$ je ke všem vektorům podprostoru L ortogonální. Přesvědčte se o tom výpočtem skalárních součinů (a_1, b) , (a_2, b) a (a_3, b) . Je tedy prvkem jádra projekce π_L a při promítání by se měl zobrazit na vektor nulový:

$$\pi_L(b) = \begin{pmatrix} -1 & 1 & i & i \end{pmatrix} \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & \frac{i}{4} & \frac{i}{4} \\ \frac{1}{4} & \frac{3}{4} & -\frac{i}{4} & -\frac{i}{4} \\ -\frac{i}{4} & \frac{i}{4} & \frac{3}{4} & -\frac{1}{4} \\ -\frac{i}{4} & \frac{i}{4} & -\frac{1}{4} & \frac{3}{4} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}.$$

Někdo si možná v tuto chvíli položí otázku: Matici P_L ortogonální projekce na podprostor L konstruujeme z vektorů ortonormální báze v tomto podprostoru. Volba báze, i když se omezuje jen na báze ortonormální, však není jednoznačná. Jak to dopadne s maticí projekce, zvolíme-li v L jinou bázi? Pochopitelně čekáme, že volba báze nemůže mít na matici P_L vliv. Je však třeba se o tom přesvědčit. Zvolme tedy v L ještě jinou bázi, například $(\bar{u}_1, \bar{u}_2, \dots, \bar{u}_k)$. Mezi oběma bázemi lze přecházet pomocí matice přechodu Q typu k/k , která je nejen regulární, ale dokonce unitární (obě báze, původní i nová, jsou ortonormální). Platí

$$\bar{u}_\alpha = \sum_{\beta=1}^k Q_\alpha^\beta u_\beta, \quad 1 \leq \alpha, \beta \leq k \implies \bar{\gamma}_\alpha^j = \sum_{\beta=1}^k Q_\alpha^\beta \gamma_\beta^j, \quad 1 \leq j \leq n,$$

$$\begin{pmatrix} \bar{\gamma}_1^1 & \bar{\gamma}_1^2 & \cdots & \bar{\gamma}_1^n \\ \bar{\gamma}_2^1 & \bar{\gamma}_2^2 & \cdots & \bar{\gamma}_2^n \\ \cdots & \cdots & \cdots & \cdots \\ \bar{\gamma}_k^1 & \bar{\gamma}_k^2 & \cdots & \bar{\gamma}_k^n \end{pmatrix} = \begin{pmatrix} Q_1^1 & Q_1^2 & \cdots & Q_1^k \\ Q_2^1 & Q_2^2 & \cdots & Q_2^k \\ \cdots & \cdots & \cdots & \cdots \\ Q_k^1 & Q_k^2 & \cdots & Q_k^k \end{pmatrix} \begin{pmatrix} \gamma_1^1 & \gamma_1^2 & \cdots & \gamma_1^n \\ \gamma_2^1 & \gamma_2^2 & \cdots & \gamma_2^n \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_k^1 & \gamma_k^2 & \cdots & \gamma_k^n \end{pmatrix} \implies \bar{C} = QC.$$

Pro matici projekce v nové bázi dostáváme

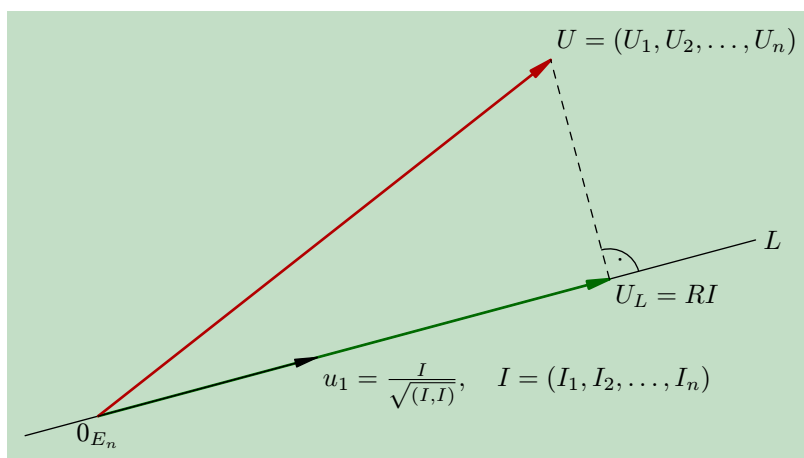
$$\bar{P}_L = \bar{C}^{T*} \bar{C} = (QC)^{T*} (QC) = C^{T*} Q^{T*} QC = C^{T*} C = P_L.$$

Využili jsme při tom unitárnosti matice Q , tj. vlastnosti $Q^{T*}Q = Q^{-1}Q = E$.

Příklad 6.16: K čemu mohou být projekce dobré? Metoda nejmenších čtverců

Věříte, že promítání není jen geometrická hříčka a může mít význam i při zpracování fyzikálních měření? V odstavci 3.3.3 kapitoly o pravděpodobnostech jsme se setkali s určováním elektrického odporu vodiče z měření dvojic odpovídajících si hodnot napětí a proudu. Provedli jsme proložení lineární závislosti těmito dvojicemi tzv. metodou nejmenších čtverců. Že musí být možné zpracovat tuto úlohu i pomocí lineární algebry, je zřejmé. Ve hře je přece úměra — lineární závislost proudu na napětí. Co to má však společného s promítáním? Uvažujme. Máme k dispozici n naměřených hodnot napětí $U = (U_1, U_2, \dots, U_n)$ a n odpovídajících hodnot proudu $I = (I_1, I_2, \dots, I_n)$. Proud I_j protékal vodičem, když jsme na voltmetru nastavili napětí U_j . Veličiny U a I můžeme chápat jako vektory v n -rozměrném vektorovém prostoru nad \mathbf{R} opatřeném skalárním součinem, tedy v prostoru euklidovském E_n . Jednotlivé hodnoty interpretujeme jako složky v ortonormální bázi tohoto prostoru (e_1, e_2, \dots, e_n) . Podle Ohmova zákona by vektory U a I měly být rovnoběžné, tj. $U = RI$. Protože však měření není přesné, vektory zcela jistě rovnoběžné nebudou. Situaci schematicky znázorňuje obrázek 6.2. Potřebujeme tedy najít takový vektor $U_L = RI$, který bude s vektorem I kolineární (rovnoběžný) a přitom bude „co nejvíce blízký“ vektoru U . Tyto podmínky splňuje ortogonální průmět vektoru U do jednorozměrného vektorového podprostoru L určeného vektorem I . Stačí tedy nalézt matici projekce. Ortonormální báze podprostoru L bude určena normovaným vektorem

$$u_1 = \frac{I}{\sqrt{(I, I)}} = \left(\frac{I_1}{\sqrt{\sum_{i=1}^n I_i^2}}, \frac{I_2}{\sqrt{\sum_{i=1}^n I_i^2}}, \dots, \frac{I_n}{\sqrt{\sum_{i=1}^n I_i^2}} \right).$$



Obr. 6.2 Promítání a Ohmův zákon.

Tímto vektorem je současně určena matice C , která je typu $1/n$. Matice projekce má tvar

$$P_L = \begin{pmatrix} \frac{I_1^2}{\sum_{i=1}^n I_i^2} & \frac{I_1 I_2}{\sum_{i=1}^n I_i^2} & \cdots & \frac{I_1 I_n}{\sum_{i=1}^n I_i^2} \\ \frac{I_1 I_2}{\sum_{i=1}^n I_i^2} & \frac{I_2^2}{\sum_{i=1}^n I_i^2} & \cdots & \frac{I_2 I_n}{\sum_{i=1}^n I_i^2} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{I_1 I_n}{\sum_{i=1}^n I_i^2} & \frac{I_2 I_n}{\sum_{i=1}^n I_i^2} & \cdots & \frac{I_n^2}{\sum_{i=1}^n I_i^2} \end{pmatrix}.$$

Průmět vektoru U do podprostoru L má složky

$$U_L = (U_1, U_2, \dots, U_n)P_L = \left(I_1 \frac{\sum_{i=1}^n I_i U_i}{\sum_{i=1}^n I_i^2}, I_2 \frac{\sum_{i=1}^n I_i U_i}{\sum_{i=1}^n I_i^2}, \dots, I_n \frac{\sum_{i=1}^n I_i U_i}{\sum_{i=1}^n I_i^2} \right).$$

Vektor U_L je úměrný vektoru I , konstantou úměrnosti je hodnota

$$R = \frac{\sum_{i=1}^n I_i U_i}{\sum_{i=1}^n I_i^2}.$$

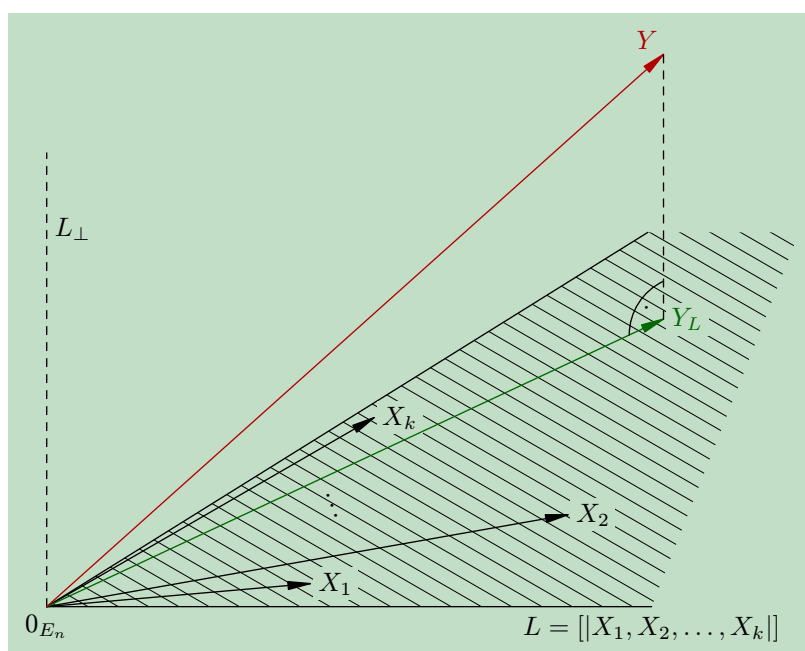
Ta představuje hodnotu elektrického odporu, která nejlépe vyhovuje naměřeným hodnotám. Výsledek souhlasí se vztahem (3.31), který jsme odvodili v odstavci 3.3.3 úplně jiným postupem. Nezdá se vám algebraická metoda snazší a názornější?

Příklad 6.17: Promítání a vícerozměrná lineární regrese

V předchozím příkladu jsme použili ortogonálního promítání pro nalezení nejlepší shody naměřených dat s modelovou lineární závislostí v nejjednodušším možném případě. Fyzikální veličina, chápaná jako závisle proměnná (napětí na vodiči), byla lineární funkcí jediné nezávisle proměnné veličiny (proudu, který vodičem protéká). Promítnutím „vektoru“ napětí, získaného jako n -tice jednotlivých měření napětí, do směru „vektoru“ proudu, tvořeného příslušnou n -ticí měření proudu, jsme získali konstantu úměrnosti R odpovídající nejlepší shodě naměřených hodnot s lineární závislostí $U = RI$, tj. elektrický odpor vodiče. Tuto myšlenku můžeme zobecnit i na případ, kdy veličina má vyhovovat lineárnímu modelu závislosti na větším počtu jiných veličin, například k , tj.

$$Y = A^1 X_1 + A^2 X_2 + \cdots + A^k X_k.$$

Všechny zúčastněné veličiny měříme n -krát, přičemž $n > k$. Máme tedy dispozici „vektory“ $X_1 = (X_1^1, X_1^2, \dots, X_1^n)$ až $X_k = (X_k^1, X_k^2, \dots, X_k^n)$ a $Y = (Y^1, Y^2, \dots, Y^n)$, hledáme neznámé konstanty A^1 až A^k . Podobně jako v případě experimentálního ověřování Ohmova zákona bude soustava n rovnic $Y^j = A^1 X_1^j + A^2 X_2^j + \dots + A^k X_k^j$, $j = 1, 2, \dots, n$, o k neznámých A^1 až A^k „přeurčená“ (více nezávislých rovnic než neznámých) a nebude tedy mít řešení. Musíme se spokojit s tím, že se nám nepodaří nalézt konstanty A^1, A^2, \dots, A^k tak, aby soustava rovnic byla splněna přesně. Můžeme je však zvolit tak, aby byla splněna „co nejlépe“. Veličiny X_1, X_2, \dots, X_k si představíme jako vektory v n -rozměrném euklidovském vektorovém prostoru E_n , které generují k -rozměrný vektorový podprostor, veličina Y je rovněž vektorem v E_n . Skalární součin v tomto prostoru bude definován tak, že naměřené hodnoty každé z veličin představují složky odpovídajícího vektoru v bázi, která je vůči tomuto skalárnímu součinu ortonormální. To proto, abychom skalární součin vektorů mohli počítat podle vztahu (6.6). Obrázek 6.3 názorně ukazuje, že nejlépe bude hledanou lineární závislost splňovat ortogonální průmět



Obr. 6.3 Promítání a vícerozměrná lineární regrese.

Y_L vektoru Y do podprostoru $L = [X_1, X_2, \dots, X_k]$,

$$Y_L = A^1 X_1 + A^2 X_2 + \dots + A^k X_k. \quad (6.13)$$

Z předchozí vektorové rovnice lze konstanty A^1 až A^k jednoznačně určit. Stačí sestavit matici ortogonální projekce na podprostor L a promítnout vektor Y . Protože však nic nemůže být tak jednoduché, jak se na první pohled zdá, i zde je malá komplikace. Nikoli však principiální, nýbrž pouze výpočetní. Matici projekce P_L nelze konstruovat z vektorů X_1 až X_k přímo, neboť obecně netvoří ortonormální bázi. Tu je třeba teprve sestavit pomocí ortogonalizačního procesu. Ortogonalizační proces však může být poněkud pracný, hlavně když je proměnných X_1, X_2, \dots, X_k větší počet. Nemůžeme se mu nějak vyhnout? Takovou možnost hned uvidíme, vynásobíme-li rovnost (6.13) libovolným vektorem X_j . Pak

$$\begin{aligned} (Y_L, X_j) &= (A^1 X_1 + A^2 X_2 + \dots + A^k X_k, X_j) = \\ &= A^1 (X_1, X_j) + A^2 (X_2, X_j) + \dots + A^k (X_k, X_j). \end{aligned}$$

434 KAPITOLA 6. LINEÁRNÍ ALGEBRA DO TŘETICE

Uvědomíme-li si, že $Y = Y_L + Y_{L^\perp}$ a tedy $(Y_L, X_j) = (Y, X_j)$, dostáváme soustavu k rovnic pro neznámé konstanty A_1 až A_k :

$$\begin{aligned} (Y, X_1) &= A^1(X_1, X_1) + A^2(X_2, X_1) + \dots + A^k(X_k, X_1), \\ (Y, X_2) &= A^1(X_1, X_2) + A^2(X_2, X_2) + \dots + A^k(X_k, X_2), \\ &\dots\dots\dots \\ (Y, X_k) &= A^1(X_1, X_k) + A^2(X_2, X_k) + \dots + A^k(X_k, X_k). \end{aligned}$$

Tato soustava se nazývá *normální systém rovnic*. S jejím vyřešením je vyřešen i problém nejlepší shody naměřených dat s modelovou lineární závislostí. Pro ilustraci uvedme ještě konkrétní číselný příklad. Nechť $Y = A^1 X_1 + A^2 X_2$ je lineární model fyzikální závislosti, tj. $k = 2$. Provedli jsme čtyři měření a získali vektory

$$X_1 = (1, 2, 3, 4), \quad X_2 = (2, 3, 4, 5), \quad Y = (2, 3, 5, 8).$$

Vypočteme potřebné skalární součiny a sestavíme normální systém rovnic:

$$\begin{aligned} (Y, X_1) &= 55, & (Y, X_2) &= 73, \\ (X_1, X_1) &= 30, & (X_1, X_2) &= (X_2, X_1) = 40, & (X_2, X_2) &= 54, \\ 30A^1 + 40A^2 &= 55, \\ 40A^1 + 54A^2 &= 73. \end{aligned}$$

Její řešení je $A^1 = 2,5$, $A^2 = -0,5$. Průmět vektoru Y do podprostoru L je $Y_L = 2,5 X_1 - 0,5 X_2 = (1,5, 3,5, 5,5, 7,5)$. „Pro pořádek“ ještě prověříme, že ke stejnému výsledku lze dospět pomocí matice projekce. Nejprve sestrojíme v L ortonormální bázi (U_1, U_2) a poté matici projekce:

$$\begin{aligned} V_1 = X_1 &= (1, 2, 3, 4), & V_2 &= -\frac{(X_2, X_1)}{(X_1, X_1)} X_1 + X_2 = \left(\frac{2}{3}, \frac{1}{3}, 0, -\frac{1}{3}\right), \\ U_1 &= \left(\frac{1}{\sqrt{30}}, \frac{2}{\sqrt{30}}, \frac{3}{\sqrt{30}}, \frac{4}{\sqrt{30}}\right), & U_2 &= \left(\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, 0, -\frac{1}{\sqrt{6}}\right), \\ P_L = C^T C &= \begin{pmatrix} \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{6}} \\ \frac{2}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ \frac{3}{\sqrt{30}} & 0 \\ \frac{4}{\sqrt{30}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{30}} & \frac{3}{\sqrt{30}} & \frac{4}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} \frac{7}{10} & \frac{4}{10} & \frac{1}{10} & -\frac{2}{10} \\ \frac{4}{10} & \frac{3}{10} & \frac{2}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{4}{10} \\ -\frac{2}{10} & \frac{1}{10} & \frac{4}{10} & \frac{7}{10} \end{pmatrix}, \\ Y_L &= \begin{pmatrix} 2 & 3 & 5 & 8 \end{pmatrix} \begin{pmatrix} \frac{7}{10} & \frac{4}{10} & \frac{1}{10} & -\frac{2}{10} \\ \frac{4}{10} & \frac{3}{10} & \frac{2}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{2}{10} & \frac{3}{10} & \frac{4}{10} \\ -\frac{2}{10} & \frac{1}{10} & \frac{4}{10} & \frac{7}{10} \end{pmatrix} = \begin{pmatrix} \frac{3}{2} & \frac{7}{2} & \frac{11}{2} & \frac{15}{2} \end{pmatrix}. \end{aligned}$$

Tento výsledek souhlasí s výpočtem pomocí normálního systému rovnic. Ze vztahu $Y_L = A^1 X_1 + A^2 X_2$ je ještě třeba určit hodnoty A^1 a A^2 . Platí

$$A^1 + 2A^2 = \frac{3}{2},$$

$$\begin{aligned}2A^1 + 3A^2 &= \frac{7}{2}, \\3A^1 + 4A^2 &= \frac{11}{2}, \\4A^1 + 5A^2 &= \frac{15}{2}.\end{aligned}$$

Vyřešte tuto soustavu a uvidíte, že vyjde $A^1 = 2,5$ a $A^2 = -0,5$.

6.1.4 Cvičení

- Rozhodněte, zda operace $(a, b) : V \times V \ni [a, b] \rightarrow (a, b) \in \mathbf{R}$ splňuje axiomy skalárního součinu ve vektorovém prostoru nad tělesem reálných čísel:
 - $V = \mathbf{R}^2$, $((x_1, y_1), (x_2, y_2)) = x_1y_1x_2y_2$,
 - $V = \mathbf{P}[n]$, $(P(x), Q(x)) = P(0)Q(0) + P(1)Q(1) + \dots + P(n)Q(n)$,
 - $V = \mathcal{M}(2/2)$, $(A, B) = \det(AB)$,
 - $V = \mathbf{R}^3$, $((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 - x_2y_2 + x_3y_3$,
 - $V = \mathbf{P}[n]$, $n \geq 1$, $(P(x), Q(x)) = P(0)Q(0)$,
 - $V = \mathcal{M}(2/2)$, $(A, B) = \operatorname{tr}A \operatorname{tr}B$, kde $\operatorname{tr}A$ značí *stopu* matice A , tj. součet jejích diagonálních prvků $a_{11} + a_{22}$,
 - $V = \mathcal{M}(n/m)$, $(A, B) = \operatorname{tr}(AB^T)$.
- Rozhodněte, zda operace $(a, b) : V \times V \ni [a, b] \rightarrow (a, b) \in \mathbf{C}$ splňuje axiomy skalárního součinu ve vektorovém prostoru nad tělesem komplexních čísel:
 - $V = \mathbf{C}$, $(x + iy, u + iv) = xu + yv - i xv + i yu$,
 - $V = \mathcal{M}(2/2)$, $(A, B) = \operatorname{tr}(AB^T)$,
 - $V = \mathbf{C}$, $(z, w) = zw^*$,
 - $V = \mathcal{M}(2/2)$, $(A, B) = \operatorname{tr}(AB^{T*})$,
 - $V = \mathbf{C}^2$, $((z_1, z_2), (w_1, w_2)) = z_1w_1 + z_2w_2$.
- Dokažte následující vlastnosti skalárního součinu a délky (zvané též *norma*) v euklidovských prostorech, vektory a, b jsou libovolné:
 - $(a, o) = 0$, kde o je nulový vektor,
 - $(a, b)^2 \leq (a, a)(b, b)$ (Cauchyova–Buňakovského nerovnost),
 - $(a, b)^2 = (a, a)(b, b) \iff a = kb$, $k \in \mathbf{R}$,
 - $(a, b) = \frac{1}{4}(|a + b|^2 - |a - b|^2)$,
 - $|ka| = |k| |a|$, $k \in \mathbf{R}$,
 - $|a + b| \leq |a| + |b|$ (trojúhelníková nerovnost),
 - $|a + b|^2 + |a - b|^2 = 2(|a|^2 + |b|^2)$ (rovnoběžníková rovnost),
 - a je ortogonální k b právě tehdy, když $|a + b| = |a - b|$,

i) $a + b$ je ortogonální k $a - b$ právě tehdy, když $|a| = |b|$.

Návod: Použijte axiomů definujících skalární součin, v části b) aplikujte axiom pozitivní definitnosti skalárního součinu na vektor $a + tb$, tj. $(a + tb, a + tb) \geq 0$ pro libovolné $t \in \mathbf{R}$ a využijte vlastností diskriminantu kvadratického trojčlenu, v části c) předpokládejte vektor b ve tvaru $b = ka + c$, kde $(a, c) = 0$, a ukažte, že c je nulový vektor, v části f) využijte platnosti b).

- *4. Které ze vztahů v předchozím příkladu nebudou obecně platit v prostorech unitárních? Nalezněte protipříklady. (V částech 3c) a 3e) uvažujte obecně $k \in \mathbf{C}$.)
5. Ve vektorovém prostoru V nad \mathbf{R} se zadaným skalárním součinem určete normu vektorů u a v , vektory normujte, vypočtěte jejich skalární součin a kosinus jejich odchylky.

a) $V = \mathbf{R}^2$,

$$G = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad u = (1, 3), \quad v = (0, 4),$$

b) $V = \mathbf{P}[3]$, $(P(x), Q(x)) = \int_0^1 P(x)Q(x) dx$, $u = x^3$, $v = x - 2$,

c) $V = \mathcal{M}(2/3)$, $(A, B) = \text{tr}(AB^T)$,

$$u = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 0 \end{pmatrix}, \quad v = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

6. Zjistěte, zda vztahy $(e_1, e_1) = 3$, $(e_2, e_1) = -2i$, $(e_1, e_2) = 2i$, $(e_2, e_2) = 2$, kde (e_1, e_2) je báze ve V^2 , je definován skalární součin. V kladném případě najděte libovolnou ortonormální bázi.
7. Určete ortogonální průmět vektoru $a = (1, 1, 0, 2)$ do směru vektoru $b = (1, -i, i, 1)$. Složky jsou zadány v ortonormální bázi.
8. Nechť $a = (-i, 0)$, $b = (1, i) \in V_2$ v bázi (e_1, e_2) . Definujte skalární součin tak, aby vektory a, b byly vzhledem k němu ortogonální. Určete matici tohoto skalárního součinu v bázi (e_1, e_2) .
9. Ortogonalizujte systém vektorů (u, v, w) v prostorech se zadaným skalárním součinem. Pořadí vektorů systému (u, v, w) můžete změnit s cílem zjednodušit výpočet.

a) $V = \mathbf{P}[2]$ se skalárním součinem $(P(x), Q(x)) = \int_0^1 P(x)Q(x)dx$, $u = 1 + x$, $v = 1 - x$, $w = x$,

b) $V = \mathbf{R}^4$, složky jsou zadány v ortonormální bázi, $u = (1, 2, 0, 0)$, $v = (0, 1, 0, -1)$, $w = (-1, 0, 0, 1)$,

c) $V = \mathbf{R}^3$,

$$G = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 3 \end{pmatrix},$$

$$u = (1, 0, 0), \quad v = (0, 1, 0), \quad w = (0, 0, 1).$$

10. Určete ortogonální doplněk k podprostoru L ve V :

a) $V = \mathbf{R}^3$, $L = [(1, 1, 0)]$, složky jsou zadány v ortonormální bázi,

b) $V = \mathbf{R}^3$, $L = [(1, 1, 0)]$,

$$G = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 3 \end{pmatrix},$$

$$c) V = \mathbf{P}[2], (P(x), Q(x)) = \int_0^1 P(x)Q(x)dx, L = [x^2 + 1, 1].$$

11. Určete matice projekce na zadané podprostory v \mathbf{C}^4 , složky vektorů jsou zadány v ortonormální bázi.

$$a) L = [(1, 0, 1, 0), (1, 1, 1, 0)],$$

$$b) L = [(1, -1, 2, 0), (-1, 1, 1, 1), (0, 0, 2, 1)],$$

$$c) L = [(0, i, i, 0)].$$

12. Skalární součin v $P[2]$ je dán vztahem $(P(x), Q(x)) = \int_0^1 P(x)Q(x) dx$. Určete ortogonální průmět vektoru $v = x^2$ do podprostoru $L = [x + 1, 9x - 5]$ a ortogonální průmět tohoto vektoru do L_{\perp} . Použijte normální systém rovnic.

13. Předpokládejte, že pro vektory $a_1, \dots, a_k \in U_n$ platí $(a_i, a_j) = 0$, $i, j = 1, \dots, n$, $i \neq j$ (vektory jsou ortogonální). Dokažte, že je-li tento systém lineárně závislý, je alespoň jeden z jeho prvků nulový.

6.2 „Fyzikální“ lineární operátory a jejich vlastní vektory

Fyzikální prostory, tedy prostory přizpůsobené popisu fyzikálních událostí, většinou vyžadují dodatečnou výbavu, kterou představuje právě vyložený skalární součin. Vzpomeňme si jen na dva nejjednodušší fyzikální příklady: elementární mechanická práce vykonaná určitou silou působící na těleso je skalárním součinem této síly a elementárního posunutí jejího působíště, elementární indukční tok magnetického pole je skalárním součinem vektoru magnetické indukce a vektoru popisujícího elementární orientovanou plošku. A našli bychom další. Má tedy jistě také smysl uvažovat o tom, zda mezi lineárními operátory působícími na vektory takového prostoru nejsou některé, jejichž chování vůči skalárnímu součinu je něčím výjimečné. Soustředíme se na dva takové typy operátorů, které mají zároveň významné fyzikální využití. Toto využití úzce souvisí s řešením problému vlastních vektorů těchto typů operátorů.

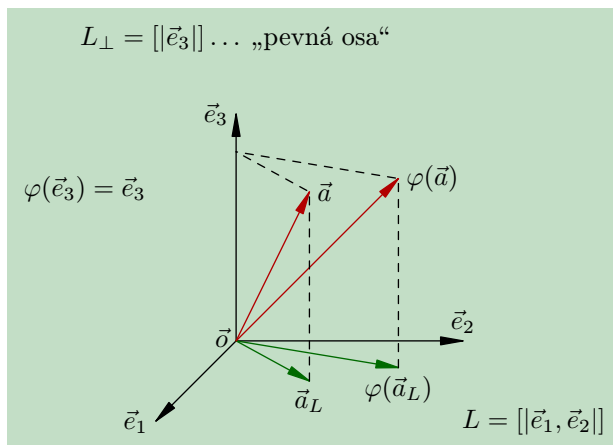
6.2.1 Unitární (ortogonální) operátory

Fyzikální, popřípadě geometrickou situaci, která navozuje definici unitárního operátoru, si dokážeme představit velmi snadno.

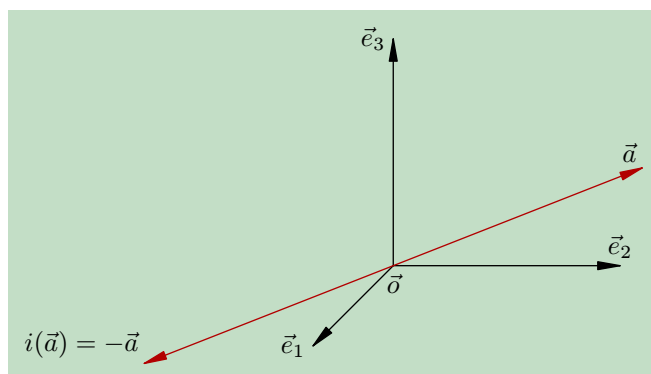
Příklad 6.18: Rotace, inverze a zrcadlení

Ve vektorovém prostoru V_3 nad polem reálných čísel tvořeném orientovanými úsečkami se společným počátkem jsme definovali skalární součin vektorů vztahem (6.1) a vytvořili tak euklidovský prostor E_3 . Tuto definici nám umožnila skutečnost, že pro orientované úsečky umíme prostřednictvím elementární geometrie definovat jejich velikosti a úhly mezi nimi. V příkladu 4.40 jsme pracovali s lineárním operátorem, který představoval otočení libovolné úsečky v rovině o pevně zvolený úhel. Tuto operaci bychom mohli zobecnit na případ trojrozměrného prostoru tak, že doplníme dvojrozměrný podprostor $L \subset E_3$ jeho ortogonálním doplňkem a budeme vektory otáčet při zachování jejich délky tak, že otáčíme jejich ortogonální průměty do L způsobem popsáním v příkladu 4.40. Ve fyzice, typicky například v mechanice rotačního pohybu tuhých těles, obvykle hovoříme o *rotaci kolem pevné osy*, která je předem zvolena. Skutečně, vektory náležející do L_{\perp} se při rotaci nezmění,

přesněji řečeno, obraz každého z nich splýne se vzorem. Obdobně si můžeme představit *rotaci kolem pevného bodu* (obecnou rotaci). Při ní se rovněž zachovává délka vektoru a úhel mezi každou dvojicí vektorů. Zdá se však, že zde již nebude obecně existovat vektorový podprostor, jehož prvky zůstávají při rotaci pevné. (Výjimkou je pochopitelně vektor nulový, pro který platí $\varphi(\vec{0}) = \vec{0}$ při libovolném lineárním zobrazení φ .) K této „názorné“ úvaze se ještě vrátíme. Ke stejným závěrům o délkách a úhlech dospějeme při studiu středové inverze, která převádí vektory ve vektory opačné, popřípadě při studiu operací zrcadlení vzhledem k různým rovinám. Oba



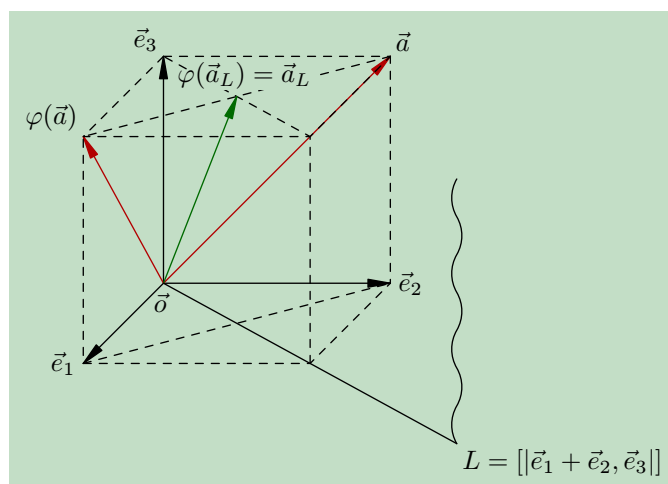
Obr. 6.4 Rotace kolem pevné osy.



Obr. 6.5 Středová inverze.

uvedené typy rotace, středovou inverzi i zrcadlení můžeme z hlediska algebry zapsat pomocí lineárních operátorů ve E_3 . Položme si otázku, jak souvisí skalární součin libovolných dvou vektorů se skalárním součinem jejich obrazů při zobrazení kterýmkoli z uvedených operátorů. Odpověď je jednoduchá. Jestliže operátor zachovává velikosti vektorů a úhly mezi nimi, zachovává nutně i skalární součin, který je pomocí velikostí a úhlů přímo definován. Je také jasné, že každý lineární operátor představující kteroukoli z těchto operací bude „převádět“ libovolnou ortonormální bázi opět v ortonormální a bude tedy v každé ortonormální bázi reprezentován ortogonální maticí. Skutečně, vyjádříme-li obrazy $\vec{f}_1 = \varphi(\vec{e}_1)$, $\vec{f}_2 = \varphi(\vec{e}_2)$, $\vec{f}_3 = \varphi(\vec{e}_3)$ vektorů báze $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ jako její lineární kombinace, dostaneme

$$\begin{aligned}\vec{f}_1 &= \alpha_1^1 \vec{e}_1 + \alpha_1^2 \vec{e}_2 + \alpha_1^3 \vec{e}_3 \sim (\alpha_1), \\ \vec{f}_2 &= \alpha_2^1 \vec{e}_1 + \alpha_2^2 \vec{e}_2 + \alpha_2^3 \vec{e}_3 \sim (\alpha_2),\end{aligned}$$



Obr. 6.6 Zrcadlení.

$$\vec{f}_3 = \alpha_1^3 \vec{e}_1 + \alpha_2^3 \vec{e}_2 + \alpha_3^3 \vec{e}_3 \sim (\alpha_3),$$

trojice (α_1) , (α_2) a (α_3) jsou tvořeny složkami obrazů vektorů báze $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ vzhledem k této bázi. Obrazy ovšem tvoří ortonormální bázi, stejně jako vzory. Pro jejich skalární součiny proto platí

$$\vec{f}_i \vec{f}_j = \begin{pmatrix} \alpha_i^1 & \alpha_i^2 & \alpha_i^3 \end{pmatrix} \begin{pmatrix} \alpha_j^1 \\ \alpha_j^2 \\ \alpha_j^3 \end{pmatrix} = \delta_{ij} \implies AA^T = E \implies A^T = A^{-1},$$

kde $A = (\alpha_i^k)$, $i, k = 1, 2, 3$. Determinant ortogonální matice je 1 (pro rotaci) nebo -1 (pro středovou inverzi resp. zrcadlení), neboť $\det(AA^T) = (\det A)^2 = 1$. Důležitou vlastností operátoru rotace, inverze či zrcadlení je jeho regulárnost (bázi převádí opět v bázi, při zobrazení nedochází k žádné „degeneraci“, obvyklé například při projekcích). Jádrem tohoto operátoru je tvořeno pouhým nulovým vektorem, obrazem je celý prostor. Pro konkrétní představu zapišme ještě matice operátorů rotace kolem osy určené vektorem \vec{e}_3 , rotace kolem bodu (počátku soustavy souřadnic, odpovídajícího vektoru $\vec{0}$), inverze a zrcadlení například vzhledem k rovině určené vektory \vec{e}_1 a \vec{e}_2 . Pro případ rotace kolem osy označme α úhel otočení, pro případ rotace kolem bodu označme γ_{ij} úhel mezi obrazem $\varphi(\vec{e}_i)$ vektoru \vec{e}_i a vektorem \vec{e}_j . Pro *směrové kosiny* platí

$$\sum_{j=1}^3 \cos \gamma_{ij} \cos \gamma_{kj} = \delta_{ik}.$$

Přijďte na to, proč? Jakou další vlastnost mají směrové kosiny, popisují-li opravdu operátor otočení a nikoli třeba zrcadlení nebo středovou inverzi? Možná vám napoví představa, že při rotaci vytvoří obrazy vektorů pravotočivé báze opět bázi pravotočivou.

Operátor rotace kolem pevné osy spojené s vektorem \vec{e}_3 , operátor obecné rotace, operátor středové inverze a operátor zrcadlení vzhledem k rovině určené vektory \vec{e}_1 a \vec{e}_2 označme postupně $\varphi_1, \varphi_2, \varphi_3, \varphi_4$. Dostáváme:

$$\begin{aligned} \varphi_1(\vec{e}_1) &= (\cos \alpha) \vec{e}_1 + (\sin \alpha) \vec{e}_2 + 0 \vec{e}_3, \\ \varphi_1(\vec{e}_2) &= -(\sin \alpha) \vec{e}_1 + (\cos \alpha) \vec{e}_2 + 0 \vec{e}_3, \\ \varphi_1(\vec{e}_3) &= 0 \vec{e}_1 + 0 \vec{e}_2 + \vec{e}_3, \end{aligned}$$

$$\begin{aligned}\varphi_2(\vec{e}_1) &= (\cos \gamma_{11})\vec{e}_1 + (\cos \gamma_{12})\vec{e}_2 + (\cos \gamma_{13})\vec{e}_3, \\ \varphi_2(\vec{e}_2) &= (\cos \gamma_{21})\vec{e}_1 + (\cos \gamma_{22})\vec{e}_2 + (\cos \gamma_{23})\vec{e}_3, \\ \varphi_2(\vec{e}_3) &= (\cos \gamma_{31})\vec{e}_1 + (\cos \gamma_{32})\vec{e}_2 + (\cos \gamma_{33})\vec{e}_3,\end{aligned}$$

$$\begin{aligned}\varphi_3(\vec{e}_1) &= -\vec{e}_1, \\ \varphi_3(\vec{e}_2) &= -\vec{e}_2, \\ \varphi_3(\vec{e}_3) &= -\vec{e}_3,\end{aligned}$$

$$\begin{aligned}\varphi_4(\vec{e}_1) &= \vec{e}_1, \\ \varphi_4(\vec{e}_2) &= \vec{e}_2, \\ \varphi_4(\vec{e}_3) &= -\vec{e}_3.\end{aligned}$$

Matice operátorů v bázi $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ jsou, opět postupně,

$$\begin{aligned}A_1 &= \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, & A_2 &= \begin{pmatrix} \cos \gamma_{11} & \cos \gamma_{12} & \cos \gamma_{13} \\ \cos \gamma_{21} & \cos \gamma_{22} & \cos \gamma_{23} \\ \cos \gamma_{31} & \cos \gamma_{32} & \cos \gamma_{33} \end{pmatrix}, \\ A_3 &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, & A_4 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.\end{aligned}$$

Všechny tyto matice jsou ortogonální (ověřte pomocí vztahu $AA^T = E$. Nepožadujeme-li, aby obecná matice $A = (\cos \gamma_{ij})$, $1 \leq i, j \leq 3$, převáděla pravotočivou bázi opět v pravotočivou, bude tato matice obsahovat ostatní tři (matici pro rotaci kolem osy, inverzi, zrcadlení) jako speciální případy. Najděte pro tyto případy *směrové úhly* γ_{ij} .

Vlastnost zachování skalárního součinu, která byla důsledkem definice geometrických operací v předchozím příkladu, se při zobecnění na případ n -rozměrného vektorového prostoru nad \mathbf{C} stane přímo součástí definice.

Uvažujme o unitárním vektorovém prostoru U_n nad \mathbf{C} .

Lineární operátor $\varphi : U_n \rightarrow U_n$ se nazývá *unitární* (v prostoru euklidovském, tj. nad \mathbf{R} , *ortogonální*), jestliže pro libovolné vektory $a, b \in U_n$ platí

$$(\varphi(a), \varphi(b)) = (a, b). \quad (6.14)$$

V příkladu 6.18 jsme na základě geometrické představy konstatovali, že operátory rotace, inverze a zrcadlení budou v ortonormálních bázích reprezentovány ortogonálními maticemi. Pro konkrétní příklady jsme tuto vlastnost ověřili. Pro případ zobecněné definice (6.14) však musíme obdobné tvrzení dokázat.

Podle základní věty o lineárních zobrazeních víme, že lineární operátor je jednoznačně určen zadáním obrazů libovolné báze. Zvolme tedy libovolnou *ortonormální* bázi v U_n , $\mathcal{B} = (e_1, e_2, \dots, e_n)$, $(e_i, e_j) = \delta_{ij}$. Označíme-li $A = (\alpha_j^i)$, $1 \leq i, j \leq n$, matici unitárního operátoru φ , která jej v této bázi reprezentuje, platí

$$\begin{aligned}\varphi(e_i) &= \alpha_i^k e_k, \\ \delta_{ij} &= (e_i, e_j) = (\varphi(e_i), \varphi(e_j)) = (\alpha_i^k e_k, \alpha_j^\ell e_\ell) = \\ &= \alpha_i^k (\alpha_j^\ell)^* (e_k, e_\ell) = \alpha_i^k (\alpha_j^\ell)^* \delta_{k\ell} = \sum_{k=1}^n \alpha_i^k (\alpha_j^k)^* \implies \\ &\implies AA^{T*} = E \implies A^{-1} = A^{T*}.\end{aligned}$$

Matrice transponovaná a komplexně sdružená je v roli matice inverzní k A . Jde tedy o unitární matici. Pro její determinant platí $(\det A)(\det A^{T*}) = |(\det A)|^2 = 1$, tj. $\det A = \exp(ia)$, $a \in \mathbf{R}$. Determinant matice unitárního operátoru je komplexní jednotkou. V případě vektorového prostoru nad \mathbf{R} je $A^T = A^{-1}$, matice A je ortogonální. Pro lineární operátory zachovávající skalární součin tedy platí:

Unitární lineární operátor je v libovolné ortonormální bázi reprezentován unitární maticí A , tj. $A^{-1} = A^{T*}$. Její determinant je komplexní jednotkou. Ortogonální lineární operátor je v libovolné ortonormální bázi reprezentován ortogonální maticí A , tj. $A^{-1} = A^T$. Její determinant je ± 1 .

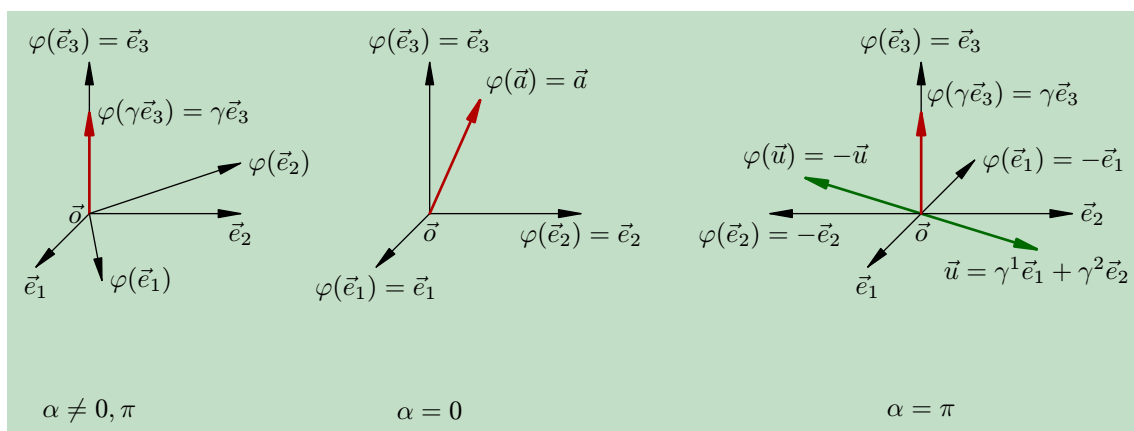
Platí také následující tvrzení:

Je-li lineární operátor $\varphi : U_n \rightarrow U_n$, resp. $\varphi : E_n \rightarrow E_n$ reprezentován některé ortonormální bázi unitární, resp. ortogonální maticí, pak je unitární, resp. ortogonální.

Všimněme si nyní problému vlastních vektorů a vlastních hodnot unitárních (ortogonálních) operátorů. Představme si nejprve opět rotaci kolem pevné osy, kolem pevného bodu, inverzi a zrcadlení v trojrozměrném euklidovském prostoru. Vektory při nich nemění svou délku. Tuto vlastnost musí mít i případný vlastní vektor \vec{u} , pro který $\varphi(\vec{u}) = \lambda\vec{u}$. Vektor $\lambda\vec{u}$ (obraz vlastního vektoru) má tedy stejnou délku jako vzor \vec{u} . Pro hodnotu λ tedy připadají v úvahu jen dvě možnosti, $\lambda = 1$, nebo $\lambda = -1$. Která z nich se u jednotlivých „geometrických“ operátorů opravdu realizuje a jaké budou vlastní vektory? Při hledání odpovědi na tuto otázku se můžeme jednak opřít o geometrickou představu, jednak řešit problém algebraicky.

Příklad 6.19: Vlastní hodnoty a vlastní vektory operátoru rotace kolem pevné osy

Uvažujme o rotaci libovolného vektoru \vec{u} (orientované úsečky umístěné v počátku soustavy souřadnic) kolem vektoru \vec{e}_3 (viz obr. 6.7). Pokud vektor \vec{u} nepatří do některého z podprostorů $[[\vec{e}_3]]$ resp. $[[\vec{e}_1, \vec{e}_2]]$, nemůže být jeho obraz při rotaci kolem \vec{e}_3 kolineární se vzorem \vec{u} . Takový vektor v žádném případě není vlastním vektorem



Obr. 6.7 Vlastní vektory operátoru rotace kolem pevné osy.

operátoru rotace. Naproti tomu libovolný vektor $\vec{u} \in [[\vec{e}_3]]$ je při rotaci totožný se svým obrazem, tj. $\varphi(\vec{u}) = \vec{u}$. Je tedy vlastním vektorem příslušným hodnotě $\lambda = 1$. Všechny prvky vektorového podprostoru vytvořeného vektory \vec{e}_1 a \vec{e}_2 (s obvyklou výjimkou nulového vektoru) jsou vlastními vektory operátoru rotace právě když otáčíme o úhel $\alpha = 0$ nebo $\alpha = \pi$ (viz také příklad 4.51). V případě obecného úhlu otočení žádné další vlastní vektory nedostáváme. Algebraický výpočet dává stejný výsledek:

$$\det \begin{pmatrix} \cos \alpha - \lambda & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha - \lambda & 0 \\ 0 & 0 & 1 - \lambda \end{pmatrix} = 0 \implies \lambda_{1,2} = \exp(\pm i\alpha), \quad \lambda_3 = 1.$$

Komplexně sdružené charakteristické kořeny $\lambda_{1,2}$ jsou (pro $\alpha \neq 0, \pi$) v prostoru nad \mathbf{R} jako vlastní hodnoty nepoužitelné. Máme jedinou vlastní hodnotu, $\lambda_3 = 1$. Dosadíme-li ji do matice $(A - \lambda E)$, dostáváme soustavu rovnic pro vlastní vektory:

$$(A - \lambda_3 E)^T = \begin{pmatrix} \cos \alpha - 1 & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha - 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} \cos \alpha - 1 & -\sin \alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \implies \vec{u} = (0, 0, \gamma) \in [[\vec{e}_3]].$$

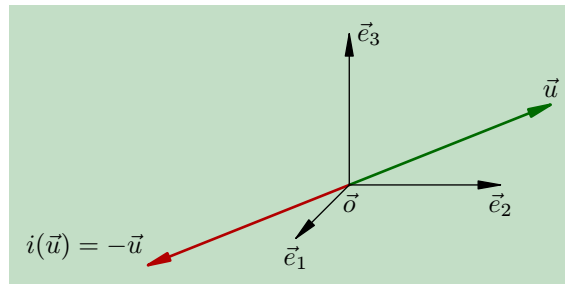
(Pamatujme, že závěr platí pro $\alpha \neq 0$. Pro $\alpha = 0$ má operátor trojnásobnou vlastní hodnotu $\lambda_{1,2,3} = 1$, matice $A - 1E$ je nulová a všechny vektory prostoru jsou vlastní. To ovšem není nijak překvapivé, neboť operátor je v tomto případě identitou, $A = E$.)

Uvažujme nyní o obecné rotaci popsané maticí směrových kosinů $(\cos \gamma_{ij})$. „Geometrická intuice“ či „představivost“ pravděpodobně napoví, že pevným zůstává opravdu jediný prvek, nulový vektor. Můžeme tedy učinit závěr, že při obecné rotaci v trojrozměrném vektorovém prostoru nad \mathbf{R} neexistují vlastní vektory? Zamysleme se nad tímto problémem ještě z algebraického hlediska. Charakteristický polynom operátoru rotace je polynomem třetího stupně s reálnými koeficienty. Polynom s reálnými koeficienty však má reálné kořeny nebo dvojice kořenů komplexně sdružených. Stupeň našeho polynomu je lichý, proto alespoň jeden kořen musí být reálný. Tomuto kořenu však bude příslušet nejméně jednorozměrný vektorový podprostor vlastních vektorů. Dospíváme tak k rozporu mezi algebraickou úvahou a „názornou“ geometrickou představou. Která z nich je správná? Na tomto příkladu vidíme, jak nás „zřejmé“ a „názorné“ geometrické úvahy, a to dokonce i v „našem“ trojrozměrném

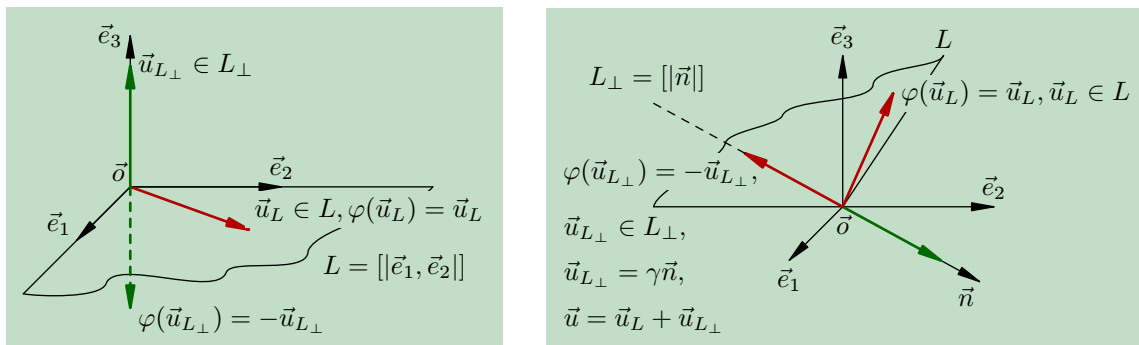
prostoru, mohou zavést k chybnému závěru. Ke správnému závěru nás bezpečně dovede úvaha algebraická, neboť pro každou rotaci určenou obecnými směrovými kosiny existuje taková osa, že rotaci lze popsat jako otočení kolem ní o určitý úhel α . Složky jednotkového vektoru této osy a úhel otočení α jsou jednoznačně určeny hodnotami $\cos \gamma_{ij}$. V řeči algebry tato „osa“ představuje jednorozměrný vektorový podprostor L_\perp vlastních vektorů příslušných reálné vlastní hodnotě operátoru rotace $\lambda = 1$, nebo $\lambda = -1$. V ortonormální bázi $(\vec{e}'_1, \vec{e}'_2, \vec{e}'_3)$, pro kterou $L_\perp = [|\vec{e}'_3\rangle]$, bude mít matice operátoru rotace „částečně“ diagonální tvar odpovídající operátoru rotace kolem pevné osy $[|\vec{e}'_3\rangle]$ o úhel α (bude tvořena dvěma bloky řádu 2 a 1 umístěnými na diagonále, ostatní prvky budou nulové).

Příklad 6.20: Vlastní hodnoty a vlastní vektory operátorů středové inverze a zrcadlení

V případě těchto operátorů funguje geometrická intuice dobře. Při středové inverzi je obrazem obecného vektoru \vec{u} opačný vektor $-\vec{u}$. Každý vektor prostoru je tedy vlastním vektorem operátoru středové inverze příslušným vlastní hodnotě $\lambda = -1$ (obrázek 6.8). Algebraický výpočet, při němž $A = -E$, dává trojnásobnou vlastní hodnotu $\lambda = -1$, matice $A - \lambda E$ je nulová a pro každý vektor prostoru \vec{u} o složkách $(u) = (u^1 \ u^2 \ u^3)$ je $(u)(A - \lambda E) = (0) = (0 \ 0 \ 0)$.



Obr. 6.8 Vlastní vektory operátoru středové inverze.



Obr. 6.9 Vlastní vektory operátoru zrcadlení.

Při zrcadlení vzhledem k rovině určené vektory \vec{e}_1 a \vec{e}_2 , tj. dvojrozměrnému vektorovému podprostoru $L = [|\vec{e}_1, \vec{e}_2\rangle]$, je obrazem každého vektoru ležícího v L přímo vzor samotný. Podprostor L je tedy tvořen samými vlastními vektory příslušnými vlastní hodnotě $\lambda_1 = 1$. Vlastním vektorem je rovněž každý vektor z ortogonálního doplňku tohoto podprostoru $L_\perp = [|\vec{e}_3\rangle]$, obrazem každého vektoru je však vektor opačný, takže příslušná vlastní hodnota je $\lambda_2 = -1$. Jiné vlastní vektory nejsou. Stejný výsledek musíme samozřejmě získat algebraickým

postupem. Operátor zrcadlení je reprezentován diagonální maticí uvedenou v příkladu 6.18 (viz též obrázek 6.9 vlevo). Charakteristický polynom je $(\lambda - 1)^2(\lambda + 1)$ a má dvojnásobný kořen $\lambda_1 = 1$ a jednonásobný $\lambda_2 = -1$. Řešením odpovídajících soustav rovnic po složky vlastních vektorů dostáváme:

$$(A - \lambda_1 E)^T = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \implies \vec{u} = (u^1, u^2, 0) \implies L_1 = [[\vec{e}_1, \vec{e}_2]] = L,$$

$$(A - \lambda_2 E)^T = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \implies \vec{u} = (0, 0, u^3) \implies L_2 = [[\vec{e}_3]] = L_\perp.$$

Co když však provádíme zrcadlení vzhledem k obecné „rovině“ (obrázek 6.9 vpravo)? Označme její normálový jednotkový vektor \vec{n} a úhly, které svírá s vektory báze \vec{e}_1 , \vec{e}_2 a \vec{e}_3 nechť jsou α , β , γ . Platí $\vec{n} = \cos \alpha \vec{e}_1 + \cos \beta \vec{e}_2 + \cos \gamma \vec{e}_3$. Ortogonální průmět vektoru \vec{u} do směru vektoru \vec{n} je $(\vec{u}, \vec{n})\vec{n}$. Zrcadlový obraz vektoru \vec{u} má tvar

$$\varphi(\vec{u}) = \vec{u} - 2(\vec{u}\vec{n})\vec{n},$$

obrazy vektorů báze jsou tedy

$$\begin{aligned} \varphi(\vec{e}_1) &= \vec{e}_1 - 2(\vec{e}_1\vec{n})\vec{n} = (1 - 2\cos^2 \alpha)\vec{e}_1 - (2\cos \alpha \cos \beta)\vec{e}_2 - (2\cos \alpha \cos \gamma)\vec{e}_3, \\ \varphi(\vec{e}_2) &= \vec{e}_2 - 2(\vec{e}_2\vec{n})\vec{n} = -(2\cos \alpha \cos \beta)\vec{e}_1 + (1 - 2\cos^2 \beta)\vec{e}_2 - (2\cos \beta \cos \gamma)\vec{e}_3, \\ \varphi(\vec{e}_3) &= \vec{e}_3 - 2(\vec{e}_3\vec{n})\vec{n} = -(2\cos \alpha \cos \gamma)\vec{e}_1 - (2\cos \beta \cos \gamma)\vec{e}_2 + (1 - 2\cos^2 \gamma)\vec{e}_3. \end{aligned}$$

Matrice operátoru zrcadlení je

$$A = \begin{pmatrix} 1 - 2\cos^2 \alpha & -2\cos \alpha \cos \beta & -2\cos \alpha \cos \gamma \\ -2\cos \alpha \cos \beta & 1 - 2\cos^2 \beta & -2\cos \beta \cos \gamma \\ -2\cos \alpha \cos \gamma & -2\cos \beta \cos \gamma & 1 - 2\cos^2 \gamma \end{pmatrix}.$$

Kdo chce, může počítat charakteristický polynom a charakteristické kořeny pomocí této obecné matice. Kdo si však vzpomene, že charakteristický polynom a tedy i charakteristické kořeny operátoru nezávisí na volbě báze, má velkou část práce hotovou. Ví totiž, že charakteristické kořeny musí vyjít $\lambda_1 = 1$ (dvojnásobný kořen) a $\lambda_2 = -1$ (jednonásobný kořen). Geometrická představa říká, že vlastní vektory příslušné hodnotě $\lambda_2 = -1$ budou tvořit jednorozměrný vektorový podprostor $[[\vec{n}]]$ a vlastní vektory příslušné hodnotě $\lambda_1 = 1$ jeho ortogonální doplněk. To snadno ověříme řešením příslušných soustav rovnic:

$$(A - \lambda_1 E)^T = \begin{pmatrix} -2\cos^2 \alpha & -2\cos \alpha \cos \beta & -2\cos \alpha \cos \gamma \\ -2\cos \alpha \cos \beta & -2\cos^2 \beta & -2\cos \beta \cos \gamma \\ -2\cos \alpha \cos \gamma & -2\cos \beta \cos \gamma & -2\cos^2 \gamma \end{pmatrix} \sim \begin{pmatrix} \cos \alpha & \cos \beta & \cos \gamma \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Této soustavě rovnic vyhovují všechny vektory $\vec{u} = (u_1, u_2, u_3)$, pro které je $u_1 \cos \alpha + u_2 \cos \beta + u_3 \cos \gamma = 0$, tj. $\vec{u} \perp \vec{n}$. Pro $\lambda_2 = -1$ dostáváme

$$(A - \lambda_2 E)^T = \begin{pmatrix} 2 - 2\cos^2 \alpha & -2\cos \alpha \cos \beta & -2\cos \alpha \cos \gamma \\ -2\cos \alpha \cos \beta & 2 - 2\cos^2 \beta & -2\cos \beta \cos \gamma \\ -2\cos \alpha \cos \gamma & -2\cos \beta \cos \gamma & 2 - 2\cos^2 \gamma \end{pmatrix}.$$

Soustavě rovnic o této matici vyhovuje vektor $\vec{n} = (\cos \alpha, \cos \beta, \cos \gamma)$ a libovolný jeho násobek, tj. vektorový podprostor $[\vec{n}]$. Snadno se o tom přesvědčíme přímým dosazením do všech rovnic soustavy, vezmeme-li v úvahu vztah $\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1$ (vektor \vec{n} je jednotkový). Vektorové podprostory vlastních vektorů příslušných různým vlastním hodnotám však, jak jsme již dříve ukázali, nemají s výjimkou nulového vektoru společné prvky. Protože hodnotě $\lambda_1 = 1$ přísluší dvojrozměrný podprostor — to jsme vypočítali před chvílí — zbývá na hodnotu $\lambda_2 = -1$ pouze podprostor jednorozměrný. Žádné další vlastní vektory než násobky vektoru \vec{n} již nedostaneme. Vidíte, jak nám opět pomohla obecná algebraická úvaha? Ani jsme nepříjemně vypadající matici $(A - \lambda_2 E)^T$ nemuseli upravovat.

V předchozích příkladech jsme studovali ortogonální operátory v trojrozměrném euklidovském prostoru. Tyto operátory zachovávaly skalární součin vektorů, tj. skalární součin dvou libovolných vektorů byl shodný se skalárním součinem odpovídajících obrazů. Vlastní hodnoty takových operátorů byly 1 nebo -1 . Je to snad náhoda? Zobecníme teď naše úvahy na unitární operátory pro případ obecné dimenze vektorového prostoru nad polem komplexních čísel a budeme se snažit zjistit nějaké obecné vlastnosti jejich vlastních vektorů a hodnot vyplývající z definičního požadavku na skalární součin. Nechť a je libovolný z vlastních vektorů operátoru $\varphi: U_n \rightarrow U_n$ příslušných vlastní hodnotě λ . Platí

$$\varphi(a) = \lambda a \implies (\varphi(a), \varphi(a)) = (\lambda a, \lambda a) = \lambda \lambda^* (a, a).$$

Protože však

$$(\varphi(a), \varphi(a)) = (a, a) \neq 0, \quad \text{je} \quad |\lambda| = 1 \implies \lambda = \exp(i\alpha), \quad \alpha \in \mathbf{R}.$$

Vlastními hodnotami unitárního operátoru jsou výhradně komplexní jednotky. Již z dřívějších úvah ve vektorových prostorech, které nemusely být vybaveny skalárním součinem, víme, že každé vlastní hodnotě přísluší, po doplnění nulového vektoru, podprostor vlastních vektorů. Podprostory příslušné různým vlastním hodnotám přitom mají společný právě jen „doplněný“ nulový vektor. Nebudou mít tyto podprostory pro případ unitárního operátoru nějakou další speciální vlastnost? Vrátime-li se k příkladům geometrických operací zachovávajících skalární součin v trojrozměrném euklidovském prostoru, určitě si vzpomeneme, že tyto podprostory byly navzájem kolmé. Libovolný vlastní vektor příslušný například u zrcadlení hodnotě 1 byl kolmý k libovolnému vlastnímu vektoru příslušnému hodnotě -1 . A opět to není náhoda. Zvolme libovolný vlastní vektor a unitárního operátoru příslušný hodnotě λ_1 a libovolný vlastní vektor b příslušný hodnotě $\lambda_2 \neq \lambda_1$. Platí

$$\begin{aligned} \varphi(a) = \lambda_1 a = \exp(i\alpha) a, \quad \varphi(b) = \lambda_2 b = \exp(i\beta) b &\implies \\ \implies (\varphi(a), \varphi(b)) = \exp([i(\alpha - \beta)]) (a, b). \end{aligned}$$

Z požadavku shodnosti skalárního součinu vektorů a obrazů dostáváme

$$(\varphi(a), \varphi(b)) = (a, b) \implies \exp[i(\alpha - \beta)] (a, b) = (a, b).$$

Splnit tuto rovnost lze pouze pro $(a, b) = 0$ nebo pro $\exp[i(\alpha - \beta)] = 1$. Druhý případ by ovšem znamenal, že $\lambda_1 = \lambda_2$. Musí tedy nastat první možnost, představující skutečnost, že vektory a a b jsou ortogonální. Protože byly vybrány libovolně, je zřejmé, že jsou ortogonální celé podprostory vlastních vektorů příslušných navzájem různým vlastním hodnotám. Označme všechny navzájem různé vlastní hodnoty unitárního operátoru λ_1 až λ_r , jejich násobnosti postupně k_1 až k_r , $k_1 + \dots + k_r = n$. V obecném případě pro dimenze vektorových podprostorů vlastních vektorů $q_j = \dim L_j$ platí $q_j \leq k_j$, takže vektory z L_1 až L_r nemusí obecně generovat celý prostor U_n . Je tomu v případě unitárního operátoru stejně? Označme $L = L_1 + \dots + L_r$ a $k = \dim L = q_1 + \dots + q_r$. Předpokládejme, že $k < n$. Označme dále L_\perp ortogonální doplněk podprostoru L v U_n , $\dim L_\perp = (n - k) \geq 1$. Z unitárnosti zobrazení φ vyplývá, že pro každý vektor $b \in L_\perp$ je také $\varphi(b) \in L_\perp$, podprostor L_\perp je takzvaně *invariantní* vzhledem k zobrazení φ (zkuste sami dokázat). Zvolíme-li nyní novou ortonormální bázi prostoru U_n tak, aby byla tvořena právě k vektory f_1, f_2 až f_k z podprostoru L a $(n - k)$ vektory f_{k+1} až f_n z ortogonálního doplňku L_\perp , bude mít operátor φ v této bázi matici, která bude zčásti diagonální, konkrétně

$$\bar{A} = \left(\begin{array}{cccc|c} \lambda_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_r & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & B \end{array} \right).$$

Každá z hodnot λ_j se přitom v diagonále vyskytne q_j -krát, B je unitární matice řádu $(n - k)$ (matice \bar{A} totiž musí být stále unitární — přešli jsme sice k jiné bázi, ale ta je rovněž ortonormální, jako byla báze původní) a reprezentuje jistý unitární operátor působící ve vektorovém prostoru L_\perp . Charakteristický polynom operátoru φ je

$$\det(\bar{A} - \lambda E) = (\lambda_1 - \lambda)^{q_1} (\lambda_2 - \lambda)^{q_2} \dots (\lambda_r - \lambda)^{q_r} \det(B - \lambda E).$$

Na druhé straně však platí

$$\det(\bar{A} - \lambda E) = (\lambda_1 - \lambda)^{k_1} (\lambda_2 - \lambda)^{k_2} \dots (\lambda_r - \lambda)^{k_r}.$$

Operátor v L_\perp reprezentovaný maticí B musí tedy mít rovněž vlastní hodnoty λ_1 až λ_r o násobnostech $(k_1 - q_1)$ až $(k_r - q_r)$. Jim musí odpovídat vlastní vektory operátoru φ ležící v L_\perp . My jsme však před chvílí označili vektorový podprostor obsahující *všechny* vlastní vektory operátoru φ jako L . Dostáváme se tak ke sporu. Výchozí předpoklad, spočívající v nerovnosti $k < n$, byl tedy chybný. Ve skutečnosti platí $k = n$, $q_j = k_j$. Dimenze vektorového podprostoru vlastních vektorů příslušných hodnotě λ_j je tedy rovna její násobnosti a vlastní vektory operátoru generují celý vektorový prostor. V ortonormální bázi tvořené právě těmito vlastními vektory má matice operátoru diagonální tvar. V diagonále jsou vlastní hodnoty, každá z nich tolikrát, kolik činí její násobnost.

Tento výsledek je velmi důležitý, neboť představuje úplné vyřešení problému vlastních hodnot a vlastních vektorů pro velkou třídu lineárních operátorů, které navíc hrají závažnou roli v geometrii, fyzice a technických aplikacích. Proto jej formulujeme v podobě matematického tvrzení.

Věta 6.1: *Pro unitární lineární operátor ve vektorovém prostoru U_n platí:*

- (1) *Všechny vlastní hodnoty operátoru jsou komplexní jednotky.*
- (2) *Dimenze podprostoru vlastních vektorů operátoru příslušných téže vlastní hodnotě je shodná s násobností této vlastní hodnoty jako kořene charakteristického polynomu operátoru.*
- (3) *Každé dva podprostory vlastních vektorů operátoru příslušných různým vlastním hodnotám jsou ortogonální.*

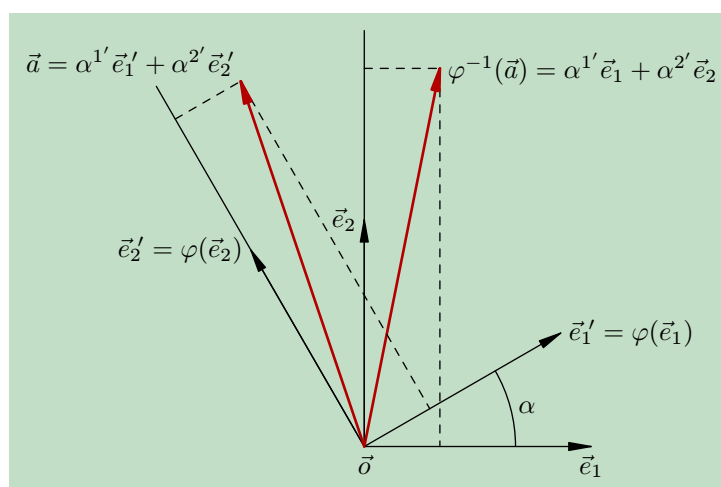
Přemýšlejme, oč je tato věta „silnější“ než byly obecné závěry o problému vlastních vektorů operátorů v n -rozměrných prostorech. Především rozvažme, zda přináší něco nového, pokud jde o možnost diagonalizace operátoru, tj. možnost nalezení takové báze vektorového prostoru U_n , v níž bude operátor reprezentován diagonální maticí. Podle věty 4.3 je každá taková báze tvořena vlastními vektory operátoru. Otázka tedy zní, zda vlastnosti shrnuté ve větě 6.1 přispějí k tomu, aby se podařilo nalézt dostatečný počet, tj. n , lineárně nezávislých vlastních vektorů. Z příkladů odstavce 4.3.2 jsme se již poučili, že obecně to vždy jít nemusí. Vektorové podprostory (například L_1, L_2) vlastních vektorů odpovídajících různým vlastním hodnotám $\lambda_1 \neq \lambda_2$ mají sice tu vlastnost, že jejich průnik obsahuje jedině nulový vektor a při jejich sečtení se „neztrácí“ dimenze ($\dim(L_1 + L_2) = \dim L_1 + \dim L_2$), jejich celková dimenze však může být nižší než dimenze celého prostoru. V případě unitárních operátorů v prostorech se skalárním součinem však je navíc zajištěno, že dimenze každého podprostoru vlastních vektorů příslušného dané vlastní hodnotě je rovna násobnosti této hodnoty. Součet násobností vlastních hodnot — kořenů charakteristického polynomu operátoru — je ovšem roven stupni polynomu, tj. n . Vektorové podprostory L_1 až L_r příslušné všem navzájem různým hodnotám λ_1 až λ_r generují celý prostor. Unitární operátor můžeme tedy diagonalizovat *vždy!* Skutečnost, že $L_i \perp L_j$, $1 \leq i, j \leq n$, $i \neq j$, navíc usnadňuje nalezení ortonormální báze z vlastních vektorů. Ortogonalizační procesy totiž stačí uplatnit jen v podprostorech vlastních vektorů. To, že vlastními hodnotami jsou zrovna a jedině komplexní jednotky, vypadá možná v tuto chvíli jako pouhá zajímavost. Ale i ona má svou významnou geometrickou a fyzikální interpretaci. Později uvidíme. Nyní však formulujeme důsledek vět 4.3 a 6.1, které jsme právě objevili.

Důsledky: Z vět 4.3 a 6.1 vyplývají tyto vlastnosti unitárních operátorů v U_n :

- Podprostory vlastních vektorů operátoru generují prostor U_n , tj. $L_1 + \dots + L_r = U_n$.
- Z vlastních vektorů operátoru lze sestavit ortonormální bázi prostoru U_n . V této bázi je operátor reprezentován diagonální maticí. Diagonála je obsazena vlastními hodnotami operátoru, z nichž každá se vyskytuje právě tolikrát, kolik činí její násobnost.

Příklad 6.21: Lineární operátory a transformace souřadnic

Pamatujete si ještě grupové panoptikum z příkladu 4.8? Vystupovaly v něm různé grupy čtvercových matic s různými vlastnostmi. Grupa $U(\mathbf{C}, n)$, je tvořena unitárními maticemi řádu n , její podgrupa $SU(\mathbf{C}, n)$ obsahuje jen ty z nich, jejichž determinant je jednotkový, grupa $O(\mathbf{R}, n)$ resp. její podgrupa $SO(\mathbf{R}, n)$ jsou tvořeny ortogonálními maticemi resp. ortogonálními maticemi s jednotkovým determinanem. Společnou vlastností všech grup matic uvedených v příkladu 4.8 ovšem byla jejich regulárnost. A takové matice vždy mohou realizovat přechody mezi bázemi ve vektorových prostorech dimenze n . Tytéž grupy čtvercových matic mohou na druhé straně hrát roli při reprezentaci lineárních operátorů určitého typu v bázích. Je mezi přechody od báze k bázi a reprezentací lineárních operátorů maticemi nějaká souvislost? Jednoduše ji ukazuje obrázek 6.10. Přechod mezi



Obr. 6.10 Transformace báze a lineární operátory.

bázemi (\vec{e}_1, \vec{e}_2) a (\vec{e}'_1, \vec{e}'_2) v rovině nakreslený na obrázku si lze představit jako otočení báze o úhel α . Lineární operátor, který otočí o úhel α všechny vektory, je v bázi (\vec{e}_1, \vec{e}_2) reprezentován ortogonální maticí T určenou obrazy báze, zpětný přechod maticí inverzní S :

$$T = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}, \quad S = T^{-1} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

Obrazy vektorů báze $\vec{e}'_1 = \varphi(\vec{e}_1)$ a $\vec{e}'_2 = \varphi(\vec{e}_2)$ opět tvoří bázi, neboť matice T reprezentující operátor φ je

regulární. Vektor \vec{a} má v bázi (\vec{e}_1, \vec{e}_2) stejné složky jako vektor $\varphi(\vec{a})$ v bázi $(\vec{e}'_1, \vec{e}'_2) = (\varphi(\vec{e}_1), \varphi(\vec{e}_2))$, tj.

$$\vec{a} = \alpha^i \vec{e}_i, \quad \varphi(\vec{a}) = \alpha^j \varphi(\vec{e}_j) = \alpha^j \tau_j^i \vec{e}_i,$$

zatímco v čárkované bázi má stejné složky jako vektor $\varphi^{-1}(\vec{a})$ v bázi původní. Operátor φ^{-1} je v bázi (\vec{e}_1, \vec{e}_2) reprezentován maticí $T^{-1} = S$, tj.

$$\alpha^{1'} = \alpha^1 \cos \alpha + \alpha^2 \sin \alpha, \quad \alpha^{2'} = -\alpha^1 \sin \alpha + \alpha^2 \cos \alpha,$$

$$(\alpha') = (\alpha)T^{-1} \implies \alpha^{i'} = \alpha^j \sigma_j^i.$$

Dostali jsme právě vztahy (4.13) pro transformaci složek vektoru při změně báze. Převod vektoru do jiné báze, jejíž vektory jsou dány jako obrazy původní báze operátorem φ , si tedy můžeme představit také jako jeho obraz inverzním operátorem φ^{-1} v bázi původní.

Příklad 6.22: Co znamenají komplexní jednotky jako vlastní hodnoty lineárních operátorů

V úvodu odstavce 4.3 jsme poněkud předběhli „fyzikálním událostem“ a řekli si, jaký význam mají lineární operátory ve vektorových prostorech pro kvantovou mechaniku. Připomeňme, že stav fyzikální soustavy, která se řídí zákony kvantové mechaniky, je vhodné reprezentovat prvkem jistého vektorového prostoru — vektorem. (Obecně se jedná o vektorový prostor určitých speciálních vlastností, Hilbertův prostor, v řadě případů však vystačíme i s unitárním vektorovým prostorem konečné dimenze.) Působením lineárního operátoru na vektor stavu dostaneme jeho obraz, tj. změním stav soustavy. Jeden z častých způsobů popisu stavů soustavy využívá takzvané *souřadnicové reprezentace*. V takovém případě je stav soustavy, například částice v trojrozměrném prostoru, popsán jistou komplexní funkcí souřadnic a času $|\text{stav}\rangle \sim \psi(x, y, z, t)$. Že i funkce tvoří vektorový prostor (s obvyklou operací sčítání funkcí a násobením funkce číslem), jsme se již dříve přesvědčili. Funkce $\psi(x, y, z, t)$ sama nemá bezprostřední názorný fyzikální význam. Má jej však čtverec její absolutní hodnoty $|\psi(x, y, z, t)|^2$. Podle Heisenbergových relací neurčitosti, o kterých jsme se v odstavci 4.3 již také zmínili, se kvantové částice chovají tak „divně“, že jim nelze přisoudit určitou polohu (ledaže bychom připustili úplnou ztrátu informace o jejich hybnosti). Vymežíme-li v prostoru malý objem s počátkem v bodě (x, y, z) , tj. $(x, x + dx) \times (y, y + dy) \times (z, z + dz)$, můžeme pouze určit elementární pravděpodobnost nalezení částice v tomto objemu. Ta je rovna

$$dP(x, y, z, t) = |\psi(x, y, z, t)|^2 dx dy dz.$$

Funkce $|\psi(x, y, z, t)|^2$ má tedy význam *hustoty pravděpodobnosti*, tj. rozdělení jakési kvantověmechanické náhodné veličiny, kterou bychom mohli nazvat „poloha částice“. Jak tyto úvahy souvisejí s vlastními hodnotami unitárních operátorů? Vynásobíme-li funkci $\psi(x, y, z, t)$ libovolnou komplexní jednotkou $\exp(i\alpha)$, dostaneme sice obecně jinou funkci — jiný stav kvantověmechanické soustavy, nezměníme však hodnotu odpovídající hustoty pravděpodobnosti. Bude-li funkce $\psi(x, y, z, t)$ jako prvek vektorového prostoru vlastním vektorem unitárního operátoru φ (unitární operátor zde může opět představovat například transformaci souřadnic nebo prostorové rotace, popřípadě časový vývoj soustavy), pak se její obraz bude od vzoru lišit pouze vynásobením příslušnou vlastní hodnotou, tedy komplexní jednotkou. Stav soustavy se změní tak, že dojde pouze ke *změně fáze*, tj.

$$\varphi(|\text{stav}\rangle) = e^{i\alpha} |\text{stav}\rangle,$$

$$\varphi(\psi(x, y, z, t)) = e^{i\alpha} \psi(x, y, z, t) \implies |\varphi(\psi(x, y, z, t))|^2 = |\psi(x, y, z, t)|^2.$$

„Fyzikálně“, z hlediska pravděpodobnostní interpretace funkce $\psi(x, y, z, t)$, tedy k žádné změně nedochází. Funkci $\psi(x, y, z, t)$ se také říká *amplituda pravděpodobnosti* nebo *vlnová funkce*, neboť za určitých okolností lze pro funkce tohoto typu použít principu superpozice obdobně jako v případě skládání vlnění.

Příklad 6.23: Kvantová mechanika a vývoj v čase

O tom, že se každá soustava mění, vyvíjí s časem, nepochybuje nikdo, ani fyzikální a matematictí laikové. Popsat tento vývoj korektně, přesně a věrohodně však již každý nedovede. Volba takového popisu totiž musí odpovídat principům, jimiž se daná soustava řídí a výsledky musí být v souladu s experimentem. Jak víme, vhodný popis soustav vyskytujících se v mikrosvětě je jiný než popis klasických mechanických soustav tvořených dostatečně velkými objekty. V klasické mechanice je stav hmotné částice určen její polohou a rychlostí. Jeho časový vývoj je dán zrychlením, a to se řídí druhým Newtonovým zákonem. Je totiž určeno výslednicí všech sil, jimiž na částici působí její okolí. Vyjádření zrychlení pomocí výslednice sil představuje *pohybovou rovnici* a jejím řešením můžeme získat stav částice v libovolném okamžiku t , známe-li stav v určitém okamžiku t_0 . V předchozím příkladu jsme si již všimli možnosti popsat vývoj stavu částice v mikrosvětě pomocí unitárního operátoru působícího na vlnovou funkci. Všimněme si nyní obecného popisu časového vývoje, kdy budeme uvažovat o stavu jako o prvku abstraktního vektorového prostoru stavů U_n , nikoli jen jako o vlnové funkci. Označme známý stav v okamžiku t_0 jako

$$|\text{stav}(t_0)\rangle = \Psi(t_0), \quad \Psi(t_0) \in U_n.$$

Stav v okamžiku t je pak

$$|\text{stav}(t)\rangle = \Psi(t), \quad \Psi(t) \in U_n.$$

Vektory stavu lišící se jen svou „délkou“ nepředstavují fyzikálně odlišné situace, reprezentují též stav. Přesněji řečeno je stav určen jednorozměrným podprostorem prostoru stavů. Nic nebrání tomu, abychom pro jednoduchost pracovali s jednotkovými vektory, tj. $(\Psi(t_0), \Psi(t_0)) = 1$, ale také $(\Psi(t), \Psi(t)) = 1$. Časový vývoj stavu, tedy pohybová rovnice částice, bude mít obecný tvar

$$\Psi(t) = \mathcal{T}_{(t_0 \rightarrow t)}(\Psi(t_0)),$$

kde $\mathcal{T}_{(t_0 \rightarrow t)}$ je vhodný lineární operátor převádějící stav v okamžiku t_0 ve stav v okamžiku t . Je zřejmé, že v případě jednotkového vektoru $\Psi(t_0)$ bude vektor $\Psi(t)$ rovněž jednotkový právě tehdy, bude-li operátor $\mathcal{T}_{(t_0 \rightarrow t)}$ zachovávat skalární součin, tj. bude-li unitární. A ještě jednu zajímavou vlastnost bude tento *operátor časového vývoje* mít. Jestliže $\mathcal{T}_{(t_0 \rightarrow t)}$ převede stav $\Psi(t_0)$ ve stav $\Psi(t)$, pak $\mathcal{T}_{(t \rightarrow t_0)}$ převede stav $\Psi(t)$ ve stav původní, $\Psi(t_0)$. Proto platí

$$\mathcal{T}_{(t_0 \rightarrow t)} = (\mathcal{T}_{(t \rightarrow t_0)})^{-1}.$$

Operátor $\mathcal{T}_{(t_0 \rightarrow t)}$ bychom mohli nazvat třeba *operátorem „TAM“* a $\mathcal{T}_{(t_0 \rightarrow t)}^{-1} = \mathcal{T}_{(t \rightarrow t_0)}$ pak *operátorem „ZPĚT“*. A co myslíte, jakou speciální vlastnost má operátor $\mathcal{T}_{(t_0 \rightarrow t_0)}$?

Pozn. 1: V souřadnicové reprezentaci je časový vývoj stavu, reprezentovaného vlnovou funkcí $\psi(x, y, z, t)$, popsán *Schrödingerovou rovnicí*

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right).$$

Pozn. 2: Pozorný čtenář si mohl všimnout, že o operátoru $\mathcal{T}_{(t_0 \rightarrow t)}$ a priori předpokládáme, že je lineární. Může se pak ptát proč. Je to proto, aby se působením operátoru „nepokazil“ princip superpozice stavů.

K vyjádření pohybové rovnice kvantové mechaniky pomocí unitárního operátoru se vrátíme ještě jednou později, v jednom z příkladů odstavce o samoadjungovaných operátorech.

6.2.2 Samoadjungované (symetrické) lineární operátory

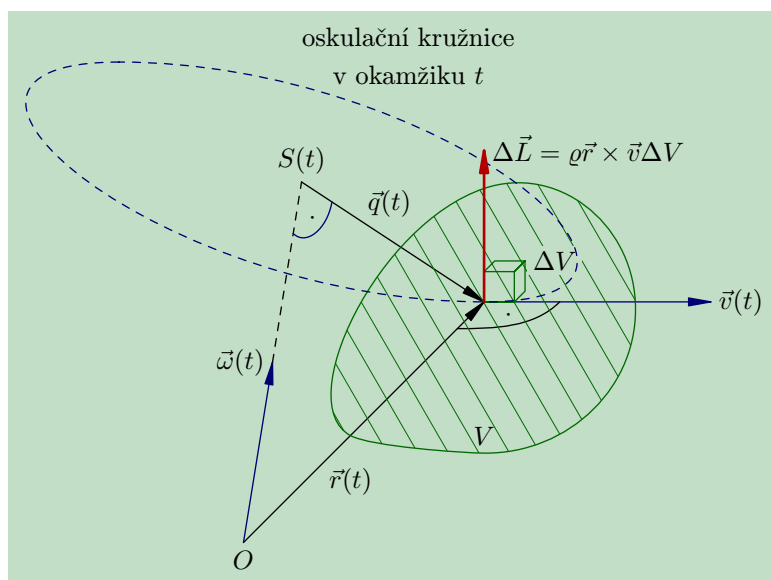
Vraťme se k úvodu odstavce 4.2.1, v němž jsme se zabývali obecným tvarem lineárních vztahů mezi vektorovými veličinami. Shrneme-li například vztah mezi momentem hybnosti tuhého tělesa \vec{L} a jeho úhlovou rychlostí $\vec{\omega}$ do jednoho zápisu, můžeme psát pro složky (s indexy podle fyzikálních zvyklostí opět dole, ale i tak s použitím Einsteinovy sčítací symboliky)

$$L_i = J_{ij}\omega_j, \quad 1 \leq i, j \leq 3.$$

Tento vztah chápeme jako působení lineárního operátoru J na vektor $\vec{\omega}$, který je vzorem, obrazem je vektor \vec{L} . Význam jednotlivých složek *tenzoru momentu setrvačnosti* (J_{ij}) snadno získáme z definice momentu hybnosti

$$\vec{L} = \int_V \rho \vec{r} \times \vec{v} dV = \int_V \rho \vec{r} \times (\vec{\omega} \times \vec{r}) dV.$$

Stejně jako v úvodu odstavce 4.2.1 je integračním oborem objem tělesa (trojný integrál), ρ je hustota, která ani nemusí být konstantní, \vec{r} je polohový vektor obecného bodu v tělese, jehož složky x , y a z jsou integrační proměnné, \vec{v} je rychlost tohoto bodu, $\vec{\omega}$ okamžitá úhlová rychlost rotace tělesa kolem pevného bodu. Z obrázku 6.11 vidíme, že pohyb obecného bodu tělesa



Obr. 6.11 K výpočtu momentu hybnosti.

můžeme v daném okamžiku interpretovat jako pohyb po jisté „okamžité“ *oskulační kružnici*, která se těsně přimyká ke křivce, po níž se bod ve skutečnosti pohybuje. Proto jsme dosadili $\vec{v} = \vec{\omega} \times \vec{q} = \vec{\omega} \times \vec{r}$. Použijeme-li vzorce pro „dvojitý“ vektorový součin (zkuste si jej odvodit)

$\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a}\vec{c})\vec{b} - (\vec{a}\vec{b})\vec{c}$, dostaneme

$$\vec{L} = \int_V \varrho [(\vec{r}^2)\vec{\omega} - (\vec{r}\vec{\omega})\vec{r}] dV.$$

Rozepíšeme-li jednotlivé složky tohoto vektorového vztahu (zkuste to), dostaneme pro J_{ij} vztahy:

$$J_{11} = \int_V \varrho(y^2 + z^2) dV, \quad J_{22} = \int_V \varrho(x^2 + z^2) dV, \quad J_{33} = \int_V \varrho(x^2 + y^2) dV,$$

$$J_{12} = J_{21} = - \int_V \varrho xy dV, \quad J_{13} = J_{31} = - \int_V \varrho xz dV, \quad J_{23} = J_{32} = - \int_V \varrho yz dV.$$

Pro pořádek poznamenejme, že je-li těleso tuhé a soustava souřadnic je s ním pevně spojena, jsou veličiny J_{ij} konstantní (nezávislé na čase). Všimněme si však jedné velmi podstatné vlastnosti. Sestavíme-li z hodnot J_{ij} matici, bude symetrická vzhledem k hlavní diagonále. Platí $J = J^T$. Tuto matici samozřejmě interpretujeme tak, že ve zvolené ortonormální bázi (soustavě souřadnic) reprezentuje lineární operátor přiřazující vektorům úhlové rychlosti odpovídající vektory momentu hybnosti. Přejdeme-li pomocí matice přechodu T k jiné ortonormální bázi, nebude již operátor reprezentovat matice $J = (J_{ij})$, ale matice podobná, $J' = TJT^{-1}$ (viz vztah (4.26)). Také tato matice je však symetrická, neboť matice přechodu mezi ortonormálními bázemi v euklidovském prostoru je ortogonální, $T^{-1} = T^T$. Platí pak

$$(J')^T = (TJT^{-1})^T = (T^{-1})^T J^T T^T = TJT^{-1} = J'.$$

Právě symetrie je typickou vlastností fyzikálních veličin, které mají charakter lineárních operátorů ve vektorových prostorech. Má tedy smysl se takovými operátory zabývat obecně a zejména se zaměřit na problém jejich vlastních vektorů. Co když vlastnost symetrie přinese pro jeho řešení něco zajímavého? Co když, podobně jako tomu bylo u unitárních operátorů, existuje i zde báze tvořená vlastními vektory? To by se pak s operátory počítalo velmi snadno, protože bychom je mohli reprezentovat diagonální maticí. Nejprve však musíme definovat „symetrii“ operátoru co nejobecněji. Ke konstrukci definice použijme ještě jednou příklad s momentem setrvačnosti.

Příklad 6.24: Vzory, obrazy a skalární součiny

Zobrazíme operátorem J dva různé vektory úhlové rychlosti a vypočteme určité skalární součiny

$$\vec{L} = J(\vec{\omega}), \quad \vec{L}' = J(\vec{\omega}'),$$

$$\vec{L}\vec{\omega}' = \begin{pmatrix} L_1 & L_2 & L_3 \end{pmatrix} \begin{pmatrix} \omega'_1 \\ \omega'_2 \\ \omega'_3 \end{pmatrix} = \begin{pmatrix} \omega_1 & \omega_2 & \omega_3 \end{pmatrix} \begin{pmatrix} J_{11} & J_{12} & J_{13} \\ J_{12} & J_{22} & J_{23} \\ J_{13} & J_{23} & J_{33} \end{pmatrix} \begin{pmatrix} \omega'_1 \\ \omega'_2 \\ \omega'_3 \end{pmatrix} =$$

6.2. „FYZIKÁLNÍ“ LINEÁRNÍ OPERÁTORY A JEJICH VLASTNÍ VEKTORY 453

$$= J_{11}\omega_1\omega'_1 + J_{12}(\omega_1\omega'_2 + \omega_2\omega'_1) + J_{13}(\omega_1\omega'_3 + \omega_3\omega'_1) + J_{22}\omega_2\omega'_2 + J_{23}(\omega_2\omega'_3 + \omega_3\omega'_2) + J_{33}\omega_3\omega'_3.$$

Při výpočtu jsme využili symetrie matice (J_{ij}) . Prohlédněte si výsledek důkladně a snažte se objevit nějakou symetrii. Jistě vám neunikne, že záměna vektorů $\vec{\omega}$ a $\vec{\omega}'$ výsledek nezmění. Platí tedy

$$\vec{L}\vec{\omega}' = \vec{\omega}\vec{L}' \implies (J(\vec{\omega}), \vec{\omega}') = (\vec{\omega}, J(\vec{\omega}')).$$

Objevená symetrie spočívá v tom, že skalární součin prvního obrazu $\vec{L} = J(\vec{\omega})$ a druhého vektoru $\vec{\omega}'$ je shodný se skalárním součinem prvního vektoru $\vec{\omega}$ a druhého obrazu $\vec{L}' = J(\vec{\omega}')$. Vzory jsme nijak speciálně nevybírali, takže zjištěná vlastnost je obecná.

Vlastnost symetrie operátoru momentu setrvačnosti se stane základem definice, kterou zobecníme na libovolný n -rozměrný unitární prostor nad polem komplexních čísel.

Nechť U_n je unitární vektorový prostor nad \mathbf{C} . Lineární operátor $\varphi : U_n \rightarrow U_n$ se nazývá *samoadjungovaný*, jestliže pro každé dva vektory $a, b \in U_n$ platí

$$(\varphi(a), b) = (a, \varphi(b)). \quad (6.15)$$

V případě euklidovského vektorového prostoru E_n , tj. nad \mathbf{R} , se operátor s vlastností (6.15) nazývá *symetrický*.

Pochopitelně očekáváme, že samoadjungovaný, resp. symetrický lineární operátor bude v ortonormálních bázích reprezentován samoadjungovanými, resp. symetrickými maticemi. Ale musíme se o tom přesvědčit. Zvolme v U_n libovolnou ortonormální bázi (e_1, e_2, \dots, e_n) . Matice operátoru φ je určena obrazy báze, tj. vektory $\varphi(e_1)$ až $\varphi(e_n)$,

$$\begin{aligned} \varphi(e_i) = \alpha_i^k e_k &\implies (\varphi(e_i), e_j) - (e_i, \varphi(e_j)) = \alpha_i^k (e_k, e_j) - \alpha_j^{\ell} (e_i, e_\ell) = \\ &= \alpha_i^k \delta_{kj} - \alpha_j^{\ell} \delta_{i\ell} = \alpha_i^j - \alpha_j^{i*}. \end{aligned}$$

Z definice samoadjungovaného operátoru je zřejmé, že výsledný výraz je nulový pro libovolné hodnoty $1 \leq i, j \leq n$, a tedy $A = A^{T*}$. Ve vektorovém prostoru nad \mathbf{R} pak $A = A^T$.

Matice, která reprezentuje samoadjungovaný lineární operátor v ortonormální bázi, je samoadjungovaná. Matice, která reprezentuje symetrický lineární operátor v ortonormální bázi, je symetrická.

Platí také následující tvrzení:

Je-li lineární operátor $\varphi : U_n \rightarrow U_n$, resp. $\varphi : E_n \rightarrow E_n$ reprezentován v některé ortonormální bázi samoadjungovanou, resp. symetrickou maticí, pak je samoadjungovaný, resp. symetrický.

Vyřešíme ještě problém vlastních vektorů a vlastních hodnot samoadjungovaného operátoru. Předpokládejme, že a je vlastní vektor samoadjungovaného operátoru φ příslušný vlastní hodnotě λ a připomeňme si, že vlastní vektor je vždy nenulový. Platí

$$0 = (\varphi(a), a) - (a, \varphi(a)) = (\lambda a, a) - (a, \lambda a) = (\lambda - \lambda^*)(a, a) \implies \lambda^* = \lambda.$$

Vlastní hodnota je komplexně sdružená sama k sobě, takže je reálná. Tento výsledek je možná poněkud překvapující — vždyť přece pracujeme v prostoru nad komplexními čísly a charakteristický polynom operátoru bude mít obecně komplexní koeficienty. Bez ohledu na to jsou však vlastní hodnoty samoadjungovaného operátoru — kořeny jeho charakteristického polynomu — reálné. Z fyzikálního hlediska je tento závěr velmi důležitý. Samoadjungovanými operátory totiž popisujeme fyzikální veličiny v kvantové fyzice a hodnoty, které můžeme měřením veličin v principu získat, jsou, jak jsme si již dříve řekli, právě vlastními hodnotami příslušných operátorů. Uvažme nyní dvě různé vlastní hodnoty λ_1 a λ_2 a vyberme k nim po jednom vlastní vektoru. Každé z hodnot přísluší pochopitelně nekonečně mnoho vlastních vektorů — celý vektorový podprostor (s výjimkou nulového vektoru, který za vlastní nepovažujeme). Vybereme však libovolně vždy pouze jeden z nich, vektor a k hodnotě λ_1 a vektor b k hodnotě λ_2 . Abychom využili definice samoadjungovaného operátoru, počítejme výraz $(\varphi(a), b) - (a, \varphi(b))$, o němž víme, že je nulový:

$$\begin{aligned} 0 &= (\varphi(a), b) - (a, \varphi(b)) = (\lambda_1 a, b) - (a, \lambda_2 b) = \\ &= (\lambda_1 - \lambda_2^*)(a, b) = (\lambda_1 - \lambda_2)(a, b). \end{aligned}$$

Vlastní hodnoty jsme zvolili různé, takže za nulovou hodnotu výsledku je zodpovědný skalární součin vektorů a a b . Tyto vektory jsou ortogonální. Protože jsme je z podprostorů vlastních vektorů L_1 a L_2 příslušejících hodnotám λ_1 a λ_2 vybrali libovolně, můžeme usoudit, že tyto prostory jsou ortogonální. Pokud jde o vektorové podprostory vlastních vektorů příslušných různým vlastním hodnotám, dospěli jsme ke stejnému závěru jako u unitárních operátorů — každé dva takové podprostory jsou ortogonální. Dále již můžeme přesně reprodukovat úvahu, pomocí níž jsme u unitárních operátorů dospěli k závěru, že všechny podprostory vlastních vektorů dohromady generují celý prostor U_n , a vyslovit následující tvrzení:

Věta 6.2: *Pro samoadjungovaný lineární operátor ve vektorovém prostoru U_n platí:*

- (1) *Všechny vlastní hodnoty operátoru jsou reálné.*
- (2) *Dimenze podprostoru vlastních vektorů operátoru příslušných téže vlastní hodnotě je shodná s násobností této vlastní hodnoty jako kořene charakteristického polynomu operátoru.*
- (3) *Každé dva podprostory vlastních vektorů operátoru příslušných různým vlastním hodnotám jsou ortogonální.*

Věta 6.2 má stejně formulované důsledky jako věta 6.1 pro unitární operátory:

Důsledky: Z vět 4.3 a 6.2 vyplývají tyto vlastnosti samoadjungovaných operátorů v U_n :

- Podprostory vlastních vektorů operátoru generují prostor U_n , tj. $L_1 + \dots + L_r = U_n$.
- Z vlastních vektorů operátoru lze sestavit ortonormální bázi prostoru U_n . V této bázi je operátor reprezentován diagonální maticí. Diagonála je obsazena vlastními hodnotami operátoru, z nichž každá se vyskytuje právě tolikrát, kolik činí její násobnost.

Vzhledem k tomu, že vlastní hodnoty jsou reálné, platí uvedené tvrzení a jeho důsledky také pro symetrické operátory v E_n !

Příklad 6.25: Opět všudypřítomné projekce

V odstavci 6.1.3 jsme se setkali s ortogonálními projekcemi do vektorových podprostorů. Zjistili jsme, že matice reprezentující takové projekce v ortonormálních bázích jsou samoadjungované ($P_L = C^{T*}C$, $P_L^{T*} = P_L$). Operátory ortogonálního promítání musí tedy být také samodjungované. Díky větě 6.2 tak předem víme, že každý operátor ortogonální projekce lze reprezentovat diagonální maticí. Vlastní hodnoty a vlastní vektory tohoto operátoru snadno určíme dokonce bez počítání. Předpokládejme, že základní unitární prostor je n -rozměrný a že promítáme do jeho k -rozměrného vektorového podprostoru L . Ortogonální doplněk L_\perp má tedy dimenzi $(n - k)$. Existují dvě základní „kategorie“ vlastních vektorů — vektory ležící v L (při promítání splývá obraz se vzorem, odpovídající vlastní hodnota je $\lambda_1 = 1$), a vektory ležící v L_\perp (obraz vektoru je vektorem nulovým, příslušná vlastní hodnota $\lambda_2 = 0$). Další vlastní vektory nejsou. Je tedy zřejmé, že vlastní hodnota $\lambda_1 = 1$ je k -násobná, hodnota $\lambda_2 = 0$ je $(n - k)$ -násobná. Podprostory L a L_\perp jsou ortogonální a dohromady generují celý prostor. V každém z nich zvolíme libovolnou bázi, kterou pak ortonormalizujeme. Získáme tak ortonormální bázi z vlastních vektorů operátoru projekce, v níž je tento operátor reprezentován diagonální maticí obsahující k jedniček a $(n - k)$ nul.

Ortogonální projekce se jeví jako velmi speciální případ samoadjungovaných operátorů. Ve skutečnosti jsou samoadjungované operátory samy natolik speciální, že se každý z nich dá vytvořit jako lineární kombinace vhodných projekcí. Zvolme libovolný samoadjungovaný operátor v U_n a označme L_1, L_2, \dots, L_r podprostory jeho vlastních vektorů příslušné všem navzájem různým vlastním hodnotám $\lambda_1, \lambda_2, \dots, \lambda_r$. Tyto podprostory generují celý vektorový prostor, jejich dimenze jsou rovny násobnostem vlastních hodnot a dohromady dávají součet n . Na základě toho, co víme o vektorových podprostorech a jejich doplňcích, si snadno uvědomíme, že libovolný vektor $a \in U_n$ lze jednoznačně rozložit do jednotlivých podprostorů, přičemž ortogonální průmět a_{L_j} vektoru a do podprostoru L_j je jeho obrazem získaným pomocí příslušného

operátoru ortogonální projekce π_j :

$$a = a_{L_1} + a_{L_2} + \cdots + a_{L_r} = \pi_1(a) + \pi_2(a) + \cdots + \pi_r(a).$$

Každý z vektorů $a_{L_j} = \pi_j(a)$ je ovšem vlastním vektorem operátoru φ příslušným vlastní hodnotě λ_j . Proto

$$\varphi(a) = \lambda_1\pi_1(a) + \lambda_2\pi_2(a) + \cdots + \lambda_r\pi_r(a).$$

Vektor a byl vybrán libovolně, předchozí rovnice proto platí nezávisle na jeho volbě, tj.

$$\varphi = \lambda_1\pi_1 + \lambda_2\pi_2 + \cdots + \lambda_r\pi_r. \quad (6.16)$$

Označíme-li A matici, která reprezentuje operátor φ v daná bázi a P_i , $i = 1, \dots, r$, matice reprezentující v téže bázi jednotlivé projekce, můžeme psát

$$A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_r P_r, \quad P_1 + P_2 + \cdots + P_r = E, \quad P_i P_j = \begin{cases} P_i & \text{pro } i = j \\ O & \text{pro } i \neq j \end{cases}$$

Vztah (6.16) se nazývá *spektrální reprezentace* operátoru φ . Určitě proto, že koeficienty ve vyjádření operátoru φ pomocí projekcí jsou prvky jeho spektra. Každý samoadjungovaný operátor je tedy lineární kombinací projekcí na podprostory svých vlastních vektorů, koeficienty této lineární kombinace jsou vlastní hodnoty operátoru φ .

Příklad 6.26: Spektrální reprezentace prakticky

Symetrický operátor v prostoru E_2 zadejme v ortonormální bázi (e_1, e_2) maticí

$$A = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix}.$$

Hledejme báze, v nichž bude tento operátor reprezentován diagonální maticí. Předem víme, že to budou všechny možné ortonormální báze tvořené vlastními vektory operátoru. Nejprve najdeme vlastní hodnoty.

$$0 = \det(A - \lambda E) = \det \begin{pmatrix} 9 - \lambda & -2 \\ -2 & 6 - \lambda \end{pmatrix} = (9 - \lambda)(6 - \lambda) - 4 \implies \lambda_1 = 10, \quad \lambda_2 = 5.$$

Získali jsme dvě různé jednonásobné vlastní hodnoty. Každé z nich podle věty 6.2 přísluší jednorozměrný podprostor vlastních vektorů. Tyto vektorové podprostory jsou ortogonální. Prozatím neznámé složky vlastních vektorů v bázi (e_1, e_2) označme $(\beta) = (\beta^1 \ \beta^2)$. Abychom je našli, musíme řešit soustavu rovnic $(\beta)(A - \lambda E) = (0)$ pro λ_1 a pro λ_2 . Tyto vlastní vektory pak budeme normovat, abychom získali ortonormální bázi u_1, u_2 :

$$(\beta)(A - \lambda_1 E) = (0) \implies \begin{pmatrix} \beta^1 & \beta^2 \end{pmatrix} \begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix},$$

$$\begin{aligned}
(A - \lambda_1 E)^T &= \begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \implies \beta^1 + 2\beta^2 = 0 \implies \\
&\implies (\beta) = \begin{pmatrix} -2\beta^2 & \beta^2 \end{pmatrix} \dots u_1 = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}, \\
(\beta)(A - \lambda_2 E) &= (0) \implies (\beta^1 \ \beta^2) \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}, \\
(A - \lambda_2 E)^T &= \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 \\ 0 & 0 \end{pmatrix} \implies 2\beta^1 - \beta^2 = 0 \implies \\
&\implies (\beta) = \begin{pmatrix} \beta^1 & 2\beta^1 \end{pmatrix} \dots u_2 = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}.
\end{aligned}$$

Vlastní vektory u_1 a u_2 tvoří ortonormální bázi v prostoru E_2 , v níž je zadaný operátor reprezentován diagonální maticí. Tato matice má tvar

$$D = TAT^{-1} = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix} \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix}.$$

Připomeňme si, že matice T je vytvořena ze složek vektorů u_1 a u_2 zapsaných do řádků. Co myslíte, existují ještě jiné ortonormální báze, v nichž je operátor reprezentován diagonální maticí? Jaké jsou to báze a jak vypadá reprezentující diagonální matice?

Pomocí vztahu (6.12) sestrojíme z vektorů získané ortonormální báze (\vec{u}_1, \vec{u}_2) matice ortogonálních projekcí na podprostory $L_1 = [u_1]$, $L_2 = [u_2]$.

$$\begin{aligned}
P_1 &= \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{pmatrix}, \\
P_2 &= \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{pmatrix}, \\
A &= \lambda_1 P_1 + \lambda_2 P_2 = 10 \begin{pmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{pmatrix} + 5 \begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{pmatrix}.
\end{aligned}$$

Poslední zápis je spektrální reprezentací operátoru v bázi (e_1, e_2) .

Příklad 6.27: Spektrum symetrického operátoru a geometrické hledisko

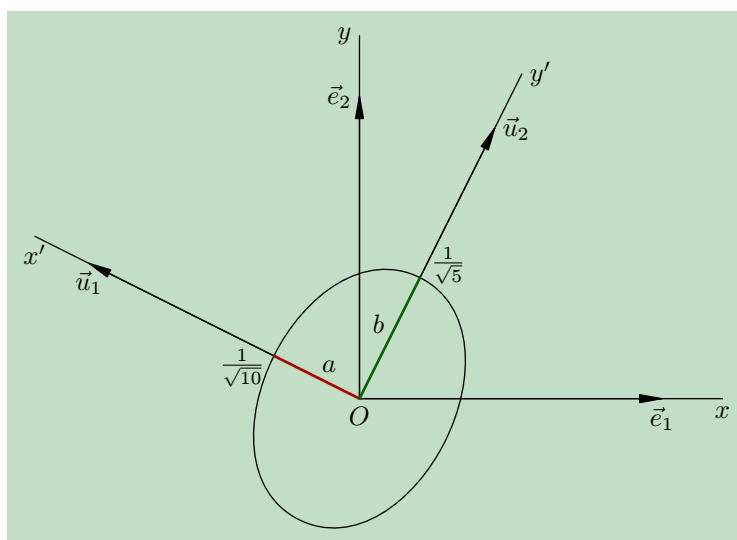
K předchozímu příkladu jsme přistupovali algebraicky. Rutinním výpočtem jsme vyřešili problém vlastních vektorů symetrického operátoru, který jsme bez bližšího popisu geometrického významu zadali maticí, a pak jsme sestrojili spektrální reprezentaci. Problém má však i jasný geometrický význam. Zvolme v euklidovské rovině počátek O a osy x a y , podél nichž jsou zvoleny navzájem kolmé jednotkové vektory \vec{e}_1 a \vec{e}_2 . Tyto vektory tvoří ortonormální bázi umístěnou v bodě O . Předpokládejme, že v rovině je zadána množina \mathcal{K} bodů, jejichž souřadnice $X = (x, y)$ zadané v soustavě $\langle O; x, y \rangle$ neboli $\langle O; \vec{e}_1, \vec{e}_2 \rangle$ splňují podmínku

$$X \in \mathcal{K} \iff 9x^2 - 4xy + 6y^2 - 1 = 0.$$

Na první pohled zřejmě těžko určíme, jak množina \mathcal{K} vypadá. Nejspíš by to mohla být nějaká křivka — ale jaká? Představme si dvojici souřadnic (x, y) jako složky vektoru \overrightarrow{OX} . Rovnici množiny \mathcal{K} zapišme, zdánlivě komplikovaně, takto:

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - 1 = 0, \quad A = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix}.$$

Přejdeme-li od ortonormální báze (\vec{e}_1, \vec{e}_2) pomocí matice přechodu T k jiné ortonormální bázi (\vec{u}_1, \vec{u}_2) , v níž



Obr. 6.12 Význačné směry elipsy.

označíme souřadnice bodu X jako (x', y') , dostaneme pro tutéž množinu \mathcal{K} formálně jinou rovnici. Platí

$$\begin{pmatrix} x & y \end{pmatrix} = \begin{pmatrix} x' & y' \end{pmatrix} T \implies \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' & y' \end{pmatrix} T A T^T \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Matice T je však ortogonální, tj. $T^T = T^{-1}$. Proto $A' = T A T^T = T A T^{-1}$. Matice A a A' můžeme interpretovat jako matice reprezentující symetrický lineární operátor φ v bázích (\vec{e}_1, \vec{e}_2) resp. (\vec{u}_1, \vec{u}_2) . Zvolíme-li bázi (\vec{u}_1, \vec{u}_2) tak, aby byla tvořena vlastními vektory tohoto operátoru (viz příklad 6.26), bude $T A T^T = D$. Rovnice množiny \mathcal{K} bude mít v této bázi tvar

$$10(x')^2 + 5(y')^2 - 1 = 0 \implies \left(\frac{x'}{\frac{1}{\sqrt{10}}}\right)^2 + \left(\frac{y'}{\frac{1}{\sqrt{5}}}\right)^2 = 1.$$

V tomto vyjádření již poznáváme rovnici elipsy o velikosti poloos $a = 1/\sqrt{10}$, $b = 1/\sqrt{5}$. Poloosa a je umístěna podél vektoru \vec{u}_1 , poloosa b podél vektoru \vec{u}_2 . Vlastní vektory operátoru φ mají tedy z hlediska elipsy \mathcal{K} význam určitých význačných směrů. Situace je znázorněna na obrázku 6.12. Zkuste si provést všechny naznačené výpočty.

Příklad 6.28: Ve vyšších dimenzích pomůže jen algebra

V prostorech vyšší dimenze a třeba i nad polem komplexních čísel se o geometrickou představu nemůžeme opřít. Problém spektrální reprezentace operátoru má však stejnou podstatu a také algebraický postup je stejný, jen o něco pracnější. Vyřešme si ještě jednu úlohu. Samoadjungovaný operátor v U_4 je v ortonormální bázi (e_1, e_2, e_3, e_4) zadán maticí

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix} \implies A - \lambda E = \begin{pmatrix} -\lambda & 1 & 0 & 0 \\ 1 & -\lambda & 0 & 0 \\ 0 & 0 & -\lambda & i \\ 0 & 0 & -i & -\lambda \end{pmatrix},$$

$$\begin{aligned} \det(A - \lambda E) &= \det \begin{pmatrix} -\lambda & 1 & 0 & 0 \\ 1 & -\lambda & 0 & 0 \\ 0 & 0 & -\lambda & i \\ 0 & 0 & -i & -\lambda \end{pmatrix} = \\ &= \det \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} \det \begin{pmatrix} -\lambda & i \\ -i & -\lambda \end{pmatrix} = (\lambda^2 - 1)^2 = (\lambda - 1)^2(\lambda + 1)^2 \implies \\ &\implies \lambda_1 = -1, k_1 = \dim L_1 = 2, \quad \lambda_2 = 1, k_2 = \dim L_2 = 2. \end{aligned}$$

Zjistili jsme, že dostaneme dva dvojrozměrné ortogonální podprostory vlastních vektorů L_1 a L_2 , matice A bude vyjádřena pomocí odpovídajících projekcí jako $A = \lambda_1 P_1 + \lambda_2 P_2 = -P_1 + P_2$. Protože spektrální reprezentace je v tomto případě dána pouze dvěma projekcemi, můžeme je určit dokonce i bez výpočtu vlastních vektorů. Platí totiž současně $P_1 + P_2 = E$, pro matice projekce tedy máme dvě rovnice. Dostaneme

$$P_1 = \frac{1}{2}(E - A), \quad P_2 = \frac{1}{2}(E + A).$$

Z „cvičných“ důvodů a pro pořádek vlastní vektory přece jen vypočteme. Označme jejich neznámé složky $(\beta) = (\beta^1 \ \beta^2 \ \beta^3 \ \beta^4)$. Platí

$$\begin{aligned} (\beta)(A - \lambda_1 E) &= (0), \\ (A - \lambda_1 E)^T &= \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -i \\ 0 & 0 & i & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ (\beta) &= (\beta^1 \ -\beta^1 \ \beta^3 \ -i\beta^3) \implies L_1 = \left[\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0, 0 \right), \left(0, 0, \frac{1}{\sqrt{2}}, -\frac{i}{\sqrt{2}} \right) \right]. \end{aligned}$$

$$\begin{aligned} (\beta)(A - \lambda_2 E) &= (0), \\ (A - \lambda_2 E)^T &= \begin{pmatrix} -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & -i \\ 0 & 0 & i & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ (\beta) &= (\beta^1 \ \beta^1 \ \beta^3 \ i\beta^3) \implies L_2 = \left[\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right), \left(0, 0, \frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}} \right) \right]. \end{aligned}$$

Vektory, které jsme prostřednictvím výběru volných neznámých zvolili jako generátory podprostorů vlastních vektorů, tvoří v prostoru U_4 ortonormální bázi (u_1, u_2, u_3, u_4) . Matice přechodu od báze (e_1, e_2, e_3, e_4) k bázi (u_1, u_2, u_3, u_4) je

$$T = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix}.$$

Ověřte, že platí $TAT^{-1} = TAT^{T*} = D$, kde D je diagonální matice s diagonálou $\text{diag } D = (-1, -1, 1, 1)$.

O použití lineární algebry v kvantové mechanice jsme se již zmínili. V příkladu 6.23 jsme použili lineární algebru při konkrétním popisu časového vývoje kvantověmechanické soustavy. V úvodu k odstavci 4.3 jsme pouze avizovali její využití pro určení možných hodnot, které můžeme při měření fyzikálních veličin na kvantověmechanické soustavě získat a stavů, do kterých soustava při měření přejde. Této úlohy si nyní všimneme podrobněji v následujícím příkladu, v němž také alespoň v obrysech vyložíme fyzikální podstatu problému.

Příklad 6.29: Lineární operátory a kvantová mechanika

Víme již, že kvantová mechanika popisuje chování mikročastic nebo jejich soustav. Tak třeba systém elektronů v elektronovém obalu atomu, nebo ve valenčním či vodivostním pásu polovodiče, se chová právě podle pravidel kvantové mechaniky. Typickým kvantovým projevem je skutečnost, že fyzikální veličiny charakterizující částice nebo soustavy nemohou někdy nabývat všech možných spojitě rozložených hodnot, ale jen některých z nich. Dochází ke *kvantování*. Kvantována je například energie elektronů v obalu atomu, kde nabývá hodnot úměrných veličině n^{-2} . Přirozené číslo n , $n = 1, 2, \dots$, přitom čísluje takzvané *energieové hladiny* neboli *slupky* elektronu. Obdobně jsou kvantovány i hodnoty jiných veličin, například velikost momentu hybnosti a jeho složky. K tomu, abychom možné hodnoty fyzikální veličiny, které budeme dostávat, budeme-li veličinu měřit, předem předpověděli, potřebujeme vhodný matematický aparát. Stav částice již nepopisujeme její polohou a rychlostí (nebo hybností), jako jsme tomu byli zvyklí v mechanice klasické. Není to totiž možné, neboť pro kvantové částice platí *Heisenbergovy relace neurčitosti*, podle nichž má částice tím neurčitější hybnost, čím určitější má polohu a naopak. (Například pro částici, pro niž je přípustný pouze jednorozměrný pohyb, platí $\Delta x \Delta p_x \sim \hbar$, $\hbar \doteq 6,63 \cdot 10^{-34} \text{ J} \cdot \text{s}$.) Těchto relací jsme se již dotkli v úvodu odstavce 4.3. Výsledky a experimenty kvantové teorie potvrzují, že stav částice, či obecně jakékoli soustavy, lze popsat vektorem z vhodně definovaného unitárního vektorového prostoru, tj. takového, v němž je zaveden skalární součin. Tento prostor má často, či spíše většinou, nekonečnou dimenzi a v souvislosti s ní je třeba požadovat ještě některé další vlastnosti. Jedná se o takzvaný *Hilbertův prostor*. V případě, že je jeho dimenze konečná, jsou dodatečné vlastnosti splněny automaticky, a proto v tuto chvíli nevádí, že o nich nebudeme hovořit. Měření fyzikální veličiny je popsáno působením samoadjungovaného lineárního operátoru, definovaného v prostoru stavů, na ten stav soustavy, v němž se bezprostředně před měřením nachází. Hodnoty, které mohou být naměřeny, jsou právě vlastní hodnoty tohoto operátoru. V principu nelze měřením získat hodnoty jiné, daná fyzikální veličina jich prostě „nenabývá“, jsou „zakázané“. Dopředu však nemůžeme říci, že danou vlastní hodnotu skutečně naměříme, můžeme pouze předem určit její pravděpodobnost. Pokud se při měření daná hodnota skutečně realizovala (naměřili jsme ji), máme již jistotu, že se soustava nyní nachází ve stavu, který je vlastním vektorem operátoru fyzikální veličiny. Při měření soustava přešla do stavu, který je popsán vlastním vektorem operátoru měřené veličiny (tzv. *vlastní stav*). Že je to „přitaženo za vlasy“ a že není jasné, proč právě takový matematický aparát máme použít? Je samozřejmé, že k formulaci matematického popisu se fyzikové dostávali postupně a že jediným kritériem správné

6.2. „FYZIKÁLNÍ“ LINEÁRNÍ OPERÁTORY A JEJICH VLASTNÍ VEKTORY 461

volby je experiment. A ten v daném případě teorii potvrzuje. Vyjděme tedy z uvedených skutečností a použijme je pro řešení konkrétní úlohy.

Uvažujme například o elektronu v atomovém obalu. Jeho stav je určen jistou obecně komplexní funkcí polohy \vec{r} . Již v prvním dílu jsme se zmínili o tom, že funkce (tehdy šlo o reálné funkce jedné reálné proměnné) s operacemi sčítání funkcí a násobení funkce (reálným) číslem tvoří vektorový prostor. Se znalostmi, které nyní máme o vektorových prostorech, můžeme snadno ověřit, že také komplexní funkce reálných proměnných tvoří vzhledem k operacím sčítání funkcí a násobení funkce komplexním číslem vektorový prostor nad \mathbf{C} . Tento prostor je sice nekonečněrozměrný, ale při řešení naší úlohy to nebude představovat komplikaci. Skalární součin dvou funkcí f a g definujeme obdobně jako v příkladu 6.5, kdy jsme jej definovali pro polynomy jedné proměnné,

$$(f, g) = \int_V f^*(\vec{r})g(\vec{r}) dV,$$

kde V je definiční obor funkcí (objem V může samozřejmě být i celý prostor \mathbf{R}^3). Předpokládá se, že „vektory“ stavu jsou normovány, tj.

$$(f, f) = \int_V f^*(\vec{r})f(\vec{r}) dV = \int_V |f(\vec{r})|^2 dV = 1.$$

Poznamenejme ještě, že pro $\Omega \subset V$ má integrál

$$P(\Omega) = \int_{\Omega} f^*(\vec{r})f(\vec{r}) dV$$

význam pravděpodobnosti nalezení částice v objemu Ω .

Pro případ elektronu v atomovém obalu vyjádříme jeho stav jako funkci sférických souřadnic (r, ϑ, φ) , které jsme zavedli v kapitole 5. Složka momentu hybnosti elektronu, vezměme například z -ovou, je představována operátorem derivace $\hat{L}_z = -i\hbar(\partial/\partial\varphi)$, $\hbar = \frac{h}{2\pi}$. Derivace je, jak víme z příkladu 4.53, lineárním operátorem. (Prověřte také jeho samoadjungovanost.) Přípustné hodnoty složky momentu hybnosti L_z jsou jeho vlastními hodnotami a vlastní stavy $\phi(r, \vartheta, \varphi)$ jsou tvaru $\phi(r, \vartheta, \varphi) = \psi(\varphi)\Theta(r, \vartheta)$, kde funkce $\psi(\varphi)$ jsou určeny rovnicí pro vlastní vektory

$$\hat{L}_z\phi(r, \vartheta, \varphi) = L_z\phi(r, \vartheta, \varphi) \implies -i\hbar\frac{\partial\psi(\varphi)}{\partial\varphi} = L_z\psi(\varphi).$$

Hned je vidět, že řešením této rovnice jsou právě všechny funkce tvaru

$$\psi(\varphi) = K \exp\left(\frac{i}{\hbar}L_z\varphi\right), \quad K \text{ je konstanta.}$$

Vzhledem k významu proměnné φ (azimutální úhel ve sférických souřadnicích) je zřejmé, že stav odpovídající hodnotě $\varphi + 2k\pi$ pro každé celé číslo k , tedy i pro $k = 1$, musí být stejný jako stav odpovídající φ , tj.

$$K \exp\left(\frac{i}{\hbar}L_z(\varphi + 2\pi)\right) = K \exp\left(\frac{i}{\hbar}L_z\varphi\right) \implies \exp\left(\frac{2\pi i}{\hbar}L_z\right) = 1 \implies L_z = m\hbar, \quad m \in \mathbf{Z}.$$

Vidíme, že složka momentu hybnosti je kvantována. (Když jsme si již jednu složku momentu hybnosti zvolili, třeba proto, že směr osy z je nějak fyzikálně významný, nemůžeme už stejně uvažovat o x -ové a y -ové složce. Je to proto, že operátory složek momentu hybnosti *nekomutují*, při jejich postupném působení na vektor stavu záleží na pořadí tohoto působení. Proto nemají shodné vlastní vektory a nejsou současně měřitelné. Má-li se totiž soustava po měření ocitnout ve stavu popsáném vlastním vektorem operátoru měřené veličiny, jsou

6.2. „FYZIKÁLNÍ“ LINEÁRNÍ OPERÁTORY A JEJICH VLASTNÍ VEKTORY 463

lze právě jen vlastní hodnoty tohoto operátoru a po měření najdeme soustavu v odpovídajícím vlastním stavu. (S tímto popisem dobře koresponduje skutečnost, že vlastní hodnoty samoadjungovaného operátoru jsou reálné.) Nyní obě tyto vlastnosti soustav podléhajících zákonitostem mikrosvěta spojíme a získáme další způsob popisu časového vývoje soustavy. Necht' \mathcal{F} je samoadjungovaný operátor představující fyzikální veličinu. Označme λ_1 až λ_r jeho navzájem různé vlastní hodnoty a L_1 až L_r odpovídající ortogonální podprostory vlastních vektorů, π_1 až π_r ortogonální projekce na tyto podprostory. Pro libovolný stav Ψ platí

$$\Psi = \pi_1(\Psi) + \pi_2(\Psi) + \cdots + \pi_r(\Psi), \quad (\pi_k(\Psi), \pi_\ell(\Psi)) = 0 \text{ pro } k \neq \ell.$$

Operátor \mathcal{F} má spektrální reprezentaci

$$\mathcal{F} = \sum_{k=1}^r \lambda_k \pi_k, \quad \text{tj.} \quad \mathcal{F}(\Psi) = \sum_{k=1}^r \lambda_k \pi_k(\Psi).$$

Platí

$$(\mathcal{F}(\Psi), \Psi) = \left(\sum_{k=1}^r \lambda_k \pi_k(\Psi), \sum_{\ell=1}^r \pi_\ell(\Psi) \right) = \sum_{k=1}^r \sum_{\ell=1}^r \lambda_k (\pi_k(\Psi), \pi_\ell(\Psi)) = \sum_{k=1}^r \lambda_k (\pi_k(\Psi), \pi_k(\Psi)).$$

Skalární součin $w_k = (\pi_k(\Psi), \pi_k(\Psi))$ představuje kvadrát „délky“ průmětu vektoru stavu Ψ do podprostoru L_k , tedy pravděpodobnost, že při měření veličiny \mathcal{F} zjistíme vlastní hodnotu λ_k a realizuje se vlastní stav ležící v L_k . Skalární součin

$$(\mathcal{F}(\Psi), \Psi) = \sum_{k=1}^r \lambda_k w_k$$

je střední hodnotou veličiny \mathcal{F} . Vyjádříme ji nyní pro případ stavu soustavy $\Psi(t)$ v čase t s použitím unitárních operátorů $TAM = \mathcal{T}_{(t_0 \rightarrow t)}$ a $ZP\check{E}T = \mathcal{T}_{(t \rightarrow t_0)} = \mathcal{T}_{(t_0 \rightarrow t)}^{-1} = TAM^{-1}$ z příkladu 6.23. Působení operátoru \mathcal{F} na stav Ψ v čase t zapíšeme jako $\mathcal{F}(t)(\Psi(t))$. Platí

$$\begin{aligned} (\mathcal{F}(t)(\Psi(t)), \Psi(t)) &= (\mathcal{F}(t)\mathcal{T}_{(t_0 \rightarrow t)}(\Psi(t_0)), \mathcal{T}_{(t_0 \rightarrow t)}(\Psi(t_0))) = \\ &= (\mathcal{T}_{(t \rightarrow t_0)}\mathcal{F}(t)\mathcal{T}_{(t_0 \rightarrow t)}(\Psi(t_0)), \mathcal{T}_{(t \rightarrow t_0)}\mathcal{T}_{(t_0 \rightarrow t)}(\Psi(t_0))) = (\mathcal{F}(t_0)(\Psi(t_0)), \Psi(t_0)), \\ \mathcal{F}(t_0) &= \mathcal{T}_{(t \rightarrow t_0)}\mathcal{F}(t)\mathcal{T}_{(t_0 \rightarrow t)} = \mathcal{T}_{(t \rightarrow t_0)}\mathcal{F}(t)\mathcal{T}_{(t \rightarrow t_0)}^{-1} \implies \mathcal{F}(t) = \mathcal{T}_{(t_0 \rightarrow t)}\mathcal{F}(t_0)\mathcal{T}_{(t_0 \rightarrow t)}^{-1}, \quad \text{tj.} \\ \mathcal{F}(t) &= (TAM)\mathcal{F}(t_0)(ZP\check{E}T). \end{aligned}$$

(Při úpravách jsme využili unitárnosti operátoru $\mathcal{T}_{(t \rightarrow t_0)}$, tj. vztahu

$$(\mathcal{T}_{(t \rightarrow t_0)}(\Psi_1(t)), \mathcal{T}_{(t \rightarrow t_0)}(\Psi_2(t))) = (\Psi_1(t), \Psi_2(t)),$$

platného pro dva libovolné stavy $\Psi_1(t)$ a $\Psi_2(t)$.) Získaný výsledek lze interpretovat také tak, že se časový vývoj připisuje operátoru \mathcal{F} , zatímco stav zůstává neměnný, $\Psi(t_0)$. Mluvíme pak o *heisenbergovské reprezentaci* časového vývoje soustavy, zatímco v příkladu 6.23 šlo o reprezentaci *schrödingerovskou*.

6.2.3 Cvičení

1. Ukažte, že pro případ euklidovského prostoru stačí definovat ortogonální operátor vztahem

$$(\varphi(a), \varphi(a)) = (a, a)$$

pro libovolný vektor $a \in E_n$. Je tato definice ekvivalentní také s definicí unitárního operátoru?

Návod: Výsledek plyne ze vztahu $(\varphi(a+b), \varphi(a+b)) = (a+b, a+b)$, platného pro libovolné vektory $a, b \in E_n$. Pro U_n je definice uvedená v zadání úlohy nedostačující.

2. Nechť $\varphi \in L(U_n, U_n)$ je unitární lineární operátor. Jakou podmínku je třeba klást na číslo $\kappa \in \mathbf{C}$, aby operátor $\kappa\varphi$ byl unitární?
3. Dokažte, že složení dvou unitárních lineárních operátorů je opět unitární.
4. Samoadjungované lineární operátory φ_1 až φ_7 jsou v ortonormální bázi zadány maticemi

$$A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 1 & 0 & -i \\ 0 & 1 & 0 \\ i & 0 & 1 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{pmatrix}, \quad A_6 = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{pmatrix}, \quad A_7 = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}.$$

Najděte jejich spektrální reprezentace.

5. Nechť φ, ψ jsou samoadjungované lineární operátory. Diskutujte, zda $\varphi \circ \psi, \varphi + \psi$ a $k\varphi$ jsou také samoadjungované.
6. Nechť φ a ψ jsou operátory ve V_n , pro něž existuje báze (u_1, u_2, \dots, u_n) prostoru V_n taková, že každý z jejích prvků je vlastním vektorem operátoru φ i ψ . Pak operátory komutují. Dokažte.
Návod: Zapište obecný vektor $a \in U_n$ v bázi u_1, \dots, u_n , tj. $a = \alpha^1 u_1 + \dots + \alpha^n u_n$. Označte λ_i , resp. μ_i vlastní hodnotu operátoru φ , resp. ψ , příslušnou vlastnímu vektoru $u_i, i = 1, \dots, n$. (Některé vlastní hodnoty daného operátoru mohou být stejné, i když jsou označeny různými indexy.) Vyjádřete vektory $\psi(\varphi(a))$ a $\varphi(\psi(a))$ a porovnejte je.
7. Uvažujte o ortogonálním operátoru $\varphi : E_n \rightarrow E_n$. Označte $L = L_1 + \dots + L_r$, kde L_1 až L_r jsou vektorové podprostory příslušné všem navzájem různým vlastním hodnotám λ_1 až λ_r operátoru φ . Dokažte, že pro každý vektor $b \in L_\perp$ platí $\varphi(b) \in L_\perp$.
Návod: Využijte ortogonálnosti operátoru φ a ukažte, že ze vztahu $(b, a) = 0$ pro libovolný vektor $a \in L$ plyne $(\varphi(b), a) = 0$. Poznámka: Úloha se vztahuje k situaci, kdy existují reálné charakteristické kořeny operátoru φ .
8. V odstavcích 6.2.1 a 6.2.2 jsme dokázali, že unitární operátor je v ortonormálních bázích reprezentován unitární maticí, samoadjungovaný operátor samoadjungovanou maticí. Formulovali jsme také tato tvrzení: Je-li lineární operátor $\varphi : U_n \rightarrow U_n$, reprezentován v některé ortonormální bázi unitární, resp. samoadjungovanou maticí, pak je unitární, resp. samoadjungovaný. Dokažte je.

6.3 Symetrické operátory v geometrii a ve fyzice

Základní vlastností samoadjungovaných, resp. symetrických operátorů je možnost jejich reprezentace diagonální maticí ve vhodné bázi. Jednu z aplikací této vlastnosti jsme poznali v

příkladu 6.27. Zatímco v obecné ortonormální bázi jsme elipsu hned nerozpoznali, po převedení do báze spojené s její velkou a malou poloosou ji její rovnice už prozradila. Obecnou rovnici elipsy jsme zapsali pomocí vhodně zvoleného symetrického lineárního operátoru a směry jejich poloos jsme pak určili jako vlastní vektory tohoto operátoru. Velikosti poloos souvisely s příslušnými vlastními hodnotami. Obdobná situace je ve fyzice. Symetrickou matici momentu setrvačnosti můžeme také interpretovat jako reprezentanta symetrického operátoru v trojrozměrném vektorovém prostoru. Směry určené vlastními vektory představují takzvané *hlavní osy momentu setrvačnosti*, jejichž určení je důležité pro popis rotačního pohybu tuhého tělesa (setrvačnicku). V tomto odstavci si všimneme právě takovýchto aplikací symetrických operátorů. Budeme uvažovat pouze o vektorových prostorech nad \mathbf{R} .

6.3.1 Kvadratické formy

S kvadratickou formou jsme se setkali již v příkladu 6.27, aniž jsme však tohoto názvu použili. Šlo o výraz $9x^2 - 4xy + 6y^2$ v rovnici elipsy, obsahující kvadráty a součiny souřadnic.

Kvadratickou formou n proměnných nazýváme zobrazení přiřazující n -tici reálných proměnných (x^1, x^2, \dots, x^n) reálné číslo podle předpisu

$$\kappa : (x^1, x^2, \dots, x^n) \longrightarrow a_{ij}x^i x^j \in \mathbf{R}, \quad a_{ij} = a_{ji}, \quad 1 \leq i, j \leq n, \quad (6.17)$$

kde alespoň jedno z čísel a_{ij} je nenulové.

Pozn.: S ohledem na praktické aplikace jsme z definice vyloučili „nulovou“ kvadratickou formu, pro níž je $a_{ij} = 0$ pro všechny hodnoty $i, j = 1, \dots, n$.

Kvadratická forma z příkladu 6.27 je definována pro dvě proměnné $x^1 = x$ a $x^2 = y$, její koeficienty jsou $a_{11} = 9$, $a_{12} = a_{21} = -2$, $a_{22} = 6$. Pro kvadratickou formu můžeme s výhodou použít maticového zápisu, jak jsme to již udělali v příkladu 6.27:

$$\begin{pmatrix} x^1 & x^2 & \dots & x^n \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix} = (x)A(x)^T.$$

Matice A je podle definice (6.17) maticí symetrickou. Pokud je kvadratická forma zadána obvyklým způsobem, třeba jako v příkladu 6.27 ve tvaru $9x^2 - 4xy + 6y^2$, je zřejmé, že koeficient -4 má význam součtu $(a_{12} + a_{21})$. Díky symetrii předepsané vztahem (6.17) je $a_{12} = a_{21} = -2$.

Obdobně jako v příkladu 6.27 můžeme n -tici $(x) = (x^1, x^2, \dots, x^n)$ chápat jako složky vektoru v dané ortonormální bázi (e_1, e_2, \dots, e_n) . Přejdeme-li pomocí matice přechodu T do jiné ortonormální báze (f_1, f_2, \dots, f_n) , v níž bude mít tento vektor složky $(y) = (y^1, y^2, \dots, y^n)$

(to abychom se nemuseli obtěžovat s čárkovanými proměnnými), platí podle vztahů (4.13) $(y) = (x)T^{-1}$ a naopak $(x) = (y)T$. Pro kvadratickou formu dostáváme v nové bázi

$$(x)A(x)^T = ((y)T)A((y)T)^T = (y)TAT^T(y)^T = (y)TAT^{-1}(y)^T.$$

Využili jsme skutečnosti, že matice přechodu mezi ortonormálními bázemi je ortogonální, proto je její inverzní matice rovna matici transponované. Pokud budeme matici A chápat jako reprezentanta symetrického lineárního operátoru φ ve vektorovém prostoru n -tic v bázi (e_1, e_2, \dots, e_n) , představuje TAT^{-1} matici (opět symetrickou) téhož operátoru v bázi čárkované. Symetrický operátor však lze podle věty 6.2 diagonalizovat, tj. reprezentovat jej ve vhodné bázi diagonální maticí. A takovou bázi je právě jakákoli ortonormální báze tvořená vlastními vektory operátoru φ . Dejme tomu, že (f_1, f_2, \dots, f_n) je právě takovou bázi. Pak platí

$$TAT^T = TAT^{-1} = D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Kvadratická forma, vyjádřená v této bázi v proměnných (y^1, y^2, \dots, y^n) , má tvar

$$(y)D(y)^T = \lambda_1(y^1)^2 + \lambda_2(y^2)^2 + \dots + \lambda_n(y^n)^2. \quad (6.18)$$

Vztah 6.18 nazýváme *kanonickým tvarem* kvadratické formy. Symboly λ_1 až λ_n jsou označeny postupně všechny charakteristické kořeny (tj. současně vlastní hodnoty) operátoru φ , pomocí něhož jsme zapsali kvadratickou formu, včetně násobnosti (znamená to, že některé mohou být stejné, například $\lambda_1 = \lambda_2 = \dots = \lambda_k$ pro k -násobnou vlastní hodnotu). Všimněme si nyní, jakých hodnot může kvadratická forma nabývat. Nejpohodlnější bude zkoumat tyto možnosti pomocí kanonického tvaru, který je velmi jednoduchý a leccos je z něj vidět na první pohled. Závěry učiněné pro kanonický tvar budou pro danou kvadratickou formu platné i obecně, neboť obor hodnot jedné a téže kvadratické formy se nemůže měnit volbou báze. A kanonický tvar je pouze vyjádřením této formy ve vhodné zvolené bázi. „Kategorie“ hodnot, kterých může kvadratická forma nabývat pro různé n -tice proměnných (y^1, y^2, \dots, y^n) , jsou dány vlastními hodnotami, nebo ještě konkrétněji, jejich znaménky. Rozebereme jednotlivé případy:

- (1) Předpokládejme, že všechny vlastní hodnoty jsou kladné. Z (6.18) plyne, že pro libovolnou volbu n -tice (y^1, y^2, \dots, y^n) je hodnota kvadratické formy $(y)D(y)^T$ kladná, s výjimkou n -tice nulové, pro niž nabývá nulové hodnoty. Taková kvadratická forma se nazývá *pozitivně definitní*. Všimněme si, že matice D je regulární a subdeterminanty všech řádů v levém horním rohu jsou kladné. Tato vlastnost zůstane zachována i po přechodu do obecné báze

(e_1, e_2, \dots, e_n) , v níž bude kvadratická forma zapsána pomocí matice $A = SDS^T$, kde S je matice přechodu od báze (f_1, f_2, \dots, f_n) , v níž má forma kanonický tvar, k obecné bázi (e_1, e_2, \dots, e_n) . Matice A je tedy rovněž regulární a její subdeterminanty všech řádů v levém horním rohu jsou kladné. Vzpomínáte si, že v odstavci o skalárním součinu jsme takovou matici nazývali pozitivně definitní? Právě symetrické pozitivně definitní matice reprezentovaly v bázích skalární součin zadaný ve vektorovém prostoru nad \mathbf{R} . Skalární součin skutečně tedy můžeme interpretovat také jako pozitivně definitní *bilineární formu* $(a, b) = (\alpha)A(\beta)^T$. Často tak bývá i definován.

- (2) V případě, že jsou všechny vlastní hodnoty záporné, je pro libovolnou volbu n -tice (y^1, y^2, \dots, y^n) hodnota kvadratické formy $(y)D(y)^T$ záporná, s výjimkou n -tice nulové. Pro ni nabývá nulové hodnoty. Taková kvadratická forma se nazývá *negativně definitní*. Matice D je opět regulární, avšak znaménka subdeterminantů v levém horním rohu se střídají (levý horní subdeterminant prvního řádu je záporný, druhého řádu kladný, atd.). Tato vlastnost zůstane, stejně jako v předchozím případě, zachována i po přechodu do obecné báze (e_1, e_2, \dots, e_n) , v níž bude kvadratická forma zapsána pomocí matice A . Matice A je tedy rovněž regulární a její subdeterminanty všech řádů v levém horním rohu střídají znaménka.
- (3) Další speciální případ nastává, když spektrum lineárního operátoru příslušného dané kvadratické formě obsahuje nulu, která je, dejme tomu, d -násobná, $1 \leq d < n$. Označme $h = n - d$. Předpokládejme, že všechny nenulové vlastní hodnoty jsou kladné. Vhodným očíslováním prvků báze (f_1, f_2, \dots, f_n) dokážeme zařídit, aby v matici D byly nuly právě na posledních d pozicích v diagonále. Z výrazu (6.18) pro kanonický tvar kvadratické formy vidíme, že

$$(y)D(y)^T = \lambda_1(y^1)^2 + \lambda_2(y^2)^2 + \dots + \lambda_h(y^h)^2$$

a forma nabývá pouze nezáporných hodnot. Pro n -tice

$$(0, 0, \dots, 0, y^{h+1}, \dots, y^n)$$

jsou to hodnoty nulové, v ostatních případech kladné. Forma se nazývá *pozitivně semidefinitní*. Její matice v libovolné bázi je singularární, má hodnotu h a její levé horní subdeterminanty jsou kladné až do řádu h včetně, další jsou nulové.

- (4) Jsou-li nenulové vlastní hodnoty naopak záporné, nabývá kvadratická forma pouze nekladných hodnot. Pro n -tice

$$(0, 0, \dots, 0, y^{h+1}, \dots, y^n)$$

jde o hodnoty nulové, v ostatních případech záporné. Forma se nazývá *negativně semidefinitní*. Její matice v libovolné bázi je singularární, má hodnotu h a její levé horní subdeterminanty střídají znaménka (první je záporný) až do řádu h včetně, další jsou nulové.

- (5) Ve zbývajících případech se forma nazývá *indefinitní*. Může nabývat jak hodnot obou znamének, tak hodnoty nulové.

Číslo h se nazývá *hodnost* kvadratické formy, dvojice $[k, z]$, kde k je počet kladných a z počet záporných vlastních hodnot (vždy včetně násobnosti), představuje její *signaturu*. Signatura pozitivně definitní formy je $[n, 0]$, negativně definitní formy $[0, n]$, signatura pozitivně, resp. negativně semidefinitní formy je $[h, 0]$, resp. $[0, h]$, $0 < h < n$. (Pro $h = 0$ by se jednalo o „nulovou“ kvadratickou formu, kterou do našich úvah na tomto místě nezahrnujeme.)

Odstavec o kvadratických formách uzavřeme jediným ukázkovým příkladem (všechny ostatní by byly analogické), který však bude velmi podrobný a poslouží hned několika účelům. Na prvním místě jako ukázka převodu kvadratické formy na kanonický tvar, dále jako praktické opakování podstatných výsledků z lineární algebry, a konečně jako potvrzení toho, že lineární algebra není nějakou samoučelnou abstraktní teorií, ale je přímo použitelná v aplikacích. Při jeho řešení použijeme vlastnosti kvadratických forem, teorii řešení soustav lineárních rovnic, vlastnosti vektorových prostorů, vlastnosti skalárního součinu a řešení problému vlastních vektorů symetrických lineárních operátorů.

Příklad 6.32: Kanonický tvar kvadratické formy prakticky

V ortonormální bázi čtyřrozměrného prostoru (e_1, e_2, e_3, e_4) je kvadratická forma κ zadána proměnnými (x^1, x^2, x^3, x^4) takto:

$$\kappa(x^1, x^2, x^3, x^4) = 2x^1x^2 + 2x^1x^3 - 2x^1x^4 - 2x^2x^3 + 2x^2x^4 + 2x^3x^4.$$

Její matice je

$$A = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}.$$

Velmi rychle zjistíme, že determinant této matice je roven -3 , tj. nenulový. Předem tedy můžeme konstatovat, že hodnost formy je rovna n a vyloučit tak, že je forma semidefinitní. Protože je determinant matice A záporný, nemůže být forma ani pozitivně definitní (to by musel být determinant kladný). Forma je tedy buď negativně definitní nebo indefinitní. Určíme-li však ještě levé horní subdeterminanty matice A prvního, druhého a třetího řádu, dostaneme čísla $0, -1, -2$. Znaménka se nestřídají, forma tedy není ani negativně definitní. Zbývá jediná možnost — forma je indefinitní. A to všechno jsme zjistili jen na základě matice A , aniž jsme kromě několika determinantů cokoli počítali. Určíme ještě kanonický tvar formy a bázi vlastních vektorů příslušného operátoru φ , determinant matice $A - \lambda E$ vypočteme pomocí Laplaceova rozvoje podle prvního sloupce:

$$\begin{aligned} \det(A - \lambda E) &= \det \begin{pmatrix} -\lambda & 1 & 1 & -1 \\ 1 & -\lambda & -1 & 1 \\ 1 & -1 & -\lambda & 1 \\ -1 & 1 & 1 & -\lambda \end{pmatrix} = \\ &= -\lambda \det \begin{pmatrix} -1 & -1 & 1 \\ -1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix} - \det \begin{pmatrix} 1 & 1 & -1 \\ -1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix} + \end{aligned}$$

$$\begin{aligned}
& +\det \begin{pmatrix} 1 & 1 & -1 \\ -\lambda & -1 & 1 \\ 1 & 1 & -\lambda \end{pmatrix} + \det \begin{pmatrix} 1 & 1 & -1 \\ -\lambda & -1 & 1 \\ -1 & -\lambda & 1 \end{pmatrix} = \\
& = \lambda^4 - 6\lambda^2 + 8\lambda - 3 = (\lambda - 1)^3(\lambda + 3).
\end{aligned}$$

Vlastní hodnoty symetrického operátoru φ odpovídajícího této kvadratické formě jsou $\lambda_1 = \lambda_2 = \lambda_3 = 1$, $\lambda_4 = -3$. Její kanonický tvar (v bázi vlastních vektorů a v proměnných (y^1, y^2, y^3, y^4)) je

$$\kappa' = (y^1)^2 + (y^2)^2 + (y^3)^2 - 3(y^4)^2.$$

Signatura formy je $[3, 1]$. Označme čtveřici neznámých složek vlastního vektoru $(\beta^1, \beta^2, \beta^3, \beta^4)$ a upravujeme matici $(A - \lambda E)^T = (A - \lambda E)$ pro $\lambda = 1$, resp. $\lambda = -3$ na schodovitý tvar:

$$A - E = \begin{pmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Dostáváme jedinou rovnici $\beta^1 - \beta^2 - \beta^3 + \beta^4 = 0$. Obecné řešení tvoří trojrozměrný vektorový podprostor L generovaný čtveřicemi $(\beta^2 + \beta^3 - \beta^4, \beta^2, \beta^3, \beta^4)$. Jeho bázi dostaneme například tak, že za volné neznámé β^2, β^3 a β^4 volíme postupně jednu jedničku a dvě nuly. Pak

$$L = [(1, 1, 0, 0), (1, 0, 1, 0), (-1, 0, 0, 1)].$$

Pro hodnotu $\lambda = -3$ dostáváme

$$A + 3E = \begin{pmatrix} 3 & 1 & 1 & -1 \\ 1 & 3 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{pmatrix} \sim \begin{pmatrix} -1 & 1 & 1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Této schodovité matici odpovídá soustava rovnic

$$\begin{aligned}
-\beta^1 + \beta^2 + \beta^3 + 3\beta^4 &= 0, \\
\beta^2 + \beta^4 &= 0, \\
\beta^3 + \beta^4 &= 0.
\end{aligned}$$

Její obecné řešení obsahuje jedinou volnou neznámou (hodnota je $h(A + 3E) = 3$) a má tvar $(\beta^4, -\beta^4, -\beta^4, \beta^4)$. Generuje ortogonální doplněk podprostoru L :

$$L_{\perp} = [(1, -1, -1, 1)].$$

Celý prostor $E_4 = L + L_{\perp}$ je generován bázi z vlastních vektorů operátoru φ ,

$$E_4 = [(1, 1, 0, 0), (1, 0, 1, 0), (-1, 0, 0, 1), (1, -1, -1, 1)].$$

Tuto bázi je ovšem třeba ortonormalizovat, neboť všechny naše úvahy jsou vedeny právě v ortonormálních bázích. Grammův-Schmidtův ortogonalizační proces (viz obecný postup nebo příklad 6.10) vede k výsledku

$E_4 = \{f_1, f_2, f_3, f_4\}$, kde

$$\begin{aligned} f_1 &= \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right), \\ f_2 &= \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \sqrt{\frac{2}{3}}, 0 \right), \\ f_3 &= \left(-\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{\sqrt{3}}{2} \right), \\ f_4 &= \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \right). \end{aligned}$$

Složky vektorů f_1 až f_4 tvoří řádky matice přechodu T od původní báze (e_1, e_2, e_3, e_4) k bázi (f_1, f_2, f_3, f_4) . Obě báze jsou ortonormální, matice přechodu je ortogonální, tj. $T^{-1} = T^T$. Pro vyjádření řádkové n -tice „starých“ proměnných $(x) = (x^1 \ x^2 \ x^3 \ x^4)$, v nichž byla forma zadána, pomocí n -tice „nových“ proměnných $(y) = (y^1 \ y^2 \ y^3 \ y^4)$, v nichž je vyjádřena v kanonickém tvaru, platí $(x) = (y)T$, tj.

$$\begin{aligned} x^1 &= \frac{1}{\sqrt{2}} y^1 + \frac{1}{\sqrt{6}} y^2 - \frac{1}{2\sqrt{3}} y^3 + \frac{1}{2} y^4, \\ x^2 &= \frac{1}{\sqrt{2}} y^1 - \frac{1}{\sqrt{6}} y^2 + \frac{1}{2\sqrt{3}} y^3 - \frac{1}{2} y^4, \\ x^3 &= \sqrt{\frac{2}{3}} y^2 + \frac{1}{2\sqrt{3}} y^3 - \frac{1}{2} y^4, \\ x^4 &= \frac{\sqrt{3}}{2} y^3 + \frac{1}{2} y^4. \end{aligned}$$

Dosaďte nyní předchozí výrazy pro x^1 až x^4 do původní rovnice kvadratické formy, zadané na začátku příkladu. Jestliže jsme při počítání neudělali chybu, měli byste dostat kanonický tvar formy.

6.3.2 Rozpoznávání kvadratických křivek a ploch

Že má dvojrozměrná a trojrozměrná lineární algebra použití v geometrii lineárních útvarů, tj. přímk a rovin, o tom již určitě nikdo nepochybuje. Další možnost aplikace se objevuje při klasifikaci takzvaných křivek (v rovině) a ploch (v trojrozměrném prostoru) druhého stupně, tedy těch, jejichž rovnice obsahují kvadratickou formu.

Zvolme v rovině kartézskou soustavu souřadnic $\langle O; x, y \rangle \sim \langle O; \vec{e}_1, \vec{e}_2 \rangle$. Bod O je její počátek, x a y jsou navzájem kolmé souřadnicové osy se stejnou délkovou jednotkou, \vec{e}_1 a \vec{e}_2 jsou vektory ortonormální báze umístěné v bodě O . Kartézské souřadnice (x, y) bodů takové křivky v rovině vyhovují rovnici obdobné té, se kterou jsme se setkali v příkladu 6.27, jen trochu obecnější,

$$(a_{11}x^2 + 2a_{12}xy + a_{22}y^2) + (b_1x + b_2y) + c = 0. \quad (6.19)$$

Aby rovnice opravdu obsahovala kvadratickou formu, je nutné, aby alespoň jedno z čísel a_{11} , a_{12} , a_{22} bylo nenulové. V první závorce je vidět kvadratická forma $a_{11}x^2 + 2a_{12}xy + a_{22}y^2$, v druhé

lineární forma $b_1x + b_2y + c$ je absolutní člen. Už víme, že taková křivka může být třeba elipsou. Mohla by být také parabolou, hyperbolou, a jsou možné ještě další útvary. Tak třeba rovnici $x^2 + y^2 = 0$ vyhovuje pouze počátek soustavy souřadnic, rovnici $x^2 + y^2 + 1 = 0$ nevyhovuje žádný bod, rovnice $x^2 - y^2 = 0$ představuje dvojici přímk $y = \pm x$. Jde o to, jak různé situace snadno rozpoznat.

Množina \mathcal{K} všech bodů v rovině, jejichž kartézské souřadnice vyhovují rovnici (6.19), v níž je alespoň jedno z čísel a_{11} , a_{12} , a_{22} nenulové, se nazývá *křivka druhého stupně* nebo též *kuželosečka*.

Je zřejmé, že druhý z názvů není pokaždé výstižný. Kuželosečkou totiž geometři rozumějí křivku, která vznikne jako průsečnice kuželové plochy nějakou rovinou. Tímto způsobem však nedokážeme získat ani prázdnou množinu (například již zmíněná rovnice $x^2 + y^2 + 1 = 0$), ani třeba dvojici přímk $x = \pm 1$, která je určena rovnicí $x^2 - 1 = 0$.

U ploch je situace poněkud složitější, jsou popsány pomocí jedné rovnice pro kartézské souřadnice (x, y, z) bodu v trojrozměrném prostoru vzhledem ke kartézské soustavě souřadnic $\langle O; x, y, z \rangle \sim \langle O; \vec{e}_1, \vec{e}_2, \vec{e}_3 \rangle$.

Množinu \mathcal{K} všech bodů v trojrozměrném prostoru, jejichž kartézské souřadnice (x, y, z) vyhovují rovnici

$$(a_{11}x^2 + 2a_{12}xy + 2a_{13}xz + a_{22}y^2 + 2a_{23}yz + a_{33}z^2) + \quad (6.20)$$

$$+(b_1x + b_2y + b_3z) + c = 0,$$

kde alespoň jedno z čísel a_{11} , a_{12} , a_{13} , a_{22} , a_{23} , a_{33} je nenulové, nazýváme *plocha druhého stupně*, nebo též *kvadrík*.

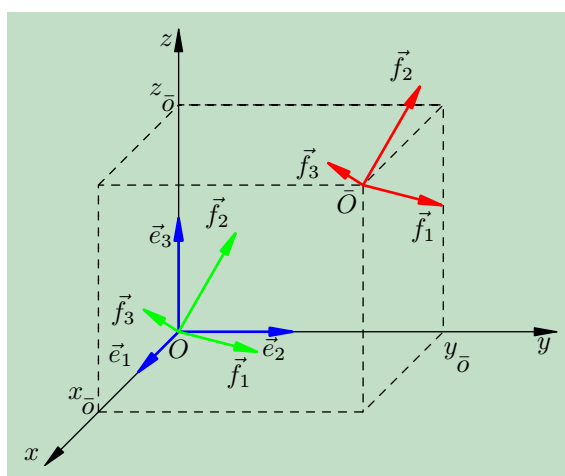
Rovnice opět obsahuje kvadratickou formu, lineární formu a absolutní člen. I když je rovnice velmi obecná a na první pohled z ní nelze přesný tvar plochy poznat, přece jen lze základní rysy odhadnout. Uvážíme-li totiž řezy této plochy souřadnicovými rovinami $(x = 0, y = 0, z = 0)$, hned vidíme, že s výjimkou speciálních případů jsou to kuželosečky. Obdobně to dopadne s řezy rovnoběžnými se souřadnicovými rovinami $(x = x_0, y = y_0, z = z_0)$. Soubor kvadrik tedy bude určitě zahrnovat elipsoidy, hyperboloidy, paraboloidy, ale také válcové a kuželové plochy. Je však třeba vypracovat metodu jejich přesné klasifikace. Klasifikaci provedeme rovnou pro kvadriky, kuželosečky jsou zvláštním případem pro $z = 0$.

Obdobně jako v příkladu 6.27 si můžeme řádkovou trojici $(X) = (x \ y \ z)$ představit jako složky vektoru s počátečním bodem v počátku soustavy souřadnic a koncovým bodem X o

souřadnicích $X = (x, y, z)$. Rovnici kvadriky zapíšeme v maticovém tvaru.

$$\begin{aligned} \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{13} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} + c = & \quad (6.21) \\ = (X)A(X)^T + (X)B + c = 0. & \end{aligned}$$

Z předchozí kapitoly víme, že přechodem k ortonormální bázi tvořené vlastními vektory operá-



Obr. 6.13 Posunutí počátku soustavy souřadnic.

toru φ reprezentovaného maticí A převedeme kvadratickou formu $(X)A(X)^T$ na kanonický tvar. Geometricky to znamená, že otočíme vektory báze při zachování počátku soustavy souřadnic. Přejdeme tedy od soustavy $\langle O; x, y, z \rangle \sim \langle O; \vec{e}_1, \vec{e}_2, \vec{e}_3 \rangle$ k soustavě $\langle O; \bar{x}, \bar{y}, \bar{z} \rangle \sim \langle O; \vec{f}_1, \vec{f}_2, \vec{f}_3 \rangle$. Označíme-li matici přechodu jako obvykle T a nové složky vektorů (kartézské souřadnice bodů) jako $(Y) = (\bar{x} \ \bar{y} \ \bar{z})$, dostáváme $(X) = (Y)T$, $(Y) = (X)T^{-1} = (X)T^T$ a

$$\begin{aligned} (X)A(X)^T + (X)B + c = (Y)TAT^T(Y)^T + (Y)TB + c = 0, \\ TAT^T = TAT^{-1} = D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad \bar{B} = TB = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{pmatrix}. \end{aligned}$$

Nová rovnice kvadriky

$$(\lambda_1 \bar{x}^2 + \lambda_2 \bar{y}^2 + \lambda_3 \bar{z}^2) + (\bar{b}_1 \bar{x} + \bar{b}_2 \bar{y} + \bar{b}_3 \bar{z}) + c = 0 \quad (6.22)$$

představuje její *seminormální tvar*. Odpovídá situaci, kdy kvadratická forma je v kanonickém tvaru, zatímco lineární forma je ještě obecná. Dalšího zjednodušení můžeme docílit posunutím

počátku soustavy souřadnic, tj. vhodnou záměnou $O \rightarrow \bar{O}$. Samozřejmě již nesmíme měnit bázi jednotkových vektorů podél souřadnicových os, neboť bychom mohli pokazit kanonický tvar kvadratické formy. Také v případě posouvání počátku lze najít obecný postup, jak to provádět. Jednodušší je však provést úpravu seminormálního tvaru pro každý konkrétní případ zvlášť. Vyžaduje to pouze proceduru takzvaného doplnění na čtverec, například pro $\lambda_1 \neq 0$ takto:

$$\lambda_1 \bar{x}^2 + \bar{b}_1 \bar{x} = \lambda_1 \left(\bar{x}^2 + \frac{\bar{b}_1}{\lambda_1} \bar{x} + \frac{\bar{b}_1^2}{4\lambda_1^2} \right) - \frac{\bar{b}_1^2}{4\lambda_1} = \lambda_1 \left(\bar{x} + \frac{\bar{b}_1}{2\lambda_1} \right)^2 - \frac{\bar{b}_1^2}{4\lambda_1}.$$

Výraz $\xi = \bar{x} + \frac{\bar{b}_1}{2\lambda_1}$ pak označíme jako novou proměnnou, která odpovídá x -ové souřadnici bodu kvadriky při posuvu počátku do polohy \bar{O} , jejíž x -ová souřadnice v soustavě $\langle O; \vec{f}_1, \vec{f}_2, \vec{f}_3 \rangle$ je $\bar{x}_{\bar{O}} = -\bar{b}_1/2\lambda_1$, přičemž $(\bar{x}_{\bar{O}}, \bar{y}_{\bar{O}}, \bar{z}_{\bar{O}}) = (x_{\bar{O}}, y_{\bar{O}}, z_{\bar{O}})T^T$. (viz obr. 6.13).

Příklad 6.33: Jak vypadá tato kvadrika?

V soustavě $\langle O; x, y, z \rangle \sim \langle O; \vec{e}_1, \vec{e}_2, \vec{e}_3 \rangle$ je kvadrika zadána rovnicí

$$x^2 + 5y^2 + z^2 + 2xy + 6xz + 2yz - 2x + 6y + 2z = 0.$$

Jak vypadá? Abychom to zjistili, musíme ji převést do vhodné soustavy souřadnic. Zapišeme kvadriku v maticovém tvaru a budeme hledat takovou bázi, v níž bude mít kvadratická forma kanonický tvar:

$$(X)A(X)^T + (X)B + c = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} -2 \\ 6 \\ 2 \end{pmatrix} = 0,$$

$$\det(A - \lambda E) = \det \begin{pmatrix} 1 - \lambda & 1 & 3 \\ 1 & 5 - \lambda & 1 \\ 3 & 1 & 1 - \lambda \end{pmatrix} = -\lambda^3 + 7\lambda^2 - 36.$$

Kořeny tohoto polynomu jsou $\lambda_1 = 3$, $\lambda_2 = 6$, $\lambda_3 = -2$. (Hned vidíme, že kvadratická forma má maximální hodnotu $h = 3$, její signatura je $[2, 1]$.) Výpočet vlastních vektorů pro jednotlivé vlastní hodnoty:

$$(A - \lambda_1 E)^T = \begin{pmatrix} -2 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 1 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$L_1 = [|\vec{f}_1|] = \left[\left[\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right] \right],$$

$$(A - \lambda_2 E)^T = \begin{pmatrix} -5 & 1 & 3 \\ 1 & -1 & 1 \\ 3 & 1 & -5 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$

$$L_2 = [|\vec{f}_2|] = \left[\left[\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right] \right],$$

$$(A - \lambda_3 E)^T = \begin{pmatrix} 3 & 1 & 3 \\ 1 & 7 & 1 \\ 3 & 1 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 7 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$L_3 = [|\vec{f}_3|] = \left[\left[-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right] \right].$$

Vektory \vec{f}_1 , \vec{f}_2 a \vec{f}_3 vyšly samy od sebe ortogonální. Víte proč? Nová matice lineární formy \bar{B} je

$$\bar{B} = TB = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} -2 \\ 6 \\ 2 \end{pmatrix} = \begin{pmatrix} -2\sqrt{3} \\ 2\sqrt{6} \\ 2\sqrt{2} \end{pmatrix}.$$

Seminormální tvar kvadriky dostáváme v soustavě souřadnic $\langle O; \vec{f}_1, \vec{f}_2, \vec{f}_3 \rangle$:

$$3\bar{x}^2 + 6\bar{y}^2 - 2\bar{z}^2 - 2\bar{x}\sqrt{3} + 2\bar{y}\sqrt{6} + 2\bar{z}\sqrt{2} = 0$$

Provedeme doplnění na čtverec:

$$3 \left(\bar{x}^2 - \frac{2}{\sqrt{3}}\bar{x} + \frac{1}{3} \right) - 1 + 6 \left(\bar{y}^2 + \frac{2}{\sqrt{6}}\bar{y} + \frac{1}{6} \right) - 1 - 2 \left(\bar{z}^2 - \frac{2}{\sqrt{2}}\bar{z} + \frac{1}{2} \right) + 1 = 0 \implies$$

$$\implies 3 \left(\bar{x} - \frac{1}{\sqrt{3}} \right)^2 + 6 \left(\bar{y} + \frac{1}{\sqrt{6}} \right)^2 - 2 \left(\bar{z} - \frac{1}{\sqrt{2}} \right)^2 - 1 = 0.$$

Nejnovější proměnné označme (ξ, η, ζ) , kde

$$\xi = \bar{x} - \frac{1}{\sqrt{3}}, \quad \eta = \bar{y} + \frac{1}{\sqrt{6}}, \quad \zeta = \bar{z} - \frac{1}{\sqrt{2}}.$$

Tyto proměnné představují souřadnice bodu kvadriky v soustavě $\langle \bar{O}; \vec{f}_1, \vec{f}_2, \vec{f}_3 \rangle$, kde

$$\bar{O} = \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{2}} \right),$$

vzhledem k soustavě souřadnic $\langle O; \vec{f}_1, \vec{f}_2, \vec{f}_3 \rangle$. V úplně původní soustavě $\langle O; \vec{e}_1, \vec{e}_2, \vec{e}_3 \rangle$ má bod \bar{O} souřadnice

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix} T = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix}.$$

V soustavě souřadnic $\langle \bar{O}; \vec{f}_1, \vec{f}_2, \vec{f}_3 \rangle$ je rovnice kvadriky v takzvaném *normálním tvaru*

$$3\xi^2 + 6\eta^2 - 2\zeta^2 - 1 = 0.$$

Všimněme si řezů kvadriky souřadnicovými rovinami nové soustavy, tj. $\xi = 0$, $\eta = 0$ a $\zeta = 0$:

$$\xi = 0 \implies 6\eta^2 - 2\zeta^2 - 1 = 0 \implies \left(\frac{\eta}{\sqrt{\frac{1}{6}}} \right)^2 - \left(\frac{\zeta}{\sqrt{\frac{1}{2}}} \right)^2 = 1,$$

$$\eta = 0 \implies 3\xi^2 - 2\zeta^2 - 1 = 0 \implies \left(\frac{\xi}{\sqrt{\frac{1}{3}}}\right)^2 - \left(\frac{\zeta}{\sqrt{\frac{1}{2}}}\right)^2 = 1,$$

$$\zeta = 0 \implies 3\xi^2 + 6\eta^2 - 1 = 0 \implies \left(\frac{\xi}{\sqrt{\frac{1}{3}}}\right)^2 + \left(\frac{\eta}{\sqrt{\frac{1}{6}}}\right)^2 = 1.$$

První dvě křivky jsou hyperboly, poslední je elipsa. Kvadrík je takzvaný hyperboloid.

Pro názornost provedeme ještě klasifikaci kuželosečky v rovině. V této situaci můžeme snadno nakreslit obrázky odpovídající přechodům mezi soustavami souřadnic.

Příklad 6.34: Co je to za kuželosečku?

Kuželosečka je v kartézské soustavě souřadnic zadána rovnicí

$$9x^2 - 4xy + 6y^2 + 6x - 8y + 2 = 0.$$

Maticový tvar je

$$(X)A(X)^T + (X)B + c = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 6 \\ -8 \end{pmatrix} + 2 = 0,$$

$$\det(A - \lambda E) = \det \begin{pmatrix} 9 - \lambda & -2 \\ -2 & 6 - \lambda \end{pmatrix} = \lambda^2 - 15\lambda - 50 = (\lambda - 10)(\lambda - 5), \quad \lambda_1 = 10, \lambda_2 = 5,$$

$$(A - \lambda_1 E) = (A - 10E) = \begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix},$$

$$L_1 = [|\vec{f}_1|] = \left[\left| \left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right) \right| \right].$$

$$(A - \lambda_2 E) = (A - 5E) = \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 \\ 0 & 0 \end{pmatrix},$$

$$L_2 = [|\vec{f}_2|] = \left[\left| \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right) \right| \right],$$

$$\bar{B} = TB = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 6 \\ -8 \end{pmatrix} = \begin{pmatrix} -\frac{20}{\sqrt{5}} \\ -\frac{10}{\sqrt{5}} \end{pmatrix}.$$

Seminormální tvar kuželosečky:

$$10\bar{x}^2 + 5\bar{y}^2 - \frac{20}{\sqrt{5}}\bar{x} - \frac{10}{\sqrt{5}}\bar{y} + 2 = 0.$$

Doplnění na čtverec:

$$10 \left(\bar{x}^2 - \frac{2}{\sqrt{5}}\bar{x} + \frac{1}{5} \right) - 2 + 5 \left(\bar{y}^2 - \frac{2}{\sqrt{5}}\bar{y} + \frac{1}{5} \right) - 1 + 2 = 0 \implies$$

$$\implies 10 \left(\bar{x} - \frac{1}{\sqrt{5}} \right)^2 + 5 \left(\bar{y} - \frac{1}{\sqrt{5}} \right)^2 - 1 = 0.$$

Nové proměnné a normální tvar:

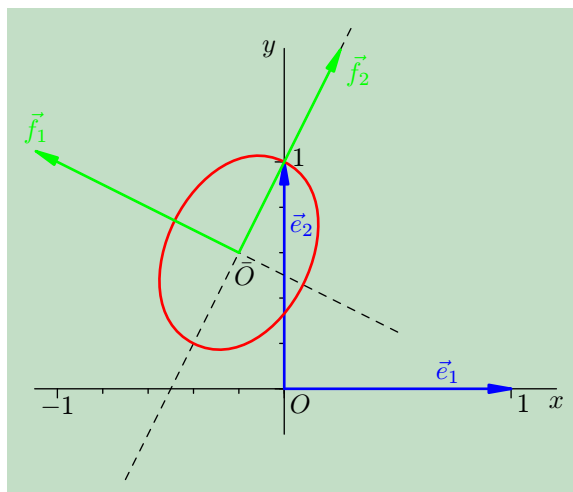
$$\xi = \bar{x} - \frac{1}{\sqrt{5}}, \quad \eta = \bar{y} - \frac{1}{\sqrt{5}}, \quad \bar{O} = \left(\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right) \text{ v soustavě } \langle O; \vec{f}_1, \vec{f}_2 \rangle,$$

$$10\xi^2 + 5\eta^2 - 1 = 0 \implies \left(\frac{\xi}{\sqrt{\frac{1}{10}}} \right)^2 + \left(\frac{\eta}{\sqrt{\frac{1}{5}}} \right)^2 = 1.$$

Souřadnice počátku \bar{O} v původní soustavě $\langle O; \vec{e}_1, \vec{e}_2 \rangle$:

$$\left(\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right) T = \left(\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right) \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} = \left(-\frac{1}{5}, \frac{3}{5} \right).$$

Situaci znázorňuje obrázek 6.14. Kuželosečkou je elipsa s poloosou $\frac{1}{\sqrt{10}}$ podél vektoru \vec{f}_1 a poloosou $\frac{1}{\sqrt{5}}$ podél



Obr. 6.14 Klasifikace kuželosečky — transformace souřadnic.

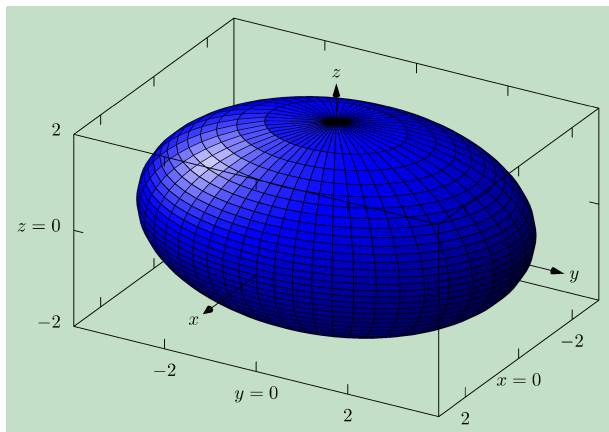
vektoru \vec{f}_2 . Z obrázku je vidět, že pro bod o souřadnicích $\xi = 0$, $\eta = \frac{1}{\sqrt{5}}$ v soustavě $\langle \bar{O}; \vec{f}_1, \vec{f}_2 \rangle$ bychom měli dostat $x = 0$, $y = 1$ v soustavě $\langle O; \vec{e}_1, \vec{e}_2 \rangle$. Zkuste to.

V závěru odstavce shrneme všechny možnosti, jak mohou vypadat tzv. *normální tvary* kuželoseček a kvadrik.

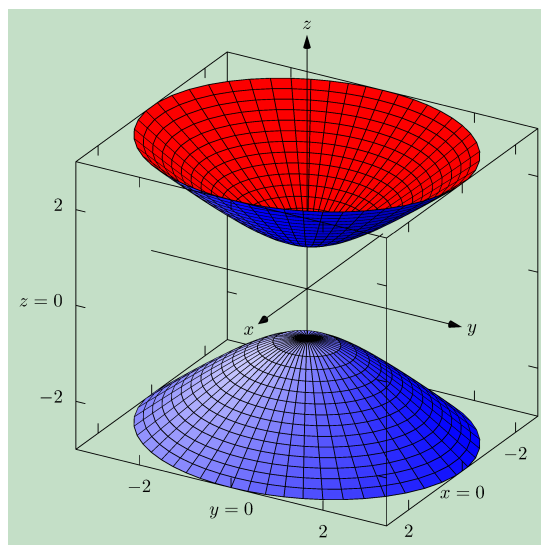
Věta 6.3: *Nechť \mathcal{K} je kvadrika v \mathbf{R}^3 . Existuje taková kartézská soustava souřadnic v \mathbf{R}^3 , že v ní má rovnice kvadriky \mathcal{K} právě jeden z následujících tvarů:*

(i1)	$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} + 1 = 0$	\emptyset (prázdná množina)
(i2)	$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0$	elipsoid
(i3)	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} + 1 = 0$	dvojdílný hyperboloid
(i4)	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} - 1 = 0$	jednodílný hyperboloid
(i5)	$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0$	$\{(0, 0, 0)\}$ (počátek soustavy souřadnic)
(i6)	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$	kužel
(ii1)	$\frac{x^2}{p} + \frac{y^2}{q} - 2z = 0, p, q > 0$	eliptický paraboloid
(ii2)	$\frac{x^2}{p} - \frac{y^2}{q} - 2z = 0, p, q > 0$	hyperbolický paraboloid
(iii1)	$\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$	\emptyset (prázdná množina)
(iii2)	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$	eliptický válec
(iii3)	$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$	hyperbolický válec
(iii4)	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$	$\{(0, 0, z) z \in \mathbf{R}\}$ (osa z)
(iii5)	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$	dvojice různoběžných rovin
(iv1)	$x^2 - 2py = 0, p > 0$	parabolický válec
(v1)	$x^2 + a^2 = 0, a \neq 0$	\emptyset (prázdná množina)
(v2)	$x^2 - a^2 = 0, a \neq 0$	dvojice rovnoběžných rovin
(v3)	$x^2 = 0$	dvojná rovina

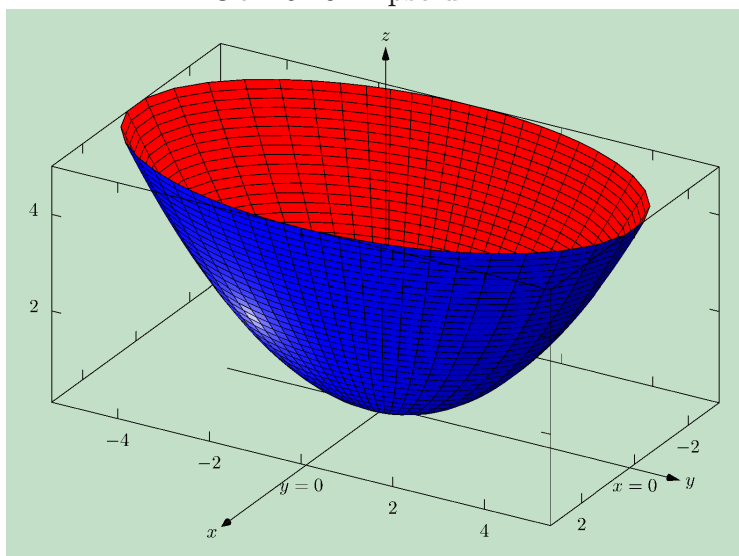
Všechny reálné situace zachycují obrázky 6.15 až ??.



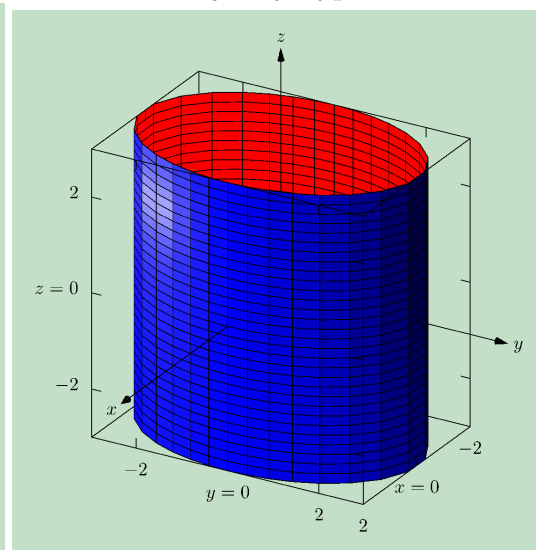
Obr. 6.15 Elipsoid.



Obr. 6.16 Dvojdílný hyperboloid.



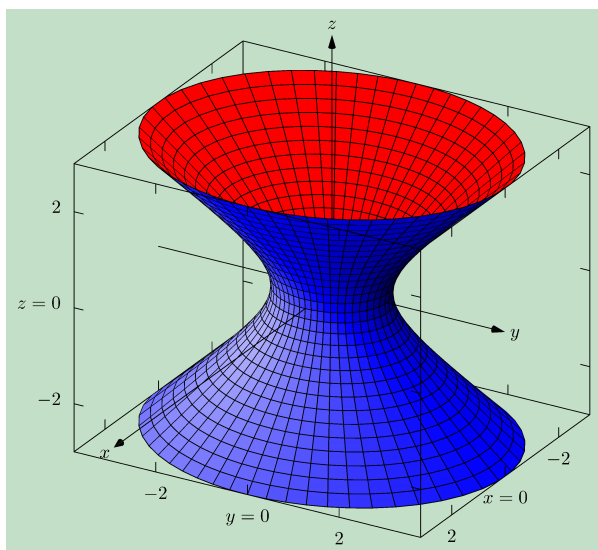
Obr. 6.17 Eliptický paraboloid.



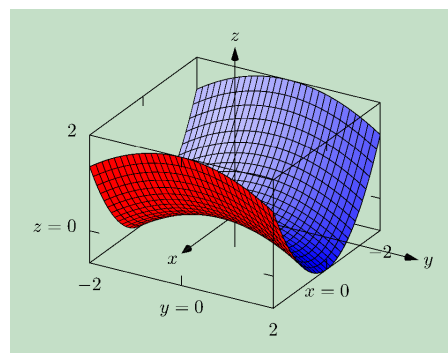
Obr. 6.18 Eliptický válec.

Věta 6.4: *Nechť \mathcal{K} je kuželosečka v \mathbf{R}^2 . Existuje taková kartézská soustava souřadnic v \mathbf{R}^2 , že v ní má rovnice kuželosečky \mathcal{K} právě jeden z následujících tvarů:*

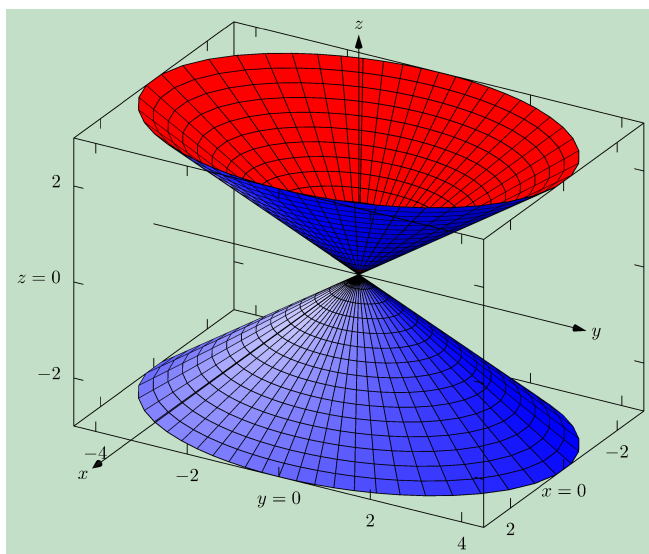
- (i1) $\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$ \emptyset (prázdná množina)
- (i2) $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$ elipsa
- (i3) $\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$ hyperbola
- (i4) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$ $\{(0, 0)\}$ (počátek soustavy souřadnic)
- (i5) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$ dvojice různoběžných přímek
- (ii1) $\frac{x^2}{p} - 2y = 0, p > 0$ parabola
- (iii1) $x^2 + a^2 = 0, a \neq 0$ \emptyset (prázdná množina)
- (iii2) $x^2 - a^2 = 0, a \neq 0$ dvojice rovnoběžných přímek
- (iii3) $x^2 = 0$ dvojnásobná přímka



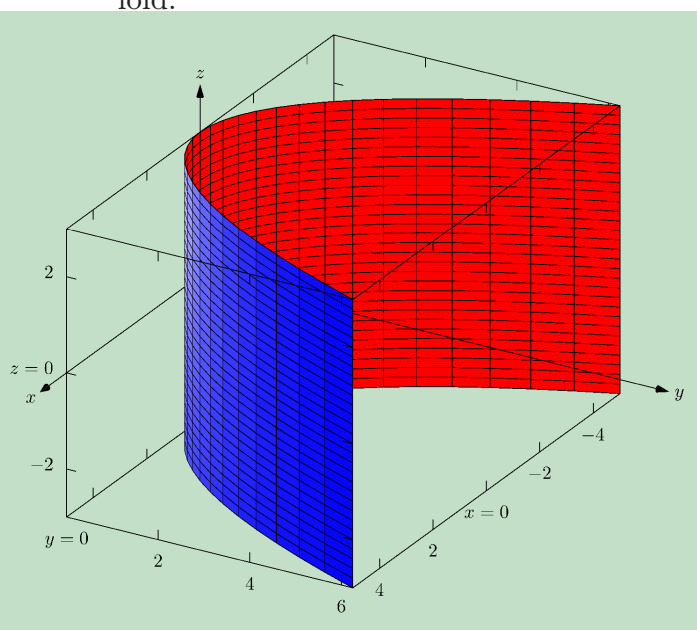
Obr. 6.19 Jednodílný hyperboloid.



Obr. 6.20 Hyperbolický paraboloid.



Obr. 6.21 Kůžel.

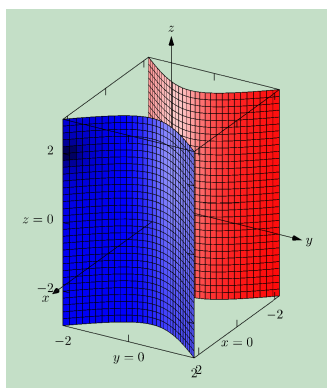


Obr. 6.22 Parabolický válec.

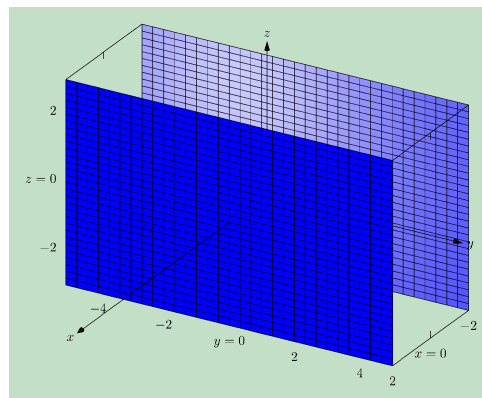
Všechny reálné situace zachycují obrázky 6.25, 6.26 a 6.27.

6.3.3 Symetrické operátory, kvadratické formy a fyzika

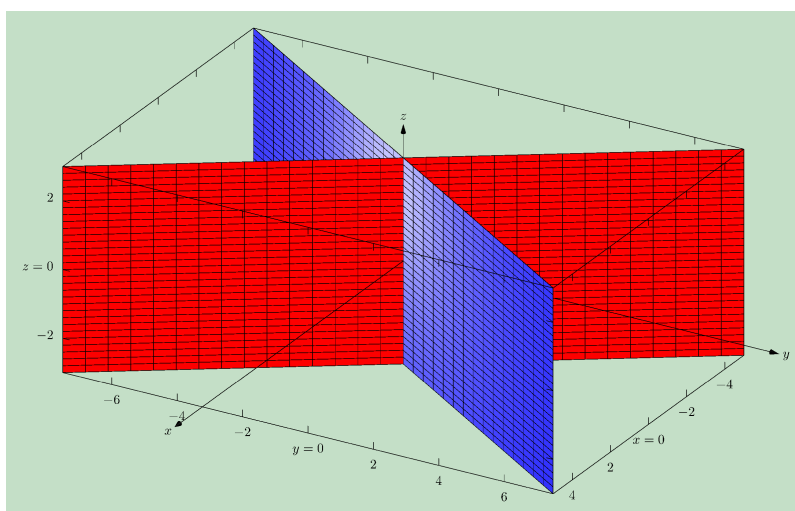
Někoho již jistě napadla otázka, k čemu potřebujeme umět klasifikovat kvadratické formy. Má to ještě jiný praktický význam než pro rozpoznávání kuželoseček a kvadrik? Pro odpověď stačí



Obr. 6.23 Hyperbolický válec.



Obr. 6.24 Rovnoběžné roviny.



Různoběžné roviny.

iii5

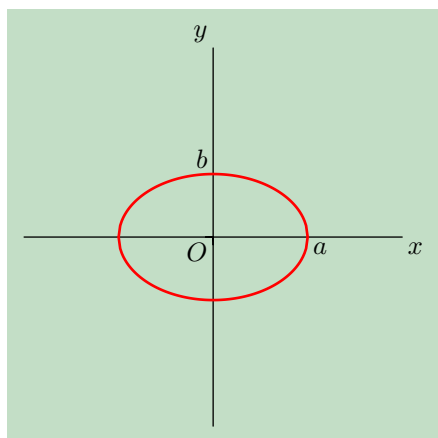
si připomenout fyzikální příklady symetrických operátorů, ať již z kapitoly 4 nebo z úvodu odstavce 6.2.2.

Příklad 6.35: Složky momentu setrvačnosti jako koeficienty kvadratické formy

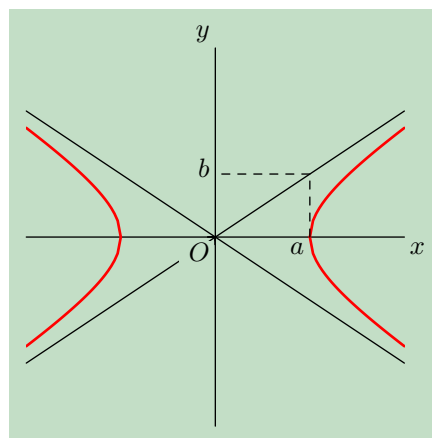
Na začátku odstavce 6.2.2 jsme se zabývali lineárním vztahem mezi momentem hybnosti \vec{L} a úhlovou rychlostí tuhého tělesa $\vec{\omega}$, $L_i = J_{ij}\omega_j$, $1 \leq i, j \leq 3$ (indexy jsou dole, jak je zvykem ve fyzice, sčítací index je j). Momenty setrvačnosti J_{ij} byly složkami symetrického operátoru. Vypočteme-li ještě rotační kinetickou energii tělesa, objeví se kvadratická forma:

$$E_k = \frac{1}{2} \int_V \rho v^2 dV = \frac{1}{2} \int_V \rho (\vec{\omega} \times \vec{r})^2 dV =$$

$$= \frac{1}{2} \int_V \rho ((\omega_2 z - \omega_3 y)^2 + (\omega_3 x - \omega_1 z)^2 + (\omega_1 y - \omega_2 x)^2) dV =$$

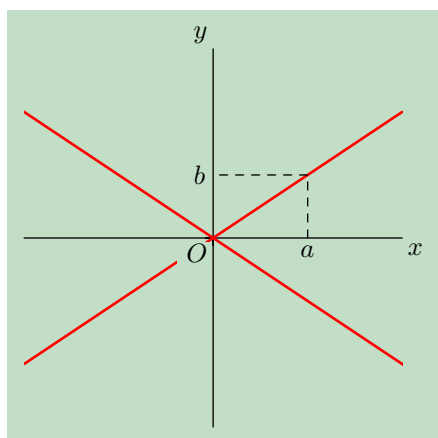


Elipsa.

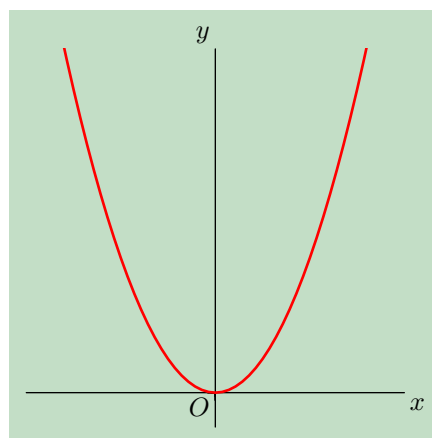


Hyperbola.

Obr. 6.25 Kuželosečky.



Různoběžky.

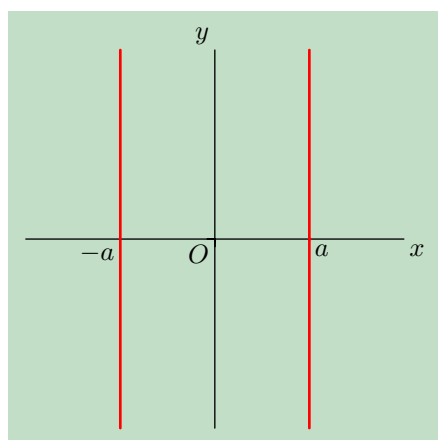


Parabola.

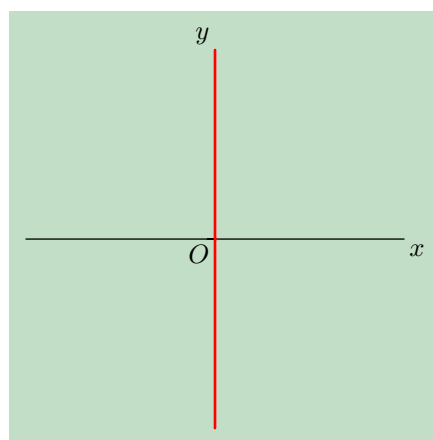
Obr. 6.26 Kuželosečky.

$$= \frac{1}{2} \left[\omega_1^2 \int_V \varrho(y^2 + z^2) dV - 2\omega_1\omega_2 \int_V \varrho xy dV - 2\omega_1\omega_3 \int_V \varrho xz dV + \right. \\ \left. + \omega_2^2 \int_V \varrho(x^2 + z^2) dV - 2\omega_2\omega_3 \int_V \varrho yz dV + \omega_3^2 \int_V \varrho(x^2 + y^2) dV \right].$$

I když úhlová rychlost $\vec{\omega}$ závisí obecně na čase, je pro všechny elementy tělesa stejná. Proto jsme složky úhlové rychlosti mohli vytknout před integrály. Poznáváte momenty setrvačnosti J_{ij} , definované v odstavci 6.2.2? Jistě ano, takže hned vidíte, že rotační kinetická energie představuje kvadratickou formu v proměnných ω_j , s



Rovnoběžky.



Dvojná přímka.

Obr. 6.27 Kuželosečky.

koefficienty J_{ij} :

$$E_k = \frac{1}{2} J_{ij} \omega_i \omega_j = \frac{1}{2} \begin{pmatrix} \omega_1 & \omega_2 & \omega_3 \end{pmatrix} \begin{pmatrix} J_{11} & J_{12} & J_{13} \\ J_{12} & J_{22} & J_{23} \\ J_{13} & J_{23} & J_{33} \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix},$$

sčítací indexy jsou i a j , výsledný výraz bude mít po rozepsání 9 sčítanců. Jak je možné, že předposlední výraz pro kinetickou energii, kde jsou jednotlivé členy explicitně rozepsány, má jen 6 sčítanců? Na tuto jednoduchou otázku jistě odpovíte bez nápovědy.

Také si hned můžeme všimnout praktické výhody toho, že už umíme něco z lineární algebry. Víme, že pro symetrický lineární operátor (v našem případě operátor momentu setrvačnosti) vždy dokážeme najít ortonormální bázi, v níž je reprezentován diagonální maticí. A v této bázi bude mít odpovídající kvadratická forma kanonický tvar. Zápis důležitých fyzikálních veličin — momentu hybnosti a kinetické energie — se tedy velice zjednoduší. Označíme-li T matici přechodu od původní báze k „výhodné“ bázi tvořené vlastními vektory operátoru, dostáváme

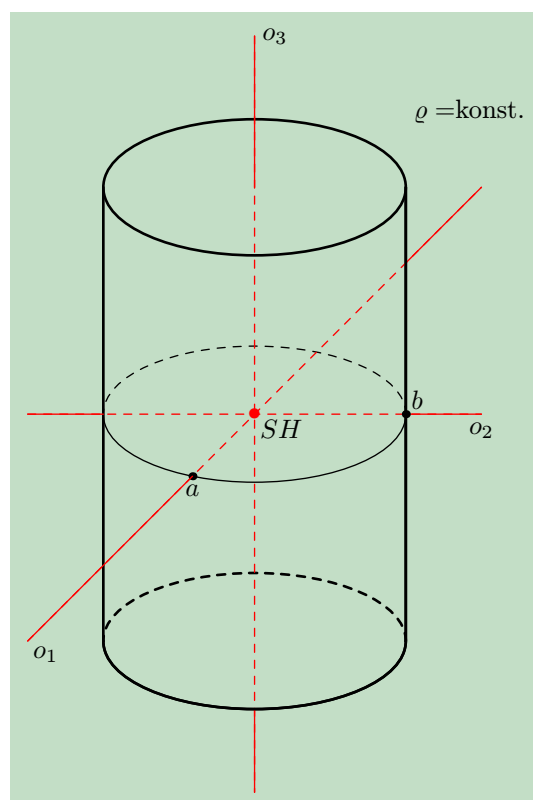
$$(\omega_1 \ \omega_2 \ \omega_3) = (\bar{\omega}_1 \ \bar{\omega}_2 \ \bar{\omega}_3) T, \quad (L_1 \ L_2 \ L_3) = (\bar{L}_1 \ \bar{L}_2 \ \bar{L}_3) T, \quad \begin{pmatrix} J_{11} & J_{12} & J_{13} \\ J_{12} & J_{22} & J_{23} \\ J_{13} & J_{23} & J_{33} \end{pmatrix} = T^{-1} \begin{pmatrix} \mathcal{J}_1 & 0 & 0 \\ 0 & \mathcal{J}_2 & 0 \\ 0 & 0 & \mathcal{J}_3 \end{pmatrix} T,$$

$$\implies \begin{pmatrix} \bar{L}_1 & \bar{L}_2 & \bar{L}_3 \end{pmatrix} = \begin{pmatrix} \bar{\omega}_1 & \bar{\omega}_2 & \bar{\omega}_3 \end{pmatrix} \begin{pmatrix} \mathcal{J}_1 & 0 & 0 \\ 0 & \mathcal{J}_2 & 0 \\ 0 & 0 & \mathcal{J}_3 \end{pmatrix} = (\mathcal{J}_1 \bar{\omega}_1, \mathcal{J}_2 \bar{\omega}_2, \mathcal{J}_3 \bar{\omega}_3),$$

$$E_k = \begin{pmatrix} \bar{\omega}_1 & \bar{\omega}_2 & \bar{\omega}_3 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \mathcal{J}_1 & 0 & 0 \\ 0 & \frac{1}{2} \mathcal{J}_2 & 0 \\ 0 & 0 & \frac{1}{2} \mathcal{J}_3 \end{pmatrix} \begin{pmatrix} \bar{\omega}_1 \\ \bar{\omega}_2 \\ \bar{\omega}_3 \end{pmatrix} = \frac{1}{2} (\mathcal{J}_1 \bar{\omega}_1^2 + \mathcal{J}_2 \bar{\omega}_2^2 + \mathcal{J}_3 \bar{\omega}_3^2).$$

Samozřejmě, že báze vybíráme ortonormální, matice přechodu je pak ortogonální. Příjemné je, že jsou-li vlastní hodnoty \mathcal{J}_1 , \mathcal{J}_2 a \mathcal{J}_3 různé, vyjde báze z vlastních vektorů ortogonální sama od sebe, stačí pak jen vektory normovat.

Fyzikové pochopitelně nehovoří o bázi vlastních vektorů. Mají pro tyto význačné směry vlastní název, související s jejich fyzikálním významem. Nazývají je *hlavní osy momentu setrvačnosti*. Algebra dokáže zjistit pouze jejich směry, nezabývá se už jejich umístěním. Z fyzikálních důvodů se však jejich průsečík ztotožňuje se středem hmotnosti tělesa SH , který se také obvykle volí za počátek soustavy souřadnic pevně spojené s tělesem. Vzhledem k tomu, že složky operátoru momentu setrvačnosti charakterizují rozložení hmotnosti v tělese, souvisejí směry hlavních os se symetrií tohoto rozložení. Nejlépe to uvidíme na případu homogenního tělesa, jehož hustota ρ je všude stejná (viz obrázek 6.28). V takovém případě je symetrie rozložení hmotnosti tělesa stejná jako jeho symetrie geometrická. Naše další úvahy související se symetrií rozložení hmotnosti tělesa budou velice názorné, neboť se budou opírat o představu geometrického tvaru. Bude-li těleso „šišaté“, třeba obecný elipsoid,



Obr. 6.28 K příkladu 6.35.

nebo nějaký úplně nepravidelný útvar, budou hodnoty \mathcal{J}_1 , \mathcal{J}_2 a \mathcal{J}_3 různé. Těm pak budou odpovídat navzájem ortogonální jednorozměrné vektorové podprostory v E_3 — směry hlavních os. Hlavní osy momentu setrvačnosti nesympetrického tělesa (fyzikové říkají asymetrického setrvačnicku) jsou tedy určeny jednoznačně. Pokud bude tělesem obecný elipsoid, jsou hlavní osy spojeny s jeho geometrickými poloosami — to nám napovídá naše geometrická a fyzikální intuice. Lze to však i korektně spočítat. A co když bude těleso válcově symetrické? Třeba rotační válec, nebo rotační kužel? Nebo jakýkoli jiný tvar s rotační symetrií, třeba cívka, nebo nějaký ozdobný pilíř? Pro fyziky je to symetrický setrvačnick. V takovém případě budou existovat dvě různé vlastní hodnoty operátoru momentu setrvačnosti, z nichž jedna bude dvojnásobná. Vlastní vektory vytvoří dva ortogonální podprostory. Ten jednorozměrný bude jednoznačně určovat jednu z hlavních os — osu rotační symetrie, dvojrozměrný pak bude obsahovat hlavní os nekonečně mnoho — všechny směry kolmé k rotační ose. Konečně nejvíce speciálním případem je těleso kulové (kulový setrvačnick), pro který bude existovat jediná trojnásobná

vlastní hodnota \mathcal{J} operátoru momentu setrvačnosti a všechny směry v prostoru budou rovnocenné. V tomto nejsymetričtějším případě dokonce platí $\vec{L} = \mathcal{J}\vec{\omega}$. A vůbec — dokážete rychle odpovědět na otázku, jakou maticí bude v *libovolné* ortonormální bázi reprezentován operátor momentu setrvačnosti, který má jednu trojnásobnou vlastní hodnotu \mathcal{J} ?

Jaký je fyzikální význam hlavních os? Když roztočíme těleso kolem hlavní osy určené třeba vlastním vektorem f_1 , tj. $\vec{\omega} = (\omega, 0, 0)$, dostaneme $\vec{L} = (\mathcal{J}_1\omega, 0, 0)$, vektor momentu hybnosti je rovnoběžný s úhlovou rychlostí, tedy s hlavní osou. Pokud již na těleso nijak silově nepůsobíme a můžeme-li zanedbat i brzdění silami tření, rotuje stále kolem hlavní osy stejnou úhlovou rychlostí. Směr osy rotace v prostoru se nemění. Osa je tzv. *pevná*, ani ji nemusíme držet. Člověk s technickým zaměřením by tento pohyb charakterizoval tak, že setrvačnick „nehází“. Co když ale chceme, aby setrvačnick rotoval kolem pevné osy, která není osou hlavní? To pak v soustavě souřadnic, která je s takovou osou spojena, nemá operátor momentu setrvačnosti diagonální tvar. Takovou rotaci nelze zařídit bez dodatečného silového působení. Osu je třeba držet, aby „necestovala“. V dlani cítíme silové působení osy — setrvačnick „hází“ a osu opotřebovává. Právě proto se například musí vyvažovat kola automobilů, aby osa, kolem které se musí otáčet, byla opravdu hlavní osou.

Pozn.: Obecně je samozřejmě třeba vést úvahy týkající se hlavních os operátoru momentu setrvačnosti i pro tělesa, která nejsou homogenní. Pak už si nemůžeme pomáhat názornou představou spojenou s geometrickým tvarem tělesa. Nehomogenní těleso s vhodně rozloženou hmotností se třeba může chovat jako symetrický setrvačnick a přitom nemusí mít žádnou geometrickou symetrii. O to více pak oceníme metody lineární algebry, neboť pomocí nich dokážeme směry hlavních os vypočítat, známe-li funkci popisující rozložení hmotnosti v tělese.

Nakonec si ještě všimněme kvadratické formy vyjadřující rotační kinetickou energii. Spojíme-li soustavu souřadnic $\langle SH; \bar{x}, \bar{y}, \bar{z} \rangle$ s hlavními osami momentu setrvačnosti protínajícími se ve středu hmotnosti SH , můžeme vyjádřit vlastní hodnoty operátoru přímo z „fyzikální“ definice:

$$\mathcal{J}_1 = \int_V \varrho(\bar{y}^2 + \bar{z}^2) dV, \quad \mathcal{J}_2 = \int_V \varrho(\bar{x}^2 + \bar{z}^2) dV, \quad \mathcal{J}_3 = \int_V \varrho(\bar{x}^2 + \bar{y}^2) dV.$$

Je zřejmé, že všechny jsou kladné a kvadratická forma pro rotační kinetickou energii je pozitivně definitní. Nic jiného jsme ani nemohli očekávat, neboť kinetická energie musí být při jakékoli nenulové úhlové rychlosti kladná, záporná není nikdy a nulová je právě tehdy, když se těleso nepohybuje.

Předchozí příklad je sice obsáhlý, ale umožnil nám pochopit souvislost algebraického popisu fyzikálních veličin s fyzikálními zákonitostmi samotnými. Jakmile jsme objevili, jak vhodně interpretovat moment setrvačnosti (jako symetrický lineární operátor v trojrozměrném euklidovském vektorovém prostoru), mohli jsme už jen pracovat s obecně platnými tvrzeními lineární algebry (problém vlastních hodnot a vektorů symetrických operátorů, vlastnosti kvadratických forem) a získat snadno a rychle zajímavé a závažné fyzikální závěry i praktické důsledky (rotace kolem pevné osy, „házení“ setrvačnicku). Pokud ještě někdo není přesvědčen o použitelnosti lineární algebry, má k dispozici další fyzikální příklad.

Příklad 6.36: Symetrické operátory a šíření světla

Že spolu tyto dvě věci nemohou souviset? Hned uvidíme. Slyšeli jste ve fyzice o dvojlomu? Určitě ano. Víte dobře, že dvojlom vzniká, když se průhledné prostředí nechová ke všem světelným paprskům, které se v něm šíří, stejně, ale rozlišuje různé směry polarizace různým indexem lomu. Jak to souvisí s operátory? Předpokládejme, že se světlo šíří v prostředí, které je nepohlcuje a které je *izotropní*, tj. nerozlišuje různé směry šíření světla — třeba sklo, nebo voda. Takové prostředí je charakterizováno jednou hodnotou indexu lomu (ta sice může záviset na vlnové délce světla, avšak tento jev nás nyní nezajímá). Pro index lomu n platí $n = \sqrt{\varepsilon}$, kde ε je relativní

dielektrická permitivita prostředí. Ta však také vystupuje ve vztahu úměrnosti mezi elektrickou indukcí \vec{D} a elektrickou intenzitou \vec{E} popisující šířící se světelnou vlnu. Platí $\vec{D} = \varepsilon_0 \varepsilon \vec{E}$, kde $\varepsilon_0 \doteq 8,85 \cdot 10^{-12} \text{ F m}^{-1}$ je permitivita vakua. Pokud však je prostředí *anizotropní*, je lineární vztah mezi vektory \vec{E} a \vec{D} obecnější, jak jsme jej popsali v úvodu odstavce 4.2.1,

$$\begin{pmatrix} D_1 & D_2 & D_3 \end{pmatrix} = \begin{pmatrix} \varepsilon_0 E_1 & \varepsilon_0 E_2 & \varepsilon_0 E_3 \end{pmatrix} \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{12} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{13} & \varepsilon_{23} & \varepsilon_{33} \end{pmatrix}.$$

Skutečnost, že je matice (ε_{ij}) symetrická, plyne z fyzikálních argumentů, jimž se zde nebudeme věnovat. Vidíme, že symetrický operátor dielektrické permitivity přiřazuje vektoru elektrické intenzity (vzoru) vektor elektrické indukce (obraz). V trojrozměrném prostoru zcela jistě existují ortonormální báze, v nichž je operátor relativní dielektrické permitivity reprezentován diagonální maticí. V diagonále vystupují, jak jinak, vlastní hodnoty operátoru. Tři možnosti — jediná trojnásobná vlastní hodnota ε , jedna jednonásobná vlastní hodnota (označme ji ε_{\parallel}) a jedna dvojnásobná vlastní hodnota (označme ε_{\perp}), tři různé (jednonásobné) vlastní hodnoty $\varepsilon_1, \varepsilon_2$ a ε_3 — rozlišují tři důležité fyzikální situace a třídí prostředí z hlediska šíření světla do tří kategorií — izotropní, opticky jednoosá a opticky dvojosá. Osy, ke kterým se názvosloví vztahuje, jsou dány směry vlastních vektorů operátoru dielektrické permitivity.

V případě izotropního prostředí je každý vektor prostoru vlastním vektorem operátoru relativní dielektrické permitivity, který má diagonální tvar ve všech ortonormálních bázích. Platí $\vec{D} = \varepsilon_0 \varepsilon \vec{E}$ a žádný směr není preferován. K izotropním prostředím se řadí například voda, sklo, ale také průhledné krystaly s krychlovou strukturou.

U prostředí, které jsme nazvali opticky jednoosým, je význačný směr určen vlastními vektory příslušnými jednonásobné vlastní hodnotě ε_{\parallel} , které, jak víme, tvoří jednorozměrný vektorový podprostor. Tento směr se ve fyzikální terminologii nazývá *optická osa*. V rovinách kolmých na tento směr jsou pak již všechny směry rovnocenné a tvoří dvojrozměrný vektorový podprostor příslušný dvojnásobné vlastní hodnotě ε_{\perp} . Platí

$$\begin{pmatrix} D_1 & D_2 & D_3 \end{pmatrix} = \begin{pmatrix} \varepsilon_0 E_1 & \varepsilon_0 E_2 & \varepsilon_0 E_3 \end{pmatrix} \begin{pmatrix} \varepsilon_{\parallel} & 0 & 0 \\ 0 & \varepsilon_{\perp} & 0 \\ 0 & 0 & \varepsilon_{\perp} \end{pmatrix}.$$

Je-li světlo polarizováno ve směru optické osy, tj. je-li vektor \vec{E} s ní rovnoběžný, šíří se světlo v prostředí rychlostí odpovídající indexu lomu $n_{\parallel} = \sqrt{\varepsilon_{\parallel}}$, $v_{\parallel} = c/n_{\parallel}$ (c je rychlost světla ve vakuu). Uvědomte si, že směr, ve kterém se světlo šíří, tedy světelný paprsek, je kolmý k vektoru \vec{E} . Paprsky vyhovující právě popsané situaci tedy vyplní dvojrozměrný podprostor kolmý k optické ose, světlo se *šíří kolmo k optické ose*. O každém z nich mluvíme jako o *mimořádném paprsku, extraordinariu*. Jestliže se naopak světlo *šíří rovnoběžně s optickou osou* (polarizováno je kolmo k ní), řídí se jeho rychlost indexem lomu $n_{\perp} = \sqrt{\varepsilon_{\perp}}$, $v_{\perp} = c/n_{\perp}$. Jedná se o *řádny paprsek, ordinarius*. Látky se dvěma různými vlastními hodnotami operátoru relativní dielektrické permitivity tedy mají dvě základní hodnoty indexu lomu a název „dvojlom“ je zcela namístě. Typicky jednoosými prostředím jsou průhledné krystaly s šesterečnou a čtverečnou strukturou (základní buňkou krystalu, jejímž opakováním vzniká celý krystal, je pravidelný šestiboký resp. pravidelný čtyřboký hranol).

V nejobecnější situaci, odpovídající navzájem různým vlastním hodnotám operátoru relativní dielektrické permitivity, existují v optickém prostředí tři preferované směry (jednorozměrné podprostory vlastních vektorů). Tyto směry představují optické osy. Proč hovoříme o dvojosém prostředí, když tyto směry jsou tři? Jednoduše proto, že taková je historická konvence. Znalost kterýchkoli dvou z těchto směrů již umožňuje určit i ten třetí jako kolmý k nim. Popis šíření světla v opticky dvojosém prostředí je velmi komplikovaný a ponechme jej fyzikům. K typicky dvojosým prostředím patří průhledné krystaly s obecnou strukturou.

Pozn.: Také v tomto příkladu najde uplatnění teorie kvadratických forem. Takzvaná *volná energie* objemové jednotky prostředí je totiž vyjádřena kvadratickou formou

$$F = \frac{1}{2} \vec{E} \vec{D} = \frac{1}{2} \varepsilon_0 \varepsilon_{ij} E_i E_j.$$

Co myslíte, jak bude vypadat její kanonický tvar?

6.3.4 Linearita ve fyzikálních a technických aplikacích

V předchozím textu jsme viděli, že jakékoli „přímé úměry“ mezi vektorovými veličinami, tj. mezi objekty, které jsou prvky nějakých vektorových prostorů, lze vyjádřit pomocí lineárních zobrazení, ve složkách ve zvolených bázích pak pomocí matic. Vzpomeňme si na fyziku, kde jsou lineární vztahy, tj. vztahy přímé úměry mezi vektorovými veličinami, běžné. Můžeme je tedy efektivně zapisovat pomocí matic? Následující příklady ukazují, že ano.

Příklad 6.37: Lineární operátory v mechanice

Uvažujme o přímé pružné středové srážce dvou částic o hmotnostech m_1 a m_2 . Jejich rychlosti před srážkou označme $\vec{v}_1 = (v_1)$ a $\vec{v}_2 = (v_2)$, rychlosti po srážce $\vec{\bar{v}}_1 = (\bar{v}_1)$ a $\vec{\bar{v}}_2 = (\bar{v}_2)$, kde $v_1, v_2, \bar{v}_1, \bar{v}_2$ jsou složky rychlostí částic podél osy x , jejíž směr je shodný se společným směrem všech rychlostí. Platí zákony zachování hybnosti a kinetické energie částic, tj.

$$m_1 v_1 + m_2 v_2 = m_1 \bar{v}_1 + m_2 \bar{v}_2, \quad \frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2 = \frac{1}{2} m_1 \bar{v}_1^2 + \frac{1}{2} m_2 \bar{v}_2^2.$$

Zdá se, že tento problém nemá s linearitou nic společného — vždyť druhý vztah je kvadratický. Jeho úpravou a poté užitím vztahu pro hybnosti však dostaneme

$$m_1(v_1 - \bar{v}_1)(v_1 + \bar{v}_1) = m_2(\bar{v}_2 - v_2)(v_2 + \bar{v}_2) \implies (v_1 + \bar{v}_1) = (v_2 + \bar{v}_2) \implies \bar{v}_1 - \bar{v}_2 = -v_1 + v_2.$$

Soustava rovnic o neznámých \bar{v}_1 a \bar{v}_2 tvořená poslední rovnicí a zákonem zachování hybnosti však již je soustavou lineárních rovnic a její řešení můžeme vyjádřit pomocí maticového zápisu

$$\begin{aligned} (\bar{v}_1 \quad \bar{v}_2) \begin{pmatrix} m_1 & 1 \\ m_2 & -1 \end{pmatrix} &= (v_1 \quad v_2) \begin{pmatrix} m_1 & -1 \\ m_2 & 1 \end{pmatrix} \implies \\ \implies (\bar{v}_1 \quad \bar{v}_2) &= (v_1 \quad v_2) \begin{pmatrix} m_1 & -1 \\ m_2 & 1 \end{pmatrix} \begin{pmatrix} m_1 & 1 \\ m_2 & -1 \end{pmatrix}^{-1}. \end{aligned}$$

Nakonec tedy můžeme vyjádřit „nové“ rychlosti jako lineární kombinace „starých“ takto:

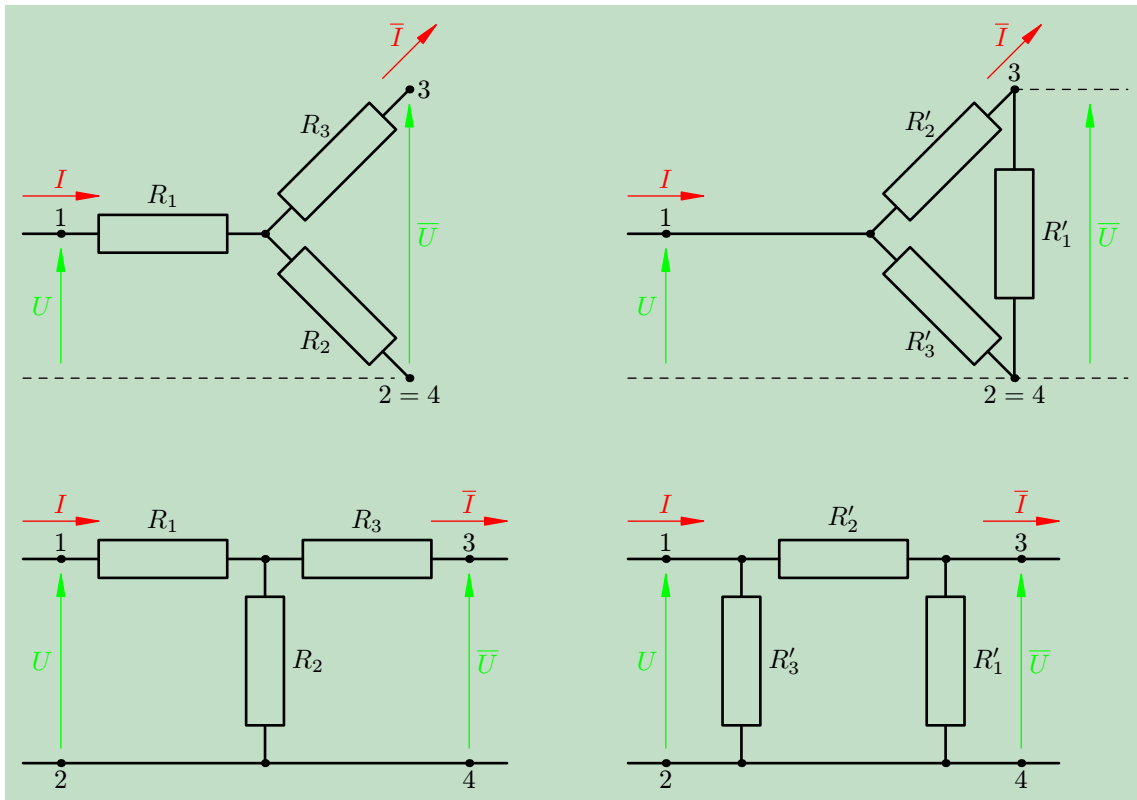
$$(\bar{v}_1 \quad \bar{v}_2) = (v_1 \quad v_2) \begin{pmatrix} \frac{m_1 - m_2}{m_1 + m_2} & \frac{2m_1}{m_1 + m_2} \\ \frac{2m_2}{m_1 + m_2} & \frac{m_2 - m_1}{m_1 + m_2} \end{pmatrix}.$$

Výpočet si proveďte. Matici vystupující v tomto vztahu, určenou jednoznačně pouze hmotnostmi částic, lze interpretovat tak, že reprezentuje lineární operátor (působící v dvojrozměrném vektorovém prostoru), který vektoru utvořenému z rychlostí částic bezprostředně před srážkou přiřadí vektor utvořený z rychlostí částic bezprostředně po srážce. Uvědomte si, že prostor, v němž se částice ve skutečnosti pohybují, je pouze jednorozměrný (jde o přímou srážku). Prostor, který jsme vytvořili „pro maticovou interpretaci“, je dvojrozměrný a jeho vektory jsou zadány dvěma složkami danými složkami dvou jednorozměrných vektorů „skutečných“ rychlostí částic.

K tomuto příkladu se ještě jednou vrátíme ve cvičení.

Příklad 6.38: Lineární operátory v elektřině a elektrotechnice

Jednou z úloh v elektrotechnice, která není považována za zcela triviální, je náhrada části obvodu zvané *trojúhelník* ekvivalentní částí obvodu zvané *hvězda*. Situaci ukazuje obrázek 6.29. Uvažujme o stejnosměrném obvodu. Úkolem je určit odpory rezistorů R'_1 , R'_2 a R'_3 v části obvodu typu trojúhelník na základě znalosti odporů R_1 , R_2 a R_3 v ekvivalentní části typu hvězda a naopak. Ekvivalence částí obvodu je dána požadavkem, že při daném vstupním napětí U mezi body 1 a 2 a daném proudu I tekoucím bodem 1 budou výstupní napětí \bar{U} mezi body 3 a 4 a proud \bar{I} protékající bodem 3 v případě hvězdy i trojúhelníka shodné. Vzhledem k možnosti názornějšího překreslení trojúhelníka a hvězdy, uvedeného rovněž na obrázku 6.29, se trojúhelník nazývá také *II-obvodem* a hvězda *T-obvodem*. Jistě si teď kladete otázku, k čemu zde mohou pomoci matice. Vzhledem k



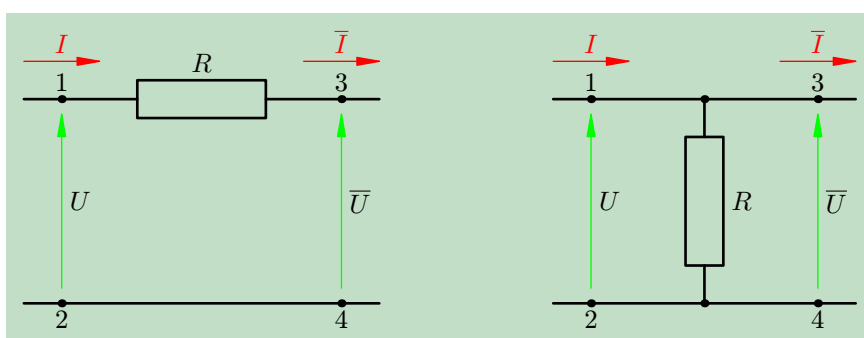
Obr. 6.29 K transformaci trojúhelníka na hvězdu.

platnosti Kirchhoffových zákonů, v nichž proudy a napětí vystupují lineárně, je možné vyjádřit dvojici (\bar{U}, \bar{I})

(výstupní napětí a výstupní proud) jako obraz dvojice (U, I) získaný působením jistého lineárního operátoru závislého na odporech. Co říkají Kirchhoffovy zákony?

- **První Kirchhoffův zákon:** Součet všech proudů v daném uzlu je nulový. (Přitom, podle konvence, má proud, který do uzlu vtéká, kladné znaménko, proud, který z uzlu vytéká, má znaménko záporné.)
- **Druhý Kirchhoffův zákon:** Součet všech úbytků napětí na všech prvcích v dané (uzavřené) větvi obvodu je nulový. (Podle konvence se úbytek napětí na daném prvku bere jako kladný, procházíme-li tímto prvkem souhlasně s předem zvolenou orientací větve, jako záporný pak v opačném případě.)

Lineární operátor složitějšího obvodu nebo jeho částí, jakými jsou třeba trojúhelník, hvězda, nebo i jiná zapojení, je složením (kompozicí) lineárních operátorů popisujících nejjednodušší části obvodu. Základní stavební kameny obvodů, zvané *elementární kvadrupóly*, znázorňuje obrázek 6.30. Kirchhoffovy zákony pro kvadrupóly mají



Obr. 6.30 Elementární kvadrupóly elektrických obvodů.

mimořádně jednoduchý tvar. Platí

$$\bar{U} = U - IR, \bar{I} = I \implies \begin{pmatrix} \bar{U} & \bar{I} \end{pmatrix} = \begin{pmatrix} U & I \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -R & 1 \end{pmatrix}$$

pro kvadrupól v levé části obrázku,

$$\bar{U} = U, \bar{I} = I - UR^{-1} \implies \begin{pmatrix} \bar{U} & \bar{I} \end{pmatrix} = \begin{pmatrix} U & I \end{pmatrix} \begin{pmatrix} 1 & -R^{-1} \\ 0 & 1 \end{pmatrix}$$

pro kvadrupól v pravé části obrázku. Na obrázku 6.29 si můžeme všimnout, že jak hvězda, tak trojúhelník, jsou složeny ze tří elementárních kvadrupólů. Platí tedy

$$\begin{pmatrix} \bar{U} & \bar{I} \end{pmatrix} = \begin{pmatrix} U & I \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -R_1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -R_2^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -R_3 & 1 \end{pmatrix}$$

pro hvězdu (T-obvod), pro trojúhelník (Π -obvod) pak

$$\begin{pmatrix} \bar{U} & \bar{I} \end{pmatrix} = \begin{pmatrix} U & I \end{pmatrix} \begin{pmatrix} 1 & -R_3'^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -R_2' & 1 \end{pmatrix} \begin{pmatrix} 1 & -R_1'^{-1} \\ 0 & 1 \end{pmatrix}.$$

Po vynásobení tří jednoduchých matic v obou případech získáme matici reprezentující lineární operátor, který dvojici (vstupní napětí, vstupní proud) přiřadí dvojici (výstupní napětí, výstupní proud). Mají-li být hvězda a trojúhelník ekvivalentní, musí být obě matice shodné. Z této podmínky dostaneme vztahy pro odpory R'_1 , R'_2 a R'_3 v trojúhelníku pomocí odporů R_1 , R_2 a R_3 ve hvězdě. Proveďte výpočet a zkontrolujte si výsledky:

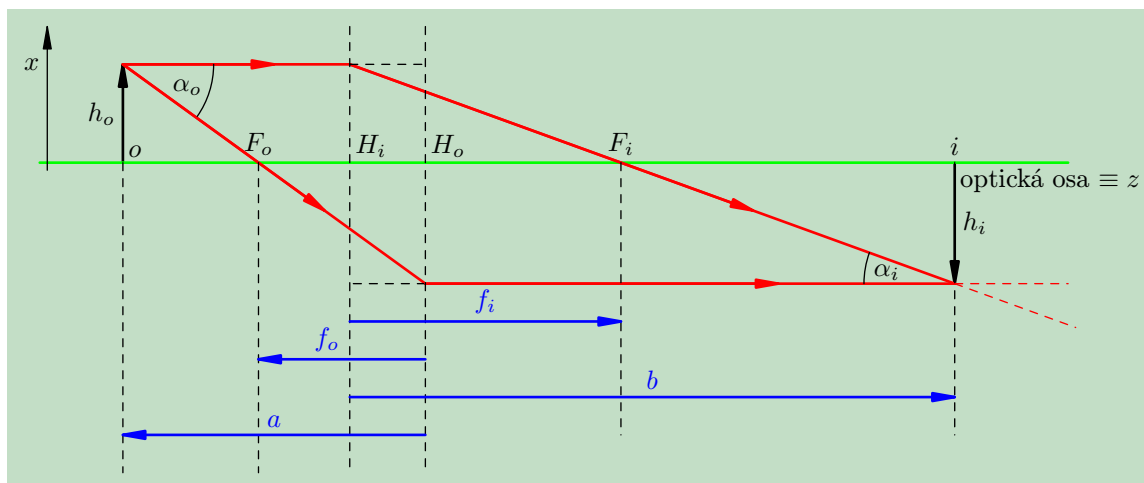
$$R'_1 = \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_1}, \quad R'_1 = \frac{R'_2 R'_3}{R'_1 + R'_2 + R'_3},$$

R'_2 , R'_3 , resp. R_2 , R_3 dostaneme z předchozích vztahů cyklickou záměnou indexů $(1, 2, 3) \rightarrow (3, 1, 2) \rightarrow (2, 3, 1)$.

Pozn.: Pro případ střídavých obvodů (úhlová frekvence proudu a napětí ω) lze předchozí postup jednoduše použít pro všechny obvyklé typy prvků — rezistory, cívky i kondenzátory. V získaných vztazích prostě nahradíme odpory příslušnými impedancemi vyjádřenými komplexními čísly (impedance rezistoru o odporu R je $Z_R = R$, impedance cívky o vlastní indukčnosti L je $Z_L = i\omega L$, impedance kondenzátoru o kapacitě C je $Z_C = 1/i\omega C$), a stejnosměrná napětí a proudy rovněž zaměníme za komplexní veličiny $U = U_0 \exp[i(\omega t + \varphi)]$, $I = I_0 \exp[i(\omega t + \psi)]$.

Příklad 6.39: Lineární operátory v geometrické optice

Také v geometrické optice je možné zapsat lineární vztahy pomocí lineárních operátorů. Lze jimi popsat šíření světla jak volným prostorem, tak optickými prvky. V tomto příkladu se budeme zabývat zobrazením tlustou čočkou, znázorněnou i s vyznačením důležitých parametrů na obrázku 6.31. Tlustá čočka dělí prostor



Obr. 6.31 Charakteristiky tlusté čočky.

na *předmětový*, který budeme charakterizovat indexem o (object), a *obrazový*, jemuž bude příslušet index i (image). Je charakterizována polohami hlavních rovin H_o a H_i , polohami ohnisek F_o a F_i a indexy lomu n_o a n_i . Pro šíření světelných paprsků platí tato pravidla:

- Paprsek, který se v předmětovém prostoru šíří rovnoběžně s optickou osou, se v průsečíku s obrazovou hlavní rovinou H_i láme do obrazového ohniska F_i .
- Paprsek, který v předmětovém prostoru projde předmětovým ohniskem F_o , se v průsečíku s předmětovou hlavní rovinou H_o láme do směru rovnoběžného s optickou osou.

Pro polohu a předmětu a polohu b obrazu platí *čočková rovnice*

$$\frac{f_o}{a} + \frac{f_i}{b} = 1, \quad \frac{f_o}{f_i} = \frac{n_o}{n_i}.$$

Vzdálenosti v předmětovém prostoru se měří od předmětové hlavní roviny, tj. $f_o = z_{H_o} - z_{F_o}$, $a = z_{H_o} - z_o$ a vzdálenosti v obrazovém prostoru se měří od obrazové hlavní roviny, tj. $f_i = z_{F_i} - z_{H_i}$, $b = z_i - z_{H_i}$. *Přímé a úhlové zvětšení* jsou definovány vztahy

$$M = \frac{h_i}{h_o} = -\frac{f_o a}{f_i b}, \quad M_A = \frac{\alpha_i}{\alpha_o} = -\frac{a}{b}.$$

Jedním ze základních vztahů tzv. *paraxiální geometrické optiky*, tj. takové aproximace, kdy se paprsky šíří v těsné blízkosti optické osy a svírají s ní velmi malé úhly, je Smithova–Helmholtzova věta, podle níž platí

$$n_o h_o \alpha_o = n_i h_i \alpha_i.$$

Pozn.: Aproximaci paraxiální optiky jsme použili již při vyjádření úhlového zvětšení, když jsme vzali v úvahu přibližné vztahy $\alpha_i \doteq \text{tg } \alpha_i$, $\alpha_o \doteq \text{tg } \alpha_o$. Těto aproximaci samozřejmě neodpovídá obrázek, v němž jsme museli použít paprsků svírajících s optickou osou větší úhly. Jinak by z něj nebylo možné geometrické vztahy vyčíst.

Pro maticový zápis chodu paprsků definujeme pojem *přenosové matice*. Označme x_1 x -ovou souřadnici průsečíku paprsku s určitou počáteční rovinou $z = z_1$ a x'_1 směrnicí paprsku v tomto bodě. Podobně označme x_2 , resp. x'_2 , x -ovou souřadnici průsečíku téhož paprsku s jistou koncovou rovinou $z = z_2$, resp. jeho směrnicí v tomto bodě. Za předpokladu, že je vztah mezi výslednými veličinami (x_2, x'_2) a počátečními veličinami (x_1, x'_1) lineární, můžeme jej hledat ve tvaru maticového zápisu

$$\begin{pmatrix} x_2 & x'_2 \end{pmatrix} = \begin{pmatrix} x_1 & x'_1 \end{pmatrix} T(z_1, z_2),$$

kde $T(z_1, z_2)$ je *přenosová matice*, nebo též *matice přenosu*.

Při šíření paprsku ve volném prostoru je zřejmé, že se směrnice nemění, tj. $x'_2 = x'_1$. Naproti tomu x narůstá úměrně se vzdáleností měřenou podél optické osy, přičemž faktorem úměrnosti je právě směrnice paprsku, tj. $x_2 = x_1 + x'_1(z_2 - z_1)$. Matice přenosu, kterou pro tento případ opatříme indexem p , má tedy tvar

$$T_p(z_1, z_2) = \begin{pmatrix} 1 & 0 \\ z_2 - z_1 & 1 \end{pmatrix}.$$

Pro přenos mezi hlavními rovinami čočky využijeme pravidla, že poloha průsečíku paprsku s rovinou H_o (tj. pro $z = z_{H_o}$) je stejná jako v rovině H_i (tj. pro $z = z_{H_i}$). Směrnice paprsku se však v prostoru mezi hlavními rovinami čočky mění. Uplatňuje se zde jednak změna úměrná poloze paprsku, jednak změna způsobená vlivem lomu světla (Snellův zákon lomu), tj.

$$x'_2 = -\frac{x_1}{f_i} + x'_1 \frac{n_o}{n_i} = -\frac{x_1}{f_i} + x'_1 \frac{f_o}{f_i}.$$

Odpovídající matice přenosu je

$$T_c(z_{H_i}, z_{H_o}) = \begin{pmatrix} 1 & -\frac{1}{f_i} \\ 0 & \frac{f_o}{f_i} \end{pmatrix}.$$

Přísluší-li index 1 paprsku před čočkou a index 2 za čočkou, platí pro výslednou matici přenosu

$$T(z_1, z_2) = T_p(z_1, z_{H_o}) T_c(z_{H_i}, z_{H_o}) T_p(z_{H_i}, z_2).$$

Proveďte výpočet a ukažte, že platí

$$T(z_1, z_2) = \begin{pmatrix} \frac{f_i - z_2 + z_{H_i}}{f_i} & -\frac{1}{f_i} \\ (z_{H_o} - z_1) + \frac{(z_1 - z_{H_o} + f_o)(z_2 - z_{H_i})}{f_i} & \frac{z_1 - z_{H_o} + f_o}{f_i} \end{pmatrix}.$$

Platí $z_{H_o} - z_1 = f_o + z_{F_o} - z_1$ a $z_2 - z_{H_i} = z_2 + f_i - z_{F_i}$, a tak nakonec pro matici přenosu dostaneme

$$T(z_1, z_2) = \begin{pmatrix} -\frac{z_2 - z_{F_i}}{f_i} & -\frac{1}{f_i} \\ f_o + \frac{(z_1 - z_{F_o})(z_2 - z_{F_i})}{f_i} & \frac{z_1 - z_{F_o}}{f_i} \end{pmatrix}.$$

Ve cvičení si všimneme dvou speciálních případů, kterým odpovídají zvlášť jednoduché matice přenosu.

Maticový (operátorový) popis šíření světla v paraxiální geometrické optice je velmi výhodný. Každý optický prvek lze popsat maticí přenosu a postupné řazení prvků do cesty světelnému paprsku znamená z výpočetního hlediska pouze násobení matic (lineární operátory působí postupně, skládají se).

6.3.5 Cvičení

1. V následujících případech najděte kartézskou soustavu souřadnic, v níž má kvadrika nebo kuželosečka normální tvar. Určete, o jakou kvadriku nebo kuželosečku se jedná.

- $6xy + 8y^2 - 12x - 26y + 11 = 0$ v \mathbf{R}^2
- $x^2 + 2xy + y^2 - 8x + 4 = 0$ v \mathbf{R}^2
- $x^2 + 5y^2 + z^2 + 2xy + 6xz + 2yz - 2x + 6y + 2z = 0$ v \mathbf{R}^3
- $4x^2 + 2xy + 2xz - 2y^2 + 5yz - 2z^2 - 22x - 19y + 8z + 1 = 0$ v \mathbf{R}^3
- $x^2 + 4xy + y^2 + 2x + 1 = 0$ v \mathbf{R}^2
- $2x^2 + 5xy - 3y^2 + x + 10y - 3 = 0$ v \mathbf{R}^2
- $4x^2 + y^2 - 4z^2 - 8x + 4y + 24z - 32 = 0$ v \mathbf{R}^3
- $6x^2 - 4xy - 4xz + 7y^2 + 5z^2 - 36 = 0$ v \mathbf{R}^3
- $2x^2 + 2xy + 2y^2 - 5z^2 - 2x - 4y - 4z + 2 = 0$ v \mathbf{R}^3
- $x^2 + 4xy - 10xz - 2y^2 + 4yz + z^2 + 2x + 4y - 10z - 1 = 0$ v \mathbf{R}^3
- $x^2 + 2xy + 6xz + 5y^2 + 2yz + z^2 - 2x + 6y + 2z = 0$ v \mathbf{R}^3
- $x^2 + 2xy + y^2 + 1 = 0$ v \mathbf{R}^2
- $4x^2 + 4xy + y^2 + 2x + y - 2 = 0$ v \mathbf{R}^2

2. Ještě jednou se vraťte k příkladu 6.37 o přímé srážce dvou částic. Dvojici rychlostí $(\bar{v}_1 \ \bar{v}_2)$ částic po srážce jsme v něm vyjádřili jako obraz dvojice rychlostí $(v_1 \ v_2)$ před srážkou operátorem φ závislým na hmotnostech částic. Určete vlastní hodnoty a vlastní vektory tohoto operátoru. Dokažte, že jako báze jednorozměrných vektorových podprostorů příslušných vypočteným vlastním hodnotám lze vzít vektory $(1, 1)$ a $\left(\frac{m_2}{m_1+m_2}, -\frac{m_1}{m_1+m_2}\right)$. Určete matici přechodu od původní báze dvojrozměrného vektorového prostoru dvojic rychlostí částic k bázi (e_1, e_2) tvořené výše specifikovanými vlastními vektory operátoru φ . (Je třeba určit matici inverzní k matici přechodu.) Zapište vztah mezi rychlostmi před srážkou $(v'_1 \ v'_2)$ a $(\bar{v}'_1 \ \bar{v}'_2)$ v bázi (e_1, e_2) a pokuste se o fyzikální interpretaci výsledků.
3. V příkladu 6.39 o tlusté čočce jsme odvodili matici přenosu, která popisuje lineární vztah mezi počátečními a koncovými charakteristikami šířícího se světelného paprsku (x_1, x'_1) , resp. (x_2, x'_2) (x je poloha a x' směrnice paprsku).
- Zapište matici přenosu pro případ tzv. *opticky konjugovaných* rovin, tj. pro $z_1 = z_o$ a $z_2 = z_i$. Vypočtěte pomocí této matice hodnoty x_i a x'_i .
 - Dále ověřte, že při dosazení $z_1 = z_{H_o}$ a $z_2 = z_{H_i}$ dostaneme z obecného tvaru matice přenosu matici čočky.