Dříve než se dostaneme k výběru vhodného typu ploch, uveďme si, co všechno budeme po těchto plochách v rámci radiačního výpočtu požadovat. Jedná se o následující operace:

- určení průsečíku plocha – úsečka
- zjištění polohy bodu, známe-li jeho souřadnice na ploše
- lokalizace bodu na ploše (tj. určení, do jakého elementu plochy daný bod patří)
- určení normály plochy v zadaném bodu
- určení hodnoty vyzařovací funkce (viz kapitola 4.4) plochy v daném bodu a požadovaném směru

Od plochy samotné pak požadujeme co nejpřesnější aproximaci povrchů těles při co možná nejmenší složitosti výše uvedených operací.

Obdobným způsobem, jakým se do počítačové grafiky zavádí čtyřúhelníkové pláty jako zobecnění obdélníků, jsou definovány i trojúhelníkové pláty. Toto zobecnění trojúhelníků má velice podobné vlastnosti jako standardní čtyřúhelníkové pláty, lze jej však považovat za přirozenější rozšíření jednodimenzionálního případu (tj. teorie křivek) na plochu. Trojúhelníkové pláty se často používají v digitálních modelech terénu či při výpočtech metodou konečných prvků, objevují se i moderních CAD systémech pro modelování tvarově složitých povrchů. V následujících odstavcích si shrneme pouze jejich nejdůležitější vlastnosti, podrobnější informace lze nalézt např. v [RJ92].

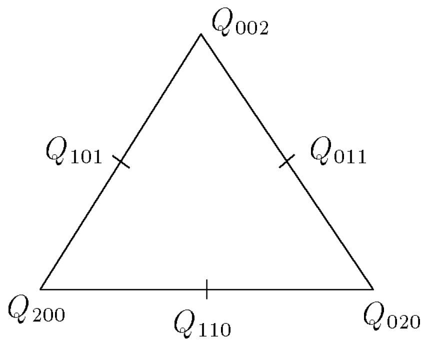
Trojúhelníkové pláty jsou algebraické plochy (tj. lze je popsat pomocí polynomů) definované soustavou tzv. definičních uzlů, jednotlivé pláty mají tři rohové body (tyto body tvoří tzv. bázový trojúhelník daného plátu) a omezují je tři okrajové křivky. Pláty se rozlišují dle typu použitých křivek, při užití Bèzierových křivek se tedy pláty nazývají *trojúhelníkové Bèzierovy pláty.* Tyto pláty jsou definovány následujícím vztahem:

$$B_n(u,v,w) = \sum_{i+j+k=n} \frac{n!}{i!\,j!\,k!} \cdot u^i v^j w^k \cdot Q_{ijk} \quad \text{kde} \;\; u+v+w=1 \tag{4.9}$$

Body Q_{ijk} se nazývají *definiční uzly* daného trojúhelníkového plátu, je jich $\frac{1}{2}(n+1)(n+2)$ a dohromady tvoří trojúhelníkovou síť, nad kterou je daný plát definován. Parametr n určuje stupeň daných plátů, v našem případě je výhodné použít kvadratické pláty, tj. když $n = 2$. Definiční rovnice (4.9) má pak tvar:

$$\begin{aligned} B_2(u,v,w) = {} & u^2 \cdot Q_{200} + v^2 \cdot Q_{020} + w^2 \cdot Q_{002} + \\ & + 2uv \cdot Q_{110} + 2vw \cdot Q_{011} + 2uw \cdot Q_{101} \end{aligned} \tag{4.10}$$

odpovídající rozložení definičních uzlů Q_{ijk} je znázorněno na obrázku 4.6.



Obr. 4.6: Kvadratické Bézierovy pláty

Jak je vidět, definiční uzly odpovídají přesně bodům A, B, C, D, E a F, které jsme použili při dělení $\triangle ABC$ na jednotlivé části (viz obrázek 4.5). Kvadratické trojúhelníkové pláty lze tedy dělit stejným způsobem jako odpovídající bázový trojúhelník, podobně probíhá i lokalizace bodu do elementárních plátů. Také ostatní požadované operace, tj. určení bodu plátu, který má zadané souřadnice, zjištění průsečíku plátu s úsečkou atd. lze vyřešit modifikací postupů uvedených pro trojúhelníky, podrobně se jimi zabývá již výše citovaný článek [RJ92].

Bèzierovy trojúhelníkové pláty aproximují značně přesně mnoho složitých povrchů, někdy však ani použití těchto plátů nemusí být postačující. V tomto případě je lze rozšířit na tzv. *racionální Bèzierovy trojúhelníkové pláty*, které jsou trojúhelníkovým ekvivalentem tzv. NURBS (Non-Uniform Rational B-Splines) plátů. Jejich podstatou je to, že jednotlivým uzlům řídící sítě jsou přiřazeny váhy, které umožňují modifikovat tvar plátu bez změny jeho řídících bodů. Pomocí racionálních plátů lze přesně popsat i povrch koule a dalších "normálním způsobem" nepopsatelných těles, tyto pláty (trojúhelníkové i čtyřúhelníkové) by se mohly díky svým schopnostem stát prostředkem pro komplexní a jednotný popis všech těles v rámci geometrického modelování (opět viz [RJ92], kde jsou uvedeny i příklady reprezentace některých objektů).

4.4 Světelný model

Světelný model použitý v popisované práci vznikl jako kombinace mnoha dosud publikovaných modelů ([Bou70] [Pho75] [CT82] [Whi80] [Str90] [Gla91] [SŽ89] [SŽ93] [SAWG91] [Kli78]). Základní vlastnosti tohoto světelného modelu jsou shrnuty v následujících odstavcích.

4.4.1 Světelné zdroje, vyzařovací funkce

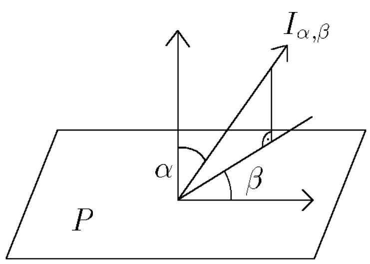
Světelné zdroje používané v počítačové grafice lze rozdělit na bodové a plošné. Bodové zdroje v reálném světě neexistují, jedná se o matematickou idealizaci, používanou pro reprezentaci skutečných světelných zdrojů v ray-tracingu a mnoha dalších algoritmech pro zobrazení umělých scén. Všechny světelné zdroje vyskytující se v přírodě jsou zdroje plošné. Standardní radiační metody dovedou korektně simulovat podmnožinu těchto zdrojů, tzv. ideální difúzní zdroje (*ideal diffuse*), jejichž základní vlastností je to, že vyzařují světlo rovnoměrně do všech směrů viditelného poloprostoru. Emise světla v požadovaném směru je pak dle Lambertova zákona přímo úměrná kosinu úhlu, který svírá tento směr s normálou povrchu daného zdroje a dále ploše tohoto zdroje. Pro emisi zdroje v daném směru lze tedy psát:

$$E_\alpha = I \cdot P \cdot \cos\alpha \tag{4.11}$$

kde:

- E_α je emise v požadovaném směru,
- I je intenzita světelné energie, kterou plocha vyzařuje
- P je velikost dané plochy
- α je úhel, který svírá požadovaný směr s normálou plochy

Směrové radiační metody musí samozřejmě postihnout i zdroje, u kterých závisí intenzita vyzařované světelné energie na směru, ze kterého tento zdroj pozorujeme. Intenzitu těchto zdrojů lze popsat pomocí tzv. *vyzařovací funkce* $I(\alpha, \beta)$, která udává intenzitu zdroje ve směru, jehož vektor svírá s normálou úhel α a s tzv. *hlavním směrem plochy* úhel β, jak je znázorněno na obrázku 4.7.



Obr. 4.7: Vyzařovací funkce

Rovnici pro emisi plochy (4.11) je nutné pro tyto zářiče přeformulovat do tvaru:

$$E_{\alpha,\beta} = I(\alpha, \beta) \cdot P \cdot \cos\alpha \tag{4.12}$$

Intenzita světelné energie vyzařované plošnými zdroji ubývá se čtvercem vzdálenosti, ve vzdálenosti d bude mít velikost:

$$I_{\alpha,\beta}(d) = \frac{I(\alpha,\beta)}{d^2} \tag{4.13}$$

Ze vztahů (4.11) a (4.13) lze určit emisi daného zdroje v požadovaném směru a vzdálenosti, odpovídající vztah má tvar:

$$E_{\alpha,\beta}(d) = \frac{I(\alpha,\beta)\cdot\cos\alpha}{d^2}\cdot P \tag{4.14}$$

Pokles intenzity se čtvercem vzdálenosti je však v praxi příliš rychlý (grafický displej totiž nedokáže zpracovat tak velký rozsah intenzit jako lidské oko, navíc pro $d < 1$ dochází k nesprávnému zvýšení intenzity), proto se často používá modifikace vztahu (4.14) ve tvaru:

$$E_{\alpha,\beta}(d) = \frac{I(\alpha,\beta)\cdot\cos\alpha}{a + b\cdot d + c\cdot d^2}\cdot P \tag{4.15}$$

kde a, b a c jsou vhodně volené konstanty (často se volí $b = 1$ a $c = 0$, někdy též $a = 1$, $b = 0$ a $c = 1$).

Při výpočtu propagace světla scénou je nutné určovat intenzity vyzařovaných energií v rámci každého elementárního kroku, bylo by tedy vhodné určovat tyto intenzity co nejrychleji. Místo vyzařovacích funkcí $I(\alpha,\beta)$ je proto velice výhodné používat tzv. *modifikovaných vyzařovacích funkcí* $I'(\alpha,\beta)$ definovaných vztahem:

$$I'(\alpha,\beta) = I(\alpha,\beta)\cdot\cos\alpha \tag{4.16}$$

Vztah pro výpočet emise (4.15) se pak zjednoduší na:

$$E_{\alpha,\beta}(d) = \frac{I'(\alpha,\beta)}{a + b\cdot d + c\cdot d^2}\cdot P \tag{4.17}$$

Zahrnutí kosinu přímo do vyzařovacích funkcí navíc zlepšuje hladkost těchto funkcí, což je velice výhodné při jejich realizaci (viz kapitola 4.7)

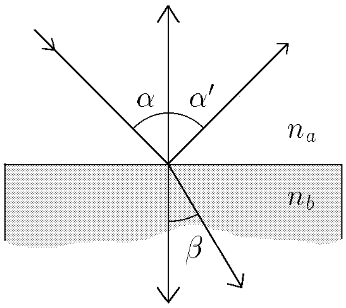
4.4.2 Odraz světla od povrchu tělesa

Světelná energie vyzařovaná jednotlivými zdroji se scénou šíří přímočaře až do té doby, pokud nedopadne na povrch nějakého tělesa. Po dopadu se světelný paprsek obecně rozdělí na paprsek odražený a paprsek lomený. Všechny tyto paprsky leží spolu s normálou povrchu v jedné rovině, odražený paprsek směřuje zpět do prostředí, lomený paprsek daným tělesem prochází. Věnujme se nyní paprsku odraženému.

Pro směr odraženého paprsku platí tzv. *zákon odrazu*:

$$\alpha' = \alpha \tag{4.18}$$

kde α' je úhel, který svírá směr odraženého paprsku s normálou plochy, od níž se daný paprsek odráží a α je obdobný úhel, ale pro původní paprsek (viz obrázek 4.8).



Obr. 4.8: Zákon odrazu a zákon lomu

4.4.3 Průhledná a průsvitná tělesa, lom paprsku

V předchozím odstavci jsme se zabývali paprskem odraženým od povrchu tělesa zpět do scény. V případě neprůhledných těles je odražené světlo jedinou složkou, která nás zajímá. U průhledných a průsvitných těles se navíc musíme zabývat světlem, které daným tělesem prochází. Toto světlo se při průchodu tělesem láme, zeslabuje a částečně přijímá barvu tělesa, u průsvitných těles (přes které vidíme rozmazaně) navíc ještě dochází k jeho rozptylu. Směr lomeného paprsku závisí (stejně jako směr paprsku odraženého) na směru paprsku dopadajícího, závislost lze vyjádřit tzv. *Snellovým zákonem lomu* (obrázek 4.8):

$$\frac{\sin\alpha}{\sin\beta} = \frac{n_b}{n_a} \tag{4.19}$$

kde n_a resp. n_b je absolutní index lomu, ve kterém se šíří dopadající resp. lomený paprsek.

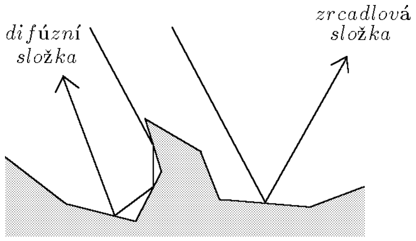
Index lomu udává poměr rychlosti světla ve vakuu a v daném prostředí. Pro jednotlivé materiály bývá určen fyzikálními měřeními, jeho hodnota se pohybuje od skoro 1 (plyny) přes 1.6 (sklo) až po skoro 3 (diamant). Tento index obecně závisí na vlnové délce, tato závislost se projevuje tzv. *disperzí* světla (např. spektrální rozklad bílého světla při jeho průchodu hranolem).

Šíří-li se paprsek z opticky hustšího do opticky řidšího prostředí ($n_a > n_b$), může nastat situace, kdy dle vztahu (4.19) bude $\sin\beta$ větší než 1. To odpovídá tzv. *totálnímu odrazu*, při němž již žádné světlo do druhého prostředí nepronikᡠa všechno se odráží. Úhel α_{tot}, od nějž výše dochází k totálnímu odrazu (tzv. *kritický úhel*) je definován vztahem:

$$\sin\alpha_{tot} = \frac{n_b}{n_a} \tag{4.20}$$

4.4.4 Charakteristika povrchu těles tzv. distribučními funkcemi

Povrch tělesa není ve skutečnosti nikdy ideálně hladký, je totiž tvořen soustavou mikroskopických náhodně orientovaných plošek, jak je znázorněno na obrázku 4.9.



Obr. 4.9: Povrch reálných těles

Pokud chceme ve světelném modelu aplikovat zákon odrazu a zákon lomu, potřebujeme vždy určit směr normály plochy v místě dopadu paprsku. Tento směr ale závisí na tom, na jakou mikroskopickou plošku paprsek dopadl. Existují světelné modely, které se směr této normály snaží postihnout co možná nejpřesněji (např. [TS67] [CT82]), častěji se však používá dále popsané řešení.

Paprsek může po dopadu opustit povrch dvěma způsoby (viz obr. 4.9). Buď se zrcadlově odrazí od některé z plošek, nebo je podroben vícenásobnému odrazu a lomu. Intenzitu odraženého světla lze tedy přirozeně vyjádřit jako součet dvou složek:

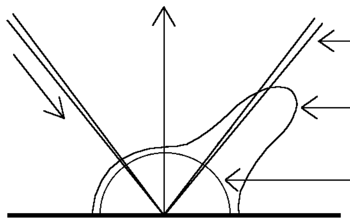
$$I = I_s + I_d \tag{4.21}$$

Složka I_s se nazývá zrcadlová (*specular*) a odpovídá přímo odraženému paprsku. Charakteristickou vlastností zrcadlové složky je její směrovost, která je tím vyšší, čím jsou plošky na povrchu méně "rozházené", neboli čím je povrch hladší. Tato složka je příčinou vzniku odlesků na zobrazovaném tělese.

Složka I_d se nazývá difúzní (*diffuse*) a odpovídá vícenásobně odraženým a lomeným paprskům. Její velikost je konstantní ve všech směrech, neboť po vícenásobném odrazu a lomu bude směr odraženého paprsku zcela náhodný.

Popsané řešení poskytuje aproximaci postačující pro jednodušší aplikace, pro popis povrchů reálných těles však selhává. K uvedeným dvěma složkám se proto přidává ještě tzv. *směrově difúzní složka*, jejíž opodstatnění potvrzují i výsledky prací z oblasti teorie odrazu světla (souhrn viz [HTSG91]). Rovnice pro intenzitu odraženého světla (4.21) pak přechází do tvaru (viz též obrázek 4.10):

$$I = I_{is} + I_{id} + I_{dd} \tag{4.22}$$



Obr. 4.10: Distribuční funkce

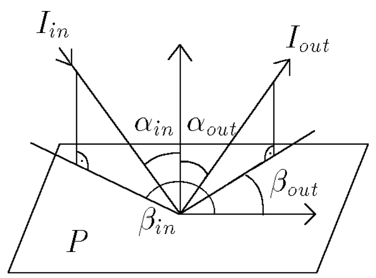
Charakteristiku povrchu je v tomto komplexním případě lepší vyjádřit pomocí tzv. *distribuční funkce* ρ. Tato funkce globálně popisuje, jak daný povrch odráží dopadající světelné paprsky a obecně závisí na následujících parametrech:

- směru dopadu světelného paprsku (tj. na úhlech α_{in} a β_{in}, které svírá dopadající paprsek s hlavním směrem resp. s normálou dané plochy)
- směru, ve kterém potřebujeme určit intenzitu odraženého paprsku (na podobně definovaných úhlech α_{out} a β_{out})
- vlnové délce dopadajícího světelného paprsku λ,

Hodnoty distribuční funkce leží v rozmezí $\langle 0, 1 \rangle$ a udávají poměrnou část energie dopadajícího paprsku, která je vyzařována v daném směru. Odraz světelné energie na povrchu tělesa lze tedy popsat následujícím vztahem (situace je zachycena na obrázku 4.11):

$$I_{out}(\alpha_{out}, \beta_{out}) = I_{in} \cdot \rho(\alpha_{in}, \beta_{in}, \alpha_{out}, \beta_{out}, \lambda) \tag{4.23}$$

obě intenzity jsou vztaženy k povrchu dané plochy (z části 4.4.1 popisující světelné zdroje již víme, že intenzita světla ubývá se vzdáleností).



Obr. 4.11: Odraz světla

Uvedený popis odrazu paprsku vychází z fyzikálních teorií a je pro aplikace v radiačních metodách zbytečně složitý. Proto se často používá jeho zjednodušení, tzv. *dvousměrové distribuční funkce (BRDF - bidirectional reflectance distribution function)*. Zjednodušení spočívá v tom, že popis pomocí BRDF neuvažuje závislost intenzity odraženého paprsku na úhlu β_{in} mezi dopadajícím paprskem a hlavním směrem dané plochy, nelze tedy reprezentovat tzv. *anizotropní povrchy* (např. soustružená tělesa či kovy broušené v jednom směru). Při použití BRDF přechází rovnice (4.23) do tvaru:

$$I_{out}(\alpha_{out}, \beta_{out}) = I_{in} \cdot \rho(\alpha_{in}, \alpha_{out}, \beta_{out}, \lambda) \tag{4.24}$$

Na distribuční funkce jednotlivých povrchů se však lze dívat i jiným způsobem. Ve své podstatě určují tyto funkce tvar vyzařovací funkce daného povrchu, která je generovaná dopadajícím paprskem. Vlastnosti povrchu lze tedy kromě výše popsaných distribučních funkcí $\rho(\alpha_{in}, \beta_{in}, \alpha_{out}, \beta_{out}, \lambda)$ popsat i pomocí tzv. *alternativních distribučních funkcí*:

$$\Psi(\alpha_{in}, \beta_{in}, \lambda) = \Phi_{\alpha_{in}, \beta_{in}, \lambda}(\alpha, \beta) \tag{4.25}$$

které každému dopadajícímu paprsku přiřazují jemu odpovídající tzv. *elementární vyzařovací funkci* $\Phi(\alpha, \beta)$. Tato vyzařovací funkce pak charakterizuje směrové rozložení světelné energie po odrazu (resp. lomu) na daném povrchu. Vztah mezi původními a alternativními distribučními funkcemi je následující:

$$\rho(\alpha_{in}, \beta_{in}, \alpha_{out}, \beta_{out}, \lambda) = \Psi(\alpha_{in}, \beta_{in}, \lambda)(\alpha_{out}, \beta_{out}) \tag{4.26}$$

V případě BRDF pak:

$$\rho(\alpha_{in}, \alpha_{out}, \beta_{out}, \lambda) = \Psi(\alpha_{in}, \lambda)(\alpha_{out}, \beta_{out}) \tag{4.27}$$

Obdobným způsobem jako pro odraz lze pomocí distribučních funkcí popsat i charakteristiku povrchu pro lom světelných paprsků. Pokud použijeme alternativní distribuční funkce, musíme v tomto případě umět k libovolnému světelnému paprsku, který dopadá na danou plochu, určit následující dvě funkce:

- elementární vyzařovací funkci pro odraz (tzv. *vnější elementární funkci*) $\Phi^{outer}_{\alpha_{in},\lambda}(\alpha,\beta)$
- elementární vyzařovací funkci pro lom (tzv. *vnitřní elementární funkci*) $\Phi^{inner}_{\alpha_{in},\lambda}(\alpha,\beta)$

Jsou použity modifikované vyzařovací funkce (viz kapitola 4.4.1), tj.

$$\Phi_{\alpha_{in},\lambda}(\alpha,\beta) = I_{\alpha_{in},\lambda}(\alpha,\beta) \cdot \cos\alpha_{in} \tag{4.28}$$

kde $I_{\alpha_{in},\lambda}(\alpha,\beta)$ je skutečný popis směrového rozložení světelné energie po jejím odrazu resp. lomu. Charakteristikou elementárních vyzařovacích funkcí v závislosti na dopadajícím světelném paprsku se zabývají lokální světelné modely, v našem případě je použita modifikace Straussova světelného modelu, viz kapitola 4.8.

Pomocí distribučních funkcí lze postihnou všechny výše citované jevy (zákon odrazu a zákon lomu, difúzní a zrcadlový odraz, směrovou difúzi, totální odraz). O tom, jak lze řešit závislost některých jevů na vlnové délce pojednává část 4.4.6. Při přenosu energie mezi ploškami tedy musíme při použití distribučních funkcí explicitně řešit pouze závislost intenzity světelné energie na vzdálenosti a dále uvedené jevy vznikající při průchodu světla prostředím (viz kapitola 4.4.7).

4.4.5 Popis radiace ploch pomocí vyzařovacích funkcí

Pro každou plochu scény potřebujeme znát její radiaci, tedy popis směrového rozložení energie, kterou daná plocha vyzařuje nebo odráží do svého okolí. Je zřejmé, že tuto radiaci lze popsat pomocí *vyzařovacích funkcí* stejně jako v případě světelných zdrojů (kapitola 4.4.1). Toto řešení navíc dobře vyhovuje základní myšlence radiačních metod, totiž že všechny plochy scény jsou chápány jako světelné zářiče a rozdíl mezi světelnými zdroji a ostatními tělesy je pouze v tom, zda se daná plocha považuje za zářič primární či sekundární. Vzhledem k tomu, že každá plocha propaguje světelnou energii na obě strany (jak dovnitř, tak i ven z odpovídajícího tělesa), jsou s jednotlivými plochami scény spojeny vždy dvě vyzařovací funkce:

- vyzařovací funkce pro radiaci plochy směrem ven z tělesa (tzv. *vnější vyzařovací funkce*) $\Phi^{outer}(\alpha,\beta)$
- vyzařovací funkce pro radiaci plochy směrem dovnitř tělesa (tzv. *vnitřní vyzařovací funkce*) $\Phi^{inner}(\alpha,\beta)$

Opět jsou použity modifikované vyzařovací funkce, tedy:

$$\Phi(\alpha, \beta) = I(\alpha, \beta) \cdot \cos \alpha \tag{4.29}$$

kde $I(\alpha, \beta)$ je charakteristika směrového rozložení světelné energie po jejím odrazu resp. lomu. Podrobněji bude způsob reprezentace vyzařovacích funkcí popsán v kapitole 4.7.

4.4.6 Modifikace modelu pro barevné scény

Až dosud jsme uvažovali pouze černobílé scény, přesněji zajímala nás pouze intenzita světla, nikoliv jeho barevný odstín. V případě barevného výstupu je jednak nutné nějakým způsobem reprezentovat barvy, jednak musíme brát zřetel na to, že mnoho jevů závisí na vlnové délce světla.

Barva tělesa odpovídá spektru světelné energie, kterou povrch tohoto tělesa vyzařuje (jedná-li se o světelný zdroj) nebo odráží. Nejlepší by tedy bylo barvu reprezentovat příslušnou *spektrální funkcí*, tento fyzikální přístup je však příliš složitý (a zbytečně dokonalý) a v oblasti počítačové grafiky se prakticky nepoužívá. Místo toho se používá drobného "triku", který umožňuje, aby se z barevnou scénou pracovalo stejně jako s několika scénami černobílými. Jeho podstata je vysvětlena v následujícím odstavci.

Je známo (viz např. [Ska92]), že jakýkoliv barevný odstín lze vyjádřit pomocí několika základních barev určitého barevného systému. V počítačové grafice se nejčastěji používají barevné systémy RGB (pro výstup na obrazovku) nebo CMYK (pro tisk obrázků na barevných tiskárnách). Barevný systém RGB, který je použit i v této práci, definuje jako základní barvy červenou (*red* - R), zelenou (*green* - G) a modrou (*blue* - B). Jedná se o systém aditivní, libovolnou barvu C lze tedy vyjádřit jako součet těchto základních barev:

$$C = c_R \cdot R + c_G \cdot G + c_B \cdot B \tag{4.30}$$

kde váhové koeficienty c_R, c_G a c_B určují množství jednotlivých základních barev, jejichž smícháním vznikne požadovaný barevný odstín. Tyto koeficienty se nazývají *trojbarví spektrální členitelé (tristimulus values)*, závislost jejich velikostí na vlnové délce lze nalézt např. v [Ska92]. Pro reprezentaci barvy tedy stačí mít u dané plochy uloženy odpovídající hodnoty váhových koeficientů c_R, c_G a c_B, nejčastěji v tzv. *RGB vektoru* $[R,\ G,\ B]^T$.

Nyní, kdy umíme reprezentovat jednotlivé barevné odstíny, vraťme se k tomu, jak modifikovat výše popsaný světelný model pro barevné scény. Jak již bylo řečeno, použití RGB vektorů umožňuje s barevnou scénou pracovat stejně jako s třemi černobílými scénami, kdy jednotlivé černobílé scény reprezentují červenou, zelenou a modrou část původní barevné scény. Všechny akce v rámci výpočtu propagace světla je tedy nutné provádět třikrát, pro každou barevnou složku zvlášť.

Použitím systému RGB lze velice jednoduše vyřešit i problém závislosti některých jevů na vlnové délce světla. Stačí, postihneme-li daný jev v jednotlivých barevných složkách R, G a B. V případě konstant a koeficientů tak místo jednoho čísla dostáváme RGB vektor, u vyzařovacích a distribučních funkcí pak strukturu, v níž jsou uloženy tvary těchto funkcí pro každou barevnou složku.

4.4.7 Průchod světla prostředím

Prochází-li světelný paprsek libovolným prostředím, postupně přijímá barvu tohoto prostředí a navíc je tímto prostředím částečně pohlcován. Intenzita světla klesá exponenciálně se vzdáleností, jakou paprsek urazil v daném prostředí a lze ji vyjádřit vztahem:

$$I(d) = I \cdot \mu(\lambda) \cdot \mathrm{e}^{-\kappa(\lambda) \cdot d} \tag{4.31}$$

kde:

- $I(d)$ je intenzita světla ve vzdálenosti d
- I je počáteční intenzita (před vstupem do daného prostředí)
- $\mu(\lambda)$ je tzv. *absorpční koeficient* daného prostředí
- $\kappa(\lambda)$ je tzv. *činitel útlumu* daného prostředí

Funkce $\mu(\lambda)$ a $\kappa(\lambda)$ obecně závisí na vlnové délce a právě tato závislost způsobuje změny barvy světelného paprsku při jeho průchodu prostředím. V popisovaném světelném modelu se však obě funkce považují pro jednotlivé materiály za konstantní a postupné přejímání barvy tělesa je postiženo následovně:

$$I_i(d) = I_i \cdot \mu \cdot C_i \cdot \mathrm{e}^{-\kappa \cdot d} \quad \text{pro} \quad i = R,\ G,\ B \tag{4.32}$$

kde:

- $I_i(d)$ je intenzita odpovídající barevné složky paprsku ve vzdálenosti d
- I_i je počáteční intenzita této složky
- C_i je odpovídající složka barvy prostředí, kterým paprsek prochází (v intervalu $\langle 0, 1 \rangle$)
- μ a κ je absorpční koeficient resp. činitel útlumu daného prostředí

4.5 Popis generované scény

Nyní, když známe použitý světelný model i prostředky pro popis geometrie scény, uveďme si, jak je generovaná scéna skutečně popsaná. Globální propagace světelné energie uvnitř scény bude určována pomocí "vystřelovacího" algoritmu (viz kapitola 4.6), tomu bude přizpůsobený i popis dané scény. Pro popis scény jsou použity celkem tři různé typy souborů:

- vstupní soubor
- mezisoubory
- výsledek (výstupní soubor)

Vstupní soubor obsahuje uživatelský popis scény, která se má generovat. Tento soubor se použije pouze na začátku radiačního výpočtu dané scény, kdy je z něj vygenerován odpovídající mezisoubor, vhodný pro vlastní výpočet. Rozmístění jednotlivých objektů ve scéně a jejich vzájemné vztahy lze postihnout libovolným způsobem, který lze převést na povrchovou reprezentaci. Výhodné je použití CSG stromů, u tvarově složitých objektů pak jejich kombinace s Bèzierovými nebo β-spline plochami. K objektům, které se mají chovat jako primární světelné zdroje, je nutné v rámci vstupního souboru připojit specifikace jejich iniciálních vyzařovacích funkcí. Vlastnosti použitých materiálů a charakteristiky jednotlivých povrchů je výhodné uchovávat odděleně v knihovnách, které pak mohou být společné pro více scén.

Mezisoubory slouží k uchování radiačního řešení mezi jednotlivými kroky výpočtu propagace světla scénou. Daná scéna je v tomto souboru popsána pomocí plošné reprezentace, pro aproximaci povrchu jsou použity trojúhelníkové plošky nebo pláty, skládající se z jednotlivých plošných elementů. S těmito elementy jsou v mezisouborech svázány jejich vyzařovací funkce a další "radiační" charakteristiky potřebné pro vlastní výpočet (optické vlastnosti, odkazy do knihoven materiálů a povrchů atd.). Vzhledem k tomu, že pro vlastní výpočet globální propagace světla je použit "vystřelovací" algoritmus (viz kapitola 4.6), jsou s každým elementem svázány vždy dvě vyzařovací funkce, a to jednak funkce charakterizující směrové rozložení energie, kterou daný element celkově vyzařuje do svého okolí, jednak obdobná funkce pro energii, která dosud nebyla propagována zpět do scény ("nevystřelená" energie). Kromě těchto údajů týkajících se vlastní scény jsou v mezisouborech uloženy i údaje charakterizující použitou směrovou metodu (např. počet a koeficienty aproximačních funkcí, kterými se reprezentují vyzařovací funkce jednotlivých elementů scény, popis, jakým způsobem se vybírá aktualizační element, hranice, kdy bude daný element rozdělen na čtyři menší, ...). Pro urychlení jednotlivých kroků výpočtu lze do mezisouborů umístit též další pomocné údaje, např. vzdálenosti mezi jednotlivými plochami, směry normál plošných elementů, informace o vzájemné viditelnosti a podobně. Velikost těchto souborů je ovlivněna hlavně efektivitou použité reprezentace jednotlivých vyzařovacích funkcí a pro složité scény může dosáhnout i několika MB. Je nutné si však uvědomit, že v těchto souborech je postižena globální propagace světelné energie v rámci celé scény.

Podobný formát jako mezisoubory mají i výstupní soubory. Tyto soubory obsahují výsledné radiační řešení dané scény a používají se při generování jednotlivých pohledů na scénu po ukončení vlastního radiačního výpočtu. Vzhledem k tomu, že při zobrazování potřebujeme znát pouze vyzařovací funkce jednotlivých plošek a umístění těchto plošek v dané scéně, obsahují výstupní soubory oproti mezisouborům pouze tyto informace (a jsou tedy výrazně kratší). Z daného výstupního souboru lze generovat libovolné množství různých pohledů na zadanou scénu, lze generovat i celé animované sekvence (např. průchod navrhovaným domem). Při kresbě jednotlivých pohledů se jedná pouze o zobrazování předem spočítané scény, vše tedy lze provádět v reálném čase.

4.6 Propagace světelné energie

Oproti ray-tracingu, který vystačí s popisem šíření jednotlivých světelných paprsků scénou tak, jak bylo popsáno v kapitole 4.4, je u radiačních metod nutné postihnout propagaci veškeré energie v rámci dané scény. Tento popis lze z uvedeného světelného modelu odvodit postupně v následujících krocích:

- přenos světelné energie z jedné diferenciální plošky na druhou
- přenos světelné energie mezi dvěma plochami
- globální propagace světla celou scénou

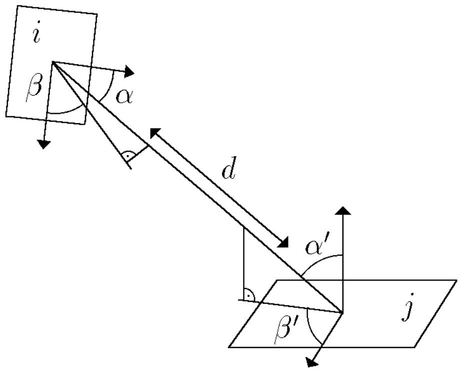
4.6.1 Přenos světelné energie z jedné diferenciální plošky na druhou

Chceme-li popsat přenos světelné energie v rámci dané scény, musíme nejprve odvodit odpovídající vztahy v rámci zanedbatelně malých (tzv. *diferenciálních*) plošek. V kapitole 3.1 jsme si uvedli odpovídající vztahy pro standardní radiační metody, nyní musíme tyto vztahy modifikovat. Modifikace spočívá v zahrnutí směrového přístupu do uvedených vztahů, tj. v přeformulování jednotlivých rovnic tak, abychom mohli použít směrový světelný model popsaný v kapitole 4.4. Úpravy jsou následující:

- náhrada konstantní intenzity vyzařované energie jejím směrovým postižením pomocí odpovídající vyzařovací funkce
- postižení poklesu intenzity světla v závislosti na vzdálenosti
- zahrnutí jevů souvisejících s průchodem paprsku prostředím
- postižení vlastností povrchů odpovídajícími distribučními funkcemi

Uvažujme situaci, kdy dochází k přenosu světelné energie z diferenciální plošky i na diferenciální plošku j (viz obrázek 4.12), kde:

- úhly α a β určují odchylku spojnice středů obou plošek od normály a hlavního směru plošky i
- úhly α' a β' určují odchylku této spojnice od normály a hlavního směru plošky j
- d je vzájemná vzdálenost obou plošek
- μ a κ jsou absorpční koeficient a činitel útlumu prostředí mezi ploškami, C je jeho barva
- $\Phi_i(\alpha, \beta)$ je vyzařovací funkce plošky i
- $\Phi_j^i(\alpha, \beta)$ je elementární vyzařovací funkce plošky j generovaná i
- $\Psi_j(\alpha', \beta')$ je distribuční funkce plošky j



Obr. 4.12: Přenos světelné energie

Intenzitu energie vyzařované z plošky i ve směru k plošce j lze pak vyjádřit následovně:

$$I_{i,out} = \Phi_i(\alpha, \beta) \tag{4.33}$$

Při dopadu na plošku j má intenzita energie přicházející z plošky i velikost (viz vztahy (4.17) a (4.32)):

$$I_{j,in} = \frac{\mu \cdot C \cdot \mathrm{e}^{-\kappa \cdot d}}{a + b \cdot d + c \cdot d^2} \cdot I_{i,out} \tag{4.34}$$

Množství energie přenášené z plošky i na plošku j závisí na velikosti obou plošek, odpovídající vztah má tvar:

$$B_j^i = I_{j,in} \cdot P_i \cdot P_j \tag{4.35}$$

V rovnici (4.35) by se normálně měly objevit i členy $\cos\alpha$ a $\cos\alpha'$, tyto členy jsou však zahrnuty již ve vyzařovací resp. distribuční funkci Φ_i a Ψ_j. Energie popsaná tímto vztahem je po dopadu na plošku j rozptylována do všech směrů, rozptyl popisuje elementární distribuční funkce této plošky odpovídající danému směru. Příspěvek plošky i k celkové vyzařovací funkci plošky j lze tedy vyjádřit takto:

$$\Phi_j^i(\alpha, \beta) = \frac{B_j^i}{P_j} \cdot \Psi_j(\alpha', \beta') \tag{4.36}$$

Celkově lze tedy působení diferenciální plošky i na diferenciální plošku j popsat vztahem:

$$\Phi_j^i(\alpha, \beta) = \frac{\mu \cdot C \cdot \mathrm{e}^{-\kappa \cdot d}}{a + b \cdot d + c \cdot d^2} \cdot \Phi_i(\alpha, \beta) \cdot P_i \cdot \Psi_j(\alpha', \beta') \tag{4.37}$$

4.6.2 Přenos světelné energie mezi dvěma plochami scény

V případě přenosu světelné energie mezi reálnými plochami (tj. mezi plochami, které již nemají zanedbatelnou velikost) A_i a A_j musíme vztah (4.36) zintegrovat přes obě plochy a dostáváme:

$$\Phi_{A_j}^{A_i} = \frac{1}{A_i} \cdot \int\limits_{A_i} \int\limits_{A_j} \Phi_j^i \; dA_i \, dA_j \tag{4.38}$$

Rozepsáním tohoto vztahu pomocí rovnice (4.37) získáme:

$$\Phi_{A_j}^{A_i}(\alpha, \beta) = \frac{1}{A_i} \cdot \int\limits_{A_i} \int\limits_{A_j} \frac{\mu \cdot C \cdot \mathrm{e}^{-\kappa \cdot d}}{a + b \cdot d + c \cdot d^2} \cdot \Phi_i(\alpha, \beta) \cdot P_i \cdot \Psi_j(\alpha', \beta') \; dA_i \, dA_j \tag{4.39}$$

Tento vztah charakterizuje přenos energie z plochy A_i na plochu A_j a popisuje vlastně elementární operaci používanou při výpočtu globální propagace světla danou scénou.

Analytické řešení dvojného integrálu ve vztahu (4.39) je obecně velice složité, proto se často používá postup, kdy se integrál přes nějakou plochu nahradí sumou přes elementární plošky, z nichž se daná plocha skládá (viz např. [CG85]). Rovnici (4.38) pak odpovídá následující aproximace:

$$\Phi_{A_j}^{A_i} \approx \frac{1}{A_j} \cdot \sum_{A_i} \sum_{A_i} \Phi_j^i \tag{4.40}$$

Vztah (4.40) již lze použít pro určení propagace energie mezi jednotlivými plochami generované scény, výpočet je ale značně časově náročný (počítá se dvojitá suma přes funkce dvou proměnných, výsledkem je opět funkce). Proto se v popisované metodě používá místo aktualizace typu plocha – plocha přímo aktualizace typu elementární ploška – elementární ploška a základní vztah (4.37). Přenos světelné energie mezi dvěma elementárními ploškami i a j pak probíhá v následujících krocích (světlo je propagováno z plošky i na plošku j):

1. urči vzájemnou polohu obou plošek; pokud se plošky navzájem nevidí, není co propagovat, konec
2. z vyzařovací funkce plošky i urči množství světelné energie, kterou tato ploška vysílá směrem k plošce j
3. urči intenzitu této energie při dopadu na plošku j (úbytek intenzity se vzdáleností a útlum světla při průchodu prostředím)
4. pomocí distribuční funkce plošky j určíme příspěvek plošky i k radiaci plošky j
5. aktualizuj vyzařovací funkci plošky j

V případě, kdy jsou distribuční a vyzařovací funkce rozděleny na vnitřní a vnější složku, použije se vždy odpovídající složka těchto funkcí. Nepřesnosti, které vznikají náhradou vztahu (4.39) vztahem:

$$\Phi_{A_j}^{A_i} \approx \Phi_j^i \tag{4.41}$$

se sledují a přesáhnou-li jistou hranici, je odpovídající element rozdělen. Podrobněji je toto adaptivní dělení elementů popsáno v kapitole 4.12.

4.6.3 Globální propagace světla celou scénou

V předchozí kapitole jsme si popsali přenos energie mezi základními stavebními prvky dané scény, nyní zbývá zobecnit daný postup na celou scénu. Jak již bylo řečeno, je použit "vystřelovací" algoritmus (viz kapitola 3.2 nebo článek [CCWG88]), který je přeformulován tak, aby uvažoval i směr propagované energie. Jednotlivé kroky modifikovaného algoritmu jsou následující:

1. Ze vstupního souboru popisujícího scénu sestav a inicializuj prvotní mezisoubor (viz kapitola 4.5)
2. Načti mezisoubor reprezentující aktuální radiační řešení pro danou scénu. Z tohoto mezisouboru je vygenerován vnitřní popis scény a inicializovány vyzařovací funkce jednotlivých elementů
3. Vyber vhodný element i, ze kterého se bude propagovat jeho dosud nevystřelená energie (viz kapitola 4.10)
4. Propaguj vyzařovanou energii elementu scénou, tj. postupem popsaným v předchozí sekci postupně urči příspěvek elementu i ke všem elementům scény j, které jsou z elementu i viditelné. Tento krok lze rozdělit na následující části:
   (a) "vnitřní" aktualizace - zde je postihnuta propagace uvnitř daného tělesa, jsou použity vnitřní vyzařovací a distribuční funkce elementu i, cílové elementy se vybírají pouze z plošek, které tvoří stejné těleso jako element i
   (b) "vnější" aktualizace - postihuje propagaci energie na ostatní tělesa scény, použijí se vnější funkce elementu i. Je-li těleso, do nějž element i patří, konvexní, nemusíme se zabývat ostatními ploškami tohoto tělesa, neboť tyto plošky jsou z elementu i neviditelné
5. U každého viditelného elementu urči příspěvek elementu i k jeho vyzařovacím funkcím (k celkové i k dosud nevystřelené)
6. U elementu i nuluj jeho vyzařovací funkci pro nevystřelenou energii (všechna tato energie byla předchozím bodem propagována do scény)
7. Došlo-li po přenosu světelné energie z elementu i na nějaký element j k nerovnoměrnostem v rozložení jeho celkově vyzařované energie, rozděl element j na čtyři menší (podrobněji viz kapitola 4.12)
8. Vyber místo pozorovatele a směr pozorování a nakresli požadovaný pohled. Tento bod lze libovolně opakovat, lze jej také přeskočit. Vlastní kresbě jednotlivých pohledů bude věnována kapitola 4.13
9. Je-li současné řešení dostatečně přesné, vygeneruj výstupní soubor, obsahující radiační charakteristiku celé scény, jinak pokračuj bodem 10 Jak lze určit přesnost daného řešení bude popsáno v kapitole 4.11
10. Chceš-li provést další "vystřelovací" krok, opakuj celý postup od bodu 3, jinak pokračuj bodem 11
11. Z vnitřního popisu scény vytvoř mezisoubor s popisem dosavadního radiačního řešení a skonči. Při příštím spuštění programu pokračuj od bodu 2

Při průběžném zobrazování dané scény může uživatel sledovat postupné zvyšování kvality generovaných obrázků a případně upravovat parametry aktuálního pohledu. Ideálním řešením je pak paralelní běh vlastního výpočtu a zobrazování (i více) pohledů současně.

Celý algoritmus lze urychlit drobnou modifikací, kdy v bodu 5. bude příspěvek elementu i přičítán pouze k dosud nevystřelené energii elementu j. Přičtení dosud nevystřelené energie k celkové energii daného elementu se pak bude provádět až před kresbou dané scény (mezi body 7 a 8) nebo pokud je daný element vybrán jako aktualizační (bod 3). Při použití této modifikace je také nutné odpovídajícím způsobem upravit bod 7.

Dalšího (a celkem výrazného) urychlení je možné dosáhnout tehdy, kdy jsou jednotlivé vyzařovací funkce rozděleny na vnější a vnitřní (jako v našem případě). Pak lze i jednotlivé aktualizace rozdělit na vnitřní a vnější, modifikovat bod 3. a z bodu 4. provádět vždy jen odpovídající část. V tomto případě se vlastně každý element scény chová jako dva na sobě zcela nezávislé objekty, každý s jednou vyzařovací funkcí.

Ani použití propagace typu plošný element – plošný element jako základního kroku celého algoritmu pro výpočet globální propagace světelné energie danou scénou není zcela ideální řešení. Jako daleko výhodnější se ukazuje použití propagace typu vrchol – vrchol, kdy se přenos světelné energie odehrává mezi jistými okolími jednotlivých vrcholů plošek, z nichž se daná scéna skládá. Vzhledem k tomu, že se i v tomto případě jedná o propagaci světla mezi nějakými ploškami, stačí pouze nepatrně upravit výše popsaný postup radiačního výpočtu. Blíže se budeme aktualizací vrcholů namísto plošných elementů zabývat v kapitole 4.9.

4.7 Realizace vyzařovacích funkcí

V předchozích kapitolách jsme si popsali, jakým způsobem se určuje množství světelné energie, která je propagována scénou. Mluvili jsme o tzv. vyzařovacích funkcích jednotlivých ploch (viz např. kapitola 4.4.1), definovali jsme, co tyto funkce vyjadřují a jak se s nimi pracuje, blíže jsme však nespecifikovali jejich skutečnou realizaci. To je náplní této kapitoly, která je rozdělena na následující části:

- požadavky na reprezentaci vyzařovacích funkcí
- aproximace řadou funkcí, ortogonální systémy

- Fourierovy řady
- výběr jednotlivých aproximačních funkcí
- operace s vyzařovacími funkcemi
- shrnutí, porovnání s dosavadními postupy

4.7.1 Požadavky na reprezentaci vyzařovacích funkcí

Jak již bylo uvedeno v popisu světelného modelu (kapitola 4.4), je propagace světelné energie v rámci dané scény postihnuta pomocí vyzařovacích funkcí. Tyto funkce jsou svázány s jednotlivými elementárními objekty dané scény (normálně s plošnými elementy, v našem případě s jednotlivými vrcholy) a popisují směrové rozložení energie, kterou dané objekty vyzařují do svého okolí. Jedná se o funkce dvou proměnných, jejichž hodnoty jsou (v případě černobílých scén) kladná reálná čísla, lze tedy psát (viz kapitola 4.4.1):

$$I(\alpha, \beta) = I_{\alpha,\beta} \tag{4.42}$$

kde

- $I(\alpha, \beta)$ je vyzařovací funkce daného objektu
- $\alpha \in \langle 0, \pi \rangle$ je úhel mezi požadovaným směrem a normálou objektu
- $\beta \in \langle -\pi, \pi)$ je odchylka tohoto směru od tzv. hlavního směru objektu (předem určený a s objektem pevně svázaný směr, kolmý na normálu)
- $I_{\alpha,\beta} \in \langle 0, 1 \rangle$ je intenzita energie vyzařované objektem v daném směru

Z praktických důvodů se vyzařovací funkce daného objektu dělí na tzv. vnější a vnitřní vyzařovací funkci (pak lze např. samostatně provádět vnější a vnitřní aktualizační krok), u obou funkcí je pak $\alpha \in \langle 0, \pi/2 \rangle$. Při použití vystřelovacího algoritmu je navíc nutné pro každý objekt kromě vyzařovací funkce pro celkovou vyzařovanou energii uchovávat ještě vyzařovací funkci pro dosud nevystřelenou energii, kterou daný objekt ještě nepropagoval do scény. Pro barevné scény musíme jednotlivé vyzařovací funkce reprezentovat jejich barevnými složkami pro červenou, zelenou a modrou část spektra (viz kapitola 4.4.6). Pokud shrneme výše uvedené požadavky, dojdeme k závěru, že pro každý základní objekt scény (tj. plošný element nebo vrchol) je nutné uchovávat celkem 12 elementárních vyzařovacích funkcí – (*celková a nevystřelená energie*) × (*vnější a vnitřní funkce*) × (*složka R, G a B*). Je tedy zřejmé, že největší vliv na paměťovou i časovou náročnost celého radiačního výpočtu bude mít právě způsob, jak úsporně a efektivně budou dané elementární funkce uloženy a jak rychle se s nimi bude dát pracovat.

V kapitole 4.1 byla mimo jiné popsána i metoda, kde je vyzařovací funkce reprezentována hodnotami v některých předem stanovených směrech. Takto uložené vyzařovací funkce lze docela jednoduše aktualizovat, směr propagované energie je však postihnut značně nedokonale. Pro určení hodnoty vyzařovací funkce v požadovaném směru je nutné průměrovat hodnoty uložené pro přibližně stejné směry, tím samozřejmě vznikají jisté nezanedbatelné chyby. Může se dokonce stát, že ostře směrové vyzařování (např. odraz světelného paprsku na lesklých površích) nebude v jistých směrech vůbec postihnuto. Paměťové nároky této metody jsou při její nízké kvalitě pro většinu aplikaci příliš vysoké. V našem případě by například bylo nutné pro reprezentaci každé vyzařovací funkce použít při kroku 5 stupňů (což je horní hranice pro alespoň trochu kvalitní popis směrové propagace energie ve scéně) celkem $12 \times 18 \times 72$, což je přibližně 15 tisíc údajů.

Opakem výše popsané metody je distribuovaný ray-tracing. V rámci tohoto postupu není nutné jednotlivé vyzařovací funkce vůbec ukládat, směrová propagace energie je simulována vrháním velkého množství světelných paprsků do stejného místa (několik desítek až stovek) a jejich rozptylem dle odpovídající distribuční funkce daného povrchu. Tentokrát je tedy paměťová náročnost zanedbatelná, použití metody však znemožňují obrovské nároky časové. Navíc nelze postihnout propagaci světelné energie v rámci celé scény, ale pouze v rámci právě generovaného obrázku, metoda také neumí rozumným způsobem postihnout ideální difúzní odraz.

Je zřejmé, že je třeba vytvořit metodu, která by při dostatečně přesném postižení směru propagované energie byla ještě pořád jak časově, tak i paměťově únosná. Poprvé se tento cíl podařilo splnit autorům metody popsané v článku [SAWG91] (viz též kapitola 4.1), základem úspěchu byl popis vyzařovacích funkcí nikoliv lokálně hodnotami v několika směrech, ale globálně její aproximací pomocí funkční řady. Jelikož je tento přístup velice zajímavý a byl použit i v popisované směrové metodě, zmíníme se o něm v následujících odstavcích trochu podrobněji.

4.7.2 Aproximace řadou funkcí, ortogonální systémy

Myšlenka aproximace zadané funkce funkční řadou je velice jednoduchá: Nechť je dána funkce f a systém funkcí $S_F = \{f_i\}_{i=1}^N$. Výraz tvaru $\sum\limits_{i=1}^{N} c_i f_i$, kde $c_i \in R$ a $f_i \in S_F$ pro $i = 1 \cdots n$ nazveme *funkční řadou* nad daným systémem funkcí S_F, čísla c_i pak *váhovými koeficienty* této řady. Úkolem aproximace funkce f je stanovit hodnoty jednotlivých váhových koeficientů c_i tak, aby bylo možné funkci f nahradit (*aproximovat*) odpovídající funkční řadou, tj. aby $f \approx \sum_{i=1}^{N} c_i f_i$. Je zřejmé, že hodnoty jednotlivých vah (případně zjištění, že funkci f nelze uvedeným systémem aproximovat) budou záviset na daném kritériu určujícím přesnost aproximace, zvoleném systému funkcí S_F a jeho velikosti N a v neposlední řadě i na samotné funkci f.

Dříve, než se budeme věnovat jednotlivým problémům spojeným s aproximací vyzařovacích funkcí, zodpovězme si na otázku, proč se vůbec aproximace funkcí provádí. Důvody jsou v zásadě dvojího druhu: Jednak nemusí být aproximovaná funkce explicitně vůbec známa (a jsou známy pouze hodnoty v několika bodech), jednak může být tato funkce pro dané účely příliš složitá (např. ji nelze jednoduše integrovat či derivovat nebo hledat jednotlivé body, v nichž funkce nabývá požadovaných vlastností). V obou případech je proto daná funkce nahrazena skupinou "jednoduchých" funkcí s takovými vlastnostmi, které jsou pro dané účely nejvýhodnější (může se např. jednat o požadavek jednoduchých a rychlých "základních" operací, spojitosti až do určitého řádu, existence derivací a podobně).

Velice výhodné je také provádět aproximaci v případě, kdy potřebujeme danou funkci reprezentovat co nejmenším počtem údajů, aby ji bylo možné efektivně uložit v počítači (což je i náš případ vyzařovacích funkcí). Po provedení aproximace $f \approx \sum_{i=1}^{N} c_i f_i$ je totiž daná funkce jednoznačně určena hodnotami jednotlivých vah c_i, neboť použité aproximační funkce f_i jsou explicitně známy. Celou funkci f lze tedy v počítači reprezentovat tzv. *vektorem vah* $\widetilde{f} = (c_1, c_2, \cdots, c_N)$, tj. N-rozměrným číselným vektorem. Popis funkce f vektorem vah $\widetilde{f}$ je při dostatečně přesné charakteristice této funkce paměťově velice efektivní (efektivnější snad již být nemůže), jistou nevýhodou jsou pouze vyšší časové nároky při práci s vektorem vah než s původní funkcí.

První, co je nutné při aproximaci funkcí udělat, je stanovit kritérium rozhodující o tom, zda je právě nalezená aproximace dané funkce vyhovující či zda je nutné určit aproximaci přesnější.

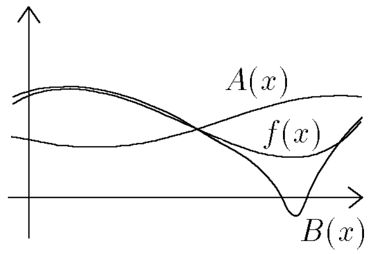
V případě tzv. *interpolace funkce* prokládáme zadanou funkci jednotlivými aproximačními funkcemi tak, abychom minimalizovali odchylky původní funkce a aproximace v jistých předem zvolených bodech. Z používaných postupů uveďme alespoň Newtonovu či Lagrangeovu interpolaci, kdy se jednotlivé body prokládají polynomy nebo tzv. spline-interpolaci, při které se používají funkce po částech polynomiální. Všechny tyto postupy umožňují danou funkci v odpovídajících bodech interpolovat zcela přesně, o chování původní funkce a přesnosti interpolace v ostatních částech intervalu však nelze říct vůbec nic.

Častěji se proto používají kritéria, která sledují přesnost aproximace nikoliv pouze v několika bodech, ale globálně na celém uvedeném intervalu. Nechť $f(x)$ je původní funkce a $A(x)$ je její aproximace na intervalu $\langle a, b\rangle$. Odchylkou aproximace v bodě $x \in \langle a, b\rangle$ nazveme výraz $\Delta_x = f(x) - A(x)$, největší odchylkou aproximace na uvedeném intervalu pak výraz $\Delta = \max\limits_{x\in\langle a,b\rangle} |\Delta_x|$. Tzv. *kritérium největší odchylky* se snaží minimalizovat hodnotu Δ, tj. pro zadanou funkci $f(x)$ a interval $\langle a, b\rangle$ je hledána aproximace $A(x)$ taková, aby hodnota Δ byla minimální. Měřítkem přesnosti konkrétní aproximace je přímo hodnota největší odchylky této aproximace, čím je tato hodnota menší, tím je daná aproximace přesnější. V případě ideální aproximace pak bude $\Delta = 0$. Existenci polynomiální aproximace při použití kritéria největší odchylky zajišťuje pro spojité funkce *Weierstrassova věta*, která tvrdí:

> Je-li funkce $f(x)$ spojitá v intervalu $\langle a, b\rangle$, existuje posloupnost polynomů $P_1(x),\ P_2(x), \cdots$ stejnoměrně konvergující k $f(x)$ na celém intervalu $\langle a, b\rangle$, tj.
>
> $$\forall \varepsilon > 0 \ \ \exists n : \ \left|f(x) - P_n(x)\right| < \varepsilon \quad na\ \langle a, b\rangle$$

Obdobná věta platí i pro aproximaci goniometrickými řadami (viz dále). Kritérium největší odchylky poskytuje oproti interpolaci podstatně lepší výsledky, z tvaru aproximované funkce jsme schopni odhadnout i tvar původní funkce, a to na celém intervalu, nikoliv pouze v několika bodech, jak tomu bylo u interpolace.

Při aproximaci funkce je však často důležitější než minimalizovat největší odchylku dosáhnout minimální "průměrné" chyby, tj. jakési pravděpodobné chyby, kterou můžeme očekávat při náhradě původní funkce její aproximací. Podívejme se na obrázek 4.13, kde je zobrazena funkce $f(x)$ a dvě její aproximace $A(x)$ a $B(x)$. Při použití kritéria největší odchylky bude vybrána aproximace $A(x)$ přesto, že aproximace $B(x)$ daleko lépe vystihuje tvar původní funkce $f(x)$ a bylo by tedy vhodné zvolit právě tuto aproximaci.



Obr. 4.13: Kritérium největší odchylky a "průměrná chyba"

Aby bylo možné postihnout "průměrnou" chybu, bylo vytvořeno *kritérium nejmenších čtverců*, kdy se pro danou funkci $f(x)$ a její aproximaci $A(x)$ na intervalu $\langle a, b\rangle$ definuje *střední kvadratická chyba* δ vztahem:

$$\delta^2 = \frac{1}{b-a} \cdot \int_a^b \Delta_x^2 \, dx = \frac{1}{b-a} \cdot \int_a^b \big(f(x) - A(x)\big)^2 \, dx \tag{4.43}$$

a cílem aproximace je tuto chybu minimalizovat. Měřítkem přesnosti konkrétní aproximace je hodnota δ, pro ideální aproximaci je opět $\delta = 0$. Dle uvedeného vztahu odpovídá čtverec střední kvadratické chyby aritmetickému průměru čtverců chyb v jednotlivých bodech celého intervalu, na němž se aproximuje, souvislost této chyby s výše uváděnou "průměrnou" chybou aproximace je tedy zjevná.

Kritérium nejmenších čtverců poskytuje z uvedených kritérií nejlepší možné výsledky, bude proto použité i v našem případě při aproximaci jednotlivých vyzařovacích funkcí. Až dosud jsme při formulaci jednotlivých kritérií uvažovali pouze funkce jedné proměnné, vyzařovací funkce jsou však funkce dvou proměnných (viz kapitola 4.7.1). Všechny uvedené kritéria lze samozřejmě zobecnit i pro aproximaci funkcí více proměnných, v případě dvojrozměrných vyzařovacích funkcí bude mít kritérium nejmenších čtverců následující tvar:

> Je dána funkce dvou proměnných $f(x, y)$ a její aproximace $A(x, y)$ na intervalu $\langle a, b\rangle \times \langle c, d\rangle$. Definujme střední kvadratickou chybu δ dané aproximace předpisem (porovnej se vztahem (4.43)):
>
> $$\delta^2 = \frac{1}{b-a} \cdot \frac{1}{d-c} \cdot \int_a^b \int_c^d \big(f(x,y) - A(x,y)\big)^2 \, dx \, dy \tag{4.44}$$
>
> Cílem aproximace metodou nejmenších čtverců je pak určit pro zadanou funkci $f(x, y)$ její aproximaci $A(x, y)$ tak, aby odpovídající střední kvadratická chyba δ byla minimální.

Po tom, co jsme si určili kritérium kvality jednotlivých aproximací, vraťme se nyní k vlastní volbě systému aproximačních funkcí $S_F = \{f_i\}_{i=1}^N$. Tento systém je charakterizován jednak svojí velikostí N (tj. počtem obsažených funkcí), jednak typem a vlastnostmi jednotlivých funkcí f_i. Při volbě N je třeba vyřešit problém spočívající v tom, že čím větší je toto číslo, tím je aproximace přesnější, s rostoucím N však také roste časová náročnost při určování hodnot aproximované funkce a hlavně velikost paměti, která je potřebná pro uložení jednotlivých váhových koeficientů. V našem případě, kdy pro každý vrchol scény potřebujeme uložit celkem 12 elementárních vyzařovacích funkcí (viz kapitola 4.7.1) je řešení tohoto problému z důvodu omezené paměťové kapacity jak hlavní paměti, tak i externích médií zřejmé – volit N co nejmenší. Hodnota N je pro praktické aplikace shora omezena hranicí 100 až 200 funkcí, což je v případě aproximace funkcí dvou proměnných velice malé číslo (této hranici odpovídá zhruba 10 až 15 funkcí pro každou souřadnici). Vlastní volbě systému S_F je proto třeba věnovat o to větší pozornost.

Jednotlivé funkce f_i je vhodné vybrat tak, aby byly co nejjednodušší a přitom měly jako celek dostatečně výhodné vlastnosti. V praxi se téměř výlučně používají následující typy aproximačních funkcí:

- polynomy – např. Hermitovy či Legenderovy polynomy
- po částech polynomiální funkce – např. B-spline, β-spline či NURBS
- trigonometrické funkce – tj. funkce typu $\cos(nx)$ a $\sin(nx)$

Výhodou prvních dvou typů aproximačních funkcí je jejich jednoduchost (polynomy se snadno a rychle vyčíslují, jsou dostatečně hladké, jednoduše se počítají jejich derivace atd.), systémy trigonometrických funkcí zase umožňují přirozený a často jednodušší popis tvarově složitých funkcí vyskytujících se při řešení problémů a popisu jevů vyskytujících se v přírodě (jak dosvědčuje téměř výhradní použití trigonometrických řad a tzv. Fourierovy transformace při zpracování zvukového či obrazového signálu nebo v nejrůznějších oblastech optiky). V našem případě popisujeme pomocí jednotlivých vyzařovacích funkcí šíření světelné energie v reálných scénách, bude proto použit systém trigonometrických funkcí.

Po volbě některého z typů aproximačních funkcí ještě zbývá určit konkrétní tvar jednotlivých funkcí, které budou použity v systému S_F. Je možné volit libovolné funkce daného typu, velice výhodné je však vybrat jednotlivé funkce tak, aby na daném intervalu tvořily tzv. *ortogonální systém.* O systému funkcí jedné proměnné $S_F = \{f_i\}_{i=1}^N$ říkáme, že je ortogonální na intervalu $\langle a, b\rangle$, jestliže platí:

$$\int_a^b f_i(x) \cdot f_j(x)\; dx = 0 \quad \Leftrightarrow \quad i \neq j \qquad \text{pro} \;\; i, j = 1, \cdots, N \tag{4.45}$$

Tuto definici lze zobecnit i pro funkce více proměnných, v našem případě, kdy jednotlivé vyzařovací funkce jsou funkcemi dvou proměnných, bude definice ortogonálního systému následující:

> Říkáme, že systém funkcí dvou proměnných $S_F = \{f_{ij}\}_{i,j=1}^N$ je ortogonální na intervalu $\langle a, b\rangle \times \langle c, d\rangle$, jestliže platí následující vztah (porovnej se vztahem (4.45)):

$$\int_a^b \int_c^d f_{ij}(x,y) \cdot f_{kl}(x,y)\, dx \neq 0 \quad \Leftrightarrow \quad i = k \;\&\; j = l \qquad (4.46)$$
$$pro \quad i,j,k,l = 1, \cdots, N$$

tj. pokud je integrál nenulový pouze tehdy, pokud v něm dvakrát vystupuje ta samá funkce. Je-li navíc:

$$\int_a^b f_i^2(x)\, dx = 1 \quad pro \;\; \forall\, i = 1, \cdots, N \qquad (4.47)$$

pro funkce jedné proměnné, respektive:

$$\int_a^b \int_c^d f_{ij}^2(x,y)\, dx = 1 \quad pro \;\; \forall\, i,j = 1, \cdots, N \qquad (4.48)$$

pro funkce dvou proměnných, nazývá se daný systém *ortonormální*.

Uveďme si ještě definici *úplného systému*:

Lze-li systémem S_F přesně aproximovat libovolnou funkci na daném intervalu, tj. pokud (v jednorozměrném případě) platí:

$$\begin{gathered} \forall f(x) \;\; \exists\, c_1, c_2, \cdots \;\; \forall x \in \langle a, b \rangle : \\ f(x) = \sum_{i=1}^{\infty} c_i f_i \qquad f_i \in S_F \end{gathered} \qquad (4.49)$$

nazýváme jej *úplný* (pak samozřejmě systém obsahuje nekonečně mnoho funkcí).

Výhody ortogonálních (resp. ortonormálních) systémů se projeví hlavně při výpočtu jednotlivých váhových koeficientů c_i pro konkrétní funkci $f(x)$. Je-li použitý systém aproximačních funkcí ortogonální, platí (viz např. [Smi56]):

1. určení hodnoty konkrétní váhy c_i nezávisí na hodnotách ostatních vah c_j (velice důležitá vlastnost, plynou z ní i některé dále uvedené vlastnosti)

2. existuje předpis, jak počítat jednotlivé váhy c_i tak, aby střední kvadratická chyba dané aproximace byla nejmenší možná, tj. při použití kritéria nejmenších čtverců umíme nalézt nejlepší možnou aproximaci (váhové koeficienty c_i se pak nazývají Fourierovy, podrobněji viz následující kapitola)

3. chyba aproximace klesá s množstvím použitých aproximačních funkcí
4. hodnoty jednotlivých vah c_i nezávisí na velikosti systému S_F (tj. na počtu funkcí N), při přidání nové funkce f_k stačí dopočítat pouze jí odpovídající váhu c_k
5. pokud je systém S_F ortonormální, mají jednotlivé váhové koeficienty stejnou "významnost", tj. stejnou měrou rozhodují o přesnosti celé aproximace (jednotlivé aproximační funkce jsou si tedy v tomto případě jistým způsobem rovnocenné)

Na závěr této kapitoly si ještě uveďme, že jak Legenderovy polynomy, tak i trigonometrické funkce tvoří úplné ortonormální systémy funkcí na intervalech $\langle 0,1\rangle$ resp. $\langle -\pi,\pi\rangle$ (důkaz je uveden např. v knize [Smi56]).

4.7.3 Fourierovy řady

Nyní, kdy jsme si uvedli základní poznatky o aproximaci funkcí funkčními řadami, věnujme se konečně vlastní realizaci vyzařovacích funkcí. Jak již bylo řečeno v předchozí kapitole, pro aproximaci těchto funkcí je použito kritérium nejmenších čtverců a systém trigonometrických funkcí.

Systém trigonometrických funkcí je v jednorozměrném případě tvořen funkcemi 1, $\cos(nx)$ a $\sin(nx)$, kde $n = 1, 2, \cdots$, jednotlivé funkce $f(x)$ jsou tedy aproximovány výrazem

$$\frac{a_0}{2} + \sum_{k=1}^{\infty}\big(a_k\cos(kx) + b_k\sin(kx)\big) \tag{4.50}$$

který se nazývá *Fourierovou řadou.* Vzhledem k tomu, že uvedený systém tvoří úplný ortonormální systém na intervalu $\langle -\pi,\pi\rangle$, lze pro libovolnou funkci f(x) definovanou na tomto intervalu psát:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}\big(a_k\cos(kx) + b_k\sin(kx)\big) \tag{4.51}$$

kde a_k a b_k jsou tzv. *Fourierovy koeficienty* funkce $f(x)$. Jejich velikosti je možné určit z následujících definičních vztahů, odvozených metodou nejmenších čtverců (viz [Smi56]):

$$
\begin{aligned}
a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\cos(kx)\,dx \quad pro \quad k \ = 0,1,2,\cdots \\
b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\sin(kx)\,dx \quad pro \quad k \ = 1,2,\cdots
\end{aligned}
\tag{4.52}
$$

Jednotlivé vyzařovací funkce, které chceme aproximovat Fourierovými řadami, jsou funkce dvou proměnných, výše uvedené vztahy je proto třeba zobecnit. Fourierovu řadu (4.50) nelze jednoduše rozšířit pro funkce více proměnných, místo ní se proto používá ekvivalentní řada definovaná v komplexním oboru:

$$
\sum_{n=-\infty}^{\infty} c_n \cdot \mathrm{e}^{inx} \tag{4.53}
$$

kde i je imaginární jednotka. Vztah (4.51) pak přejde do tvaru:

$$
f(x) = \sum_{n=-\infty}^{\infty} c_n \cdot \mathrm{e}^{inx} \tag{4.54}
$$

a jednotlivé Fourierovy koeficienty c_n budou nyní definovány předpisem:

$$
c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cdot \mathrm{e}^{-inx}\,dx \tag{4.55}
$$

Poznamenejme ještě, že c_n jsou nyní komplexní čísla, jejich vztah k reálným koeficientům a_n a b_n použitých ve vztahu (4.51) je následující:

$$
\begin{aligned}
c_0 &= \frac{a0}{2} \\
c_n &= \frac{a_n - ib_n}{2} \\
c_{-n} &= \frac{a_n + ib_n}{2}
\end{aligned}
\tag{4.56}
$$

Řadu (4.53) lze velice přirozeně rozšířit pro libovolný počet proměnných, mluvíme pak o tzv. *násobných Fourierových řadách*. V případě funkcí dvou proměnných bude mít odpovídající řada tvar (viz [Smi56]):

$$
\sum_{k,l=-\infty}^{\infty} c_{kl} \cdot \mathrm{e}^{ikx} \cdot \mathrm{e}^{ily} = \sum_{k,l=-\infty}^{\infty} c_{kl} \cdot \mathrm{e}^{i(kx+ly)} \tag{4.57}
$$

a aproximace libovolné funkce $f(x,y)$ na intervalu $\langle -\pi, \pi\rangle \times \langle -\pi, \pi\rangle$ pak bude vypadat následovně:

$$f(x,y) = \sum_{k,l=-\infty}^{\infty} c_{kl} \cdot \mathrm{e}^{i(kx+ly)} \tag{4.58}$$

kde

$$c_{kl} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(x,y) \cdot \mathrm{e}^{-i(kx+ly)}\, dx\, dy \tag{4.59}$$

Výše uvedené aproximaci odpovídá následující systém funkcí:

$$S_F = \left\{\mathrm{e}^{i(kx+ly)}\right\}_{k,l=-\infty}^{\infty} \tag{4.60}$$

Tento systém je nekonečný, pro praktické aplikace je z něj třeba vybrat nějaký konečný (a značně malý) podsystém. Výběrem nejvhodnějšího podsystému se budeme zabývat v následující kapitole, nyní se na chvíli zastavme u výpočtu jednotlivých Fourierových koeficientů c_{kl}, pokud již tento podsystém známe. Je zřejmé, že ve většině aplikací nelze vztah (4.59) použít přímo, hodnoty jednotlivých koeficientů je nutné určovat numericky. Obsahuje-li daný systém N aproximačních funkcí, musíme pro každou vyzařovací funkci určit N koeficientů. Hodnoty jednotlivých koeficientů lze počítat pomocí tzv. *Fourierovy transformace*, časová náročnost speciálního algoritmu pro tuto transformaci (tzv. rychlé Fourierovy transformace – *Fast Fourier Transform*) je řádu $O(N \log N)$ pro jeden koeficient, pro všechny koeficienty jediné vyzařovací funkce je tedy zapotřebí řádově $O(N^2 \log N)$ operací sečítání, násobení a hlavně časově velice náročných výpočtů sinů a kosinů. Obě trigonometrické funkce lze sice tabelovat, pro $N = 100$ až 200 (což je horní hranice daná celkovými paměťovými nároky nutnými pro uložení Fourierových koeficientů všech vyzařovacích funkcí) však budou výpočty stále značně pomalé. Řešením je použití tzv. *Fourierovských procesorů*, tj. speciálních zařízení provádějících FFT hardwarově (a tedy řádově rychleji).

Vraťme se k vlastní realizaci vyzařovacích funkcí, zbývá totiž vyřešit ještě jeden problém. Pomocí vztahů (4.58) a (4.59) umíme aproximovat libovolné dvojrozměrné funkce, ovšem pouze na intervalu $\langle -\pi, \pi\rangle \times \langle -\pi, \pi\rangle$, námi použité vyzařovací funkce jsou však definované na intervalu $\langle 0, \pi/2\rangle \times \langle -\pi, \pi\rangle$ (viz kapitola 4.7.1). Samozřejmě by šlo před aproximací danou funkci na zbývající části intervalu $\langle -\pi, \pi\rangle \times \langle -\pi, \pi\rangle$ fiktivně dodefinovat (např. jako sudou funkci), v tomto případě by však docházelo ke zbytečné ztrátě přesnosti provedené aproximace (neboť některé koeficienty c_{kl} by byly stále nulové). Daleko výhodnější řešení je jednotlivé vyzařovací funkce "roztáhnout" na celý požadovaný interval, tj. použít následující transformaci souřadnic:

$$\begin{aligned} \alpha' &= 4 \cdot \alpha - \pi \\ \beta' &= \beta \end{aligned} \tag{4.61}$$

a pak provést aproximaci tak, jak byla popsána v předchozích odstavcích.

4.7.4 Výběr jednotlivých aproximačních funkcí

Systém S_F (viz definiční vztah (4.60)), který je použit pro aproximaci jednotlivých vyzařovacích funkcí, obsahuje obecně nekonečně mnoho funkcí. Pro praktické aplikace je třeba z tohoto systému vybrat konečný (a často značně malý) podsystém $S_{F'}$. Jeho velikost je většinou určena předem (např. v našem případě je shora omezena hranicí 100 – 200 funkcí), zbývá tedy pouze určit jednotlivé funkce, které použijeme pro vlastní aproximaci.

Velice častý postup pro obecné systémy funkcí je ten, že se zvolí prvních N funkcí daného systému, tj. v našem případě by byl vybrán podsystém $S_{F'} = \{\mathrm{e}^{i(kx+ly)}\}_{k,l=-M}^{M}$, kde číslo M je zvoleno tak, aby $|S_{F'}| \approx N$. Tímto způsobem by bylo možné zcela přesně aproximovat skoro konstantní a značně hladké funkce, hodně "zvlněné" funkce by však byly aproximovány značně nepřesně. Podívejme se, proč tomu tak je. Jak již bylo řečeno, systém S_F je dvojrozměrnou obdobou systému trigonometrických funkcí, proveďme proto své úvahy nejprve pro tento jednorozměrný systém.

Nechť je tedy pro aproximaci použito prvních N funkcí jednorozměrného trigonometrického systému, tedy $S_{F'} = \{1, \cos(kx), \sin(kx)\}_{k=1}^{M}$ pro vhodné M. Tento systém je ekvivalentní systému $S_{F''} = \{\sin(kx + \phi_k)\}_{k=1}^{M}$, který je pro naše úvahy výhodnější. Pokud si představíme grafy jednotlivých funkcí systému $S_{F''}$ na aproximačním intervalu $\langle -\pi, \pi \rangle$, jedná se o sinusovky se stále vyšším počtem vln, tj. o funkce s postupně se zvyšujícím kmitočtem. Systém $S_{F''}$ tvoří (pro $M = +\infty$) úplný ortogonální systém na $\langle -\pi, \pi \rangle$, každou funkci $f(x)$ na tomto intervalu lze tedy vyjádřit odpovídající trigonometrickou řadou:

$$f(x) = \sum_{k=0}^{\infty} A_k \cdot \sin(kx + \phi_k) \tag{4.62}$$

Danou funkci lze tedy "poskládat" z jednotlivých sinusovek s postupně vzrůstajícím kmitočtem, koeficienty A_k určují amplitudu a ϕ_k fázi daného kmitočtu k. Hladké a skoro konstantní funkce budou obsahovat pouze nízké kmitočty, čím víc bude funkce "divočejší", tím víc bude obsahovat vyšších kmitočtů. Pokud ze systému $S_{F''}$ vybereme pouze prvních N funkcí, budeme vlastně provádět aproximaci dané funkce $f(x)$ pouze v oblasti nízkých kmitočtů. Čím více bude dané funkce složitější, tím více bude obsahovat vyšších kmitočtů, a tím tedy bude hůř aproximovaná.

Výše popsaná situace je zcela analogická i v případě funkcí více proměnných, kdy sinusovky ve směru jednotlivých os nahrazují funkce $e^{ik\xi}$, kde ξ je odpovídající souřadnice. Pojem "kmitočet" nelze v tomto případě jednoznačně definovat, často se používá součet kmitočtů v jednotlivých osách.

Jak bylo ukázáno v předchozích odstavcích, volba prvních N funkcí daného aproximačního systému příliš výhodná, umožňuje sice zcela přesně aproximovat funkce nižších kmitočtů, pro vyšší kmitočty však selhává. V případě vyzařovacích funkcí nízkým kmitočtům odpovídá případ ideálního difúzního odrazu, čím je však odraz světelné energie na povrchu daného tělesa více směrový (tj. čím víc je daný povrch vyleštěný), tím více vysokofrekvenčních složek daná funkce obsahuje. Pokud tedy chceme dostatečně postihnout právě směr propagované energie, musíme do systému použitých aproximačních funkcí vybrat i funkce vyšších kmitočtů.

Jednou z možností je použít následující aproximační systém:

$$S_{F'} = \left\{ e^{i(kx+ly)} \right\}_{k,l=-M,\cdots,-K+1,-K,K,K+1,\cdots M}$$

tj. vybrat pouze funkce vyšších kmitočtů. Tento systém však neumožňuje postihnout funkce nižších kmitočtů, v případě jeho použití pro aproximaci vyzařovacích funkcí bychom se tak dostali na pozice ray-tracingu – značně přesně by byl postihnut úzce směrový odraz na vyleštěných površích, zcela nemožná by ale byla reprezentace povrchů s difúzním rozptylem. Navíc je zde ještě problém spočívající v určení dolního mezního kmitočtu K. Ani tento systém tedy nelze použít.

Použitý aproximační systém by měl umožňovat zhruba stejně přesnou aproximaci všech možných funkcí nezávisle na jejich tvaru, je tedy nutné, aby zahrnoval co největší rozsah kmitočtů. Celý možný rozsah kmitočtů lze pravidelně "rozsekat" způsobem, kdy vybereme funkce s kmitočty $c \cdot 2^i$, kde $i = 0, 1, \cdots, N$ a c je celočíselná konstanta určující maximální možný kmitočet ($f_{max} = c \cdot 2^N$). Aby bylo možné postihnout i ideální difúzi, je nutné k těmto kmitočtům přidat ještě kmitočet 0, tj. do daného systému přidat konstantní funkci 1. Nevýhoda uvedeného systému je ta, že kromě prvních dvou kmitočtů jsou všechny zbývající kmitočty buď sudé nebo liché, což způsobuje zbytečné nepřesnosti při aproximaci některých typů funkcí (viz libovolná kniha zabývající se teorií Fourierovy transformace). Navíc jsou jednotlivé kmitočty od sebe značně vzdáleny (každý je dvojnásobek předchozího).

Vzájemného přiblížení jednotlivých kmitočtů lze dosáhnout volbou systému kmitočtů K^i, $i = 0, 1, \cdots, N$, pro $K \in (1, 2)$. Nevýhodou tohoto postupu je ta, že jednotlivé kmitočty nevycházejí celočíselně a je nutné je zaokrouhlovat, neboť celočíselnost kmitočtů je nutnou podmínkou pro ortogonalitu daného aproximačního systému.

Ještě výhodnější rozložení jednotlivých kmitočtů dostaneme v případě, kdy použijeme tzv. *Fibonacciho posloupnost*, jejíž definice je následující:

$$\begin{aligned} F_0 &= 0 \\ F_1 &= 1 \\ F_{n+1} &= F_n + F_{n-1} \quad \text{pro} \quad n \geq 2 \end{aligned} \tag{4.63}$$

Pokud vynecháme duplicitní 1 z počátku posloupnosti, dostaneme téměř ideální posloupnost kmitočtů jednotlivých aproximačních funkcí. Prvních 16 členů této posloupnosti vypadá následovně:

$$F = \{0, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987\} \tag{4.64}$$

Uveďme si alespoň některé výhody posloupnosti (4.64):

- posloupnost do značné míry odpovídá řadě K^i, $i = 0, 1, \cdots, N$, kde K má hodnotu *zlatého řezu* (přibližně 1.618), jednotlivé členy jsou však celočíselné, není nutné je zaokrouhlovat
- rovnoměrné rozložení jednotlivých kmitočtů umožňuje stejně přesnou aproximaci jak nízkofrekvenčních, tak i vysokofrekvenčních funkcí
- v posloupnosti se pravidelně střídají sudé a liché členy (vždy dva liché a jeden sudý), nemůže tedy nastat situace, kdy je nějaký interval kmitočtů zastoupen pouze jedním typem prvků

Je zřejmé, že pro svoje výhody bude Fibonacciho posloupnost kmitočtů použita i pro výběr systému funkcí pro aproximaci jednotlivých vyzařovacích funkcí v rámci směrového radiačního výpočtu. Vzhledem k tomu, že vyzařovací funkce jsou funkce dvou proměnných, zbývá ještě určit, kolik aproximačních funkcí se použije pro aproximaci podél osy x a kolik podél osy y (připomeňme si, že úplný systém tvoří funkce e^{ikx} a e^{ily}, viz definiční vztah (4.60)). Je sice možné v obou směrech zvolit stejný počet funkcí, v našem případě, kdy jednotlivé vyzařovací funkce před aproximací "roztahujeme" podél osy x (viz vztah pro transformaci souřadnic (4.61)) však bude výhodnější pro tuto x použít méně aproximačních funkcí než pro osu y (jinak by docházelo ke zbytečně přesné aproximaci dané vyzařovací funkce podél odchylky od normály v porovnání se směrem ve smyslu odchylky od hlavního směru dané plochy). Jelikož souřadnici x "natahujeme" $4\times$, kmitočet aproximované funkce se v tomto směru sníží také $4\times$. Pokud tedy pro osu y zvolíme prvních M kmitočtů z Fibonacciovy řady (4.64), bude pro osu x stačit o tři nejvyšší kmitočty méně (neboť $F_{M-3} \approx \frac{F_M}{4}$). Celkově tedy bude použitý systém aproximačních funkcí S_F vypadat následovně:

$$S_F = \left\{ \mathrm{e}^{iF_k x} \right\}_{k=-M}^{M} \times \left\{ \mathrm{e}^{iF_l y} \right\}_{l=-M+3}^{M-3} \tag{4.65}$$

kde i je imaginární jednotka, F_k je k-tý člen Fibonacciho posloupnosti (4.64) (pro $k < 0$ je $F_k = -F_{|k|}$) a $M > 3$ je uživatelem volený počet funkcí použitých pro aproximaci podél osy y. Velikost tohoto systému závisí na hodnotě M, pro dané M platí:

$$N = |S_F| = (2M-1) \cdot (2 \cdot (M-3) - 1) = (2M-1) \cdot (2M-7) \tag{4.66}$$

4.7.5 Operace s vyzařovacími funkcemi

Jak již bylo řečeno, aproximace systémem trigonometrických funkcí $_F$ nám umožňuje jednotlivé vyzařovací funkce vyjádřit vztahem $\Phi \approx \sum_{i=1}^{N} c_i f_i$, kde f_i je i-tá funkce daného aproximačního systému a číslo c_i je její váha. Vzhledem k tomu, že systém S_F je explicitně znám, je pak možné každou vyzařovací funkci Φ uchovávat ve formě N-rozměrného číselného vektoru vah $\widetilde{\Phi} = (c_1, c_2, \cdots, c_N)$. Pokud použijeme tuto náhradu vyzařovacích funkcí Φ odpovídajícími vektory vah $\widetilde{\Phi}$, musíme nad vektory vah definovat operace, které bychom normálně prováděli přímo s původními funkcemi. V případě radiačního výpočtu se jedná o tyto operace:

- reprezentace vyzařovací funkce odpovídajícím vektorem vah, tedy přechod $\Phi \rightarrow \widetilde{\Phi}$
- zjištění hodnoty vyzařovací funkce v daném směru, tj. určení $\Phi(\alpha, \beta)$ (ve své podstatě opak předchozí operace)
- sečtení dvou vyzařovacích funkcí $\Phi = \Phi_1 + \Phi_2$
- vynásobení vyzařovací funkce konstantou $\Phi' = c \cdot \Phi$
- rotace vyzařovací funkce, tj. určení funkce Φ' tak, aby platil následující vztah: $\Phi'(\alpha, \beta) = \Phi(\alpha + \Delta\alpha, \beta + \Delta\beta)$

První dvě operace (tj. přechody mezi původní funkcí a jí odpovídajícím vektorem vah) byly již vlastně definovány. Přechodu $\Phi \rightarrow \widetilde{\Phi}$ odpovídá výpočet jednotlivých vah c_i Fourierovou transformací (numericky ze vztahu (4.59)), přechodu $\widetilde{\Phi} \rightarrow \Phi$ pak výpočet hodnoty $\Phi(\alpha, \beta) = \sum_{i=1}^{N} c_i \cdot f_i(\alpha, \beta)$ pomocí tzv. inverzní Fourierovy transformace. Obě operace jsou časově značně náročné (hlavně v případě, kdy aproximační systém S_F obsahuje větší množství funkcí), výrazně jej lze urychlit použitím Fourierovských procesorů.

Realizace operací součtu dvou vyzařovacích funkcí a vynásobení dané funkce konstantou je velice jednoduchá, stačí požadovanou operaci provést postupně po jednotlivých složkách odpovídajících váhových vektorů. Součtu vyzařovacích funkcí tedy odpovídá součet jejich váhových vektorů:

$$\left.\begin{aligned} \Phi &= \Phi_1 + \Phi_2 \\ \widetilde{\Phi}_1 &= (c_1^1, \cdots, c_N^1) \\ \widetilde{\Phi}_2 &= (c_1^2, \cdots, c_N^2) \end{aligned}\right\} \Rightarrow \begin{aligned} &\widetilde{\Phi} = (c_1, \cdots, c_N) \\ &\text{kde} \;\; c_i = c_i^1 + c_i^2 \end{aligned} \tag{4.67}$$

a obdobná je i situace pro vynásobení vyzařovací funkce konstantou:

$$\left.\begin{aligned} \Phi' &= c \cdot \Phi \\ \widetilde{\Phi} &= (c_1, \cdots, c_N) \\ c &\in C \end{aligned}\right\} \Rightarrow \begin{aligned} &\widetilde{\Phi}' = (c_1', \cdots, c_N') \\ &\text{kde} \;\; c_i' = c \cdot c_i' \end{aligned} \tag{4.68}$$

Složitější je pouze realizace operace rotace vyzařovacích funkcí. Pomocí této operace se v případě použití alternativních distribučních funkcí určuje elementární vyzařovací funkce objektu generovaná dopadajícím paprskem, blíže viz kapitola 4.8. Nechť je tedy dána funkce $f(x,y)$, s maticí vah $\widetilde{f}(x,y) = (c_{00}, \cdots, c_{KL})$, tj. $f(x,y) = \sum\limits_k \sum\limits_l c_{kl} \cdot \mathrm{e}^{i(kx+ly)}$. Pro rotovanou funkci $f'(x,y)$ lze pak psát $f'(x,y) = f(x+\alpha, y+\beta) = \sum\limits_k \sum\limits_l c_{kl} \cdot \mathrm{e}^{i(k(x+\alpha)+l(y+\beta))}$. Z uvedeného vztahu rozpisem dostáváme $f'(x,y) = \sum\limits_k \sum\limits_l c_{kl} \cdot \mathrm{e}^{i(kx+k\alpha+ly+l\beta)} = \sum\limits_k \sum\limits_l c_{kl} \cdot \mathrm{e}^{i(k\alpha+l\beta)} \cdot \mathrm{e}^{i(kx+ly)}$. Nechť rotované funkci $f'(x,y)$ odpovídá matice vah $\widetilde{f}'(x,y) = (c_{00}', \cdots, c_{KL}')$, tedy $f'(x,y) = \sum\limits_k \sum\limits_l c_{kl}' \cdot \mathrm{e}^{i(kx+ly)}$. Porovnáním obou vztahů pro $f'(x,y)$ získáme výsledný vztah $c_{kl}' = c_{kl} \cdot \mathrm{e}^{i(k\alpha+l\beta)}$. Celkově lze tedy rotaci vyzařovací funkce $\Phi(\alpha,\beta)$ realizovat následovně:

$$\left.\begin{aligned} \Phi'(\alpha,\beta) &= \Phi(\alpha+\Delta\alpha, \beta+\Delta\beta) \\ \widetilde{\Phi} &= (c_{00}, \cdots, c_{KL}) \end{aligned}\right\} \Rightarrow \begin{aligned} &\widetilde{\Phi}' = (c_{00}', \cdots, c_{KL}') \\ &\text{kde} \;\; c_{kl}' = c_{kl} \cdot \mathrm{e}^{i(k\Delta\alpha+l\Delta\beta)} \end{aligned} \tag{4.69}$$

4.7.6 Shrnutí, porovnání s dosavadními postupy

V předchozích částech této kapitoly byla kompletně popsána realizace jednotlivých vyzařovacích funkcí, nyní si ji nejprve shrneme do několika bodů a poté použitý postup porovnáme s dosud publikovanými metodami.

Nejprve tedy shrnutí celé realizace vyzařovacích funkcí:

1. s každým elementárním objektem zobrazované scény (normálně se jedná o plošné elementy, v našem případě pak o jednotlivé vrcholy) je spojena jeho vyzařovací funkce Φ, tato funkce popisuje směrové rozložení světelné energie, kterou daný objekt vysílá do svého okolí
2. každá vyzařovací funkce Φ se skládá celkem z 12 elementárních funkcí Φ_{elem}, jedná se o následující funkce: $(celkov\acute{a}\ a\ nevyst\check{r}elen\acute{a}\ energie) \times (vn\check{e}j\check{s}\acute{\imath}\ a\ vnit\check{r}n\acute{\imath}\ funkce) \times (slo\check{z}ka\ R,\ G\ a\ B)$
3. elementární vyzařovací funkce $\Phi_{elem}(\alpha, \beta)$ jsou funkce dvou proměnných definované na intervalu $\langle 0, \pi/2\rangle \times \langle -\pi, \pi\rangle\rangle$, jejich hodnoty jsou z intervalu $\langle 0, 1\rangle$
4. každá elementární vyzařovací funkce je aproximována řadou předem zvolených aproximačních funkcí, parametry použitého aproximačního systému jsou shodné pro všechny vyzařovací funkce
5. při aproximaci je použito kritérium nejmenších čtverců
6. aproximační systém $S_{F'}$ tvoří podmnožina dvojrozměrného trigonometrického systému $S_F = \left\{ \mathrm{e}^{i(kx+ly)} \right\}_{k,l=-\infty}^{\infty}$; systém S_F tvoří úplný ortogonální systém na intervalu $\langle -\pi, \pi\rangle \times \langle -\pi, \pi\rangle$ a lze jej snadno ortonormalizovat
7. v systému S_F jsou použity funkce, jejichž kmitočty odpovídají jednotlivým členům Fibonacciho řady (4.64)
8. aproximační systém $S_{F'}$ je definován následovně:

   $$S_{F'} = \left\{ \mathrm{e}^{i(F_k x + F_l y)} \right\}_{\substack{k=-M+3,\cdots,M-3 \\ l=-M,\cdots,M}}$$

   kde F_k je k-tý člen Fibonacciho posloupnosti (pro $k < 0$ je $F_k = -F_{|k|}$). Číslo $M > 3$ volí uživatel, pro dané M obsahuje systém celkem $N = (2M-1)\cdot(2M-7)$ funkcí
9. přesnost aproximace je skoro stejná v obou souřadnicích
10. každou elementární vyzařovací funkci Φ_{elem} lze vyjádřit jako:

    $$\Phi_{elem} = \sum_{k=-M+3}^{M-3} \sum_{l=-M}^{M} c_{kl} \cdot \mathrm{e}^{i(F_k x' + F_l y)}$$

    kde $x' = 4x - \pi$
11. váhové koeficienty c_{kl} se určují numericky (Fourierovou transformací) ze vztahu:

$$c_{kl} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(x', y) \cdot \mathrm{e}^{-i(F_k x' + F_l y)} \; dx \, dy$$

kde opět $x' = 4x - \pi$

12. váhové koeficienty c_{kl} jsou symetrické podle počátku, tedy $c_{-k,-l} = c_{k,l}$
13. každá elementární vyzařovací funkce Φ_{elem} je reprezentována maticí vah $\widetilde{\Phi}_{elem} = \{c_{kl}\}_{\substack{k=-M+3,\cdots,M-3 \\ l=-M,\cdots,M}}$, jedná se o matici komplexních čísel rozměrů $(2M-1) \times (2M-7)$ funkcí; tuto matici lze v případě potřeby linearizovat na vektor vah
14. s vektorem (resp. maticí) vah $\widehat{\Phi}_{elem}$ dané elementární vyzařovací funkce Φ_{elem} lze efektivně provádět všechny požadované operace (převod funkce $\rightarrow$ vektor vah, zjištění hodnoty v daném směru, sečtení dvou funkcí, vynásobení funkce konstantou, její rotace)
15. pro uložení celé vyzařovací funkce Φ daného elementu je zapotřebí celkem $96 \cdot M \cdot (2M-7)$ bytů paměti
16. nejvhodnější hodnoty pro M jsou 9 nebo 11

Ještě než se dostaneme k porovnání popisované metody s ostatními dosud publikovanými metodami realizace vyzařovacích funkcí, podívejme se, jak byl získán vztah pro potřebné množství paměti uvedený v bodu 15. Systém S_F obsahuje celkem $(2M-1) \cdot (2M-7)$ funkcí (bod 13), pro reprezentaci každé elementární funkce potřebujeme tedy stejné množství komplexních koeficientů. Pokud zvolíme M liché a využijeme symetrie váhových koeficientů (bod 12), postačí uchovávat pouze $M \cdot (2M-7)$ komplexních koeficientů pro každou elementární funkci. Pro celou vyzařovací funkci potřebujeme tedy $12 \cdot M \cdot (2M-7)$ komplexních koeficientů, což je $24 \cdot M \cdot (2M-7)$ reálných čísel. Použijeme-li reálná čísla délky 4 byty (typ $float$), bude pro uchování jedné vyzařovací funkce celkem zapotřebí $96 \cdot M \cdot (2M-7)$ bytů paměti, jak bylo uvedeno výše.

Hodnotu $M > 3$ ovlivňující počet funkcí použitých pro aproximaci (viz bod 8) volí uživatel. Je vhodné volit liché číslo, pak lze při použití symetrie váhových koeficientů ukládat pouze polovinu všech vah. Otázkou zůstává, jak velké M zvolit. Podívejme se na tabulku 4.1, ve které jsou shrnuty vlastnosti aproximace vyzařovacích funkcí v závislosti na velikosti M. První sloupec obsahuje hodnotu M, ve druhém je odpovídající počet aproximačních funkcí N, v následujícím je obsažen maximální použitý kmitočet F_{max} a v posledním pak paměťové nároky pro realizaci vyzařovací funkce jednoho elementu scény.

Jak již bylo řečeno v bodu 16, nejvhodnější hodnoty M jsou 9 nebo 11. Matice vah má při postačujícím maximálním kmitočtu rozměry dostatečně malé na